bookboon
LEARNING

CMOS Analog IC Design: Fundamentals

Erik Bruun

ERIK BRUUN

CMOS ANALOG IC DESIGN
FUNDAMENTALS

CMOS Analog IC Design: Fundamentals
3rd edition

ISBN 978-87-403-4247-5

CONTENTS

Preface

This book has been written with the specific goal to serve as a textbook for a first course in CMOS analog integrated circuit design. It is intended for electronics engineering students at bachelor level who have followed basic courses in mathematics, physics, circuit theory, electronics and signal processing.

A distinct feature of this book is the emphasis on the interaction between analytical methods and simulation methods. Traditionally, the teaching of circuit design is based on mathematical models of circuit elements and on basic circuit theorems. Often, simulations are just referred to as additional examples. In this book, simulations using the universally accepted program SPICE (Simulation Program with Integrated Circuit Emphasis) are used throughout. The simulations are integrated into the text, both in order to illustrate and verify the mathematically based theory and in order to develop design methods using the simulator for more accurate design of integrated circuits.

The material presented in this book has been adapted from material used by the author for many years of teaching an introductory one-semester course (5 ECTS credits) in CMOS analog integrated circuit design at the Technical University of Denmark.

Chapter 1 is an introduction to the subject, outlining the basic method for integrated circuit design with close interaction between hand calculations and simulations, and providing a motivation for the study of analog integrated circuit design.

Chapter 2 is a recapitulation of some of the basic theory which you are expected to know already. This recapitulation is used for setting the scene for the following chapters which form the bulk of the course.

Chapter 3 presents the models used for MOS transistors in the book. For analytical purposes, hand calculations, the Shichman-Hodges model is used. For simulations, both the Shichman-Hodges model and BSIM transistor models are presented.

Chapter 4 describes the basic gain stages in analog CMOS design, i.e., the common-source stage, the common-drain stage, the common-gate stage and the differential pair. A thorough review of the concept of small-signal analysis is given with several examples showing how to build a small-signal model of a gain stage from the transistor schematic of the stage.

Chapter 5 is about multistage amplifiers. It is shown how a simple opamp can be designed by combining the stages introduced in Chapter 3.

Chapter 6 is a brief review of feedback theory. The use of feedback to improve many system parameters in an amplifier is explained. Additionally, conditions for stability in a feedback system are presented. The chapter is also included in order to be able to present a design of the two-stage opamp, considering that the main application of an opamp is as a gain element in a feedback system where the stability of the system is a critical parameter.

Chapter 7 reviews the two-stage opamp and presents a design where the issue of stability in a feedback system is a major consideration.

Chapter 8 gives a brief presentation of some circuits used for biasing of the amplifier circuits presented in the previous chapter.

Finally, Chapter 9 summarizes the essentials from the previous chapters. It is intended as a quick reference to the most important equations and conclusions from the book.

All chapters (except Chapter 1 and Chapter 9) include end-of-chapter problems, and answers to the questions are included in Appendix B. Fully worked-out solutions can be found in the companion book:

Bruun, E., *CMOS Analog IC Design: Problems and Solutions*, bookboon.
Available from: http://bookboon.com/en/cmos-analog-ic-design-problems-and-solutions-ebook

Also a selection of multiple-choice problems is included after each chapter as a quick practice test. This is meant as a quick tool for testing if the learning objectives defined in the beginning of the chapter have been accomplished. However, it is not a replacement for the end-of-chapter problems. The practical experience obtained by solving a substantial number of problems is essential for the learning process in a field such as integrated circuit design. More practical experience can be achieved from a supplementary book comprising a selection of problems covering many aspects of the design methods from the present book:

Bruun, E. & Jørgensen, I., 2022, *CMOS Analog IC Design: Learning by Problem Solving*,
bookboon. Available from:
http://bookboon.com/en/cmos-analog-ic-design-ebook

The material presented in the present book has been found suitable for a course comprising 12 lectures (with problem solving sessions) and (optionally) an additional final wrap-up lecture. The duration of each lecture is about 2 hours including a 15 minutes break.

The lectures may be organized as follows:

Lecture 1: Introduction and recapitulation of basic theory (Chapters 1 and 2).

Lecture 2: MOS transistor basics, Shichman-Hodges transistor model (Chapter 3.1 – 3.4).

Lecture 3: Small-signal transistor models and advanced transistor models (Chapter 3.5 – 3.7).

Lecture 4: Common-source stage and common-drain stage at low frequencies (Chapter 4.1 – 4.2).

Lecture 5: Common-gate stage and cascode stage at low frequencies (Chapter 4.3).

Lecture 6: The differential gain stage (Chapter 4.4).

Lecture 7: Frequency response of the basic gain stages (Chapter 4.5).

Lecture 8: Introduction to multistage amplifies (Chapter 5).

Lecture 9: Feedback theory, basic properties of systems with feedback (Chapter 6.1 – 6.4).

Lecture 10: Stability in feedback systems (Chapter 6.5 – 6.7).

Lecture 11: Two-stage opamp design example (Chapter 7).

Lecture 12: Biasing circuits (Chapter 8).

Prerequisites: This book is aimed at a fundamental level in analog CMOS circuit design. The prerequisites needed are only a general background in mathematics and physics and a background in the fundamentals of signal processing, circuit theory and electronics. There are numerous courses and textbooks which will provide an adequate background. Without this being an exhaustive list, it can be mentioned that the textbooks used in the introductory courses at the Technical University of Denmark in signal processing, circuit theory and electronics are the following:

Lathi, BP. 2009, *Signal Processing & Linear Systems*, International Second Edition, Oxford University Press, New York, USA.

Hambley, AR. 2018, *Electrical Engineering, Principles and Applications*, Seventh Edition, Pearson Education Ltd., Harlow, UK.

As already mentioned, an important feature of the present book is the extensive use of SPICE for verification of the mathematically based theory and for providing designs with properties which are closer to the properties obtained through actual experimental circuits. A very useful book for learning about LTspice (the version of SPICE used in this book) is the following, part of which is also used in an introductory course in circuit theory and electronics:

Bruun, E. 2020, *CMOS Integrated Circuit Simulation with LTspice*, Third Edition, bookboon.
Available from: http://bookboon.com/en/cmos-integrated-circuit-simulation-with-ltspice-ebook

If you are not already familiar with LTspice, it is highly recommended to complete Tutorials 1 and 2 from this book concurrently with reading Chapter 2 in the present book.

Further reading: Hopefully, some readers of the present book find the subject of CMOS analog integrated circuit design so interesting that they wish to study the subject further. There are many excellent textbooks about analog CMOS design covering both the topics from the present book in more depth and covering many other subjects such as noise, data converters and filters. Also, details of semiconductor physics and VLSI technology are not dealt with in the present book. For such subjects, the reader is referred to more advanced textbooks, some of which are referred to in the following chapters. Again, without this being an exhaustive list, the following books are used in advanced courses at the Technical University of Denmark:

Chan Carusone, T., Johns, D. & Martin, K. 2012, *Analog Integrated Circuit Design*, Second Edition, International Student Version, John Wiley & Sons, Inc., Hoboken, USA.

Plummer, JD., Deal, M. & Griffin, PD. 2000, *Silicon VLSI Technology: Fundamentals, Practice, and Modeling*, First Edition, Pearson Education Ltd., Harlow, UK.

Acknowledgements: The author would like to acknowledge Dr. Lars Stenberg, former director of IC research and development at GN Hearing, for a careful review of the book, providing many suggestions for improving the text. Also, a particular acknowledgement goes to my colleagues Ivan Jørgensen, Pere Llimós Muntal and Nicolai Jerram Dahl who reviewed the entire manuscript of the book.

Erik Bruun
Department of Electrical Engineering, Technical University of Denmark

Chapter 1 – Introduction

This chapter gives an introduction to the subject of design of analog integrated circuits (ICs) in CMOS technology. A short background explaining the reason for focusing on CMOS technology is given, and the design methodology presented in this book is motivated. After having studied the chapter, you should be able to

- describe a few basic properties of CMOS technology.
- describe Moore's law.
- know the orders of magnitude of device geometries in a modern CMOS circuit.
- describe the basic steps in the development of a CMOS integrated circuit.
- explain the impact of device scaling on digital circuits and on analog circuits.
- describe and explain a generic model for the design of an electronic system.

1.1 CMOS technology

CMOS technology is the dominant fabrication technology used for today's electronic systems including computers, radios, TV sets, smartphones and also electronics embedded into other systems such as cars and household appliances.

CMOS is short for complementary metal-oxide-semiconductor and it is a technology for fabricating integrated circuits using both p-channel field-effect transistors (PMOS transistors) and n-channel field-effect transistors (NMOS transistors), the two dominant device types in most integrated circuits. In Chapter 3, we investigate the properties of MOS transistors in detail. The majority of CMOS integrated circuits are fabricated using silicon as the basic material.

The combination of PMOS and NMOS transistors makes it very easy to design digital circuit functions, and the transistors for digital circuit functions scale down in size as the fabrication technology improves and the manufacturers are able to control progressively smaller dimensions. The downscaling of device sizes implies that still more functions can be put into a single integrated circuit while maintaining or improving speed/power ratio for digital circuitry. In the beginning of the IC era (mid 1960s), device dimensions were counted in tens of μm. Today, device dimensions are counted in nm, i.e., a reduction in linear device size by more than three orders of magnitude. Contributing to the development towards more devices in a single IC is also an increase in the die size of an integrated circuit which can be manufactured with a reasonable yield (percentage of functional dies from a fabrication lot). In the 1960s, die sizes were counted in mm^2. Today's ICs reach die sizes of more than 100 mm^2, allowing tens of billions of devices to be integrated in a single IC (Altera's 30 billion transistor FPGA 2015).

This scaling of semiconductor devices was predicted by Gordon Moore, one of the founders of Intel, already in 1965. In a famous paper published in *Electronics* (Moore 1965), he presented a graph showing that the number of components in an IC could be expected to increase by a factor of two per year, at least until 1975. This development has proved to be even more long-lived than originally foreseen by Gordon Moore, and in 1975, he revised the prediction and modified the rate of development to a factor of two in device count per die for every second year (Moore 1975), a prediction known as Moore's law.

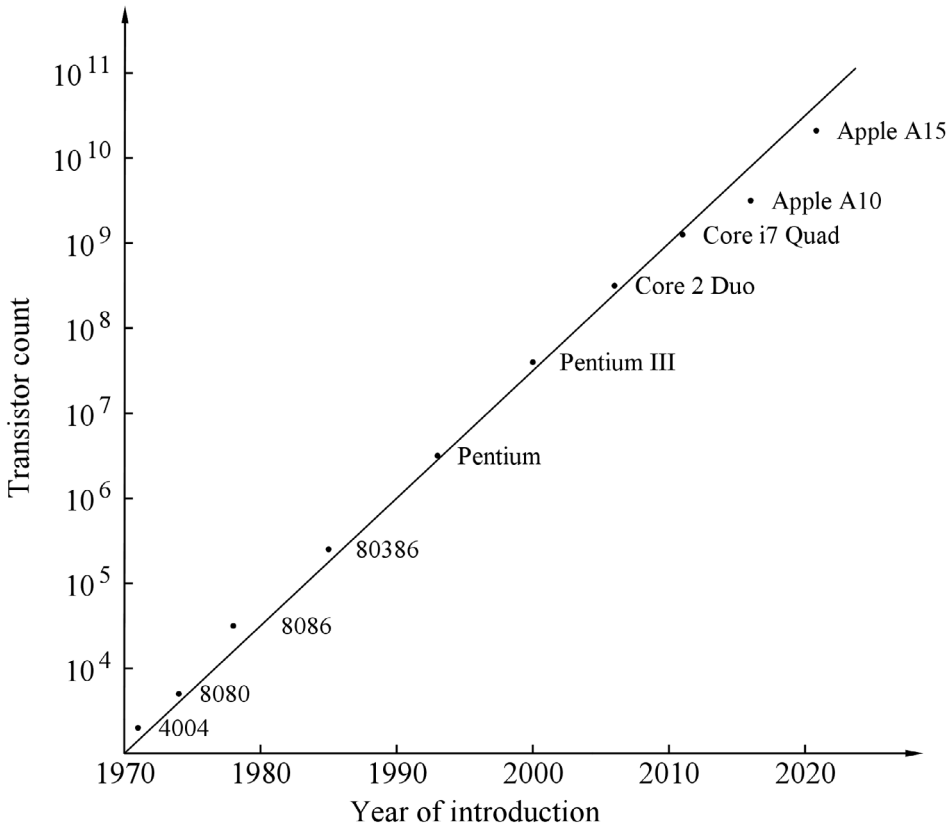


Figure 1.1: Development in number of transistors per chip versus year of introduction. Adapted from Transistor count (2021).

This rate has been maintained for many years as shown in Fig. 1.1 for transistor count for selected microprocessors. The development is closely followed by IEEE IRDS, an organization which monitors the development in the semiconductor industry and regularly issues reports concerning this development (International Roadmap for Devices and Systems (IRDS^{TM}) 2021).

1.2 Why analog circuit design?

The development towards smaller device dimensions is particularly beneficial for digital circuit functions which can be implemented using minimum-size devices. Closely linked to the downscaling of device dimensions is a downscaling of the supply voltage. This is advantageous for digital circuits as it leads to reduced power consumption, but it provides a challenge for analog circuit design because it means reduced signal swing without a corresponding reduction in noise voltages. Hence, analog circuit functions often need device geometries larger than the minimum size offered by the technology in order to achieve a useful performance with respect to noise properties, signal voltage swings, power consumption, etc. A consequence of this is that the growth rate of the complexity of digital systems is higher than that of analog systems, so today's electronic systems are mostly digital.

However, digital systems often need an analog interface to the physical world, for instance in the form of a transducer interface to a microphone, a loudspeaker, an ultrasound transponder, a light-sensitive device, a pressure sensor, or some other type of sensing device. Therefore, even though the growth of complexity in a digital system is higher than that in an analog system and an increasing number of system functions are handled by digital circuits, analog circuits remain an essential part of many electronic systems.

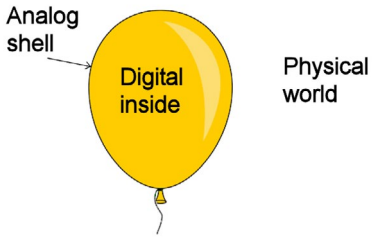


Figure 1.2: The analog balloon.

The situation may be illustrated as a balloon where the inside represents the digital functions, increasing in size as air is blown into the balloon, and the shell of the balloon, representing the analog interface, gets thinner and thinner as the balloon expands but remains to be an essential part in which a flaw would cause the balloon to collapse. If too much air is blown into the balloon, the analog shell may no longer be able to withstand the pressure and the balloon will explode. To avoid this situation, new developments in analog circuit design are needed for a stronger analog shell, including new fabrication technologies, new device types and new circuit techniques. This field of research and development is sometimes denoted 'More-than-Moore' (Arden et al. 2010). Thus, there is good reason to study the design of CMOS analog integrated circuits. This book is an introduction to this subject.

1.3 Design methodology

The design of an integrated circuit often follows a design flow as illustrated in Fig. 1.3. There are five major steps involved: (1) The schematic design using hand calculation and simplified analytical models, (2) the schematic simulation and design iteration using computer simulation, (3) the layout and layout verification using computer tools, (4) the fabrication of experimental test circuits, and (5) the experimental verification using measurements on the test circuits from the fabrication. After each of the steps in the design flow, it may be necessary to return to an earlier design step in order to correct deviations from the design goal. It may even be necessary to go back to start and revise the specifications if they turn out to be impossible to fulfill.

One of the characteristics of integrated circuit design is that the fabrication of test circuits is very time consuming and very expensive due to the fabrication of the IC wafers. This fabrication is carried out by specialized companies, and the normal turn-around time is counted in months while the cost is counted in thousands of dollars, even when achieving favorable university prices. For this reason, in IC design there is always a strong focus on achieving a first-time-right design, and this means that the verification steps in the design process (i.e., simulation, design rule check, layout versus schematic check, parasitic extraction, post-layout simulation) are very important, and a circuit design is not submitted for fabrication until these steps have been completed to perfection. Another reason for the extensive use of simulation is that debugging a circuit after fabrication is often much more difficult than debugging by simulation.

The phases illustrated in Fig. 1.3 can be treated in a generic development model as shown in Fig. 1.4 which is applicable to the development of electronic systems in general (Jørgensen 2015). In a commercial setting, it is absolutely vital to complete all three phases of development prior to the introduction of new products. In a university environment, all of the phases in the design flow cannot be covered in a single, one-semester university course.

This book is aimed at an introductory course, and for this, the focus is on the two initial steps in the design flow, the circuit design using hand calculations on the basis of a theoretical foundation and the circuit design aided by the use of computer simulation.

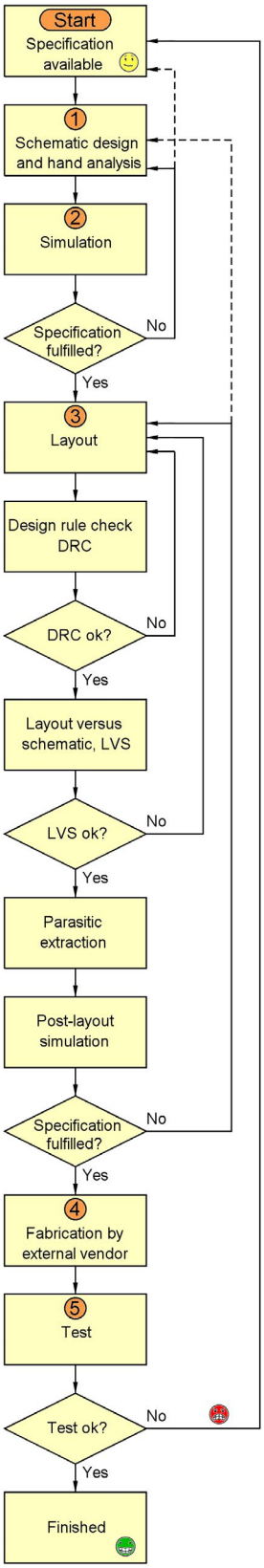


Figure 1.3: Development flow for IC design.

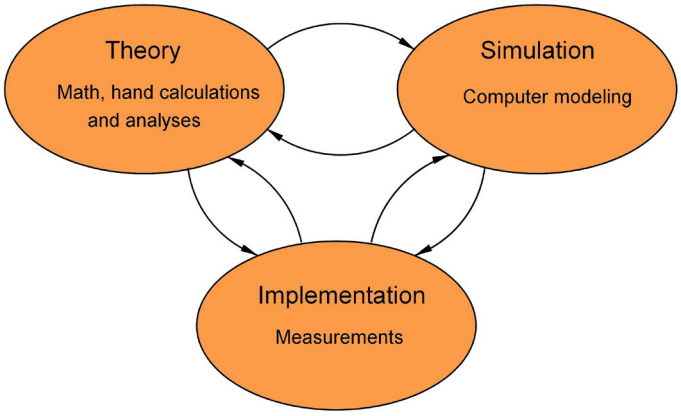


Figure 1.4: A generic model for the development of an electronic system.

The two phases and the interaction between them are visualized in Fig. 1.5. The simulation serves not only as a design tool but also as a tool for the designer's self-evaluation of the design. As an introduction to the subject, the book covers only some basic circuit functions and circuit blocks which appear in almost any analog design, the fundamental gain stages, a basic design of an operational amplifier, and some additional circuits for providing bias currents and voltages to the analog circuit blocks.

Along the way, we review some of the fundamental issues which are required for the analytical design, including a brief repetition of basic principles for signals, electronic devices, circuit theorems and circuit analysis (Chapter 2). Also, an introduction to feedback theory is included (Chapter 6), and throughout the book, the analytical methods are complemented by simulations using the simulation program SPICE (Simulation Program with Integrated Circuit Emphasis) which is the de facto standard for simulating analog integrated circuits.

The SPICE simulator used throughout this book is LTspice. You may use alternative circuit simulators but LTspice has been chosen here because it is easily available from Analog Devices in a free version, (http://www.analog.com/en/design-center/design-tools-and-calculators/ltspice-simulator.html), and also it is fairly easy to learn how to use it for CMOS design. Contrary to most commercial IC design tools, it runs on both Windows PCs and Macs, and it does not require any sort of licenses. LTspice is optimized for simulating systems using integrated circuits from Analog Devices, but it is a general SPICE simulator, and it is also perfectly suited as a design tool for simulating CMOS integrated circuits. A free guide to the use of LTspice specifically for CMOS integrated circuits is available on the web (Bruun 2020).

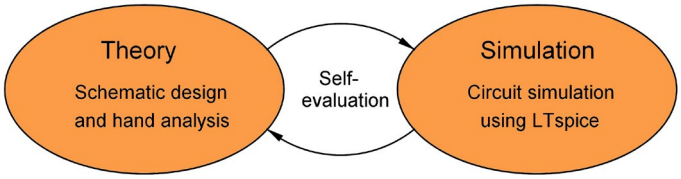


Figure 1.5: The design phases illustrated in this book.

References

Altera's 30 billion transistor FPGA 2015. *Gazettabyte*, 28 June 2015.
Retrieved from http://www.gazettabyte.com/home/2015/6/28/alteras-30-billion-transistor-fpga.html

Arden, W., Brillouët, M., Cogez P., Graef, M., Huizing, B. & Mahnkopf, R. (Editors) 2010, *"More-than-Moore" White Paper*, ITRS.
Available from: http://www.itrs2.net/uploads/4/9/7/7/49775221/irc-itrs-mtm-v2_3.pdf

Bruun, E. 2020, *CMOS Integrated Circuit Simulation with LTspice*, Third Edition, bookboon.
Available from: http://bookboon.com/en/cmos-integrated-circuit-simulation-with-ltspice-ebook

International Roadmap for Devices and Systems (IRDS™) 2021 Edition, *IEEE IRDS*.
Retrieved from https://irds.ieee.org/editions/2021

Jørgensen, IHH. 2015, 'Three Step Development Model used for Active Learning in Electrical Engineering', *Proceedings of the 43rd Annual SEFI Conference*, Orléans, France.

Moore, GE. 1965, 'Cramming more components onto integrated circuits', *Electronics*, vol. 38, no. 4, pp. 114-117, 19 April 1965.

Moore, GE. 1975, 'Progress in Digital Integrated Electronics', *Technical Digest. International Electron Devices Meeting 1975*, IEEE pp. 11-13.

Transistor count 2021, *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Transistor_count

Multiple-choice test

1. Complete the following statements by selecting the appropriate continuation from the table below.

A:	CMOS is an abbreviation for ...
B:	Modern CMOS processes normally use minimum device dimensions in range ...
C:	The most commonly used material for CMOS technology is ...
D:	Gordon Moore presented his first prediction of semiconductor device scaling in ...
E:	According to Moore's law, the number of devices per die increases ...
F:	Device downscaling is mostly beneficial for ...
G:	In analog integrated circuit design, the schematic design and analysis is verified through ...
H:	SPICE is an abbreviation for ...

Continuation:

1:	circuit simulation.
2:	1 - 100 mm.
3:	breadboarding.
4:	silicon.
5:	germanium.
6:	1947.
7:	linearly with time.
8:	1 - 100 μm.
9:	analog circuit.
10:	digital circuits.
11:	complementary metal-oxide-semiconductor.
12:	carbon.
13:	1965.
14:	Simulation Program with Integrated Circuit Emphasis.
15:	current-mode systems.
16:	exponentially with time.
17:	1 - 1000 nm.
18:	logarithmically with time.
19:	power electronics.

2. According to Moore's law from 1975, what is the expected factor of increase in device count per die per 10 years?

   A: 10
   B: 32
   C: 1024

3. Assuming a constant die area and a device count increase by a factor of two in two years, what is the scaling factor per year of the linear device dimensions?

   A: 0.84
   B: 0.71
   C: 0.50

4. Assuming a linear scaling factor of 0.84 per year, what would be a typical device dimension in 2024 if 28 nm is a typical device dimension in 2019?

   A: 7 nm
   B: 12 nm
   C: 20 nm

Chapter 2 – Basic Concepts

This chapter provides some of the information which you are expected to have as a background knowledge from previous studies. Here, the emphasis is on issues related to CMOS integrated circuit design. After having studied the chapter, you should be able to

- explain and describe analog signals in the time domain and in the frequency domain.
- explain and use device relations for passive components (resistors, capacitors and inductors).
- explain and use device relations for ideal voltage sources and current sources, both independent sources and controlled sources.
- explain and use basic circuit theorems (Kirchhoff's laws, Thévenin equivalents and Norton equivalents, superposition).
- calculate and simulate the step response in the time domain for simple *RC* networks.
- calculate and simulate transfer functions in the frequency domain for simple *RC* networks.
- explain the concepts of small-signal analysis and large-signal analysis and derive small-signal parameters from large-signal characteristics.

2.1 Signals

MOS integrated circuits are designed to process voltages and currents in an electronic system. Examples include the processing of information in a computer system, acoustic signals in an audio system, and power in a system for controlling LED lamps. Thus, one of the fundamental issues is the mathematical representation of voltages and currents or – more generally – signals. We start by considering a classification of signals.

Analog signals. Analog signals are signals with a value varying continuously in time. An example is a voltage which is a function of the time t. Figure 2.1 shows a sinusoidal voltage signal. It is described by its amplitude V_a and its angular frequency ω using the relation

$$v_a(t) = V_a \sin(\omega t) \tag{2.1}$$

It is a periodic signal with the period $T = 2\pi/\omega = 1/f$ where f is the frequency. In general, a sinusoidal signal is described by $v_a(t) = V_a \cos(\omega t + \theta)$ where θ is the phase angle. Using $\sin(x) = \cos(x - 90°)$, we note that $\theta = -90°$ for the signal given by Eq. (2.1). The mean value of the signal over an integer number of periods is 0. A signal as the one shown in Fig. 2.1 is called an ac signal. The term 'ac' is short for 'alternating current', and in the short form, it is used for both currents and voltages.

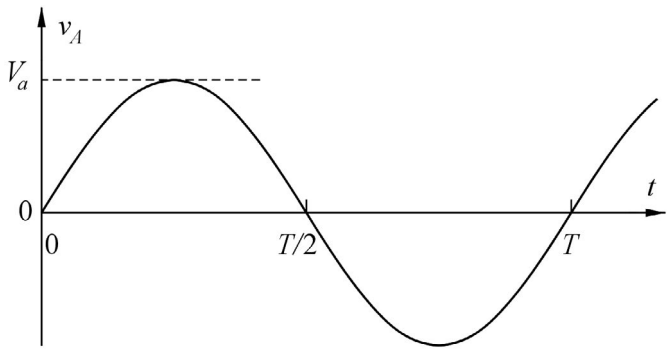


Figure 2.1: A sinusoidal voltage signal.

Often a sinusoidal signal is described by its effective value or rms value (root-mean-square) rather than its amplitude. This is a concept derived from the power which can be delivered by a sinusoidal signal. A sinusoidal voltage with the rms value $V_{R,\text{rms}}$ delivers the same average power to a load resistor with the value R as a constant voltage with the value $V_{R,\text{dc}}$. A constant voltage is also called a dc voltage, where the term 'dc' is short for 'direct current', and just as for the term 'ac', in the short form, it is used for both currents and voltages.

For the dc voltage, the power delivered to the resistor R is

$$P_{\text{dc}} = \frac{V_{R,\text{dc}}^2}{R} \tag{2.2}$$

For the ac voltage with an amplitude V_a, the average power delivered to the resistor R is

$$P_{\text{avg}} = \frac{1}{R}\frac{1}{T}\int_0^T V_a^2 \sin^2(\omega t)dt \tag{2.3}$$

Using $\int_0^T \sin^2(\omega t)dt = \int_0^T \cos^2(\omega t)dt$ and $(\sin^2(\omega t)+\cos^2(\omega t)) = 1$, we can rewrite Eq. (2.3):

$$\begin{aligned} P_{\text{avg}} &= \frac{1}{R}\frac{1}{T}\int_0^T V_a^2 \sin^2(\omega t)dt = \frac{1}{R}\frac{1}{T}\int_0^T V_a^2 \cos^2(\omega t)dt \\ &= \frac{1}{2R}\frac{1}{T}\int_0^T V_a^2(\sin^2(\omega t)+\cos^2(\omega t))dt = \frac{1}{2R}\frac{1}{T}\int_0^T V_a^2\, dt = \frac{V_a^2}{2R} \end{aligned} \tag{2.4}$$

From Eqs. (2.2) and (2.4), it follows that the same average power is obtained from the dc source and the ac source if

$$V_{R,\text{rms}}^2 = V_{R,\text{dc}}^2 = \frac{V_a^2}{2} \Rightarrow V_a = \sqrt{2}\, V_{R,\text{rms}} \tag{2.5}$$

Figure 2.2 shows another analog signal, a triangular signal. Like the sinusoidal signal from Fig. 2.1, this is a periodic signal with a period T. It can be shown that any periodic signal can be decomposed into sine-wave signals with frequencies which are multiples of the basic frequency $f = 1/T$. The basic

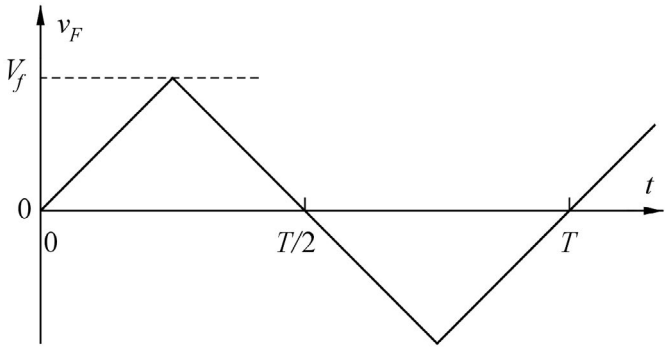


Figure 2.2: A triangular voltage signal.

frequency is also called the fundamental frequency and the multiples of the fundamental frequency are the harmonic frequencies with $2f$ being the second harmonic, $3f$ the third harmonic, and so on. The decomposition of the time-varying signal into components with different frequencies is obtained using the Fourier transformation (Lathi 2009), and for the triangular signal shown in Fig. 2.2, the decomposition results in

$$\begin{aligned} v(t) &= \frac{8V_f}{\pi^2}\sum_{n=0}^{\infty}(-1)^n\frac{\sin((2n+1)\omega t)}{(2n+1)^2} \\ &= \frac{8V_f}{\pi^2}\left(\sin(\omega t)-\frac{1}{9}\sin(3\omega t)+\frac{1}{25}\sin(5\omega t)-\cdots\right) \end{aligned} \tag{2.6}$$

We note that the triangular waveform contains only odd harmonics of the fundamental frequency and that the amplitude of the harmonics decreases with frequency.

Thus, using the Fourier transform, we can describe the signal in the frequency domain, rather than in the time domain.

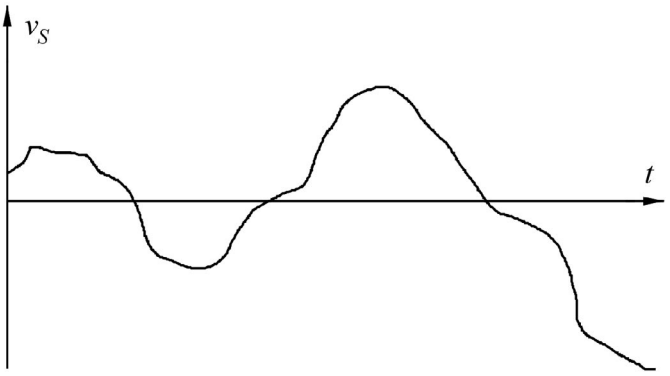


Figure 2.3: An arbitrary time-varying voltage signal.

Figure 2.3 shows an analog signal with an arbitrary, nonperiodic waveform. Also such a signal can be described in the frequency domain through the use of the Fourier transformation or the Laplace transformation (Lathi 2009). In this case, the resulting frequency spectrum is not discrete as for the periodic

signals but contains in principle all frequencies. For signals used in practice in electronic systems, the frequency spectrum is limited. Thus, for audio signals, the frequency range of interest is from about 20 Hz to 20 kHz, the audible frequency band. For radio transmission, several frequency bands are defined, ranging all the way from kHz-frequencies to GHz-frequencies.

The average value of the signals shown in the previous figures is 0. This is not always the case. Often, a signal can be treated as a sum of a constant dc value and an ac signal which is a sum of sinusoidal signals with an average value of 0. Transforming a time-varying signal into the frequency domain thus results in a dc value (a frequency component with $f = 0$) plus a spectrum of ac sinusoids.

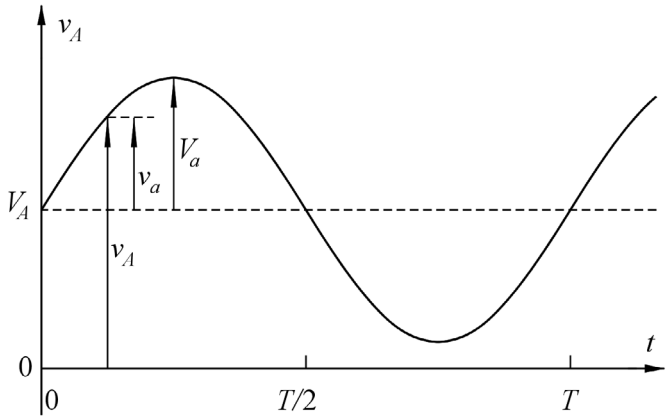


Figure 2.4: Signal notation.

When denoting a signal by a letter and a subscript, the standard convention used in many textbooks, including the present book, is as shown in Fig. 2.4 for a signal represented by a voltage. The total instantaneous value is denoted by a lowercase letter and an uppercase subscript. The dc value of the signal is denoted by an uppercase letter and an uppercase subscript. The ac value of the signal is denoted by a lowercase letter and a lowercase subscript. This means that

$$v_A(t) = V_A + v_a(t) \tag{2.7}$$

If the ac signal is a sinusoid, the amplitude of the sinusoid is denoted by an uppercase letter and a lowercase subscript, so

$$v_a(t) = V_a \sin(\omega t) \tag{2.8}$$

for the signal shown in Fig. 2.4.

Finally, when describing a signal in the frequency domain, we also use an uppercase letter and a lowercase subscript, $V_a(j\omega)$ or $V_a(s)$ where the complex frequency $s = \sigma + j\omega$ is $s = j\omega$ for physical frequencies (Lathi 2009), see examples in Section 2.4.

When using LTspice for simulating a circuit, an important issue is that LTspice is case-insensitive. This implies that you cannot use the conventions described above for distinguishing between dc values, ac

values, amplitudes, etc. You have to learn the signal concepts to a level where you can see from the context if a certain signal is a dc value, an ac signal, a sum of a dc value and an ac signal, or a signal in the frequency domain.

The distinction between (constant) dc values and time-varying signals superimposed on the dc values is extremely useful when analyzing electronic systems, especially if the variations in the ac signals are small enough that the electronic system can be described by linear equations with a reasonable accuracy. This leads to the concept of small-signal analysis which we will use extensively in this book.

A special category of time-varying signals is signals varying in a stochastic manner as a function of time. A stochastic signal is also called a noise signal and noise signals play an important role in all electronic systems. However, the treatment of noise signals is beyond the scope of this book.

Sampled analog signals. Often in electronic systems, an analog signal is sampled at specific points in time as shown in Fig. 2.5. In this way, the signal is described by a series of signal values. Each of the samples may assume any value (within a limited range of signal values), and the series of values is called a sampled analog signal, discrete in time but continuous in value.

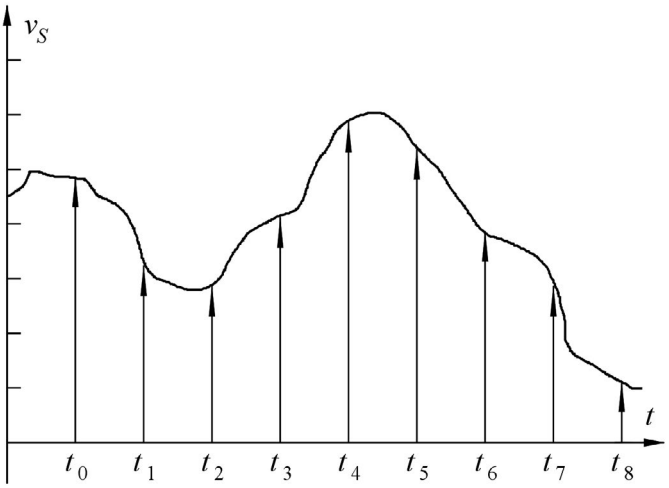


Figure 2.5: Sampling of an analog signal.

Also the values of each sample may be truncated into a finite number of values where each value is typically represented in a binary notation. The number of bits used for this notation determines the number of different signal values and the precision in the representation of the sampled signal values. If the signal shown in Fig. 2.5 is sampled in a system where each sample is discretized into one of 8 allowed values represented by 3 bits, the resulting sampled and digitized signal may be shown as the graph in Fig. 2.6.

Clearly, there is a difference between the original signal and the sampled signal with only 8 possible signal levels. The difference between the original signal and the quantized, sampled signal is an error

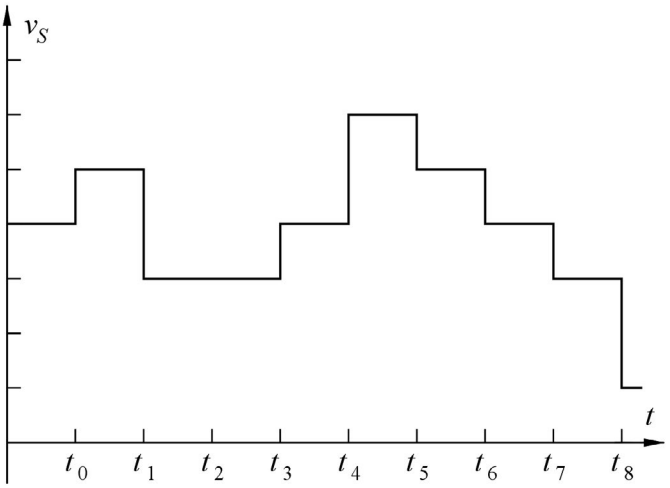


Figure 2.6: Sampled values truncated to 8 levels for the analog signal from Fig. 2.5.

signal called the quantization error or quantization noise since it can often be treated in the same way as a stochastic noise signal.

An extreme case of the quantized, sampled signal is a signal quantized into only two levels, represented by the logic values 0 and 1. Two different values of voltage (or current) correspond to the logic values 0 and 1. This is a binary digital signal and it is described by a stream of bits which can be processed in a computer or digital signal processing system.

The topic of the present book is analog circuits and systems for the processing of signals which are continuous in time and amplitude but the processing in digital systems is becoming ever more important, and often the purpose of analog electronic systems is to perform some sort of pre-processing of a signal before it can be sampled and quantized for further processing using digital signal processing.

2.2 Circuit elements

In CMOS integrated circuits, a number of passive and active circuit elements are available. They include the standard passive devices, i.e., resistors, capacitors and inductors, but the main feature of the CMOS technology is the availability of both p-channel and n-channel MOS transistors. In this section, we give a brief review of passive devices and ideal active devices (voltage sources and current sources) while the description of the real physical, active devices (MOS transistors) is the topic in Chapter 3.

Resistors, capacitors and inductors. An ideal resistor is characterized by the resistance value R and the relation Ohm's law, named after Georg Simon Ohm (1789-1854), a German physicist:

$$v = Ri \tag{2.9}$$

where v is the voltage across the resistor and i is the current flowing in the resistor. This relation applies both to a description in the time domain and in the frequency domain.

In most CMOS technologies, resistors can be designed to have values ranging from less than a kΩ to several tens of kΩ but they occupy a large area compared to the area of a typical MOS transistor. Therefore, for most designs, it is good practice to limit the number of resistors, or rather, the total value of resistance included in the design. Also, resistors in integrated circuits have parasitic components associated with them. Often resistors are implemented in a polysilicon resistive layer separated from the silicon surface by an insulating oxide layer, so a parasitic capacitance from the resistive layer to the silicon surface cannot be avoided.

An ideal capacitor is characterized by the capacitance value C. In the time domain, the following relations between capacitor charge $Q(t)$, current and voltage apply:

$$Q(t) = C\,v(t) \tag{2.10}$$

$$i(t) = C\frac{dv(t)}{dt} \tag{2.11}$$

In the frequency domain, the relation is

$$I(s) = sCV(s);\;\; s = j\omega \tag{2.12}$$

In CMOS technology, capacitors can be designed to have values ranging from fF to tens of pF. Since capacitors are often implemented using the plate capacitance between two conducting layers (or a conducting layer and the silicon substrate) separated by an insulating oxide layer, the capacitor value is proportional to the area of the capacitor. Therefore, it is important to keep the total capacitor value in a circuit small in order to avoid excessive use of silicon area for capacitors. The plate capacitors implemented using conducting layers separated by an insulating oxide layer have a parasitic capacitance from the bottom plate to the silicon surface which cannot be avoided.

Capacitors may also be implemented using a pn-junction biased in the reverse direction, see Chapter 3.1. In this case, the capacitor shows a nonlinear behavior, and the linear relations given by Eqs. (2.11) and (2.12) apply only to small signal variations from a bias point.

An ideal inductor is characterized by the inductance value L. In the time domain, the following relation between voltage and current applies:

$$v(t) = L\frac{di(t)}{dt} \tag{2.13}$$

In the frequency domain, the relation is

$$V(s) = sL\,I(s);\;\; s = j\omega \tag{2.14}$$

In CMOS technology, only rather poor inductors can be implemented. They are typically realized in one or more metal layers as spiral inductors with a planar layout, see Fig. 2.7. In practice, inductors can be designed to have values up to a few nH, and they are suffering from several parasitic elements, including capacitance to the substrate, series resistance, and loss due to induced currents in the substrate. Integrated

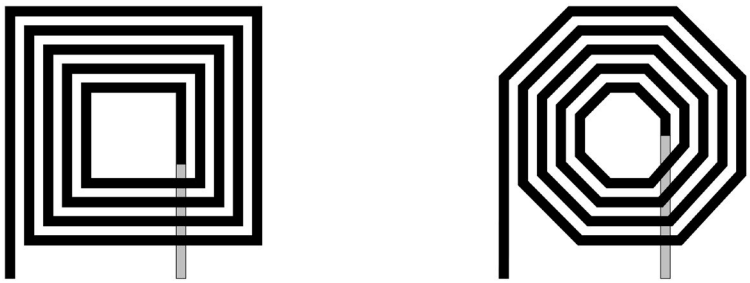


Figure 2.7: Layout examples of planar inductors in CMOS technology.

inductors are used mainly in circuits operating at GHz frequencies, and they will not be treated further in this book.

Except for the pn-junction capacitor, the devices described above are all linear devices (the relation between voltage and current is a linear equation), and in the frequency domain, they are all described by the relation

$$V(s) = Z(s)\,I(s) \tag{2.15}$$

where $Z(s)$ with $s = j\omega$ is the complex impedance of the device. For a resistor, $Z(s) = R$, for a capacitor, $Z(s) = 1/(sC)$, and for an inductor, $Z(s) = sL$.

Figure 2.8 shows the standard textbook diagram symbols for the devices and the basic device equations. In LTspice, the device names are '`R`', '`C`' and '`L`', respectively, and the symbols resemble those shown in Fig. 2.8.

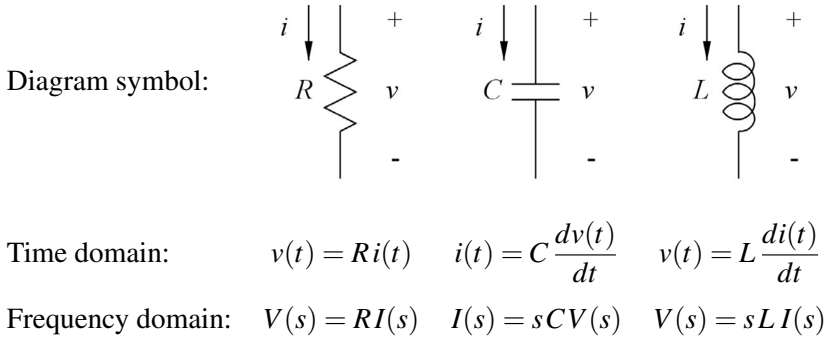
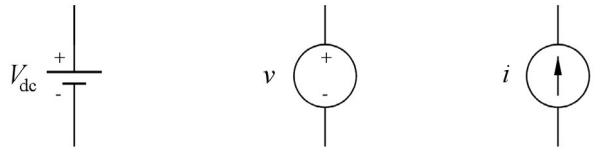


Figure 2.8: Diagram symbols and basic device equations for resistors, capacitors and inductors.

Ideal active devices. An active device is able to deliver energy or control the flow of energy in an electronic system. The ideal active devices include the independent voltage source, the independent current source, the controlled voltage source and the controlled current source.

The independent sources are not available in integrated circuit technology but they are used extensively as signal sources and power supply sources for CMOS integrated electronic systems. Physically, an example of an independent voltage source is a battery, and another example is a voltage regulator, generating a

constant voltage for supplying and electronic system with the dc power needed for its proper operation. The standard diagram symbols used in textbooks are shown in Fig. 2.9. In LTspice, the device names are 'V' and 'I', respectively, and the symbols resemble the circular symbols shown in Fig. 2.9. In LTspice, independent sources may be specified as dc sources, time-varying sources or sources in the frequency domain. To open a specification window for the source, you right-click on the symbol and left-click on 'Advanced' in the dialogue window (Bruun 2020, Tutorial 2.1).



Dc voltage source Voltage source Current source

Figure 2.9: Diagram symbols for independent sources. The leftmost symbol is only used for dc voltage sources (batteries).

The controlled sources include four types: The voltage-controlled voltage source, the current-controlled voltage source, the voltage-controlled current source and the current-controlled current source. They are characterized by a voltage gain A_{voc}, a transresistance R_{moc}, a transconductance G_{msc} and a current gain A_{isc}, respectively (Hambley 2018). The subscript 'oc' denotes 'open circuit' and the subscript 'sc' denotes 'short circuit'. The symbols normally used in textbooks for controlled sources are shown in Fig. 2.10. Notice the diamond-shaped symbols. They serve to distinguish the controlled sources from the independent sources.

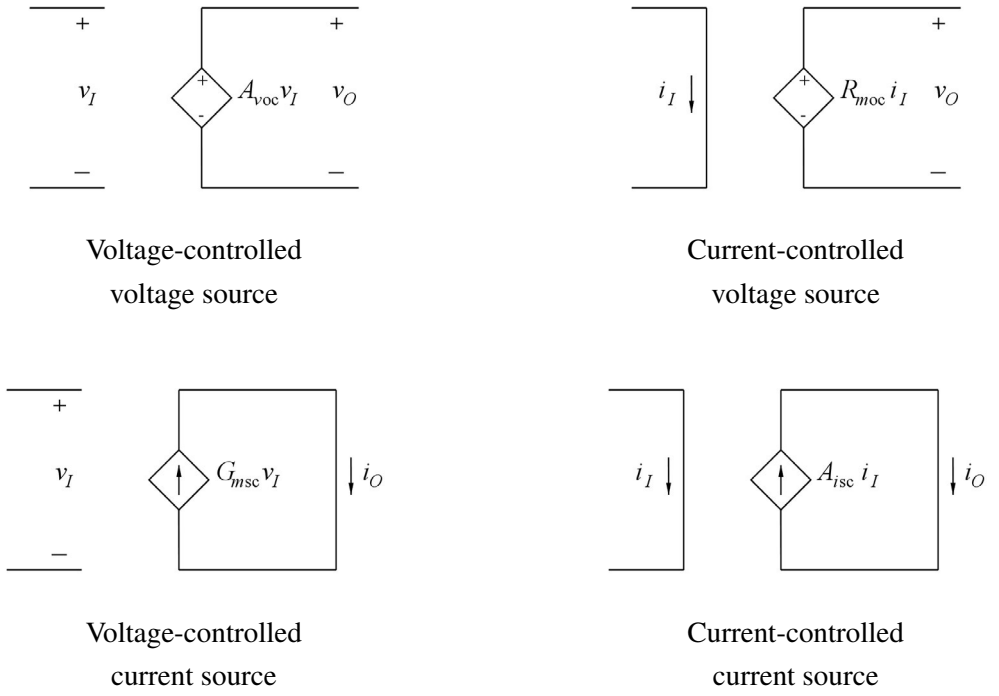
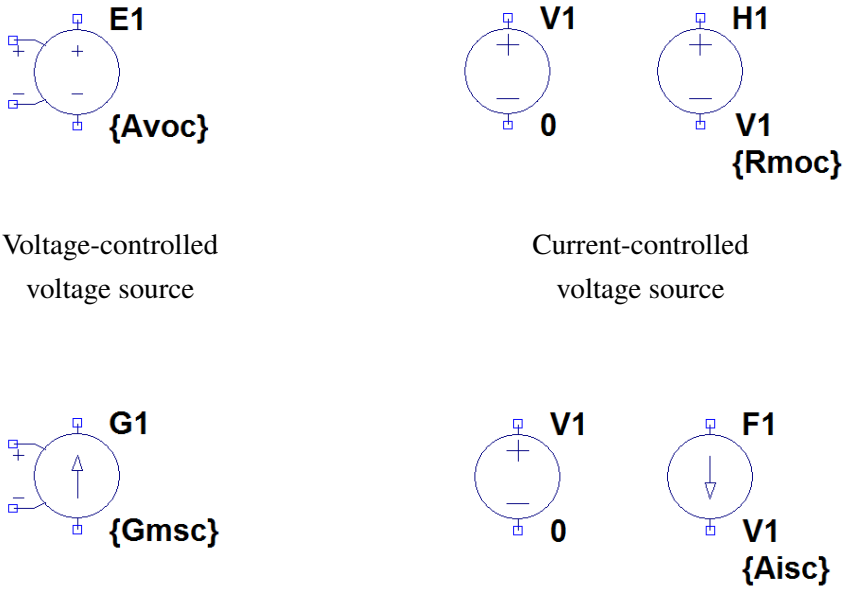


Figure 2.10: Diagram symbols for controlled sources.

The controlled sources are also available in LTspice with the device names '`E`', '`H`', '`G`' and '`F`', respectively. The LTspice symbols are shown in Fig. 2.11. By default, the direction of the current in the current-controlled current source is downwards as shown in Fig. 2.11. In the figure, the values of gain in the different sources are shown as parameters, '`Avoc`', '`Rmoc`', '`Gmsc`' and '`Aisc`'. For the voltage-controlled sources ('`E`' and '`G`'), the terminals for the controlling voltage are specifically shown in the symbols. For the current-controlled sources, the controlling current must be specified as the current through an independent voltage source with a value of 0 (Bruun 2020, Tutorial 1.3), and the name of this voltage source is specified in the 'Value'-line using a specification window opened by a right-click on the symbol. The gain ('`{Rmoc}`' or '`{Aisc}`') is specified in the 'Value2'-line.

Also notice that LTspice does not use a diamond-shaped symbol for controlled sources. You may actually modify the default LTspice symbol into a diamond-shaped symbol using the symbol editor in LTspice (Bruun 2020, Tutorial 1.2). LTspice also has available an 'arbitrary behavioral voltage source' (device type '`BV`') and an 'arbitrary behavioral current source' (device type '`BI`') where the relation between the controlled signal and the controlling signal is specified by an equation. It may be a linear equation as for the ideal amplifiers, or it may be a nonlinear equation describing a nonlinear relation between the controlled signal and the controlling signal.



Voltage-controlled voltage source

Current-controlled voltage source

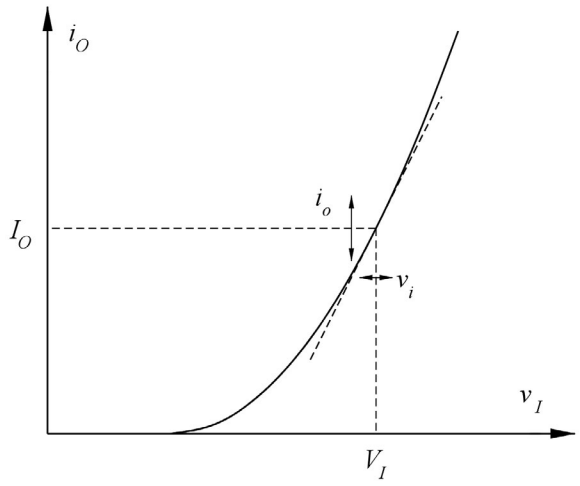
Voltage-controlled current source

Current-controlled current source

Figure 2.11: LTspice symbols for controlled sources.

In CMOS technology, ideal controlled sources are not available, but as we shall see in Chapter 3, a MOS transistor provides an approximation to a voltage-controlled current source. However, it is nonlinear, and it has a fairly complicated relation between the current and the voltages applied to the transistor.

The fact that the relation between current and voltage is nonlinear for real devices in integrated circuits means that the equations for analyzing the behavior of a circuit become nonlinear, and in general, nonlinear equations are more difficult to solve than linear equations, especially when performing calculations by hand. In order to simplify circuit analysis, the nonlinear relation for a device can be linearized by calculating the partial derivative of the nonlinear relation and using this as a linear approximation for small signal variations from a bias point. Figure 2.12 illustrates this concept for a voltage-controlled current source where small signal variations around the bias point (V_I, I_O) can be analyzed using the tangent to the nonlinear relation as an approximation to the actual relation. This approach is called small-signal analysis, and it is extremely important in the analysis of analog circuits.



$$i_o = g_m v_i$$

$$g_m = \left.\frac{\partial i_O}{\partial v_I}\right|_{V_I, I_O}$$

Figure 2.12: Small-signal modeling of a voltage-controlled current source.

2.3 Circuit theorems

The electrical behavior of a circuit is described mathematically by a few fundamental circuit theorems in combination with relations specifying the behavior of each of the circuit elements included in the circuit. In the previous section, we reviewed the relations describing some simple, ideal circuit elements. In the next chapter, we go into detail with the modeling of the MOS transistor which is the predominant device type in a CMOS integrated circuit. In this section, we briefly review some basic circuit theorems, and we describe some additional theorems which are useful for simplifying circuit structures when analyzing circuits by hand calculations.

Kirchhoff's laws. The two laws first described in 1845 by Gustav Kirchhoff (1824-87), a German physicist, are the fundamental, general equations in circuit analysis.

Kirchhoff's current law (KCL) states that the algebraic sum of currents flowing into a node equals zero. This is illustrated in Fig. 2.13, showing a small part (four devices) of a circuit. A node is a point where two ore more circuit elements are connected together. In Fig. 2.13, we can identify two nodes labeled A and B, respectively. Also, the currents in the circuit elements are labeled as shown in the figure with the arrow defining the positive direction of current. Applying KCL to the two nodes, we find the two node equations

$$\text{Node A:} \quad I_1 + I_2 - I_3 = 0 \tag{2.16}$$

$$\text{Node B:} \quad I_3 - I_4 = 0 \tag{2.17}$$

We note specifically that node B is a node where two circuit elements are connected in series, i.e., there are two and only two circuit elements connected to the node. Thus, KCL implies that the current in series-connected circuit elements is the same for all elements.

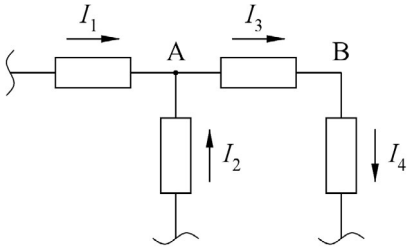


Figure 2.13: Circuit for illustrating Kirchhoff's current law, KCL.

Kirchhoff's voltage law (KVL) states that the algebraic sum of voltages across circuit elements connected in a closed loop equals zero. This is illustrated in Fig. 2.14. In the circuit shown in the figure, we can identify three closed loops, loop 1, loop 2 and loop 3 with the positive directions defined by the arrows shown in the figure and the voltage across each element defined with signs as shown in the figure.

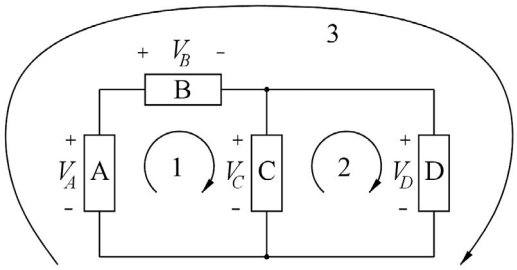


Figure 2.14: Circuit for illustrating Kirchhoff's voltage law, KVL.

Applying KVL to each of the loops, we find the loop equations

$$\text{Loop 1:} \quad -V_A + V_B + V_C = 0 \tag{2.18}$$

$$\text{Loop 2:} \quad -V_C + V_D = 0 \tag{2.19}$$

$$\text{Loop 3:} \quad -V_A + V_B + V_D = 0 \tag{2.20}$$

We note specifically that loop 2 contains two circuit elements, C and D, connected in parallel, i.e., both ends are connected together. From Eq. (2.19), we see that $V_D = V_C$. KVL implies that the voltage across parallel-connected circuit elements is the same for all elements.

Series connection of resistors, inductors and capacitors. While Kirchhoff's laws are the fundamental theorems for analyzing circuits, there are several other rules derived from KCL and KVL which are very useful for reducing the complexity of a circuit by combining different circuit elements into simpler equivalents. This is particularly useful when performing circuit analysis by hand calculations. The first of such rules which we consider is the series connection of passive elements of identical type.

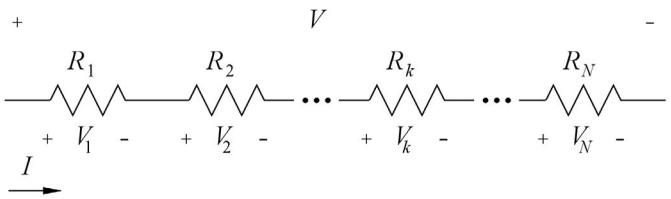


Figure 2.15: A series connection of N resistors.

Figure 2.15 shows a series connection of N resistors, R_1, R_2, ... R_k, ... R_N. Applying KCL, KVL and Ohm's law to the series connection results in

$$I = I_1 = I_2 = ... = I_k = ... = I_N \tag{2.21}$$

and

$$\begin{aligned} V &= V_1 + V_2 + ... + V_k + ... + V_N \\ &= R_1 I_1 + R_2 I_2 + ... + R_k I_k + ... + R_N I_N \\ &= (R_1 + R_2 + ... + R_k + ... + R_N) I \end{aligned} \tag{2.22}$$

From Eq. (2.22), we see that the series connection of resistors R_1 to R_N is equivalent to a single resistor with the value

$$R_{\text{eq}} = R_1 + R_2 + ... + R_N \tag{2.23}$$

We also find that the voltage across any resistor R_k in the series connection is given by

$$V_k = R_k I = \frac{R_k}{R_1 + R_2 + ... + R_N} V \tag{2.24}$$

This is the voltage-divider rule for resistors.

For an inductor, we have $V(s) = sLI(s)$, so using equations similar to Eqs. (2.21) and (2.22), we find that the equivalent inductance for a series connection of inductors is the sum of the inductances. Likewise, the voltage divider rule Eq. (2.24) applies to inductors as well as to resistors.

For a capacitor, we have

$$I(s) = sCV(s) \Rightarrow V(s) = \frac{1}{sC} I(s) \tag{2.25}$$

From Eq. (2.25), it follows – using equations similar to Eqs. (2.21) and (2.22) – that the equivalent capacitance C_{eq} for a series connection of capacitors is given by

$$\frac{1}{C_{\text{eq}}} = \frac{1}{C_1} + \frac{1}{C_2} + ... + \frac{1}{C_N} \tag{2.26}$$

Also, the voltage divider rule states that

$$V_k = \frac{1/C_k}{1/C_1 + 1/C_2 + ... + 1/C_N} V \tag{2.27}$$

The expressions given by Eqs. (2.23) and (2.24) may be generalized for impedances in series:

$$Z_{\text{eq}} = Z_1 + Z_2 + ... + Z_N \tag{2.28}$$

$$V_k = \frac{Z_k}{Z_1 + Z_2 + ... + Z_N} V \tag{2.29}$$

Parallel connection of resistors, inductors and capacitors. Applying KCL, KVL and Ohm's law to a parallel connection of resistors R_1, R_2, ... R_k, ... R_N as shown in Fig. 2.16 results in

$$V = V_1 = V_2 = ... = V_k = ... = V_N \tag{2.30}$$

$$\begin{aligned} I &= I_1 + I_2 + ... + I_k + ... + I_N \\ &= \frac{V_1}{R_1} + \frac{V_2}{R_2} + ... + \frac{V_k}{R_k} + ... + \frac{V_N}{R_N} \\ &= V\left(\frac{1}{R_1} + \frac{1}{R_2} + ... + \frac{1}{R_k} + ... + \frac{1}{R_N}\right) \end{aligned} \tag{2.31}$$

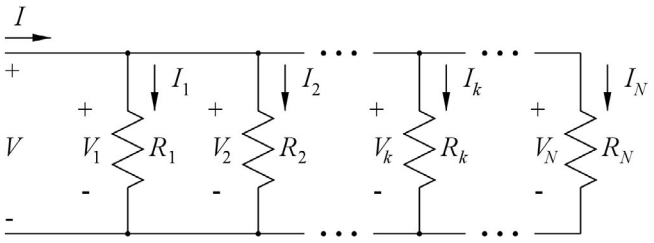


Figure 2.16: A parallel connection of N resistors.

From Eq. (2.31), we see that the parallel connection of resistors R_1 to R_N is equivalent to a single resistor R_{eq} given by

$$\frac{1}{R_{\text{eq}}} = \frac{1}{R_1} + \frac{1}{R_2} + ... + \frac{1}{R_N} \tag{2.32}$$

For the parallel connection, we often use the notation

$$R_{\text{eq}} = R_1 \parallel R_2 \parallel ... \parallel R_N \tag{2.33}$$

We also find that the current I_k in any resistor R_k in the parallel connection is given by

$$I_k = \frac{1/R_k}{1/R_1 + 1/R_2 + ... + 1/R_N} I \tag{2.34}$$

This is the current-divider rule for resistors.

For a parallel connection of inductors, similar equations apply:

$$\frac{1}{L_{\text{eq}}} = \frac{1}{L_1} + \frac{1}{L_2} + ... + \frac{1}{L_N} \tag{2.35}$$

and

$$I_k = \frac{1/L_k}{1/L_1 + 1/L_2 + ... + 1/L_N} I \tag{2.36}$$

For a parallel connection of capacitors, equations similar to Eqs. (2.30) and (2.31) lead to

$$C_{\text{eq}} = C_1 + C_2 + ... + C_N \tag{2.37}$$

and

$$I_k = \frac{C_k}{C_1 + C_2 + ... + C_N} I \tag{2.38}$$

The expressions given by Eqs. (2.32), (2.33) and (2.34) may be generalized for impedances in parallel:

$$Z_{\text{eq}} = Z_1 \parallel Z_2 \parallel ... \parallel Z_N \tag{2.39}$$

$$\frac{1}{Z_{\text{eq}}} = \frac{1}{Z_1} + \frac{1}{Z_2} + ... + \frac{1}{Z_N} \tag{2.40}$$

$$I_k = \frac{1/Z_k}{1/Z_1 + 1/Z_2 + ... + 1/Z_N} I \tag{2.41}$$

Thévenin – Norton equivalent circuits. In order to simplify circuits with both passive devices and active voltage sources and/or current sources, Thévenin and/or Norton equivalent circuits may be applied. The Thévenin theorem states that a linear two-terminal circuit consisting of resistances, voltage sources and current sources can be replaced by an equivalent circuit consisting of an independent voltage source in series with a resistor as shown in Fig. 2.17.

The Thévenin voltage V_t is found as the open-circuit voltage of the original network. The Thévenin resistance R_t is found from the Thévenin voltage V_t and the short-circuit current I_{sc} flowing in a short circuit connected between the two terminals, i.e., $R_t = V_t / I_{\text{sc}}$. Alternatively, the Thévenin resistance can be found as the resistance between the two terminals with all independent voltage sources and current sources in the original circuit reset. Observe that resetting an independent voltage source means replacing it by a short circuit. Resetting an independent current source means replacing it by an open circuit, i.e., removing it completely.

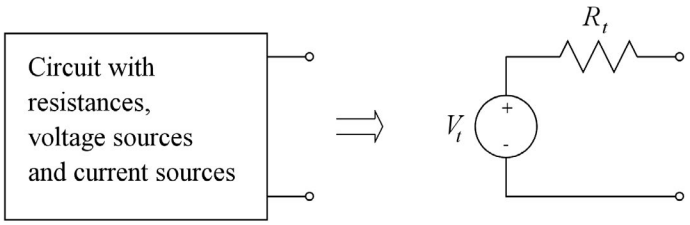


Figure 2.17: Thévenin equivalent circuit.

In the original version of the theorem formulated in 1883 by Léon Thévenin, a French telegraph engineer (1857-1926), the theorem was stated for resistive circuits only. However, it also applies to linear circuits with capacitors and inductors, in which case R_t is replaced by a complex impedance Z_t.

An alternative to the Thévenin equivalent is the Norton equivalent described in 1926 by Edward Norton (1898-1983), an American electrical engineer. The Norton theorem states that a linear two-terminal circuit consisting of resistances, voltage sources and current sources can be replaced by an equivalent circuit consisting of an independent current source in parallel with a resistor as shown in Fig. 2.18.

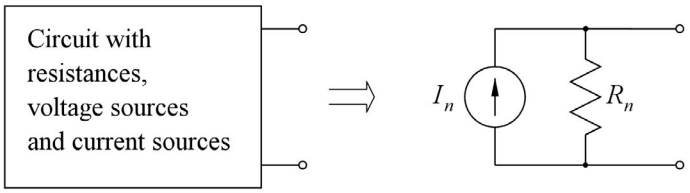


Figure 2.18: Norton equivalent circuit.

The Norton current I_n is equal to the short-circuit current flowing in a short circuit connected between the two terminals. The Norton resistance is found from the Norton current and the open-circuit voltage V_{oc} between the two terminals, i.e., $R_n = V_{\text{oc}}/I_n$. Alternatively, the Norton resistance can be found as the resistance between the two terminals with all independent voltage sources and current sources in the original circuit reset. By comparing the Thévenin equivalent and the Norton equivalent, we see that $R_t = R_n = V_t/I_n$.

Just as the Thévenin theorem, the Norton theorem is also applicable to linear circuits with capacitors and inductors.

Superposition. The superposition principle states that in a linear circuit with more independent sources, the total response is the sum of the responses to each of the independent sources acting alone with all other independent sources being reset.

In order to illustrate both the superposition principle, the Thévenin equivalent and the Norton equivalent, we examine the circuit shown in Fig. 2.19.

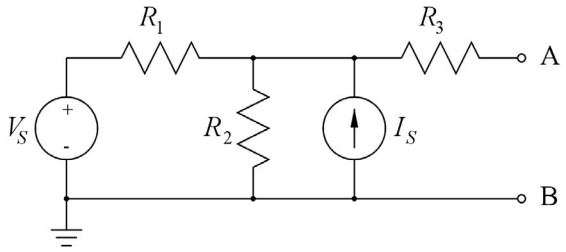


Figure 2.19: Circuit example for illustrating superposition.

For finding the Thévenin voltage, we find the open-circuit voltage between the terminals A and B. The contribution from V_S with I_S reset, i.e., $I_S = 0$, is

$$V_{\text{oc},1} = V_S \frac{R_2}{R_1 + R_2} \tag{2.42}$$

The contribution from I_S with V_S reset, i.e., $V_S = 0$, is

$$V_{\text{oc},2} = I_S (R_1 \parallel R_2) \tag{2.43}$$

Thus, the total open-circuit voltage (and the Thévenin voltage) is

$$V_t = V_{\text{oc}} = V_{\text{oc},1} + V_{\text{oc},2} = V_S \frac{R_2}{R_1 + R_2} + I_S (R_1 \parallel R_2) = (V_S + I_S R_1) \frac{R_2}{R_1 + R_2} \tag{2.44}$$

The Thévenin resistance is found by resetting both V_S and I_S and finding the resistance between the terminals, i.e., V_S is replaced by a short circuit and I_S is removed. By inspection, we find

$$R_t = R_3 + (R_1 \parallel R_2) \tag{2.45}$$

For finding the Norton equivalent, we place a short circuit between the terminals A and B and find the short-circuit current I_{sc}. The contribution from V_S with I_S reset, i.e., $I_S = 0$, is

$$I_{\text{sc},1} = V_S \left(\frac{R_2 \parallel R_3}{R_1 + R_2 \parallel R_3} \right) \left(\frac{1}{R_3} \right) = V_S \frac{R_2}{R_1R_2 + R_1R_3 + R_2R_3} \tag{2.46}$$

The contribution from I_S with V_S reset, i.e., $V_S = 0$, is

$$I_{\text{sc},2} = I_S \frac{1/R_3}{1/R_3 + 1/(R_1 \parallel R_2)} = I_S \frac{R_1R_2}{R_1R_2 + R_1R_3 + R_2R_3} \tag{2.47}$$

Thus, the total short-circuit current (and the Norton current) is

$$I_n = I_{\text{sc}} = I_{\text{sc},1} + I_{\text{sc},2} = (V_S + I_S R_1) \frac{R_2}{R_1R_2 + R_1R_3 + R_2R_3} \tag{2.48}$$

We may verify

$$R_n = R_t = \frac{V_t}{I_n} = \frac{R_1R_2 + R_1R_3 + R_2R_3}{R_1 + R_2} = \frac{R_1R_2}{R_1 + R_2} + R_3 = (R_1 \parallel R_2) + R_3 \tag{2.49}$$

which is also the result given by Eq. (2.45).

2.4 Circuit analysis

In the design of analog CMOS integrated circuits, we apply both circuit analysis using hand calculations and circuit analysis using simulations. Hand calculations are often based on simplified circuit and device models while simulations by computer can be performed with complex device models. There are three types of analysis and simulation which are used throughout this book: dc steady-state analysis, time-domain analysis and frequency-domain analysis.

Dc steady-state analysis. The dc analysis serves the purpose of finding dc voltages and currents in a circuit with all time-varying signals reset. Once the values of dc voltages and currents have been found, linearized small-signal device models can be calculated using the approach illustrated in Fig. 2.12. The dc analysis is based on Kirchhoff's laws in combination with device models describing the behavior of devices in the circuit. For nonlinear devices, the nonlinear models must be used.

In LTspice, a dc analysis for fixed values of dc voltages and currents is performed using the simulation directive '.op'. It results in an output file with values for all node voltages and all device currents, and the error log file (opened with 'Ctrl-L') contains small-signal parameters, device voltages and device currents for the transistors and diodes in the circuit. The small-signal concept was illustrated in Fig. 2.12, and in Chapter 3, we investigate in detail the small-signal parameters for MOS transistors.

Often a dc steady-state analysis with one or more independent sources varied over a suitable range is performed. This analysis serves to find static (time-invariant) relations between voltages and currents in a circuit, for example the static relation between input voltage and output voltage for an amplifier. It is very useful for finding a suitable bias point for the operation of a circuit when time-varying signals are applied. In a dc analysis, there are no time-varying signals, implying that $di(t)/dt$ and $dv(t)/dt$ are both equal to zero for all voltages and currents. From Eqs. (2.11) and (2.13), it follows that all capacitor currents are zero and that all inductor voltages are zero, so in a dc steady-state analysis, capacitors are treated as open circuits and inductors are treated as short circuits.

In LTspice, a dc steady-state simulation with a variation of independent voltage sources or current sources is specified by the simulation directive '.dc' where one or more (up to three) signal sources can be specified with their range of variation and the increment between each value of the input signal. The simulation results in a plot window where the varying signal is by default the x-axis and node voltages and device currents can be selected for the y-axis (Bruun 2020, Tutorial 1.1).

Time-domain analysis. The time-domain analysis resembles the dc analysis with a sweep of the signal sources. However, in the time-domain analysis, signals are varied as a function of time, implying that both capacitors and inductors must be taken into account using the Eqs. (2.11) and (2.13). This complicates the analysis to a considerable extent, and hand calculations are normally performed only for simple circuits with linear circuit elements while circuit simulations can be performed for complex circuits including nonlinear circuit elements.

In order to illustrate a time domain analysis, consider the very simple *RC* network shown in Fig. 2.20.

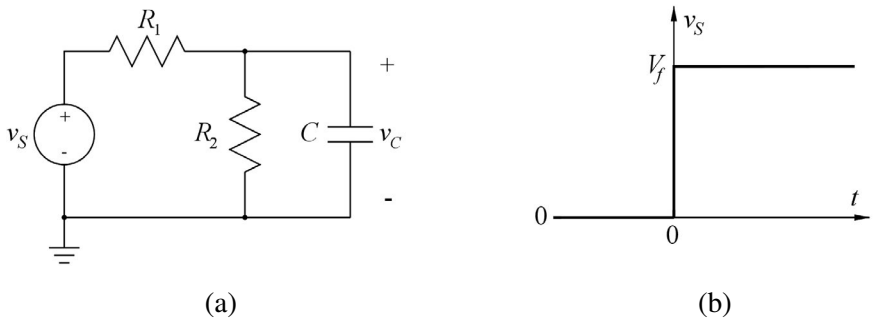


Figure 2.20: *RC* network (a) and input voltage $v_S(t)$ to the network (b).

For this, we may find the output voltage $v_C(t)$ across the capacitor C when the input voltage changes as a step from 0 V to a value V_f at time $t = 0$ as also shown in Fig. 2.20. We first perform a hand calculation. For time $t < 0$, there are no changing signals, and we have the dc solution $V_C = 0$ V. When the input voltage changes to $v_S = V_f$, Kirchhoff's current law at the output node gives in combination with the device equations for R_1, R_2 and C:

$$\frac{V_f - v_C(t)}{R_1} = \frac{v_C(t)}{R_2} + C\frac{dv_C(t)}{dt} \tag{2.50}$$

This is a differential equation for finding $v_C(t)$ and the solution is:

$$v_C(t) = V_f \frac{R_2}{R_1 + R_2}(1 - \exp(-t/\tau)) \tag{2.51}$$

where

$$\tau = \frac{1}{(R_1 \parallel R_2)C} \tag{2.52}$$

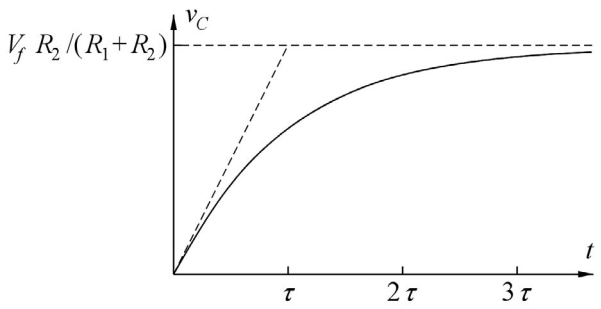


Figure 2.21: The capacitor voltage $v_C(t)$ for the *RC* network from Fig. 2.20 with a step function applied to the input.

The voltage $v_C(t)$ can be sketched as a function of time as shown in Fig. 2.21. We see that the capacitor voltage increases exponentially and approaches asymptotically a value of $V_f R_2/(R_1+R_2)$.

In LTspice, a time-domain analysis is specified by the simulation directive '.tran' and the input signals must be specified as time-varying voltages and/or currents. For analyzing the circuit from Fig. 2.20, we may use the transient simulation in LTspice but this requires that we have numerical values for the

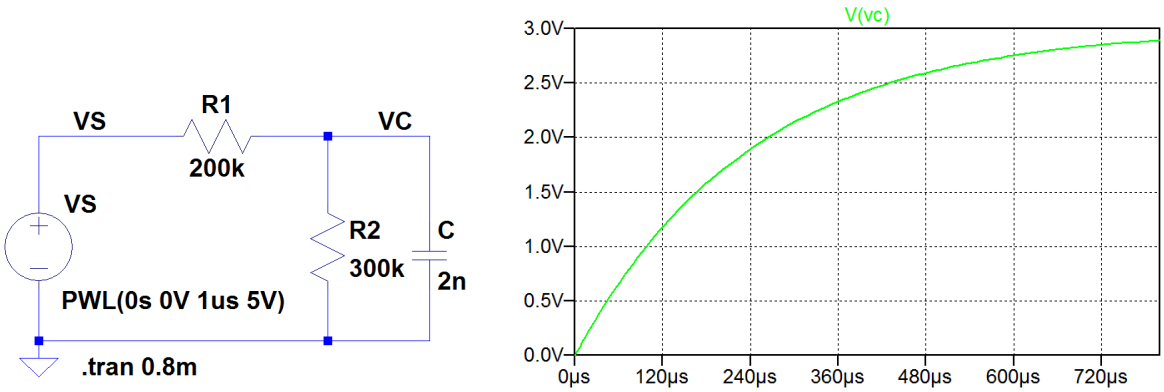


Figure 2.22: LTspice schematic and simulation results for the circuit in Fig. 2.20.

devices and voltages in the circuit. As an example, let us assume $R_1 = 200\ \text{k}\Omega$, $R_2 = 300\ \text{k}\Omega$, $C_1 = 2\ \text{nF}$ and $V_f = 5\ \text{V}$. With these values, we can draw the schematic in LTspice as shown in Fig. 2.22, and we can specify the input signal as a piecewise-linear voltage.

Notice that we cannot have a rise time of 0 for the input specification. Instead, we select a rise time which is much smaller than the time constant $\tau = 1/((R_1 \parallel R_2)C) = 240\ \mu\text{s}$. In order to show the capacitor voltage approaching $V_f R_2(R_1+R_2) = 3\ \text{V}$, we select a simulation time of 0.8 ms which is slightly more than 3τ.

Figure 2.22 also shows the plot of the capacitor voltage $v_C(t)$ resulting from the simulation. The resemblance with the analytically derived plot, Fig. 2.21, is apparent, so here we have an example of verification by simulation of the results achieved by hand analysis.

Frequency-domain analysis. In the frequency-domain analysis, a circuit is analyzed using the device equations valid in the frequency domain. For nonlinear devices, a linearization of the device equations is done first, so the frequency-domain analysis is always performed using linear device models, i.e., small-signal models. In general, the first step in a frequency-domain analysis by hand calculations is to establish a small-signal diagram for the circuit. For the circuit shown in Fig. 2.20, all devices are linear already, so the small-signal diagram is identical to the circuit shown in Fig. 2.20 with the input voltage defined as an ac voltage $V_s(s)$ in the frequency domain. In later chapters, we examine several examples of circuits with nonlinear devices, mostly transistors, and we investigate in detail how the small-signal diagram for such circuits can be established, but for the circuit in Fig. 2.20, we can directly apply KCL at the output node $V_c(s)$ using the device equations in the frequency domain with $s = j\omega$:

$$\frac{V_s(j\omega) - V_c(j\omega)}{R_1} = \frac{V_c(j\omega)}{R_2} + j\omega C V_c(j\omega) \tag{2.53}$$

This results in a transfer function $H(j\omega) = V_c(j\omega)/V_s(j\omega)$ given by

$$H(j\omega) = \frac{V_c(j\omega)}{V_s(j\omega)} = \left(\frac{R_2}{R_1 + R_2}\right)\left(\frac{1}{1 + j(\omega/\omega_0)}\right) \tag{2.54}$$

where $\omega_0 = 1/\tau = 1/((R_1 \parallel R_2)C)$.

Often the result of a frequency domain analysis is shown graphically in a Bode plot where a logarithmic frequency axis is used and the transfer function is shown as a gain function in dB, i.e., $|H(j\omega)|_{\text{dB}} = 20\log(|H(j\omega)|)$ and a phase in degrees, $\phi = \angle H(j\omega)$. The Bode plot is named after the American engineer Hendrik Bode (1905-82). For the transfer function given by Eq. (2.54), we find:

$$|H(j\omega)| = \left(\frac{R_2}{R_1 + R_2}\right)\left(\frac{1}{\sqrt{1 + (\omega/\omega_0)^2}}\right) \tag{2.55}$$

$$\angle H(j\omega) = -\arctan(\omega/\omega_0) \tag{2.56}$$

This results in the plots shown in Fig. 2.23.

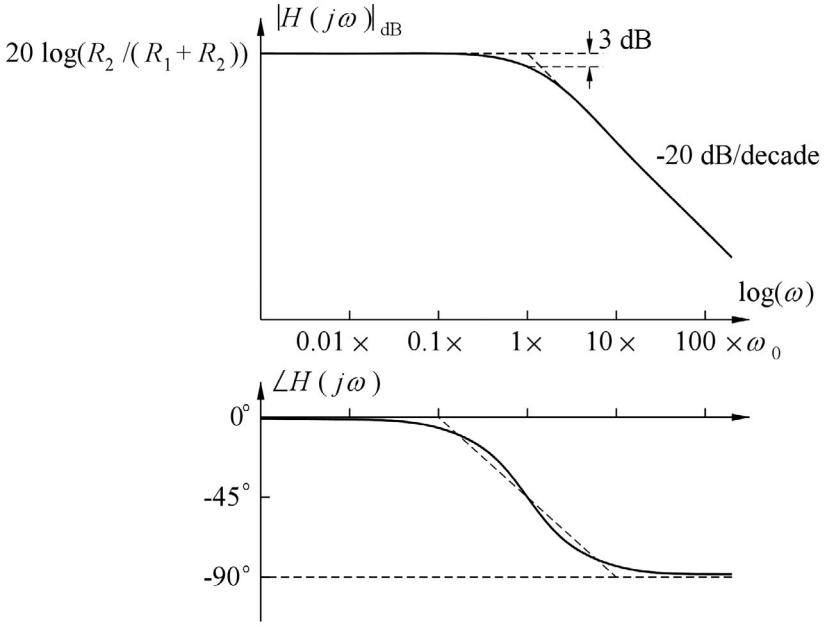


Figure 2.23: Bode plot for the lowpass filter function given by Eq. (2.54).

Examining Eq. (2.55), we see that when ω is very small, $|H(j\omega)|$ approaches the constant value $R_2/(R_1+R_2)$ whereas when ω is very large, the transfer function is inversely proportional to the frequency. Thus, a piecewise-linear approximation to the gain plot is a horizontal line for small values of ω and a line with a slope of -20 dB per decade of frequency for large values of ω. The intersection between the two line segments is given by $\omega=\omega_0$, and for this frequency, the -3 dB bandwidth, the gain has dropped by a factor of $\sqrt{2}$ from the low-frequency value, corresponding to -3 dB.

From Eq. (2.56), we find that the phase of $H(j\omega)$ is 0° for very small frequencies, $-90°$ for very high frequencies and $-45°$ for $\omega=\omega_0$. For the phase characteristics, a piecewise-linear approximation may also be applied as shown in Fig. 2.23. For frequencies below $\omega_0/10$ and above $10\,\omega_0$, horizontal lines are used, and for frequencies between $\omega_0/10$ and $10\,\omega_0$, the phase is approximated by a straight line with a slope of $-45°$ per decade of frequency.

In LTspice, a frequency-domain analysis is specified by the simulation directive '.ac' and the input signal must be specified as an ac signal. For the '.ac' directive, the upper and lower limit of the frequency range is specified, and also the number of points per decade (or octave) must be specified. Using an amplitude of 1 directly gives the transfer function in the frequency domain. The simulation results in a Bode plot with the frequency $f=\omega/(2\pi)$ as the x-axis, see the simulation for the circuit from Fig. 2.22 in Fig. 2.24. Again, we notice that the simulation results provide a verification of the analytical results achieved using math and hand calculations. We also note that the simulation provides a very easy way of achieving the exact gain and phase plots, rather than the asymptotic piecewise-linear approximations.

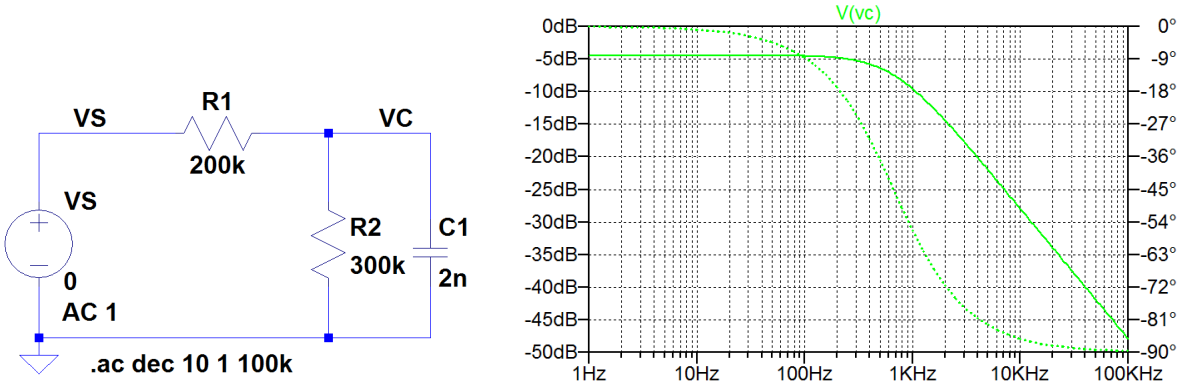


Figure 2.24: LTspice schematic and ac simulation results for the circuit in Fig. 2.20.

As the final example in this chapter, we show the frequency-domain analysis for the circuit shown in Fig. 2.25. It is derived from the circuit in Fig. 2.20 by replacing R_1 with a coupling capacitor C_1. Also, the capacitor in parallel with R_2 has been renamed to C_2.

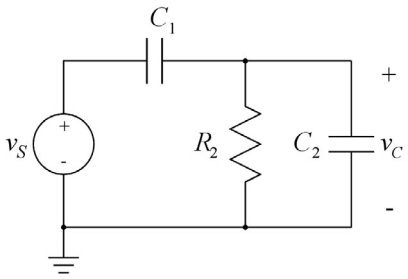


Figure 2.25: *RC* highpass filter.

Using KCL, we find

$$(V_s(j\omega)-V_c(j\omega))j\omega C_1 = \frac{V_c(j\omega)}{R_2} + j\omega C_2 V_c(j\omega) \tag{2.57}$$

From this, we find the transfer function $H(j\omega) = V_c(j\omega)/V_s(j\omega)$

$$H(j\omega) = \frac{V_c(j\omega)}{V_s(j\omega)} = \frac{j\omega C_1 R_2}{1+j\omega R_2(C_1+C_2)} \tag{2.58}$$

For this transfer function, we find

$$|H(j\omega)| = \frac{\omega C_1 R_2}{\sqrt{1+(\omega R_2(C_1+C_2))^2}} \tag{2.59}$$

$$\angle H(j\omega) = 90^\circ - \arctan(\omega R_2(C_1+C_2)) \tag{2.60}$$

We see that at low frequencies, the gain is proportional to the frequency, giving a positive slope of 20 dB/decade in the Bode plot, and for high frequencies, the gain approaches a constant value of $C_1/(C_1+C_2)$, corresponding to a capacitive voltage division between C_2 and C_1. The intersection between the two asymptotic line segments in the Bode plot is given by the frequency $\omega_0 = 1/(R_2(C_1+C_2))$.

From Eq. (2.60), we find that the phase of $H(j\omega)$ is 90° for very small frequencies, 0° for very high frequencies and 45° for $\omega = \omega_0$.

Figure 2.26 shows the Bode plot corresponding to this analysis.

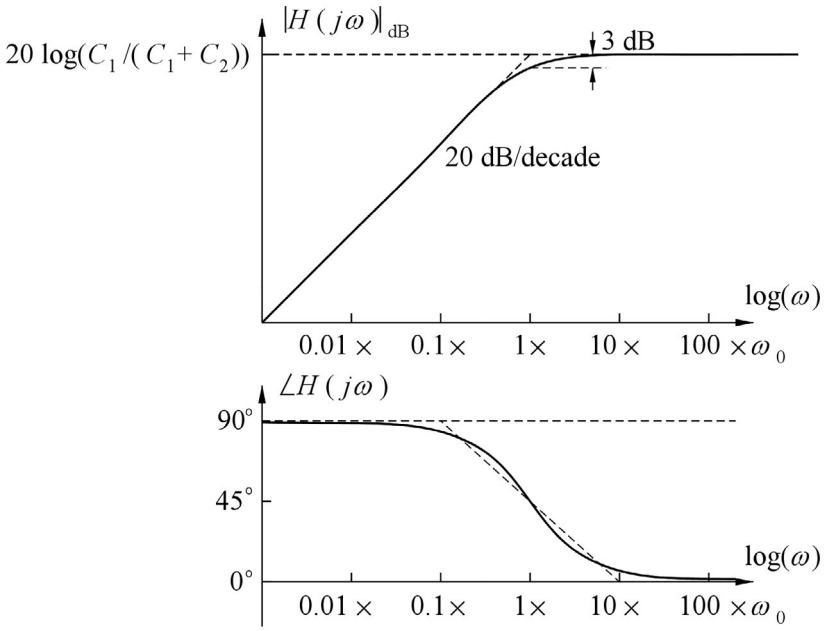


Figure 2.26: Bode plot for the highpass filter function given by Eq. (2.58).

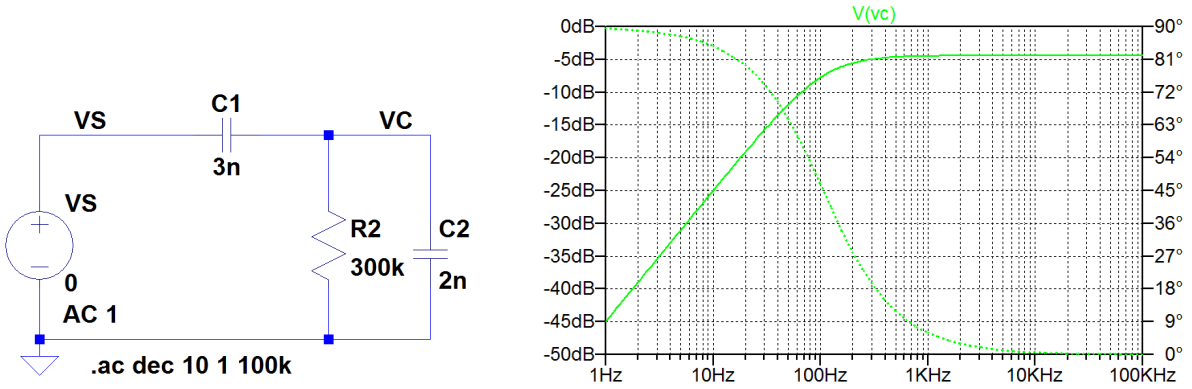


Figure 2.27: LTspice schematic and ac simulation results for the circuit in Fig. 2.25.

Again, for a numerical example, the circuit can be simulated using LTspice. The schematic for this and the resulting Bode plot is shown in Fig. 2.27 where we have used $C_1 = 3$ nF.

The circuit examples shown here contain only linear circuit elements, so the frequency-domain analysis can be performed directly on the schematics shown. For circuits including nonlinear devices, LTspice automatically derives small-signal models for all devices and performs the ac simulation using a linearized small-signal diagram with a bias point as found from a '.op' simulation of the nonlinear circuit. It is always very important to check that the ac simulation is performed using reasonable values of bias voltages and currents. Often, an analysis of a circuit by simulation will start by using a dc analysis ('.dc') for finding a suitable bias point, then an operating point analysis ('.op') to verify the bias point, and then an ac analysis ('.ac') for investigation of the frequency response.

References

Bruun, E. 2020, *CMOS Integrated Circuit Simulation with LTspice*, Third Edition, bookboon. Available from: http://bookboon.com/en/cmos-integrated-circuit-simulation-with-ltspice-ebook

Hambley, AR. 2018, *Electrical Engineering, Principles and Applications*, Seventh Edition, Pearson Education Ltd., Harlow, UK.

Lathi, BP. 2009, *Signal Processing & Linear Systems*, International Second Edition, Oxford University Press, New York, USA.

Multiple-choice test

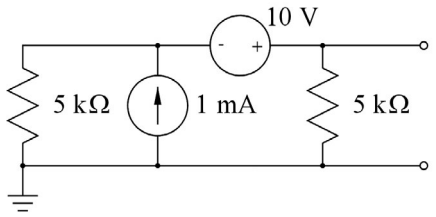
1. Complete the following statements by selecting the appropriate continuation from the table below.

A:	Kirchhoff's current law (KCL) states that ...
B:	Kirchhoff's voltage law (KVL) states that ...
C:	Kirchhoff's laws are valid for ...
D:	The device equation for a resistor is known as ...
E:	The device equation for a capacitor is ...
F:	In a dc analysis, an inductor is replaced by ...
G:	A Thévenin equivalent circuit is ...
H:	A small-signal device model is ...
I:	When an independent current source is reset, it corresponds to ...
J:	The superposition principle is valid for ...

Continuation:

1:	the algebraic sum of voltages across circuit elements equals zero for any closed path (loop) in an electric circuit.
2:	an independent current source in series with an impedance.
3:	the algebraic sum of currents flowing into a node equals zero.
4:	when v is the voltage across a resistor R, the current in the resistor is $i = v/R$.
5:	when i is the current in a resistor R, the voltage across the resistor is $v = Ri$.
6:	Ohm's law.
7:	$v = L(di/dt)$.
8:	$i = C(dv/dt)$.
9:	both linear and nonlinear circuits.
10:	linear circuits only.
11:	an independent voltage source in series with an impedance.
12:	an independent voltage source in parallel with an impedance.
13:	an independent current source in parallel with an impedance.
14:	an open circuit.
15:	a short circuit.
16:	a linearized device model.
17:	the superposition principle applies to linear circuits.
18:	$v = Ri$.
19:	a resistor.
20:	a nonlinear device model.

2. The Thévenin resistance for the circuit shown below is



A: 2.5 kΩ
B: 5 kΩ
C: 10 kΩ

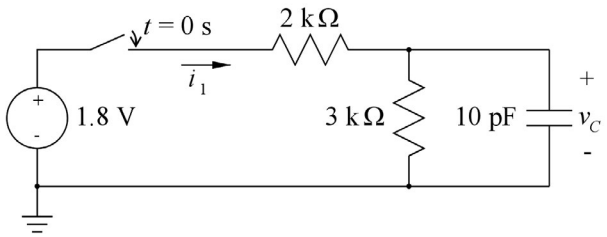
3. The contribution to the Thévenin voltage from the 1 mA current source in the circuit above is

A: 2.5 V
B: 5 V
C: 10 V

4. The contribution to the Thévenin voltage from the 10 V voltage source in the circuit above is

A: 2.5 V
B: 5 V
C: 10 V

5. In the circuit shown below, the switch is closed at time $t = 0$. Immediately after the switch is closed, the current i_1 is



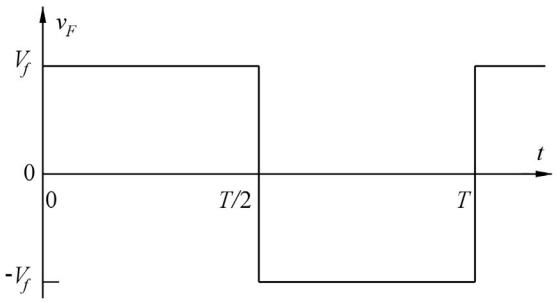
A: 0.36 mA
B: 0.90 mA
C: 1.5 mA

6. When the switch is re-opened in the circuit above, the capacitor discharges with a time constant of

A: 12 ns
B: 30 ns
C: 50ns

Problems

Problem 2.1



A symmetric square-wave voltage signal $v_F(t)$ as shown above is described by the Fourier series

$$v_F(t) = \frac{4V_f}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} \sin((2n+1)\omega t)$$

where $\omega = (2\pi)/T$. Assume that $v_F(t)$ is connected across a resistor R. Find an expression for the total average power dissipated in the resistor. What fraction of the power is in the fundamental frequency? What fraction of the power is in the 3rd harmonic and in the 5th harmonic?

Problem 2.2

A sinusoidal voltage signal v_A is characterized as follows: The frequency is 3 kHz. The maximum value is 700 mV. The minimum value is 300 mV. The value for time $t = 0$ is 400 mV. Find V_A, V_a, ω and θ for describing the signal in the time domain using the expression $v_A = V_A + V_a \cos(\omega t + \theta)$.

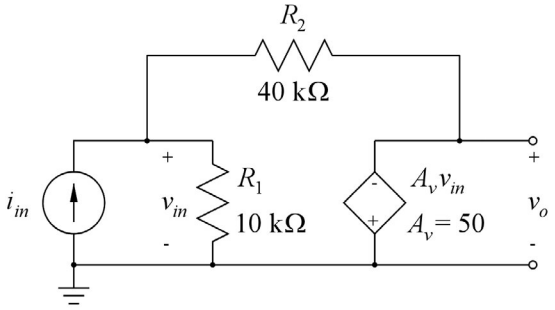
Problem 2.3

A signal $V_o(j\omega)$ is described in the frequency domain by the following the relation:

$$V_o(j\omega) = \frac{V_i(j\omega)}{1 + j\omega/\omega_0}$$

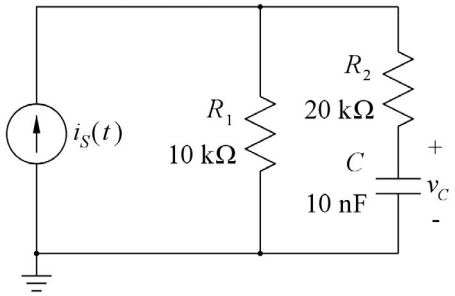
The signal $V_i(j\omega)$ is a sinusoid with an amplitude $|V_i(j\omega)| = V_a$ and a phase angle $\theta = 0°$.
For a frequency $\omega = 2\,\omega_0$, find the amplitude and phase of $V_o(j\omega)$ and find an expression for $v_O(t)$ in the time domain.

Problem 2.4



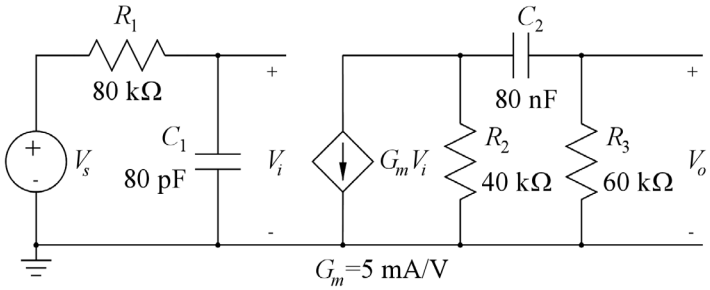
The circuit shown above is a transresistance amplifier built from an inverting voltage amplifier with an input resistance of $R_1 = 10\,\text{k}\Omega$ and an inverting voltage gain of $A_v = 50$ V/V. The feedback resistor $R_2 = 40\,\text{k}\Omega$ turns it into a transresistance amplifier with a transresistance $R_m = v_o/i_{in}$.
Find the transresistance R_m and the input resistance $R_{in} = v_{in}/i_{in}$ as functions of R_1, R_2 and A_v and calculate numerical values for the specified values of R_1, R_2 and A_v.
Verify your results by simulating the circuit with LTspice.

Problem 2.5



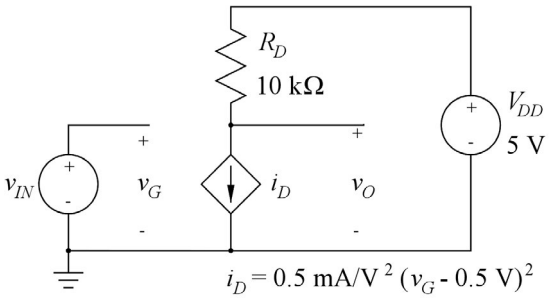
For the circuit shown above, $i_S(t)$ is a current source with the value 0 for $t \leq 0$ and 1 mA for $t > 0$. Find the value of v_C at time $t = 0$. Find the value of v_C for $t \to \infty$. Find the time constant for the charging of the capacitor C. Sketch the voltage $v_C(t)$ as a function of time t for $0 \leq t \leq 1$ ms.
Verify your results by simulating the circuit with LTspice.

Problem 2.6



For the circuit shown above, find the transfer function $H(s) = V_o(s)/V_s(s)$ and find the -3 dB cutoff frequencies. Sketch the magnitude of the transfer function in a Bode plot.
Hint: Find $H(s)$ as the product of $(V_i(s)/V_s(s))$ and $(V_o(s)/V_i(s))$.
Verify your results by simulating the circuit with LTspice.

Problem 2.7



Shown above is a nonlinear transconductance amplifier. Find the values of bias voltages and currents for an input bias voltage (quiescent voltage) of $V_{IN} = 1.0$ V. Sketch the output voltage v_O versus the input voltage v_{IN} for 0.5 V $\leq v_{IN} \leq$ 1.5 V. Find the small-signal voltage gain v_o/v_{in} versus the input bias voltage V_{IN} for 0.5 V $\leq V_{IN} \leq$ 1.5 V and calculate the value of the small-signal gain for $V_{IN} = 1.0$ V.
Verify your results by simulating the circuit with LTspice.

Problem 2.8

Derive an expression for the small-signal voltage gain v_o/v_{in} for an amplifier with the nonlinear transfer function

$$v_O = \frac{10\text{ V}}{1+\exp(-5\text{ V}^{-1}\cdot v_{IN})}.$$

Calculate the small-signal voltage gain for $V_{IN} = -0.3$ V, 0 V and +0.3 V.

Problem 2.9

An amplifier with a right-half-plane zero at the frequency f_z and two left-half-plane poles at the frequencies f_{p1} and f_{p2} has the transfer function shown below where the gain A_0 at low frequencies is assumed to be positive.

$$H(jf) = \frac{A_0(1 - jf/f_z)}{(1 + jf/f_{p1})(1 + jf/f_{p2})}$$

Find an expression for the gain function $|H(jf)|$ and the phase function $\angle H(jf)$ and sketch a Bode plot using piecewise-linear approximations for the frequency range 10 kHz - 1 GHz, assuming that $f_z = 200$ MHz, $f_{p1} = 200$ kHz, $f_{p2} = 40$ MHz and $A_0 = 100$ V/V.

Calculate the gain and the phase at the frequencies $f = 0.4$ MHz, $f = 4$ MHz, $f = 40$ MHz and $f = 400$ MHz.

Chapter 3 – The MOS Transistor

This chapter is an introduction to the MOS transistor. We derive a simple analytical model for the transistor, and we show how this can be linearized into a small-signal model. Also, we introduce more advanced transistor models, unsuitable for hand calculations but very useful with computer simulations. After having studied the chapter, you should be able to

- explain the basic physical structure of the MOS transistor.
- explain the difference between an NMOS transistor and a PMOS transistor.
- use the Shichman-Hodges transistor model for nonlinear modeling of the MOS transistor.
- derive a small-signal model for the MOS transistor.
- derive expressions for the small-signal parameters of the MOS transistor using the Shichman-Hodges model.
- analyze very simple circuits using MOS transistors.
- explain some properties of advanced transistor models.
- simulate circuits with MOS transistors using LTspice, both with the Shichman-Hodges model and with BSIM transistor models.

Before we can start explaining the operation of the MOS transistor, we need to review the pn junction used to form a diode in a semiconductor crystal. This is an essential component in the structure of the MOS transistor.

3.1 Fundamentals of pn diodes

The majority of CMOS integrated circuits are fabricated using silicon as the basic material. Silicon is a semiconductor, number 14 in the periodic table. It has 4 electrons in the outer shell of electrons which are available to form bonds with other atoms in a crystal. This is a fairly stable structure and only a small fraction of the electrons will gain sufficient energy (due to thermal agitation) to move freely in the crystal.

However, if some of the silicon atoms in the crystal lattice are replaced with atoms having 5 electrons in the outer shell (for example phosphorus), this extra electron easily escapes from the phosphorus atom and is free to move around as a carrier of a negative charge. A region of the silicon lattice where a fraction of the silicon atoms has been replaced by atoms with 5 electrons in the outer shell is called an n-doped region or n-type silicon, ‘n’ for negative. The atoms used for the doping are called donor atoms.

Conversely, if some of the silicon atoms in the crystal lattice are replaced with atoms having only 3 electrons in the outer shell (for example boron), these atoms tend to fill the outer shell with an electron from an adjacent silicon atom. Thus, the missing electron appears as a 'hole' which is free to move around as a carrier of positive charge. A region of the silicon lattice where a fraction of the silicon atoms has been replaced by atoms with 3 electrons in the outer shell is called a p-doped region or p-type silicon, 'p' for positive. The atoms used for the doping are called acceptor atoms.

It should be noted that when doping a silicon lattice with dopants such as phosphorus or boron only a very small fraction of the atoms in the lattice are dopant atoms. The concentration of the dopants is typically in the range of 10^{14} to 10^{21} atoms/cm^3 whereas the number of silicon atoms in a single-crystal silicon lattice is about 5×10^{22} atoms/cm^3. In pure silicon, the number of free carriers, the intrinsic carrier concentration n_i, is about 1.5×10^{10} cm^{-3} at room temperature and it is strongly temperature dependent. In p-doped silicon, the concentration p of positive carriers (holes) is approximately equal to the concentration N_A of acceptor atoms and in n-doped silicon, the concentration n of negative carriers (electrons) is approximately equal to the concentration N_D of donor atoms. It can be shown that in a doped silicon structure, the product of n and p is equal to n_i^2, so a high concentration of electrons automatically implies a low concentration of holes and vice versa (Grove 1967). The dominant carrier type is called the majority carrier while the non-dominant type is the minority carrier.

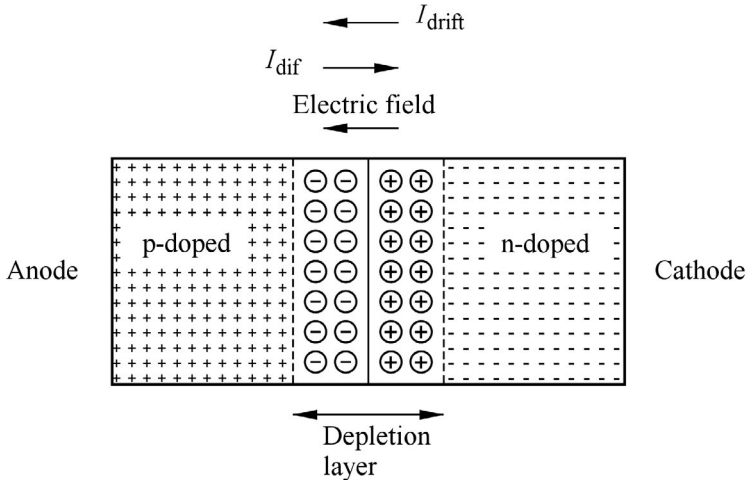


Figure 3.1: A cross-sectional view of a pn junction.

Now, let us consider a silicon structure where a part of the silicon is n-doped and an adjacent part is p-doped as shown in Fig. 3.1. In the n-type region, there is a high concentration of negative free carriers while the p-type region has a high concentration of positive free carriers. The free carriers will tend to diffuse in the silicon crystal in order to obtain an even distribution of free carriers throughout the crystal. When the electrons leave the n-type region, they fill holes in the p-type region. This is called recombination. However, when the holes and electrons recombine, they leave behind the fixed charges from the atomic nuclei, and this creates an electric field near the junction between the regions, pulling the free carriers back towards their region of origin.

When the semiconductor junction is in equilibrium, the flow of free carriers by diffusion, the diffusion current I_{dif}, is exactly balanced by the flow of free carriers caused by the electric field, the drift current I_{drift}, i.e., $I_{\text{dif}} = I_{\text{drift}}$. Hence, there is no net current flowing in the diode, and in the region around the junction, there are no free carriers. This region is called the depletion region or the space-charge region. Due to the electric field, there is a built-in potential barrier in the junction determined by the magnitude of the field and the thickness of the depletion region. In this situation, the diode may be considered as a plate capacitor with the n-doped region and the p-doped region forming the plates and the depletion region forming the insulating layer between the plates.

The p-type region is called the anode and the n-type region is called the cathode. With this convention, we note that a diffusion current of holes flows from the anode to the cathode while a drift current of holes flows from the cathode to the anode. If the anode is biased positively with respect to the cathode, the pn junction is said to be forward-biased. Conversely, if the cathode is biased positively with respect to the anode, the pn junction is said to be reverse-biased.

Reverse bias. When the diode is reverse-biased, this increases the electric field and the corresponding potential barrier in the depletion region between the n-type region and the p-type region, making it even more difficult for the free carriers to diffuse away from their region of origin, so the diffusion current is reduced. The increase in potential barrier does not cause a significant change in the drift current because this current is limited by the number of carriers available for it, and this is the very limited number of minority carriers from each of the doped regions. Thus, the resulting net current I_D in the diode under reverse bias is $I_D \simeq -I_{\text{drift}}$, and this is a very small current, often negligible, so the diode is considered to be off. In the off-state, the thickness of the depletion layer is increased with increasing reverse-bias voltage, so considering the diode as a capacitor, the value C_j of the capacitance depends on the reverse-bias voltage V_R, following the relation (Chan Carusone, Johns & Martin 2012)

$$C_j = C_{j0}(1 + V_R/\phi_0)^{-m_j} \tag{3.1}$$

where C_{j0} is the capacitance with zero bias, ϕ_0 is the diode built-in potential and m_j is the grading coefficient. These parameters are technology-dependent with typical values of $\phi_0 = 0.8$ V and $m_j = 0.5$. Thus, the diode in the reverse-bias direction may be utilized as a voltage-dependent capacitor, also called a varactor.

Forward bias. When the diode is biased in the forward direction, this decreases the electric field and the potential barrier in the depletion region, making it easier for the free carriers to diffuse away from their region of origin, so the diffusion current is increased. Even with a reduced field and potential barrier, the drift current remains almost unchanged because it is limited by the availability of minority carriers and not by the strength of the field. With the increase in diffusion current, there is a net current $I_D = I_{\text{dif}} - I_{\text{drift}} > 0$ in the diode. It has been shown by William Shockley (1949), an American physicist

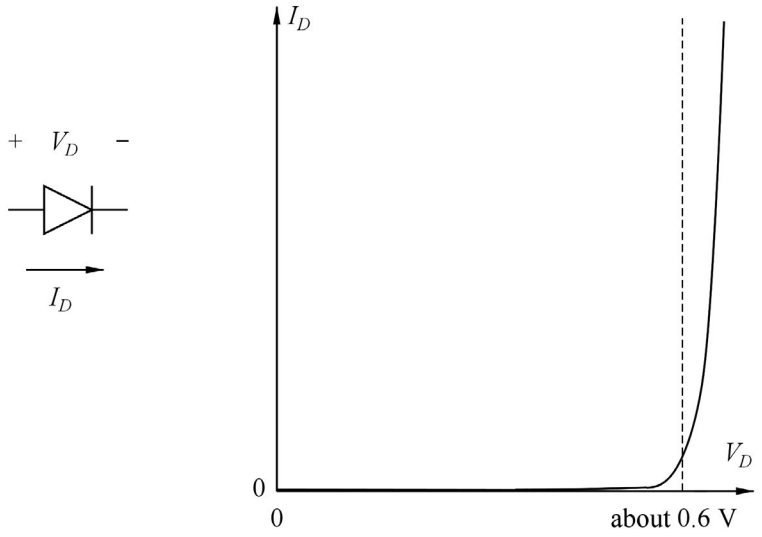


Figure 3.2: The diode symbol and the Shockley diode characteristics.

(1910-1989), that this current follows the relation

$$I_D = I_S \exp\left(\frac{V_D}{nV_T} - 1\right) \tag{3.2}$$

where V_D is the voltage across the diode as shown in Fig. 3.2. The current I_S is called the saturation current, n is called the emission coefficient and $V_T = kT/q$ is called the thermal voltage. T is the absolute temperature, $k = 1.38 \times 10^{-23}$ J/K is Boltzmann's constant and $q = 1.602 \times 10^{-19}$ C is the magnitude of the electron charge. At room temperature, $V_T \simeq 26$ mV. The saturation current I_S is typically in the range of fA to pA. It is proportional to the area of the diode, and it depends on the doping levels in the doped semiconductor regions. It is strongly temperature dependent, increasing about 15%/°C with temperature. The emission coefficient n is normally between 1 and 2. The value of n depends on the details of the doping profiles of the semiconductor regions (Grove 1967).

This relation is shown graphically in Fig. 3.2. We see that the current increases very abruptly as a function of the forward bias voltage, and often the forward bias voltage can be considered as almost constant over a large range of currents. For a silicon diode, the forward voltage is often in the range 0.5 V to 0.8 V. A simple model for a diode is an open circuit (or a capacitor) when it is reverse-biased and a dc voltage source of some hundreds of mV when it is forward-biased, i.e., a piecewise-linear model.

Both the Shockley diode model and piecewise-linear models are available in LTspice (Bruun 2020, Tutorial 2.2). For the Shockley model, the default value of the emission coefficient is $n = 1$ and the default value of the saturation current at 27°C is $I_S = 10$ fA.

Physical structure. The structure shown in Fig. 3.1 is a simplified structure. In modern semiconductor processes, a diode is often physically formed in the surface of a silicon wafer, so it has a planar structure as shown in the cross-sectional view in Fig. 3.3 where an n^+-region is formed by ion implantation or diffusion of donor atoms into a substrate with a majority of acceptor atoms. The notation n+ is used to specify a highly doped n-region. In order to obtain a good contact to the p-region, a highly doped p^+-region may be diffused into the p-region, ensuring a good contact to the metallic connection layers at the surface of the wafer.

For the MOS transistors which we consider next, the main purpose of the pn junction is to establish an insulating depletion region between the active regions of the MOS transistor and an underlying substrate.

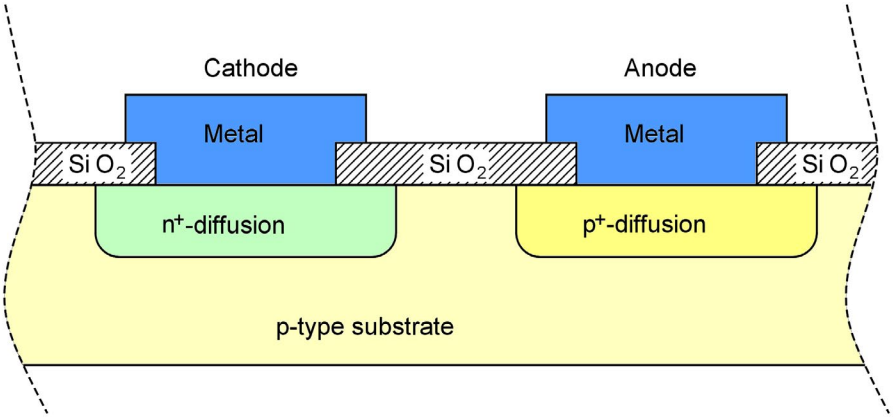


Figure 3.3: Cross-sectional view of a planar diode.

3.2 Physical characteristics of the MOS transistor

The pn diode is a two-terminal device with a nonlinear relationship between current and voltage. This makes it useful in applications such as rectifiers but the diode is not capable of providing a voltage gain or a current gain. For this, we need a device with an additional terminal to be used for an input signal which controls the current flow between two other terminals of the device. The transistor is such a device, 'transistor' being a combination of 'transfer' and 'resistor' (The Transistor 2014).

The first type of transistor to come into practical use was the bipolar transistor invented at Bell Labs in 1947 by the three American physicists John Bardeen, Walter Brattain, and William Shockley (Brinkman, Haggan & Troutman 1997). The bipolar transistor was followed by the field-effect transistor which was conceptually described by Julius Lilienfeld already in 1925 (Lilienfeld 1925) and 1928 (Lilienfeld 1928) but did not become realizable until about 1960 when Kahng and Attala demonstrated the first practical realization of a MOS transistor (Sah 1988).

Physical structure. The MOS transistor is a device with a structure as shown in Fig. 3.4. In its simplest form shown in Fig. 3.4, it is a planar device in the surface of a silicon wafer. The MOS transistor has two terminals which are doped regions, diffused or ion implanted into a substrate or bulk. In Fig. 3.4, these two terminals are n-doped regions in a p-doped substrate. The silicon surface between the two n-regions

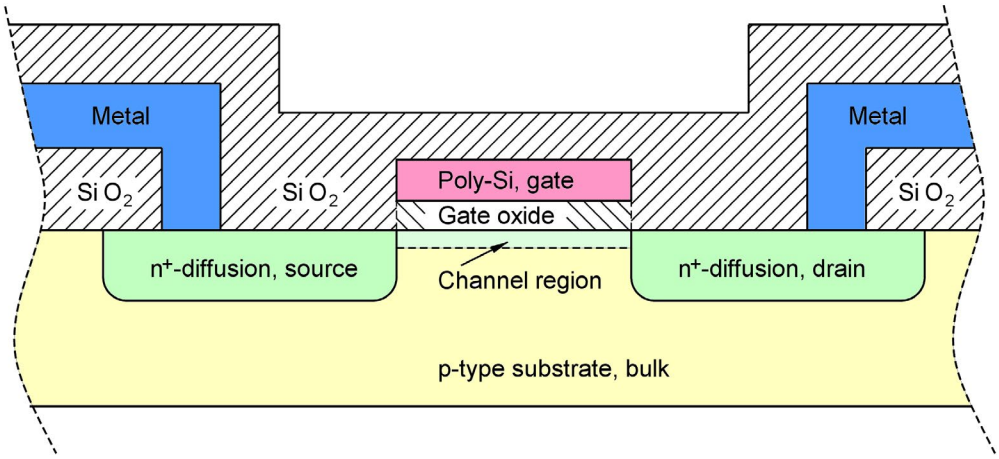


Figure 3.4: Cross-sectional view of a MOS transistor.

is covered by a thin layer of silicon dioxide, SiO_2, and on top of this layer, there is an electrode, the gate electrode, of conducting material.

In the early days of MOS technology, the gate electrode was made from metal, hence the name MOS transistor, MOS being an acronym for Metal-Oxide-Semiconductor. In present-day's technology, the gate material is polysilicon. The term 'polysilicon' or just 'poly' indicates that it is a polycrystalline form of silicon whereas the silicon substrate is monocrystalline or single-crystal silicon. Also, in the early days of MOS technology, the doped regions were formed by diffusion of dopants into the substrate. Nowadays, the doped regions are normally formed by ion implantation but the term 'diffused region' is still in common use (Plummer 2000).

The entire device is covered by a passivation layer of SiO_2 serving as an insulator between additional conductive layers (metal or poly) used to connect devices together. Contact to the doped n-type regions and to the gate is obtained via metal layers through contact holes (vias) in the insulating SiO_2 layer.

In the cross-sectional view in Fig. 3.4, the metal contacts to the n-type regions are shown whereas the metal contact to the gate electrode is not shown.

The basic idea of the MOS transistor is that by applying a positive voltage to the gate with respect to the p-type substrate, an electric field is created which repels the positive mobile carriers from the region beneath the gate electrode, and when the electric field is high enough, electrons are attracted to the region beneath the gate from the adjacent n-regions. This forms a channel for electrons between the two n-regions, and with a voltage difference between the two n-regions, a current will flow in the channel. The two n-type regions are called source and drain, respectively, with the drain being the region with the highest voltage so that the current in the channel flows from drain to source. Notice that the electrons in the channel flow in the opposite direction, i.e., from source to drain. Also note that the pn junctions between the substrate (or bulk) and the drain and the source must be reverse-biased so that the depletion layer of these junctions serves to isolate the MOS transistor from the substrate and from other transistors in the same substrate.

For the transistor shown in Fig. 3.4, the free carriers in the channel are electrons, hence the transistor is called an n-channel MOS transistor or simply NMOS transistor. A complementary version of the transistor is also feasible with an n-type bulk and diffused or ion implanted p-regions for source and drain. For this, the conducting channel is formed by a voltage at the gate which is negative with respect to the bulk so that the channel formed in this transistor has a majority of free positive carriers, hence it is called a p-channel transistor or PMOS transistor.

The two transistor types can be combined in a structure as shown in Fig. 3.5. In order to obtain the correct dopant type for the bulk for both NMOS transistors and PMOS transistors, the PMOS transistor is placed in an n-region which is diffused or implanted into a p-type substrate. The n-type bulk region for a PMOS transistor is called the n-well. Thus, the PMOS transistors have four terminals: gate, drain, source, and bulk. Since the n-well is implanted into the substrate, it is feasible to have separate n-type bulk regions for each PMOS transistor but also several PMOS transistors may share a common well. For the NMOS transistors, the bulk is the common p-type substrate, so all NMOS transistors in a circuit using the fundamental structure shown in Fig. 3.5 will have their bulk contacts connected together to the same common potential.

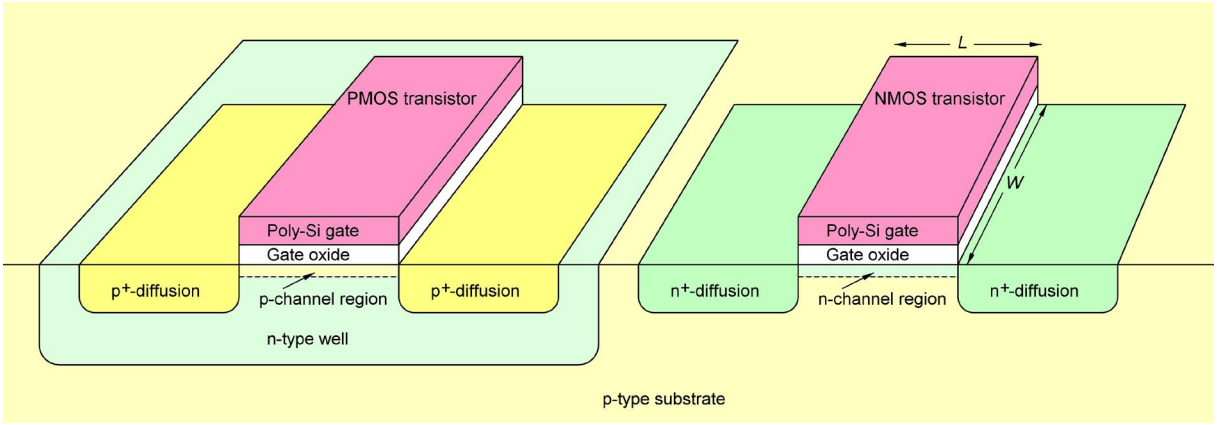


Figure 3.5: A CMOS device structure. Metal layers and insulating oxide layers are not shown. For the NMOS transistor, the channel length L and the channel width W are indicated on the figure.

The structure shown in Fig. 3.5 is the basic CMOS structure commonly used today. The CMOS structure can also be found in a version with an n-type substrate and p-wells for the NMOS transistors. This was actually the structure used for early CMOS circuit families such as the RCA 4000 series of logic circuits (Pujol 1975). More advanced twin-well technologies provide the possibility of both p-wells and n-wells separated by insulating layers. This gives added flexibility for the electrical connection of the bulk contacts for both PMOS transistors and NMOS transistors.

In Fig. 3.5, the metal layers for establishing contact to gate, drain, source and bulk are not shown. Nor are the thick insulating oxide layers and contact holes shown but the figure shows the definition of the transistor parameters channel length L and channel width W, indicated for the NMOS transistor. These two parameters are the physical design parameters available to the circuit designer when a suitable CMOS process has been selected for the design.

Observe that all the pn junctions in the structure must be reverse-biased. This implies that the p-substrate should be connected to the lowest voltage in the circuit, typically ground or the negative supply voltage.

3.3 Electrical characteristics of the MOS transistor

The basic idea of the MOS transistor is that the current flowing between drain and source can be controlled by the voltage applied between gate and bulk. For modern submicron MOS transistors, the relationship between current and voltage is complicated, depending on many design parameters and technology parameters. A realistic device model for MOS transistors with channel dimensions (channel length L and channel width W) of submicron magnitude can only be handled by computer simulation but in order to achieve a first-order mathematical transistor model, we examine a transistor model generally known as the Shichman-Hodges model (Shichman & Hodges 1968). This model was preceded by the Sah model described already in 1964 (Sah 1964). These models were developed when transistor dimensions were on the order of 10 μm and a simple analytical model provided a reasonable fit to the actual device behavior.

The Shichman-Hodges transistor model. We start by considering an NMOS transistor as shown in Fig. 3.6. It is a simplified cross-sectional view showing the electrical contacts to drain, source and gate simply as wires connected to ideal voltage sources or to ground.

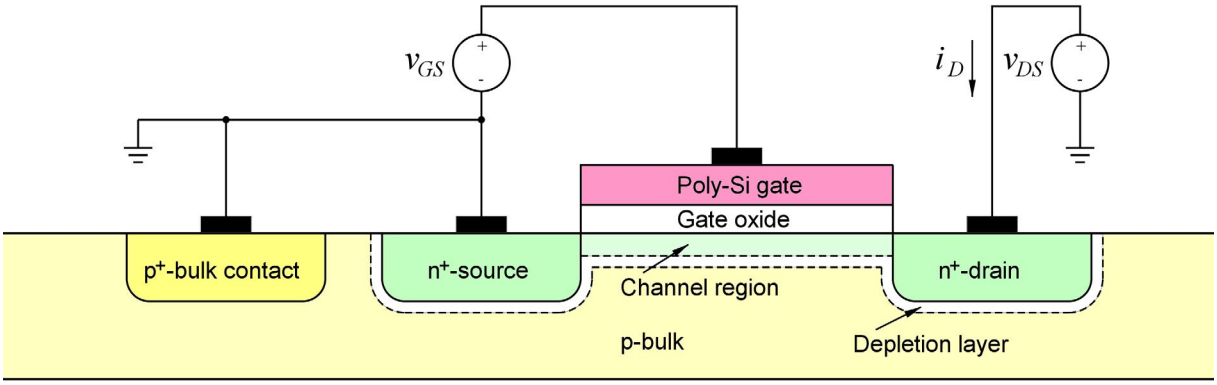


Figure 3.6: Cross-sectional view of an NMOS transistor with electrical connections to gate, drain, source and bulk and a uniform channel depth corresponding to $v_{DS} \ll v_{GS} - V_t$, implying $v_{DS} \simeq 0$, and $v_{GS} > V_t$.

For the initial investigation, we assume that bulk and source are both connected to ground. Also, we assume that the drain-source voltage v_{DS} is very small.

When no gate voltage is applied, i.e., $v_{GS} = 0$, both the source n-type region and the drain n-type region are surrounded by a depletion layer, compare to Fig. 3.1, and the free carriers in the region underneath the gate are holes from the p-type bulk. When a positive voltage is applied to the gate, i.e., $v_{GS} > 0$, an electric field is generated in the gate oxide which repels the positive free carriers in the region under the gate. When the gate-source voltage is increased, at a certain value of v_{GS}, all the free carriers underneath the gate are repelled, so a depletion region is formed underneath the gate. The value of v_{GS} for which this occurs is called the threshold voltage V_t and it is dependent on the characteristics of the silicon substrate and the oxide. It is a technology parameter for the manufacturing process.

Assuming that the voltage in the surface region beneath the gate is 0, we find from Eq. (2.10) that for $v_{GS} = V_t$, the magnitude of the charge density (charge per unit area) in the silicon surface region underneath the gate is $Q_B = C_{ox} V_t$ where C_{ox} is the gate capacitance per unit area. For a depleted surface region, this charge originates from the negative fixed charge from the atomic nuclei left behind in the depletion layer when the positive free carriers are repelled from the region.

For a further increase in v_{GS}, i.e., $v_{GS} > V_t$, the electric field in the interface between the gate oxide and the silicon substrate causes electrons from the drain and source to enter the region underneath the gate, thus forming a channel with mobile negative carriers as shown in Fig. 3.6. This channel may be considered as a voltage-controlled resistor R_{channel}. The value of R_{channel} can be assumed to be proportional to the length L of the channel and inversely proportional to the channel width W. Also, it can be assumed to be inversely proportional to the density of charge available for the conduction in the channel and to the mobility of the carriers of the charge.

For a voltage of 0 in the channel region, the total magnitude of the density of charge in the surface region is $Q_s = C_{ox} v_{GS}$ but a charge density of $Q_B = C_{ox} V_t$ originates from the negative fixed charge from the atomic nuclei in the channel, so the density of mobile charge (electrons) is $Q_n = Q_s - Q_B = C_{ox}(v_{GS} - V_t)$. The voltage $v_{GS} - V_t$ is called the effective gate voltage or the overdrive voltage, $v_{\text{eff}} = v_{\text{ov}} = v_{GS} - V_t$. The gate capacitance per unit area is given by $C_{ox} = \varepsilon_{ox}/t_{ox}$ where ε_{ox} is the permittivity of the gate oxide and t_{ox} is the thickness of the gate oxide layer.

The mobility μ_n of the electrons in the channel is a property of the silicon substrate. Just as the oxide capacitance C_{ox}, it is dependent on the manufacturing process and cannot be controlled by the circuit designer, once a technology and a fabrication provider has been selected.

Combining these dependencies, we find

$$R_{\text{channel}} = \left(\frac{L}{W}\right)\left(\frac{1}{\mu_n C_{ox} v_{\text{eff}}}\right) \tag{3.3}$$

With a small voltage v_{DS} applied to the drain, we find

$$i_D = v_{DS}/R_{\text{channel}} = \mu_n C_{ox}(W/L) v_{\text{eff}} v_{DS} = \mu_n C_{ox}(W/L)(v_{GS} - V_t) v_{DS} \tag{3.4}$$

Notice that in the derivation of Eq. (3.4), we have assumed that the effective electric field v_{eff}/t_{ox} is the same in the entire channel region. This is an approximation since the voltage in the channel is $v_{DS} > 0$ in the drain end of the channel and 0 in the source end of the channel. Hence, Eq. (3.4) is only valid for $v_{DS} \ll v_{GS} - V_t$.

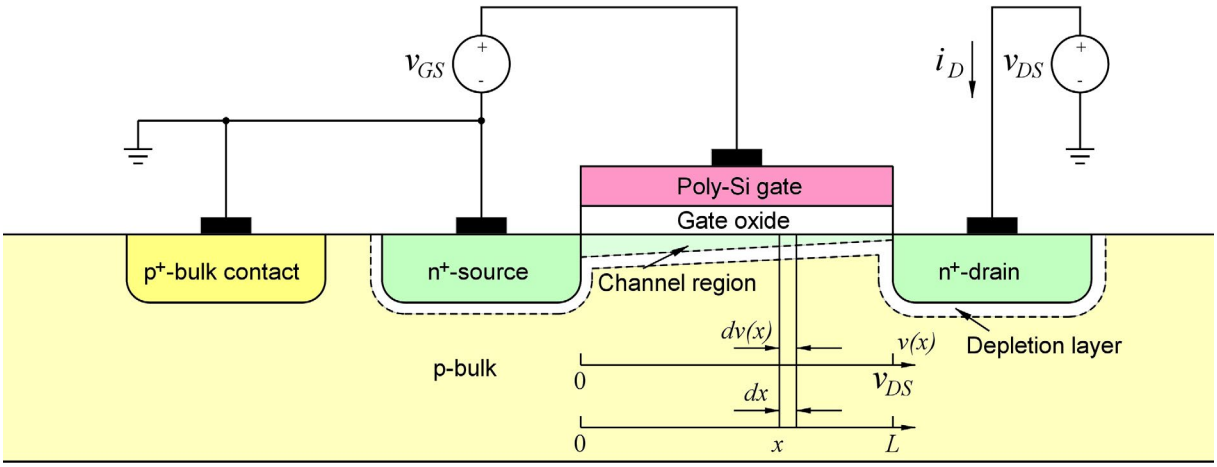


Figure 3.7: Cross-sectional view of an NMOS transistor with electrical connections to gate, drain, source and bulk and a tapered channel corresponding to $v_{DS} < v_{GS} - V_t$.

If v_{DS} is not much smaller than $v_{GS} - V_t$, the voltage across the gate oxide cannot be assumed to be constant along the channel. In the source end of the channel, the voltage across the gate oxide is v_{GS} but in the drain end, it is only $v_{GS} - v_{DS}$. This means that the depth of the channel is reduced when going from source to drain as illustrated in Fig. 3.7. When considering a small section of the channel from distance x to $x + dx$ from the source, see Fig. 3.7, we can utilize that Kirchhoff's current law requires the current i_D to be the same for any value of x and re-write Eq. (3.4) to

$$i_D = \mu_n C_{ox} (W/dx) (v_{GS} - v(x) - V_t) dv(x) \tag{3.5}$$

where $v(x)$ is the voltage in the channel at position x and $dv(x)$ is the voltage drop across the channel segment dx. From Eq. (3.5), we find

$$i_D dx = \mu_n C_{ox} W (v_{GS} - v(x) - V_t) dv(x) \tag{3.6}$$

and integrating with integration limits of 0 and L for x and the corresponding limits of 0 and v_{DS} for $v(x)$, we find

$$\int_0^L i_D dx = \mu_n C_{ox} W \int_0^{v_{DS}} (v_{GS} - v(x) - V_t) dv(x) \tag{3.7}$$

Inserting the integration limits, Eq. (3.7) results in

$$\begin{aligned} i_D L &= \mu_n C_{ox} W \left[(v_{GS} - V_t) v_{DS} - v_{DS}^2/2 \right] \\ \Rightarrow i_D &= \mu_n C_{ox} \left(\frac{W}{L} \right) \left[(v_{GS} - V_t) v_{DS} - v_{DS}^2/2 \right] \end{aligned} \tag{3.8}$$

Notice that in the derivation of Eq. (3.8), we have assumed that there is an electric field to generate a channel all along the region between the source and the drain. This implies that $v_{GS} - v(x) - V_t$ must be positive for all values of x, and with $v(x)$ reaching a maximum value of v_{DS}, we find

$$v_{GS} - v_{DS} - V_t \geq 0 \Rightarrow v_{DS} \leq v_{GS} - V_t \tag{3.9}$$

When $v_{DS} = v_{GS} - V_t$, the channel depth at the drain end of the channel is 0, and inserting $v_{DS} = v_{GS} - V_t$ in Eq. (3.8) results in

$$i_D = \frac{\mu_n C_{ox}}{2} \left(\frac{W}{L} \right) (v_{GS} - V_t)^2 \tag{3.10}$$

When v_{DS} is increased beyond $v_{GS} - V_t$, the channel is pinched off at the drain. The voltage in the end of the channel is $v_{GS} - V_t$ and the additional voltage $v_{DS} - (v_{GS} - V_t)$ appears across the depletion region surrounding the drain. Thus, the current i_D flows through the pinch-off region due to the electric field appearing across the region. With the channel profile remaining the same as for $v_{DS} = v_{GS} - V_t$, the current in the channel follows Eq. (3.10), so to a first-order approximation, Eq. (3.10) is valid for $v_{DS} \geq v_{GS} - V_t$. Thus, Eqs. (3.8) and (3.10) describe the relation between the drain current, the gate-source voltage and the drain-source voltage for $v_{GS} > V_t$ and $v_{DS} \geq 0$.

For $v_{GS} \leq V_t$, the drain current is assumed to be 0. This is called the cut-off region.

For $v_{DS} \leq v_{GS} - V_t$, the drain current follows Eq. (3.8). This is called the triode region. Sometimes the term linear region is used to indicate the linear relation between i_D and v_{DS} for very small values of v_{DS}.

For $v_{DS} > v_{GS} - V_t$, the drain current follows Eq. (3.10). This is called the active region or the saturation region of the MOS transistor.

The drain-source voltage at the border between the triode region and the active region is called the saturation voltage $v_{DS\,\mathrm{sat}}$. The drain current for $v_{DS} = v_{DS\,\mathrm{sat}}$ is given by Eq. (3.10) with $v_{GS} - V_t = v_{DS}$, i.e.,

$$i_D = \frac{\mu_n C_{ox}}{2} \left(\frac{W}{L} \right) (v_{DS\,\mathrm{sat}})^2 \tag{3.11}$$

Equation (3.10) describes the drain current as a function of the gate-source voltage when the transistor is in the active region. For the active region, the drain current does not depend on the drain voltage, and Eq. (3.10) can be shown graphically as a single curve, valid for all values of $v_{DS} \geq v_{GS} - V_t$, see the black curve Fig. 3.8. For $v_{DS} < v_{GS} - V_t$, the transistor is in the triode region for large values of v_{GS}, see the colored curves in Fig. 3.8. The characteristics shown in Fig. 3.8 are called the input characteristics of the transistor.

The relations (3.8) and (3.10) may also be used to show the drain current as a function of the drain-source voltage for different values of the gate-source voltage, see Fig. 3.9. The characteristics shown in Fig. 3.9 are called the output characteristics of the transistor.

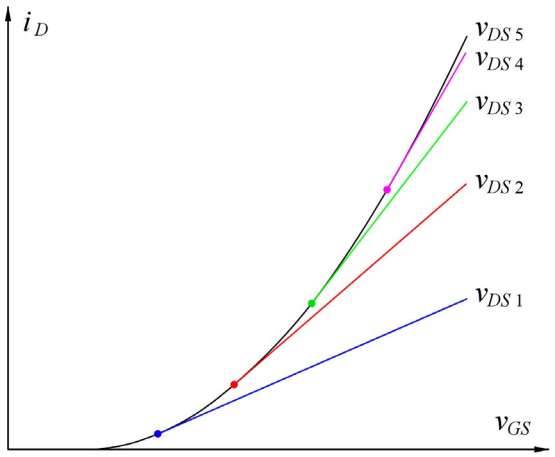


Figure 3.8: Drain current versus gate-source voltage for an n-channel MOS transistor. The colored dots show the border between the active region (small v_{GS}) and the triode region (large v_{GS}). For the largest value of v_{DS}, i.e., v_{DS5}, the transistor is in the active region for the range of v_{GS} shown in the figure.

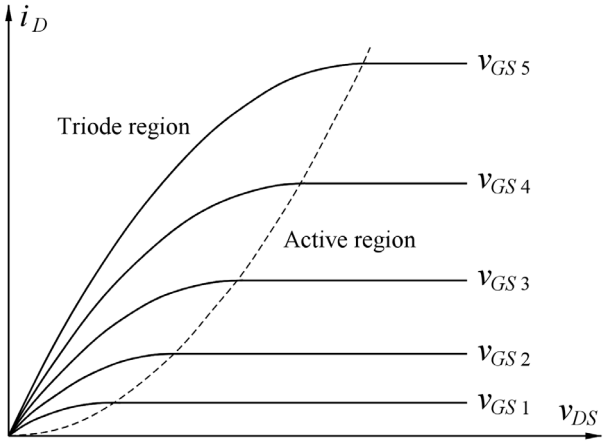


Figure 3.9: Drain current versus drain-source voltage for an n-channel MOS transistor. The dashed curve shows the border between the triode region and the active region given by Eq. (3.11).

The relations (3.8) and (3.10) are the simplest relations describing the electrical behavior of the NMOS transistor. Often they are sufficient for rough hand calculations. However, in many situations, the assumption that i_D does not depend on v_{DS} when the transistor is in the active region turns out to be too inaccurate. In reality, the drain current increases somewhat with increasing drain-source voltage, even when the transistor is in the active region. This can be explained by the phenomenon called the channel-length modulation. When the transistor is in the active region, the channel is pinched off at the drain and a depletion region is formed between the drain and the channel, see Fig. 3.10. This depletion region pushes the channel back towards the source and reduces the effective length of the channel by a small amount ΔL as shown in Fig. 3.10, so that the effective channel length is $L_{\text{eff}} = L - \Delta L$. The channel-length modulation ΔL increases with increasing drain voltage.

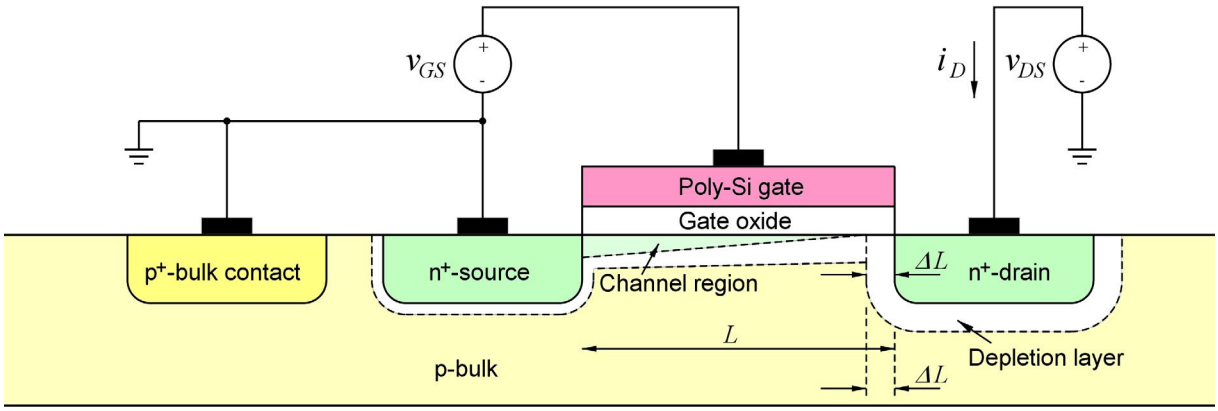


Figure 3.10: Cross-sectional view of an NMOS transistor with electrical connections to gate, drain, source and bulk and a tapered channel corresponding to $v_{DS} > v_{GS} - V_t$ where channel-length modulation occurs.

Using the very course simplification $\Delta L \simeq \lambda' v_{DS}$ and assuming $\Delta L \ll L$, we can modify Eq. (3.10) as follows:

$$\begin{aligned} i_D &= \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L_{\text{eff}}}\right)(v_{GS}-V_t)^2 \qquad (3.12)\\ \Rightarrow i_D &= \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)\left(\frac{1}{1-\Delta L/L}\right)(v_{GS}-V_t)^2 \\ \Rightarrow i_D &\simeq \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(1+\Delta L/L)(v_{GS}-V_t)^2 \\ \Rightarrow i_D &\simeq \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)\left(1+(\lambda'/L)v_{DS}\right)(v_{GS}-V_t)^2 \\ \Rightarrow i_D &\simeq \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(1+\lambda v_{DS})(v_{GS}-V_t)^2 \qquad (3.13)\end{aligned}$$

In Eq. (3.13), we have introduced the channel-length modulation parameter $\lambda = \lambda'/L$ where λ' is a technology parameter. Notice that λ is inversely proportional to the channel length L, reflecting that the channel-length modulation is more pronounced for transistors with a very short channel than for transistors with a long channel. Some textbooks use the parameter $V_A = 1/\lambda$ for describing the channel-length modulation. The parameter V_A is called the Early voltage, named after James M. Early (1922-2004), an American engineer.

When using Eq. (3.13) instead of Eq. (3.10) for the active region, there is a discontinuity between the characteristics given by Eq. (3.8) for the triode region and the characteristics for the active region. Clearly, this is not acceptable for an accurate device modeling. A simple way to solve the problem is to use the factor $(1+\lambda v_{DS})$ also for the triode region and this leads to the following set of equations, the Shichman-Hodges model:

$$i_D = 0;\ v_{GS} \le V_t \qquad (3.14)$$

$$i_D = \mu_n C_{ox}\left(\frac{W}{L}\right)[(v_{GS}-V_t)v_{DS} - v_{DS}^2/2](1+\lambda v_{DS});\ 0 \le v_{DS} \le v_{GS}-V_t \qquad (3.15)$$

$$i_D = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(v_{GS}-V_t)^2(1+\lambda v_{DS});\ 0 \le v_{GS}-V_t \le v_{DS} \qquad (3.16)$$

An alternative way to eliminate the discontinuity is to use a factor $(1+\lambda(v_{DS}-(v_{GS}-V_t)))$ rather than $(1+\lambda v_{DS})$ (Chan Carusone, Johns & Martin 2012) but since the expressions given in Eqs. (3.14) - (3.16) are used by SPICE for the Spice level 1 transistor model, we use these expressions in this book in order to be able to compare hand calculations directly to LTspice simulations using the Shichman-Hodges model.

The body effect. In the derivation of the Shichman-Hodges transistor model, we have assumed that source and bulk are connected. This may not always be the case. Consider for instance the NMOS transistors in a CMOS process like shown in Fig. 3.5. They share a common bulk (or body), that is the p-type substrate which has to be connected to the most negative voltage in the circuit. Hence, NMOS transistors which do not have their source connected to the most negative voltage may have a positive source-bulk voltage v_{SB}, reverse-biasing the source-bulk pn junction. With an increase in v_{SB}, the width of the depletion layer between the channel and the substrate increases, so the channel is pushed towards the silicon surface and the channel depth is reduced. A higher value of v_{GS} is required to maintain the same channel depth. This effect is called the body effect or the bulk effect, and it can be modeled as an increase in the threshold voltage of the transistor. From the device physics, it can be shown that the threshold voltage depends on the source-bulk voltage v_{SB} following the relation

$$V_t = V_{to} + \gamma(\sqrt{v_{SB}+|2\Phi_F|} - \sqrt{|2\Phi_F|}) \tag{3.17}$$

where V_{to} is the threshold voltage with $v_{SB}=0$, γ is the bulk threshold parameter (or body effect constant) and $|\Phi_F|$ is the Fermi potential of the body (Chan Carusone, Johns & Martin 2012). These parameters are all related to the manufacturing process and cannot be controlled by the circuit designer.

PMOS transistors. The transistor used for deriving the Shichman-Hodges model in the previous section was an NMOS transistor. The NMOS transistor normally operates with positive values of v_{GS} and v_{DS} and the drain current is positive when defined with the positive direction *into* the drain terminal. Also, normally the threshold voltage is positive for an NMOS transistor.

As shown in Fig. 3.5, also PMOS transistors are available in a CMOS process. For the PMOS transistor, the relation between drain current, gate-source voltage and drain-source voltage can be derived in the same way as for the NMOS transistor and the resulting equations are the same as Eqs. (3.14) - (3.16), provided that the drain current is defined with the positive direction *out of* the drain terminal, and provided that absolute values are used for all voltages. For a PMOS transistor, normally the threshold voltage is negative, and also v_{GS}, v_{DS} and v_{SB} are negative in the normal regions of operation which are defined as follows:

Cut-off region: $V_t \le V_{GS}$ (or $|V_{GS}| \le |V_t|$). (3.18)

Triode region: $V_{GS}-V_t \le V_{DS} \le 0$ (or $0 \le |V_{DS}| \le |V_{GS}-V_t| = |V_{GS}|-|V_t|$). (3.19)

Active region: $V_{DS} \le V_{GS}-V_t \le 0$ (or $0 \le |V_{GS}-V_t| = |V_{GS}|-|V_t| \le |V_{DS}|$). (3.20)

The fact that the voltages are negative for a PMOS transistor seems quite obvious. Nevertheless, it is the cause of many sign errors when analyzing circuits with PMOS transistors, so be careful with the signs.

Transistor symbols. For drawing schematics with MOS transistors, we need symbols for both PMOS transistors and NMOS transistors. Figure 3.11 shows the symbols normally used in textbooks together with the normal sign conventions for voltages and currents. The direction of current is indicated also for the gate current i_G and the bulk current i_B, but at dc, these currents are 0. The gate is isolated from the bulk and channel by the gate oxide which is a perfect insulator, and the bulk current i_B is the current for a reverse-biased pn junction, and that is an extremely small current which can normally be neglected. At high frequencies, these currents are not 0. They must be calculated taking the capacitive effects of the transistor into account, see Fig. 3.33.

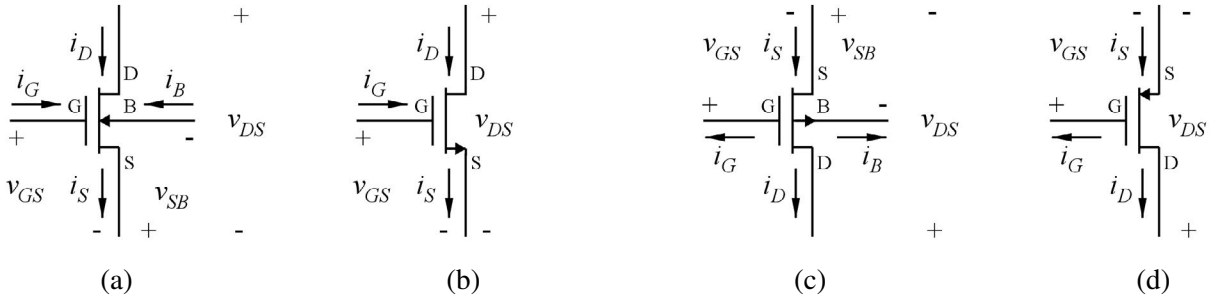


Figure 3.11: Normal textbook definitions of sign conventions for transistor currents and voltages. NMOS transistor (a) and (b). PMOS transistor (c) and (d).

For each of the transistors, we have a symbol showing both gate, drain, source and bulk connections, and the arrow on the bulk connection illustrates the pn junction between channel and bulk in the same way as the arrow in the diode symbol, Fig. 3.2. Occasionally, we may use a simplified symbol without explicitly showing the bulk contact. This is useful if all bulk contacts are connected to the same node, normally the most negative voltage for NMOS transistors and the most positive voltage for PMOS transistors (Chan Carusone, Johns & Martin 2012). The simplified symbol may also be used for transistors where source and bulk are connected or where the body effect can be neglected (Sedra & Smith 2016). For the simplified symbols, the arrow on the source terminal shows the positive direction of the source current. Notice that the NMOS symbols have the source terminal pointing downwards while the PMOS symbols have the source terminal as the upper terminal. This ensures that the current flow in the transistor is downwards for both NMOS and PMOS transistors. Also, the voltage levels are higher in the upper terminal of the transistors than in the lower terminal, and this corresponds to the normal conventions for schematic drawings.

In LTspice, similar transistor symbols are found. For transistors for integrated circuit design, the symbols ‘nmos4’ and ‘pmos4’ are used since they permit specification of layout parameters and device models

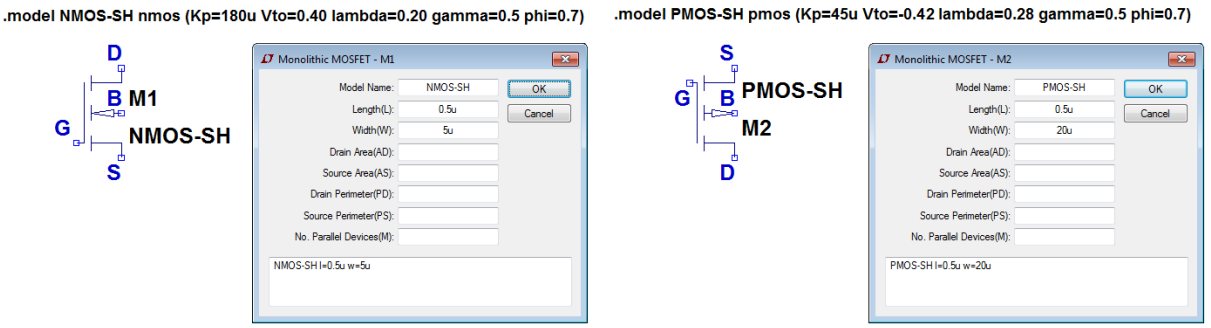


Figure 3.12: MOS transistor symbols and specification windows in LTspice.

used for the transistors. The symbols are shown in Fig. 3.12 together with the specification windows for the transistors and model specifications for both the NMOS transistor and the PMOS transistor using parameters for the Shichman-Hodges model. The window for specifying a transistor is opened by a right-click on the transistor symbol. Here, you must specify a transistor model name corresponding to the name in model specification, and you must specify channel length and width. You may also specify areas and perimeters for the source and drain regions. They are used for calculating the capacitances of the reverse-biased junctions from source/drain to substrate.

Also for the LTspice symbols, the NMOS transistor is shown with the source pointing downwards while the source for the PMOS transistor is pointing upwards. This requires that the PMOS symbol is rotated and mirrored when inserted.

Notice that the sign conventions for LTspice dictate that the positive direction of current flow is *into* the transistor, regardless of the terminal and transistor type. For more details concerning the specification of transistor models and parameters, see Bruun (2020, Tutorial 3.1). In particular, note that the model parameters for the Shichman-Hodges model given by Eqs. (3.14) - (3.16) are '`Kp`', '`Vto`', '`Lambda`', '`Gamma`' and '`Phi`', corresponding to μC_{ox}, V_{to}, λ, γ and $|2\Phi_F|$, respectively. Since LTspice is case-insensitive, it does not matter if you specify the parameters with capital letters (e.g., '`Lambda`') or small letters (e.g., '`lambda`').

Shichman-Hodges model parameters. The parameters for the Shichman-Hodges transistor model depend on the fabrication technology. Process technologies presently in use typically have minimum dimensions in the range from 28 nm to 350 nm for processes used for analog CMOS circuits. Generally, the smaller the dimensions, the smaller is $|V_{to}|$, and the larger is μC_{ox} and λ. Normally, the mobility μ_n for an NMOS transistor is larger by a factor of 3 to 4 compared to the mobility μ_p for a PMOS transistor. In this book, we apply the model parameters shown in Table 3.1 for the Shichman-Hodges model, unless otherwise indicated. Be aware that in Table 3.1, the channel-length modulation is specified by the parameter $\lambda' = \lambda L$.

Parameter:	μC_{ox}	V_{to}	$\lambda' = \lambda L$	γ	$\lvert 2\Phi_F \rvert$
NMOS:	180 µA/V^2	0.40 V	0.10 µm/V	0.5 $\sqrt{\text{V}}$	0.7 V
PMOS:	45 µA/V^2	−0.42 V	0.14 µm/V	0.5 $\sqrt{\text{V}}$	0.7 V

Table 3.1: Shichman-Hodges transistor parameters for a generic 0.18 µm CMOS process.

The parameters are selected to be representative of a generic 0.18 µm process, i.e., a process with 0.18 µm as the minimum dimension for L. For analog circuits, you would normally use a channel length which is at least 3 to 10 times the minimum length. A comprehensive discussion concerning the selection of L can be found in Binkley (2008). The maximum supply voltage for a 0.18 µm process is typically 1.8 V. Many textbooks provide a selection of model parameters for different technologies, e.g., Baker (2010), Chan Carusone, Johns & Martin (2012) and Sedra & Smith (2016), but for the purpose of an introductory course, the parameters given in Table 3.1 will suffice.

3.4 Examples of the use of the Shichman-Hodges transistor model

Before proceeding with small-signal transistor models and more advanced models for simulation, we will examine a few simple circuits in order to gain some experience with the Shichman-Hodges model. Initially, we examine the circuits using hand calculations based on the models given by Eqs. (3.14) - (3.16), and subsequently, we use LTspice to verify the hand calculations and to provide more accurate results and insight into the properties of the circuits. Since LTspice is using exactly the same transistor models, this provides a way of evaluating the hand calculations as illustrated in Fig. 1.5.

A diode-connected NMOS transistor. The first example illustrates the use of transistor model Eq. (3.16) in combination with Ohm's law and the general circuit theorems from Chapter 2. A diode-connected transistor is a transistor with a direct connection between drain and gate as shown in Fig. 3.13.

For this circuit, we will design the resistor R_1 so that the drain current of transistor M_1 is $I_D = 200$ µA. For the transistor, we assume the parameters given in Table 3.1, and we assume $L = 1$ µm and $W = 25$ µm. The supply voltage is $V_{DD} = 1.8$ V.

We notice that Kirchhoff's current law implies that the drain current I_D is the same as the current in the resistor R_1. The approach to the solution is to find the gate voltage V_G resulting in a drain current of 200 µA and then use Kirchhoff's voltage law and Ohm's law to find R_1 as $R_1 = (V_{DD} - V_G)/I_D$.

Since the transistor has $V_{DS} = V_{GS}$, the condition $V_{DS} > V_{GS} - V_t$ is fulfilled, so the transistor is in the active region and Eq. (3.16) applies. With $V_{GS} = V_{DS} = V_G$, this gives

$$I_D = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_G - V_t)^2(1 + \lambda V_G) \tag{3.21}$$

This is a cubic equation from which V_G can be found. However, closed-form solutions to a cubic equation are quite complicated and not suited for hand calculations, so assuming that the channel-length

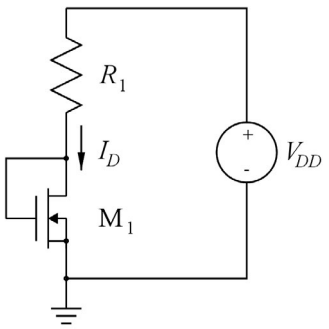


Figure 3.13: A diode-connected NMOS transistor.

modulation has only a minor influence on I_D, we can simplify Eq. (3.21) by using $\lambda = 0$.

This leads to a quadratic equation:

$$I_D = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_G - V_t)^2 \tag{3.22}$$

Inserting the specified values, we find:

$$\begin{aligned} 200\ \mu\text{A} &= \frac{180\ \mu\text{A/V}^2}{2}\left(\frac{25\ \mu\text{m}}{1\ \mu\text{m}}\right)(V_G - 0.40\ \text{V})^2 \\ \Rightarrow (V_G - 0.40\ \text{V})^2 &= 0.0889\ \text{V}^2 \\ \Rightarrow\ V_G - 0.40\ \text{V} &= \pm 0.298\ \text{V} \end{aligned} \tag{3.23}$$

We notice that Eq. (3.23) shows two solutions, a positive value and a negative value for $V_G - V_t$. Since M_1 is an NMOS transistor, we know that $V_{GS} - V_t$ must be positive when the transistor is in the active region, so we select the positive solution, i.e.,

$$V_G - 0.40\ \text{V} = 0.298\ \text{V} \Rightarrow V_G = 0.698\ \text{V} \tag{3.24}$$

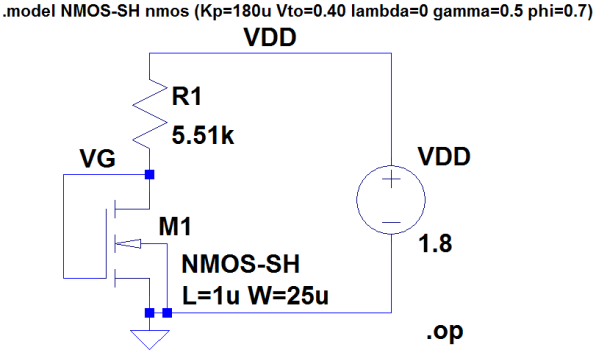
Finally, we calculate

$$R_1 = \frac{V_{DD} - V_G}{I_D} = 5.51\ \text{k}\Omega \tag{3.25}$$

Next, we may use LTspice to solve the problem. In order to verify the solution derived by hand calculations, we use $\lambda = 0$ in the transistor model, and we insert the value for R_1 calculated above.

The resulting schematic is shown in Fig. 3.14 together with the output file resulting from a 'DC operating point analysis', SPICE directive '.op'. From the output file, we see that the simulated values for '`V(vg)`' and '`Id(M1)`' (rounded off to three digits) are the same as the values found from the hand calculations, thus confirming the results of the hand calculation.

Finally, we may use LTspice to find a more exact value for R_1, taking the channel-length modulation into account. The channel-length modulation will cause the drain current to be slightly larger than expected



Output file

```
        --- Operating Point ---

V(vg):          0.698126        voltage
V(vdd):         1.8             voltage
Id(M1):         0.000199978     device_current
Ig(M1):         0               device_current
Ib(M1):         -7.08126e-013   device_current
Is(M1):         -0.000199978    device_current
I(R1):          0.000199977     device_current
I(Vdd):         -0.000199977    device_current
```

Figure 3.14: LTspice schematic and output file for the circuit from Fig. 3.13.

from the simple calculation because of the factor $(1 + \lambda V_G)$ in Eq. (3.21). In order to compensate for this, R_1 should be increased. We can find R_1 by stepping the value of R_1 through a suitable range using a '.step' directive in LTspice with the value of R_1 defined as a parameter '`RD`', see Fig. 3.15.

In this figure, we also show the result of the simulation, i.e., a plot showing '`Id(M1)`' versus '`RD`'. From the plot, we find $R_1 = 5.55\ \text{k}\Omega$ which is (as expected) a value slightly larger than the approximate value found before.

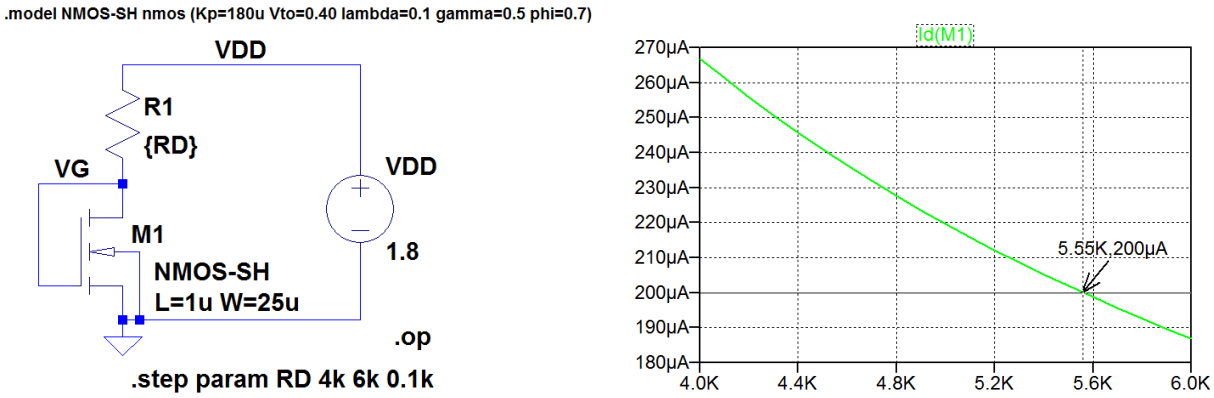


Figure 3.15: LTspice schematic and simulation plot for a sweep of I_D versus R_1.

An NMOS current mirror. Now we connect an additional NMOS transistor to the diode-connected transistor as shown in Fig. 3.16. Transistor M_2 is assumed to be identical to M_1, i.e., $L = 1$ µm and $W = 25$ µm. We notice that V_{GS} is the same for M_1 and M_2 since they have their gates connected together and they both have their sources connected to ground. Assuming that both transistors are in the active region, an approximate calculation neglecting the channel-length modulation, i.e., using Eq. (3.22), leads to the same drain current for M_2 as for M_1, so $I_{D2} = I_{D1}$.

The circuit is called a current mirror since it mirrors the current from the input transistor M_1 to the output transistor M_2. Using Eq. (3.22), we find that the output is an ideal current source delivering the same current regardless of the value of the drain resistor R_2 connected from the drain of M_2 to the supply

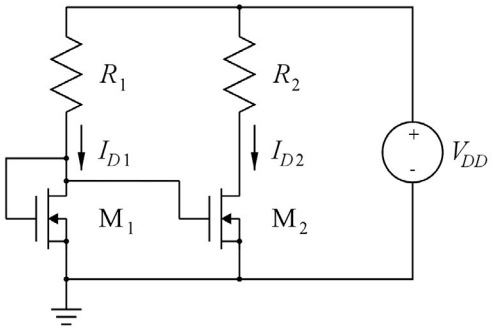


Figure 3.16: An NMOS current mirror.

voltage V_{DD}. This makes the circuit very useful as a current source for supplying bias currents to other circuit configurations, for example gain stages. In Chapter 4, we will see several examples of gain stages requiring a constant current to bias a transistor providing a voltage gain.

With two identical transistors, the output current is the same as the input current. However, the output current is easily scaled with respect to the input current by a scaling of W/L. The current scaling factor is $A_I = (W/L)_2/(W/L)_1$. The precision of the scaling depends on the matching of the transistors. For a good matching, it is advisable to scale W rather than L. For an integer scaling factor, the scaling can be obtained simply and accurately by connecting identical transistors in parallel, i.e., transistors where all drain terminals are connected together, all source terminals are connected together, and all gate terminals are connected together. For example, $A_I = 3$ is obtained by designing M_2 as three identical parallel-connected transistors of the same size as M_1.

However, the circuit has some limitations. If the drain resistor R_2 is too large, the voltage drop $V_{R2} = R_2 I_{D2}$ across R_2 becomes too large to fit within the supply voltage range V_{DD}.

This implies that the drain voltage V_{DS2} falls below the value required to keep M_2 in the active region. From Kirchhoff's voltage law, we have

$$V_{DD} = V_{D2} + V_{R2} = V_{D2} + R_2 I_{D2} \tag{3.26}$$

For small values of R_2, the drain current I_{D2} is given by Eq. (3.22) but for large values of R_2, the drain current drops because M_2 is no longer in the active region. Rather, it is in the triode region where the current is given by Eq. (3.8), neglecting the factor $(1+\lambda v_{DS})$. Combining Eqs. (3.8) and (3.26), we find

$$I_{D2} = \mu_n C_{ox} \left(\frac{W}{L}\right) \left[(V_G - V_t) V_{D2} - V_{D2}^2/2\right] = \frac{V_{DD} - V_{D2}}{R_2} \tag{3.27}$$

where V_G is found from Eq. (3.22).

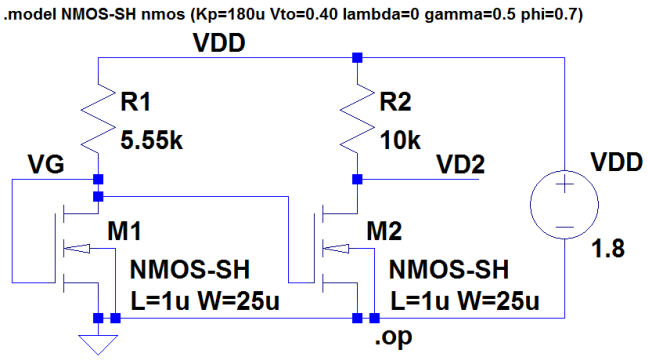
This is a quadratic equation for finding V_{D2}. As an example, we will find V_{D2} and I_{D2} for $R_2 = 10\ \text{k}\Omega$, using the value of V_G found in Eq. (3.24):

$$\begin{aligned} 0.18\ \text{mA/V}^2 \times 25 \times (0.298\ \text{V} \times V_{D2} - V_{D2}^2/2) &= (1.8\ \text{V} - V_{D2})/10\ \text{k}\Omega \\ \Rightarrow V_{D2} &= \begin{cases} 0.471\ \text{V} \\ 0.171\ \text{V} \end{cases} \end{aligned} \tag{3.28}$$

We find two solutions to the equation but as we know that $V_{D2} < V_G - V_t = 0.298$ V when M_2 is in the triode region, we select the solution $V_{D2} = 0.171$ V from Eq. (3.28). With this value of V_{D2}, we may calculate

$$I_{D2} = \frac{V_{DD} - V_{D2}}{R_2} = 163\ \mu\text{A} \tag{3.29}$$

We may verify the result using LTspice. Figure 3.17 shows the LTspice schematic with the Shichman-Hodges model with $\lambda = 0$, and from the resulting output file, we find the same values for V_{D2} and I_{D2} as found by hand calculations.



Output file

```
      --- Operating Point ---

V(vg):          0.697177        voltage
V(vdd):         1.8             voltage
V(vd2):         0.171022        voltage
Id(M2):         0.000162898     device_current
Ig(M2):         0               device_current
Ib(M2):         -1.81008e-013   device_current
Is(M2):         -0.000162898    device_current
Id(M1):         0.000198707     device_current
Ig(M1):         0               device_current
Ib(M1):         -7.07177e-013   device_current
Is(M1):         -0.000198707    device_current
I(R2):          0.000162898     device_current
I(R1):          0.000198707     device_current
I(Vdd):         -0.000361605    device_current
```

Figure 3.17: LTspice schematic and output file for the circuit from Fig. 3.16.

Using the requirement for M_2 that V_{DS2} must be larger than $V_{GS2}-V_t$, we can easily calculate the maximum value of R_2 for which M_2 operates as a constant current source with $I_D = I_{D2} = I_{D1}$:

$$\begin{aligned} V_{DS2} &> V_{GS2} - V_t \\ \Rightarrow V_{DD} - R_2 I_{D2} &> V_{DD} - R_1 I_{D1} - V_t \\ \Rightarrow R_2 &< R_1 + V_t/I_D \\ \Rightarrow R_{2\max} &= R_1 + V_t/I_D = 5.55\ \text{k}\Omega + 0.4\ \text{V}/0.2\ \text{mA} = 7.55\ \text{k}\Omega \end{aligned} \tag{3.30}$$

The result obtained in Eq. (3.30) may also be verified by LTspice: In the schematic shown in Fig. 3.17, we define 'R2' as a parameter 'RD2' to be stepped from 1 Ω (LTspice does not accept 0) to 20 kΩ and run the '.op' simulation. The resulting plot of I_{D2} versus R_2 is shown in Fig. 3.18. Also shown in Fig. 3.18 is a plot of I_{D2} versus V_{D2} obtained from the same simulation by changing the x-axis variable to 'V(VD2)', using a right-click on the axis to open the window for specifying the variable.

Another limitation of the circuit is caused by the channel-length modulation. Taking this into account, the output current is given by

$$I_{D2} = \frac{\mu_n C_{ox}}{2}\left(\frac{W_2}{L_2}\right)(V_G - V_t)^2(1+\lambda V_{D2}) \tag{3.31}$$

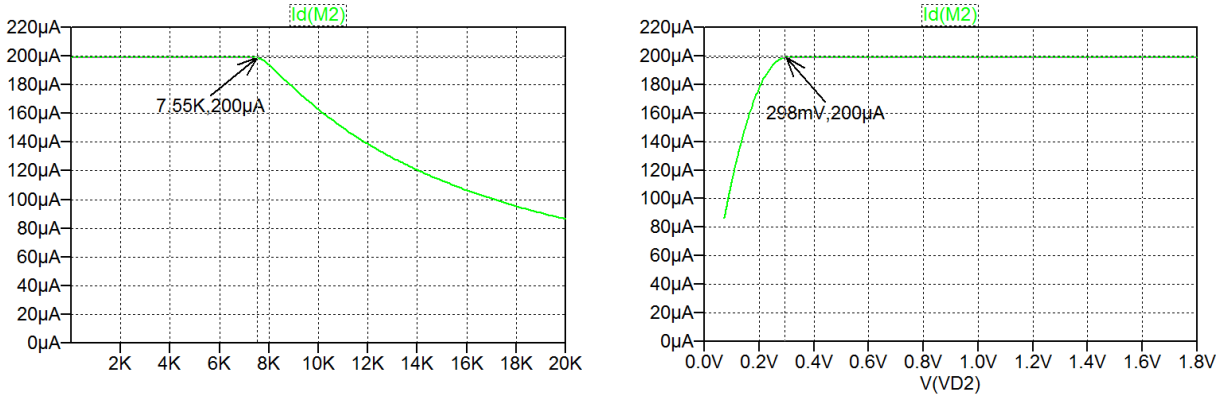


Figure 3.18: Simulation of current-mirror output current I_{D2} versus drain resistor R_2 and output voltage V_{D2}.

where V_G has the value found for the diode-connected transistor with the channel-length modulation taken into account, i.e., $V_G = 1.80\ \text{V} - 5.55\ \text{k}\Omega \times 0.2\ \text{mA} = 0.69\ \text{V}$. The channel-length modulation causes the output current to change slightly with a changing value of V_{D2}. In the voltage range where M_2 is in the active region, Eq. (3.31) shows that I_{D2} is a linear function of V_{D2}.

The change may be characterized by a load regulation defined as $\Delta I_{D2}/\Delta V_{D2}$. The load regulation can be found as the derivative of i_{D2} with respect to v_{D2}, calculated for a bias point $v_{D2} = V_{D2}$. For the bias point, we may select $V_{D2} = V_G$ since this is the value of V_{D2} resulting in $I_{D2} = I_{D1}$ as M_1 and M_2 has identical gate-source voltages and drain-source voltages for this value of V_{D2}.

From Eq. (3.31), we find

$$\frac{\Delta I_{D2}}{\Delta V_{D2}} = \frac{\mu_n C_{ox}}{2}\left(\frac{W_2}{L_2}\right)(V_G - V_t)^2\lambda = \frac{180\ \mu\text{A/V}^2}{2}\left(\frac{25\ \mu\text{m}}{1\ \mu\text{m}}\right)(0.69\ \text{V} - 0.4\ \text{V})^2 \times 0.1\ \text{V}^{-1} = 18.7\ \mu\text{A/V} \tag{3.32}$$

The load regulation may be interpreted as the small-signal output conductance of the current mirror. In Section 3.5, we discuss the concept of small-signal modeling in detail.

The characteristics of the current mirror with channel-length modulation may also be illustrated using LTspice. Figure 3.19 shows an LTspice schematic where the load resistor R_{D2} has been replaced by a

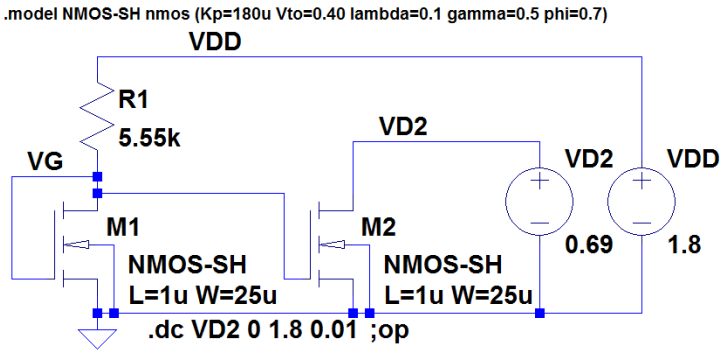


Figure 3.19: LTspice schematic for simulation of the NMOS current mirror with channel-length modulation.

voltage source V_{D2} so that the output current can be simulated directly as a function of the output voltage. Also, the value of R_{D1} has been specified to 5.55 kΩ as found for the diode-connected transistor with channel-length modulation. In Fig. 3.19, a dc sweep of V_{D2} from 0 to 1.8 V is specified. The resulting plot is shown in Fig. 3.20. From the plot, we see that the relation between I_{D2} and V_{D2} is indeed linear as expected for $V_{D2} > 0.298$ V, and using the cursors, we can verify $\Delta I_{D2}/\Delta V_{D2} = 18.7$ µA/V.

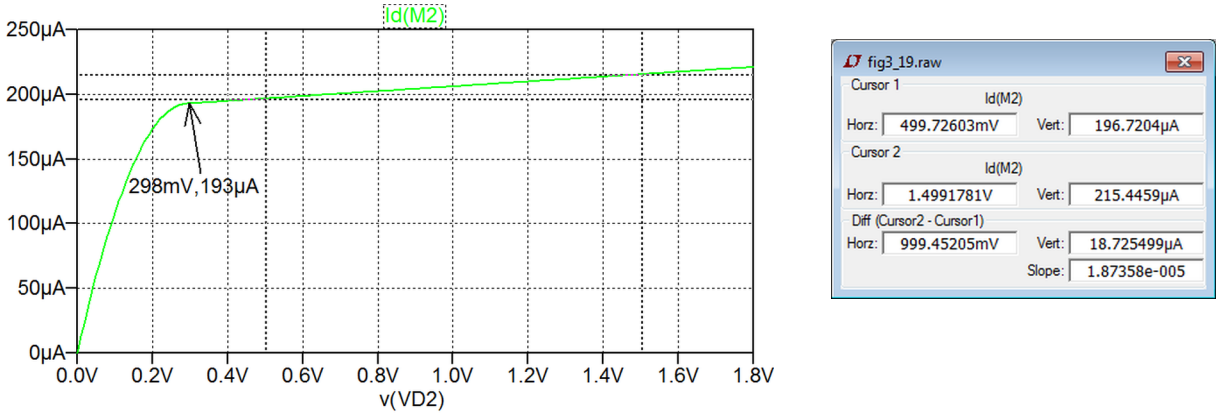


Figure 3.20: Simulation of current-mirror output current I_{D2} versus output voltage V_{D2} for the current mirror with channel-length modulation. From the cursor positions, we find a load regulation of $\Delta I_{D2}/\Delta V_{D2} = 18.7$ µA/V.

A PMOS current mirror. The circuit analyzed in the previous example was an NMOS current mirror. Of course, a current mirror can also be implemented with PMOS transistors. Figure 3.21 shows a PMOS current mirror similar to the NMOS current mirror shown in Fig. 3.16. In Fig. 3.21, I_{D1} is the input current to the mirror and I_{D2} is the output current.

This example is shown primarily to illustrate the selection of the correct signs for solutions to quadratic equations and equations involving absolute values.

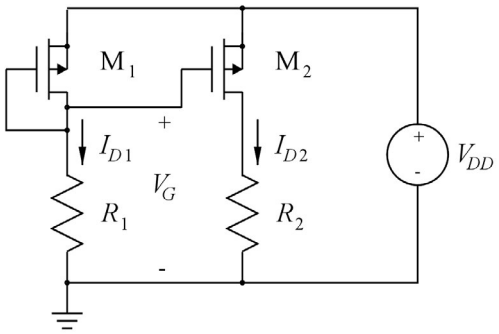


Figure 3.21: A PMOS current mirror.

The analysis of the PMOS current mirror is very similar to the analysis of the NMOS current mirror, so we will not complete all details of the analysis. Rather, we will show the hand calculation of the input side of the current mirror, i.e., we will find the gate voltage V_G resulting in an input current $I_{D1} = 200$ µA.

For the transistors, we assume the Shichman-Hodges parameters from Table 3.1 and $W = 100$ µm and $L = 1$ µm. Notice that we use a larger channel width for the PMOS transistor than for the NMOS transistors in the previous examples. This is in order to compensate for the difference in mobility between NMOS transistors and PMOS transistors. With a transistor width which is 4 times the width of the NMOS transistor, we can expect an effective gate voltage with the same absolute value as the value found for the NMOS transistor since the current is the same as in the previous examples.

Just as for the diode-connected NMOS transistor, we will neglect the channel-length modulation in the hand calculations in order to avoid dealing with a cubic equation. This means that the equation for the input current is:

$$I_{D1} = \frac{\mu_p C_{ox}}{2}\left(\frac{W}{L}\right)(|V_{GS}| - |V_t|)^2 \tag{3.33}$$

We are using absolute values in the equation in order to avoid confusion concerning the signs of the voltages when dealing with PMOS transistors.

Inserting the specified values, we find:

$$\begin{aligned} 200\ \mu\text{A} &= \frac{45\ \mu\text{A/V}^2}{2}\left(\frac{100\ \mu\text{m}}{1\ \mu\text{m}}\right)(|V_{GS}| - 0.42\ \text{V})^2 \\ \Rightarrow (|V_{GS}| - 0.42\ \text{V})^2 &= 0.0889\ \text{V}^2 \\ \Rightarrow\ |V_{GS}| - 0.42\ \text{V} &= \pm 0.298\ \text{V} \end{aligned} \tag{3.34}$$

We see that Eq. (3.34) shows two solutions, a positive value and a negative value for $|V_{GS}| - |V_t|$. For both NMOS transistors and PMOS transistors, the absolute value of the gate-source-voltage is larger than the absolute value of the threshold voltage when the transistor is in the active region. Hence, we select the positive solution

$$|V_{GS}| - 0.42\ \text{V} = 0.298\ \text{V} \Rightarrow |V_{GS}| = 0.718\ \text{V} \tag{3.35}$$

From Eq. (3.35), we find

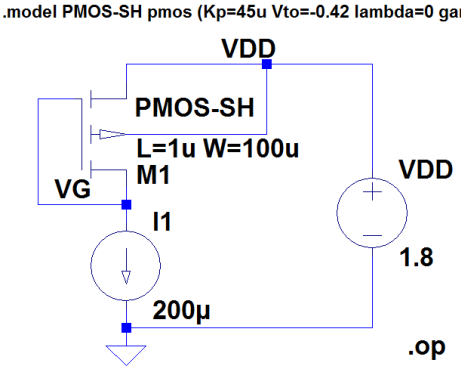
$$V_{GS} = \pm 0.718\ \text{V} \tag{3.36}$$

Now, since M_1 is a PMOS transistor in the active region, the voltage V_{GS} is negative, and more negative than V_t which is also negative. This implies that we must select $V_{GS} = -0.718$ V.

From Kirchhoff's voltage law, we have $V_G = V_{DD} + V_{GS}$, and with $V_{GS} = -0.718$ V, we find $V_G = 1.082$ V.

Incorrect use of the absolute value signs in equations like Eqs. (3.34) - (3.36) is one of the most common errors when analyzing circuits with PMOS transistors, so be careful with circuits including PMOS transistors. Always ensure that V_{GS} is negative and more negative than V_t when there is a current flowing in the transistor.

We may verify the solution for V_G using LTspice with a transistor model with $\lambda = 0$. Figure 3.22 shows an LTspice schematic where the input current is set to 200 µA using a dc current source instead of R_1. The PMOS transistor is inserted using the 'pmos4' symbol rotated and mirrored to have the source terminal upwards. From the output file, we find $V_G = 1.08186$ V, confirming the hand calculation.



Output file

```
        --- Operating Point ---

V(vdd):         1.8             voltage
V(vg):          1.08186         voltage
Id(M1):         -0.000200001    device_current
Ig(M1):         -0              device_current
Ib(M1):         7.28143e-013    device_current
Is(M1):         0.000200001     device_current
I(I1):          0.0002          device_current
I(Vdd):         -0.0002         device_current
```

Figure 3.22: LTspice schematic and output file for a diode-connected PMOS transistor.

3.5 Small-signal models

For the MOS transistor, the relations between gate-source voltage, drain-source voltage, source-bulk voltage and drain current are nonlinear. We even have different nonlinear relations for different regions of operation. This causes the circuit equations based on Kirchhoff's laws to be nonlinear, typically quadratic or cubic as we saw in the examples in Section 3.4. Such equations are difficult to solve analytically using hand calculations, and this not only complicates the circuit analysis. It also makes it difficult to achieve an insight into which parameters are essential for the performance of a circuit.

An approach to ease hand calculations and to simplify ways of getting insight into the operation of a circuit is the use of small-signal analysis. As already mentioned in Chapter 2, the small-signal analysis is based on a linearization of all nonlinear device models so that the equations resulting from Kirchhoff's laws turn out as linear equations. The small-signal analysis always has a specific bias point, i.e., specific values of device currents and voltages, as the starting point for the linearization of the device models, see Fig. 2.12.

By solving the small-signal equations, we can determine the influence of the small-signal model parameters on the circuit performance, and by relating the small-signal parameters to the basic design parameters, i.e., transistor geometry and bias conditions, we can – hopefully – design the circuit to achieve the desired performance, provided this can be achieved using design parameters within a feasible range of variation.

The small-signal analysis is an approximate analysis. Nonlinearities are neglected, and this is only reasonable for signal variations where the linear model is a good approximation to the actual device model. Also, higher-order effects such as harmonic distortion cannot be analyzed from small-signal models since distortion is caused by the nonlinearities of the circuit elements. Before proceeding to

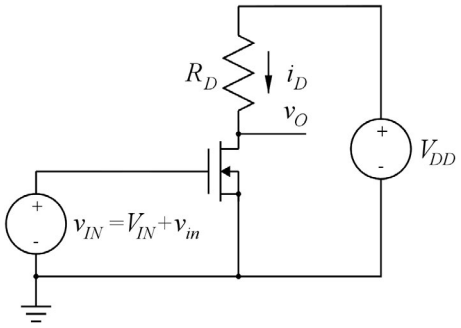


Figure 3.23: An inverting amplifier.

derive a small-signal transistor model, we will examine a simple amplifier circuit in order to illustrate both the virtues and the limitations of the small-signal analysis.

An inverting amplifier. An inverting amplifier can be implemented by an NMOS transistor and a drain resistor as shown in Fig. 3.23. For the amplifier to work, the bias value V_{IN} for the input voltage must be selected to be so much larger than the threshold voltage V_t of the transistor that the total instantaneous value of the input signal $v_{IN} = V_{IN} + v_{in}$ is larger than the threshold voltage for the range of input signals v_{in} to be applied. Here we are using the notation for the input voltage defined by Eq. (2.7).

The basic operation of the amplifier is as follows: When the input voltage v_{IN} is increased, the drain current i_D increases, and this causes an increase in the voltage drop across R_D so that the output voltage $v_O = V_{DD} - i_D R_D$ is decreased. Conversely, if v_{IN} is decreased, i_D is reduced and $v_O = V_{DD} - i_D R_D$ is increased. Hence, the amplifier is called an inverting amplifier.

For the analysis of the amplifier, we assume that the output voltage is large enough that the transistor is in the active region. This requires $v_O > v_{IN} - V_t$. For this range of operation, we find when neglecting the channel-length modulation

$$\begin{aligned} v_O = V_O + v_o &= V_{DD} - i_D R_D \\ &= V_{DD} - \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(v_{IN} - V_t)^2 R_D \\ &= V_{DD} - \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_{IN} + v_{in} - V_t)^2 R_D \end{aligned}$$

$$\begin{aligned} &= V_{DD} - \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_{IN}-V_t)^2 R_D \\ &\quad - \mu_n C_{ox}\left(\frac{W}{L}\right)(V_{IN}-V_t)\, v_{in} R_D - \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right) v_{in}^2 R_D \end{aligned} \tag{3.37}$$

From Eq. (3.37), we notice that the dc part of v_O, the bias value, is

$$V_O = V_{DD} - \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_{IN}-V_t)^2 R_D = V_{DD} - I_D R_D \tag{3.38}$$

The signal part v_o has a linear term

$$v_{o\,\mathrm{lin}} = -\mu_n C_{ox}\left(\frac{W}{L}\right)(V_{IN}-V_t)\, R_D\, v_{in} \tag{3.39}$$

and a nonlinear term

$$v_{o\,\mathrm{nonlin}} = -\frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right) R_D\, v_{in}^2 \tag{3.40}$$

The term given by Eq. (3.39) is the desired output signal which is proportional to the input signal v_{in}.

In a linear analysis, the nonlinear term given by Eq. (3.40) will be neglected. This is only reasonable if the nonlinear term is much smaller than the linear term, i.e.,

$$\frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right) R_D\, v_{in}^2 \ll \mu_n C_{ox}\left(\frac{W}{L}\right)(V_{IN}-V_t)\, R_D\, v_{in} \tag{3.41}$$

$$\Rightarrow v_{in} \ll 2(V_{IN}-V_t) \tag{3.42}$$

If this is the case, an approximate expression for the output signal swing is

$$v_o = -\frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_{IN}-V_t)\, R_D\, v_{in} = -\frac{2I_D}{V_{IN}-V_t} R_D\, v_{in} \tag{3.43}$$

and we find a voltage gain of

$$A_v = -\frac{2I_D}{V_{IN}-V_t} R_D \tag{3.44}$$

Having derived also the nonlinear term, we may examine the impact of the nonlinearity. From Eq. (3.40), we see that the nonlinear term depends on v_{in}^2. This will cause a dc shift of the output bias voltage and – more important – a second-harmonic distortion if v_{in} is a sinusoid. With $v_{in} = V_{in}\sin(\omega t)$ we find

$$\begin{aligned} v_{o\,\mathrm{nonlin}} &= -\frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right) R_D V_{in}^2 \sin^2(\omega t) \\ &= -\frac{\mu_n C_{ox}}{4}\left(\frac{W}{L}\right) R_D V_{in}^2 (1-\cos(2\omega t)) \end{aligned} \tag{3.45}$$

From Eq. (3.45), we see that the nonlinearity produces a second harmonic output signal with an amplitude

$$V_{o\,\mathrm{dist}} = \frac{\mu_n C_{ox}}{4}\left(\frac{W}{L}\right) R_D V_{in}^2 \tag{3.46}$$

This may be compared to the amplitude of the fundamental-frequency output

$$V_{o\,\mathrm{fund}} = \mu_n C_{ox}\left(\frac{W}{L}\right)(V_{IN}-V_t)\, V_{in} R_D \tag{3.47}$$

and a distortion $D = V_{o\,\mathrm{dist}}/V_{o\,\mathrm{fund}}$ may be calculated:

$$D = \frac{V_{in}}{4(V_{IN} - V_t)} \tag{3.48}$$

It should be realized that the calculations above are performed using some simplifying assumptions: The channel-length modulation is ignored, and the transistor is assumed to be in the active region which is only true for output voltages larger than $v_{IN} - V_t$. Clearly, without these simplifications, the analysis becomes even more complicated and hand calculations using the complete set of nonlinear equations turn out to be if not impossible, then at least time consuming. If the issue is not an analysis of phenomena related to the nonlinearities, but rather an analysis of fundamental linear properties, an analysis using linearization is definitely the way to proceed with hand calculations.

The more complex second-order effects may be analyzed through the use of simulations, and to show this, we investigate the amplifier from Fig. 3.23 using LTspice. But first, we need to find some values for R_D and the transistor dimensions. We assume that the supply voltage is $V_{DD} = 1.8$ V and that the dc value of i_D is $I_D = 100$ µA. Also, the amplifier should be designed to provide a gain of $A_v = -6$ V/V.

In order to allow a large voltage swing at the output, we select a bias value of the output voltage which is half the supply voltage.

We can then calculate R_D using Ohm's law:

$$R_D = \frac{V_{DD}}{2I_D} = 9.0\ \mathrm{k\Omega} \tag{3.49}$$

Inserting this in Eq. (3.44), we find

$$A_v = -\frac{V_{DD}}{V_{IN} - V_t} \tag{3.50}$$

Using Eq. (3.50) with $A_v = -6$ V/V, we find an input bias voltage of

$$V_{IN} = V_t - V_{DD}/A_v = 0.4\ \mathrm{V} - (1.8\ \mathrm{V})/(-6) = 0.7\ \mathrm{V} \tag{3.51}$$

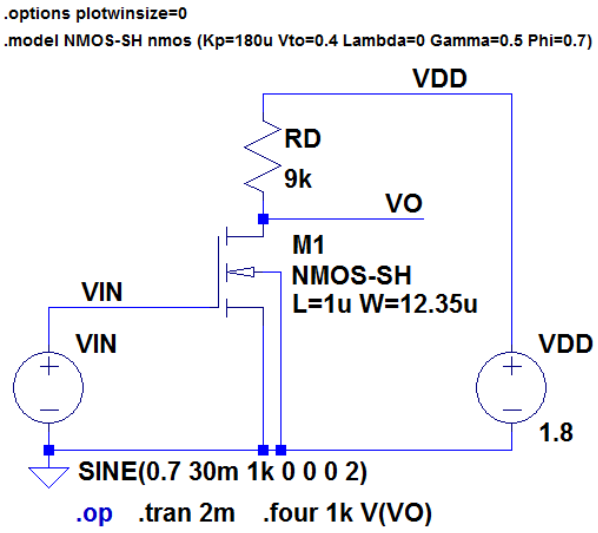
The remaining parameters to be determined are the transistor geometries W and L. As mentioned earlier, for analog circuits you would normally use a channel length which is 3 to 10 times the minimum length specified for the technology. Using a 0.18 µm technology with the parameters shown in Table 3.1, a reasonable value for L would be $L = 1$ µm. Neglecting the channel-length modulation, the width W can then be calculated from

$$I_D = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_{IN} - V_t)^2 \tag{3.52}$$

$$\Rightarrow W = L\frac{2I_D}{\mu_n C_{ox}(V_{IN} - V_t)^2} = 1\ \mu\mathrm{m} \times \frac{2 \times 100\ \mu\mathrm{A}}{180\ \mu\mathrm{A/V^2} \times (0.3\ \mathrm{V})^2} = 12.35\ \mu\mathrm{m} \tag{3.53}$$

These values are then inserted in an LTspice schematic shown in Fig. 3.24. The input voltage is defined as a sinusoid with a dc bias value of 0.7 V, an amplitude of 30 mV and a frequency of 1 kHz. The amplitude is much smaller (20 times) than two times the effective gate voltage of 300 mV as requested by Eq. (3.42) if the linear approximations should provide reasonable results. The first simulation to run is a '.op' analysis in order to check the bias point. The '.op' directive is shown as a comment in Fig. 3.24.

This is run with the value of v_{IN} equal to the initial value in the transient specification for v_{IN}, i.e., 0.7 V. The resulting output file is also shown in Fig. 3.24, and we see that all voltage levels and current levels are exactly as expected.



Output file

```
        --- Operating Point ---

V(vo):          0.899685        voltage
V(vin):         0.7             voltage
V(vdd):         1.8             voltage
Id(M1):         0.000100035     device_current
Ig(M1):         0               device_current
Ib(M1):         -9.09685e-013   device_current
Is(M1):         -0.000100035    device_current
I(Rd):          0.000100035     device_current
I(Vdd):         -0.000100035    device_current
I(Vin):         0               device_current
```

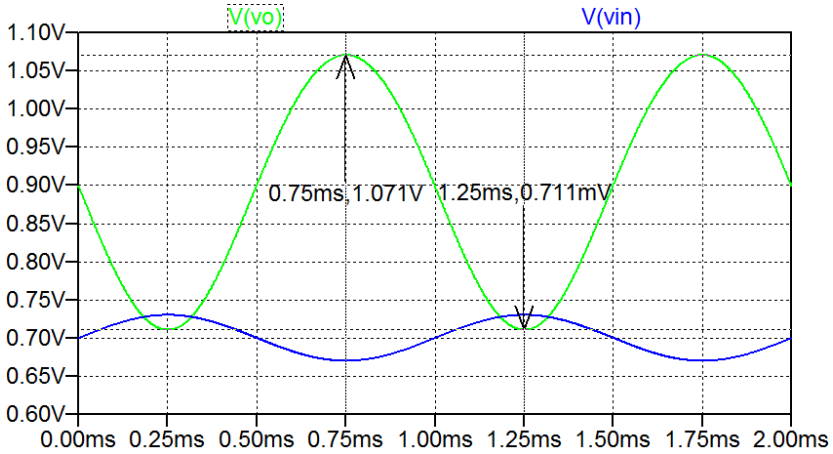
Figure 3.24: LTspice schematic and output file from a '.op' simulation for an inverting amplifier.

Next, we run a transient simulation. For this, we have also specified a '.four' SPICE directive in order to find the distortion. The SPICE directive '`.options plotwinsize=0`' is included to disable the data compression normally performed for a transient simulation (Bruun 2020, Tutorial 4.1). This is needed in order to get accurate results for the distortion analysis (Brocard 2013, page 261). Figure 3.25 shows the simulation plot and the results of the distortion analysis. These results are found in the error log file which is opened by 'Ctrl-L'.

From the simulation plot, we see that the output signal is inverted compared to the input signal. We also find the peak-to-peak value of the output signal to be 0.360 V, corresponding to an amplitude $V_o = 0.180$ V, i.e., 6 times the input amplitude of $V_{in} = 30$ mV.

From the error log file, we see that the amplitude of harmonic number 1, the fundamental frequency, is 0.180 V as expected. Also, we find that only the second harmonic contributes significantly to the distortion with a normalized value of 2.50×10^{-2}. From Eq. (3.48), we can calculate the distortion to be $D = 30\text{ mV}/(4 \times (0.7\text{ V} - 0.4\text{ V})) = 2.50 \times 10^{-2}$, i.e., exactly the same value as found from the simulation. Thus, the simulations by LTspice confirm the results obtained by hand calculations.

We may also use LTspice to see what happens if we violate the assumption that the transistor is in the active region. For instance, with an input amplitude of more than 300 mV, the transistor reaches the cut-off region for $v_{in} = -V_{in}$. Figure 3.26 shows a simulation plot for a simulation with $V_{in} = 310$ mV. From the plot, it is evident that the transistor is cut off for low values of the input signal and is in the triode region for high values of the input signal. This leads to a clipping of the output signal and a severe distortion.



Error log file

```
Fourier components of V(vo)
DC component:0.895183

Harmonic          Frequency          Fourier          Normalized
 Number             [Hz]            Component         Component
    1           1.000e+03          1.801e-01          1.000e+00
    2           2.000e+03          4.501e-03          2.500e-02
    3           3.000e+03          1.093e-09          6.071e-09
    4           4.000e+03          5.770e-09          3.205e-08
    5           5.000e+03          4.855e-09          2.696e-08
    6           6.000e+03          4.510e-09          2.505e-08
    7           7.000e+03          4.351e-09          2.416e-08
    8           8.000e+03          3.770e-09          2.094e-08
    9           9.000e+03          3.442e-09          1.912e-08
Total Harmonic Distortion: 2.499622%(2.499619%)
```

Figure 3.25: Simulation plot from transient simulation and results of distortion analysis.

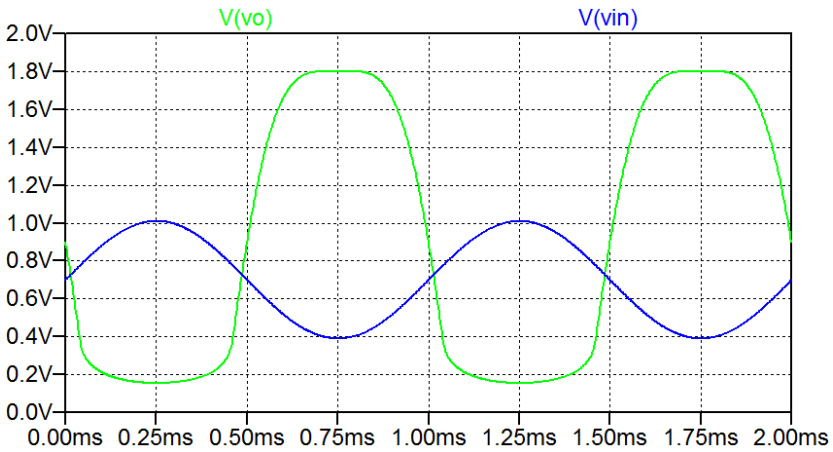


Figure 3.26: Simulation plot from a transient simulation showing the clipping resulting from an overdrive on the input.

In this example, the linearization was done after deriving the nonlinear expression for the output signal. However, if we know already before the analysis that we are only interested in the linearized result, it is much easier to perform the linearization before writing the circuit equations, and for this, we need linearized transistor models.

The low-frequency small-signal model for an NMOS transistor. The NMOS transistor can be seen as a controlled current source from drain to source with v_{GS}, v_{DS} and v_{BS} as the controlling voltages. For the initial investigation, we consider a transistor with source and bulk connected, i.e., $v_{BS} = 0$.

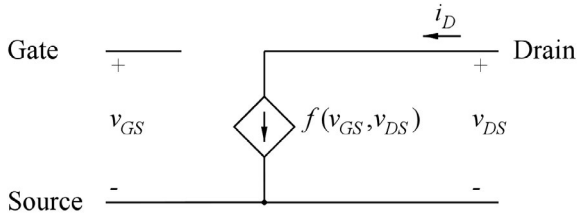


Figure 3.27: Nonlinear large-signal model for an NMOS transistor.

This can be shown using the symbol for a controlled current source, see Fig. 3.27. The relation describing i_D is nonlinear, and using the Shichman-Hodges model, the nonlinear relation $i_D = f(v_{GS}, v_{DS})$ is given by Eqs. (3.14) - (3.16). In order to linearize this relation, we apply a Taylor expansion of $f(v_{GS}, v_{DS})$ using only the linear terms. The expansion must be calculated from a specific bias point, i.e., a value of i_D corresponding to the bias values of v_{GS} and v_{DS}. Using the normal notation shown in Fig. 2.4 with capital letters and capital subscripts for bias values, we have $I_D = f(V_{GS}, V_{DS})$. The Taylor series with only linear terms is

$$f(v_{GS}, v_{DS}) = f(V_{GS}, V_{DS}) + \left(\frac{\partial f(v_{GS}, v_{DS})}{\partial v_{GS}} \bigg|_{\substack{\text{bias} \\ \text{point}}} \right) \Delta v_{GS} + \left(\frac{\partial f(v_{GS}, v_{DS})}{\partial v_{DS}} \bigg|_{\substack{\text{bias} \\ \text{point}}} \right) \Delta v_{DS} \tag{3.54}$$

With $i_D = f(v_{GS}, v_{DS})$, $\Delta v_{GS} = v_{GS} - V_{GS} = v_{gs}$, $\Delta v_{DS} = v_{DS} - V_{DS} = v_{ds}$ and the small-signal value i_d defined as $i_d = i_D - I_D$, we get

$$i_d = f(v_{GS}, v_{DS}) - f(V_{GS}, V_{DS}) = \left(\frac{\partial i_D}{\partial v_{GS}} \bigg|_{\substack{\text{bias} \\ \text{point}}} \right) v_{gs} + \left(\frac{\partial i_D}{\partial v_{DS}} \bigg|_{\substack{\text{bias} \\ \text{point}}} \right) v_{ds} \tag{3.55}$$

The term $(\partial i_D / \partial v_{GS})|_{\substack{\text{bias} \\ \text{point}}} \cdot v_{gs}$ shows that the current i_d has a part i_{d1} which is controlled by an input voltage v_{gs}. This corresponds to the voltage-controlled current source shown in Fig. 2.11, and the parameter describing the controlled source is a transconductance. Traditionally, for a MOS transistor, this is called the gate transconductance (or just transconductance) and is denoted by g_m, so

$$g_m = \frac{\partial i_D}{\partial v_{GS}} \bigg|_{\substack{\text{bias} \\ \text{point}}} \tag{3.56}$$

The term $(\partial i_D / \partial v_{DS})|_{\substack{\text{bias} \\ \text{point}}} \cdot v_{ds}$ shows that the current i_d has a part i_{d2} which is controlled by the voltage v_{ds}. The current i_{d2} is proportional to the voltage between the two terminals of the controlled current source, and a linear relation between current and voltage for two terminals of a device is Ohm's law, so $i_{d2} = v_{ds}/r_{ds}$ where r_{ds} is called the small-signal output resistance of the transistor. A corresponding output conductance is defined as $g_{ds} = 1/r_{ds}$, and from Eq. (3.55), we find

$$g_{ds} = \frac{\partial i_D}{\partial v_{DS}} \bigg|_{\substack{\text{bias} \\ \text{point}}} \tag{3.57}$$

Equation (3.55) shows that the two parts i_{d1} and i_{d2} are added to form the total small-signal drain current, and in a circuit model, it follows from Kirchhoff's current law that this addition corresponds to a parallel connection. Thus, the linearized circuit becomes a parallel connection of a voltage-controlled current source and a resistor as shown in Fig. 3.28

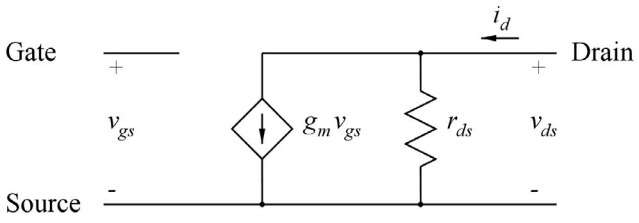


Figure 3.28: Linear small-signal model for an NMOS transistor.

The parameters characterizing the small-signal model are g_m and r_{ds}. For a graphical interpretation of g_m, let us examine the relation between i_D and v_{GS}. For a fixed value of v_{DS}, this may look like shown in Fig. 3.29. Assuming that the bias value of v_{GS} is V_{GS1}, we find that g_m is the slope of the straight line A shown in the figure, i.e., the tangent to the curve for $v_{GS} = V_{GS1}$. With a bias value of V_{GS2}, we find another value of g_m, corresponding to the slope of line B in the figure.

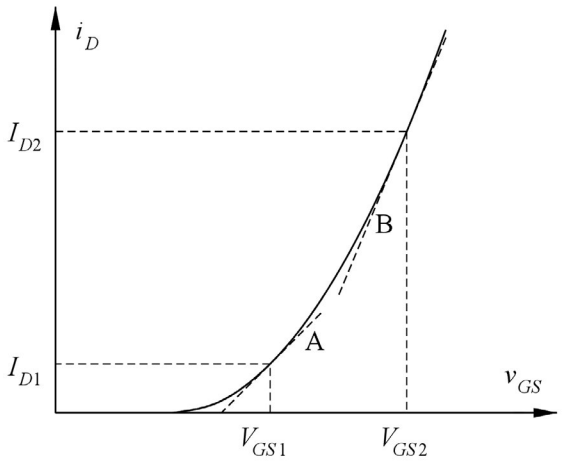


Figure 3.29: Graphical interpretation of g_m.

Clearly, the value of g_m depends strongly on the bias point. We can derive an expression for g_m by differentiation. We need to consider the triode region and the active region separately.

For the triode region, we find

$$g_m = \left.\frac{\partial i_D}{\partial v_{GS}}\right|_{V_{DS}} = \mu_n C_{ox}\left(\frac{W}{L}\right) V_{DS}(1+\lambda V_{DS});\ 0 \le V_{DS} \le V_{GS} - V_t \tag{3.58}$$

We notice that in the triode region, g_m is independent of V_{GS}. Neglecting the channel-length modulation, g_m increases linearly with V_{DS}.

For the active region, we find

$$g_m = \left.\frac{\partial i_D}{\partial v_{GS}}\right|_{V_{GS},V_{DS}} = \mu_n C_{ox}\left(\frac{W}{L}\right)(V_{GS} - V_t)(1+\lambda V_{DS});\ 0 \le V_{GS} - V_t \le V_{DS} \tag{3.59}$$

In the active region, g_m depends both on V_{GS} and V_{DS}, but for approximate hand calculations, we can often neglect the channel-length modulation, and in this case, g_m is independent of V_{DS}.

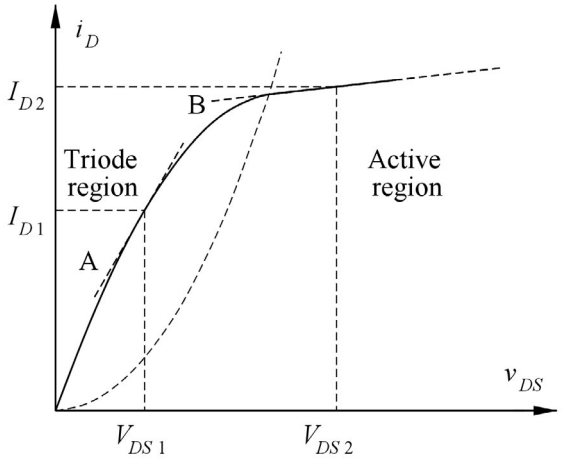


Figure 3.30: Graphical interpretation of g_{ds}.

Next, we consider a graphical interpretation of g_{ds} or $1/r_{ds}$. This is derived from the relation between i_D and v_{DS}. For a fixed value of v_{GS}, this may look like shown in Fig. 3.30. The parameter $g_{ds} = (\partial i_D/\partial v_{DS})|_{\substack{\text{bias}\\\text{point}}}$ is the slope of the curve, calculated for the bias values of v_{DS} and v_{GS}.

For the transistor in the triode region, for example V_{DS1} in Fig. 3.30, line A, we find

$$\begin{aligned} g_{ds} = \left.\frac{\partial i_D}{\partial v_{DS}}\right|_{V_{GS},V_{DS}} &= \mu_n C_{ox}\left(\frac{W}{L}\right)\left[(V_{GS}-V_t-V_{DS})(1+\lambda V_{DS})+[(V_{GS}-V_t)v_{DS}-V_{DS}^2/2]\,\lambda\right] \\ &= \mu_n C_{ox}\left(\frac{W}{L}\right)\left[(V_{GS}-V_t-V_{DS})(1+2\lambda V_{DS})+\lambda V_{DS}^2/2\right] \end{aligned} \tag{3.60}$$

For $\lambda V_{DS} \ll 1$ which is normally the case in the triode region where V_{DS} is small, $V_{DS} < V_{GS}-V_t$, this expression can be simplified to

$$g_{ds} \simeq \mu_n C_{ox}\left(\frac{W}{L}\right)(V_{GS}-V_t-V_{DS}) \tag{3.61}$$

Often when the transistor is in the triode region, it is used for very small values of v_{DS}, $V_{DS} \ll V_{GS}-V_t$, and in this situation, the transistor can be treated as a voltage-controlled resistor where the value of the resistance is controlled by V_{GS} and is found from Eq. (3.61) where the approximation $(V_{GS}-V_t-V_{DS}) \simeq (V_{GS}-V_t)$ is used, i.e.,

$$r_{ds} \simeq \left[\mu_n C_{ox}\left(\frac{W}{L}\right)(V_{GS}-V_t)\right]^{-1} \tag{3.62}$$

For the transistor in the active region, for example V_{DS2} in Fig. 3.30, line B, the relation between i_D and v_{DS} is already linear, and we find the slope as

$$g_{ds} = \left.\frac{\partial i_D}{\partial v_{DS}}\right|_{V_{GS}} = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_{GS}-V_t)^2\,\lambda \tag{3.63}$$

We notice that in the active region, g_{ds} depends only on the bias voltage V_{GS}.

In the equations above for g_m and g_{ds}, the bias point was defined by the bias voltages V_{GS} and V_{DS}. However, when they are changed, also the bias current I_D changes. In practice, we often use MOS transistors in the active region, and it can be helpful to have different expressions for g_m and g_{ds} depending on which parameters are available to define the bias conditions. Sometimes, it is the bias current I_D. Sometimes, it is the bias voltages V_{GS} and V_{DS}, and sometimes, it is the geometry parameters L and W. The following equations show alternative expressions for the small-signal parameters.

$$g_m = \frac{\partial i_D}{\partial v_{GS}} = \mu_n C_{ox}\left(\frac{W}{L}\right)(V_{GS}-V_t)(1+\lambda V_{DS}) \tag{3.64}$$

$$= \frac{2\,I_D}{V_{GS}-V_t} \tag{3.65}$$

$$= \sqrt{2\mu_n C_{ox}\left(\frac{W}{L}\right) I_D(1+\lambda V_{DS})} \tag{3.66}$$

$$g_{ds} = 1/r_{ds} = \frac{\partial i_D}{\partial v_{DS}} = \lambda\frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_{GS}-V_t)^2 \tag{3.67}$$

$$= \frac{\lambda I_D}{1+\lambda V_{DS}} \tag{3.68}$$

Often, for hand calculations, a precision of 10% to 15% in the numerical values of g_m and g_{ds} is sufficient for an evaluation of a circuit. This can often be obtained, even when using a value of 0 for λV_{DS}. For $\lambda V_{DS} < 0.15$, substituting the term $1+\lambda V_{DS}$ by 1 leads to errors of less than 15%. For a 0.18 µm CMOS process, the maximum supply voltage is about 1.8 V, and using channel lengths of about 1 µm or more, λ is smaller than 0.14 V^{-1}, so with drain-source bias voltages of up to approximately half the supply voltage, the condition $\lambda V_{DS} < 0.15$ is fulfilled. So for hand calculations, the following simplified equations are often used

$$g_m = \frac{2\,I_D}{V_{GS}-V_t} \tag{3.69}$$

$$\simeq \mu_n C_{ox}\left(\frac{W}{L}\right)(V_{GS}-V_t) \tag{3.70}$$

$$\simeq \sqrt{2\mu_n C_{ox}\left(\frac{W}{L}\right) I_D} \tag{3.71}$$

$$g_{ds} = 1/r_{ds} = \lambda\frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_{GS}-V_t)^2 \simeq \lambda I_D \tag{3.72}$$

When verifying approximate solutions based on these equations using LTspice, it is necessary to use $\lambda = 0$ in the Shichman-Hodges model. Subsequently, simulations may be run with the actual value of the channel-length parameter in order to check that the approximations are acceptable.

An important parameter often used as a figure of merit for a MOS transistor is the intrinsic voltage gain $A_{v\mathrm{i}} = g_m/g_{ds}$. With an input voltage $v_{GS} = V_{GS} + v_{in}$ connected between gate and source, we have $v_{gs} = v_{in}$, and with an ideal dc current source connected to the drain, the small-signal output voltage is $v_o = -g_m r_{ds} = -g_m/g_{ds}$. This gives the maximum small-signal voltage gain, the open-circuit voltage gain, which can be obtained from the transistor. From Eqs. (3.65) and (3.68), we find

$$A_{v\mathrm{i}} = \frac{g_m}{g_{ds}} = \frac{2(1+\lambda V_{DS})}{\lambda(V_{GS}-V_t)} \tag{3.73}$$

Using the simplified equations (3.70) - (3.72), we find

$$A_{v\mathrm{i}} \simeq \frac{2}{\lambda(V_{GS}-V_t)} = \frac{2L}{\lambda'(V_{GS}-V_t)} \tag{3.74}$$

and

$$A_{v\mathrm{i}} \simeq \frac{1}{\lambda}\sqrt{2\mu_n C_{ox}\left(\frac{W}{L}\right)\frac{1}{I_D}} = \frac{1}{\lambda'}\sqrt{2\mu_n C_{ox} WL\frac{1}{I_D}} \tag{3.75}$$

where we have used $\lambda' = \lambda L$.

From Eq. (3.74), we see that for the 0.18 µm process with the transistor parameters from Table 3.1, we achieve an intrinsic gain in the range from 5.7 to 200 with a channel length in the range from 0.2 µm to 1 µm and an effective gate voltage in the range from 100 mV to 500 mV. A large value of $A_{v\mathrm{i}}$ is obtained with a small effective gate voltage and a long transistor.

The bulk transconductance. In all of the previous calculations, we have assumed that bulk and source are connected, i.e., $v_{BS} = 0$. When this is not the case, i_D depends also on v_{BS}, so Eq. (3.55) must be expanded with a term $(\partial i_D/\partial v_{BS})|_{\substack{\text{bias}\\\text{point}}} \cdot v_{bs}$ accounting for the bulk effect. This is a voltage-controlled current source where the controlling voltage is the small-signal part of the bulk-source voltage v_{BS}. The parameter describing the controlled source is a transconductance, and traditionally, this is called the bulk transconductance or body transconductance g_{mb} or g_s. Examining Eqs. (3.15) and (3.16), we note that $(\partial i_D/\partial v_{GS}) = -(\partial i_D/\partial V_t)$, and using $v_{BS} = -v_{SB}$ in Eq. (3.17), we find

$$\begin{aligned} g_{mb} = \frac{\partial i_D}{\partial v_{BS}} &= \left(\frac{\partial i_D}{\partial V_t}\right)\left(\frac{\partial V_t}{\partial v_{BS}}\right) = \left(-\frac{\partial i_D}{\partial v_{GS}}\right)\left(-\frac{\partial V_t}{\partial v_{SB}}\right) \\ &= (-g_m)\left(-\gamma\frac{1}{2\sqrt{V_{SB}+|2\Phi_F|}}\right) = g_m\frac{\gamma}{2\sqrt{V_{SB}+|2\Phi_F|}} = \chi g_m \end{aligned} \tag{3.76}$$

Typically, the bulk transconductance is 10% – 30% of the gate transconductance (Sedra & Smith 2016). In the small-signal circuit model, the current source controlled by v_{bs} appears in parallel with the gate transconductance current source, and the complete low-frequency small-signal model is shown in Fig. 3.31.

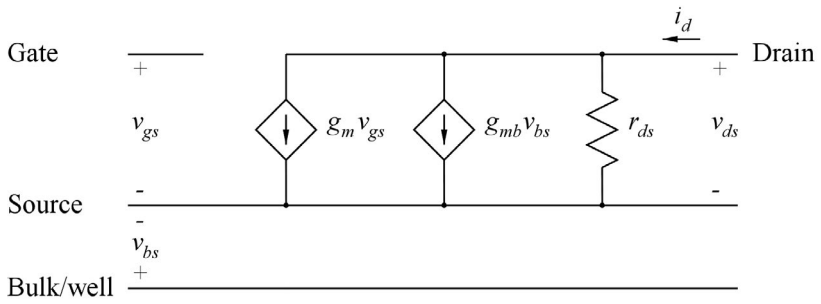


Figure 3.31: Low-frequency small-signal for an NMOS transistor with body effect.

Often the bulk is connected to a dc voltage, typically the negative supply voltage for an NMOS transistor, so the small-signal value of the bulk voltage is 0, leading to $v_{bs} = 0 - v_s = -v_s$. For this case, the small-signal model can be drawn as shown in Fig. 3.32.

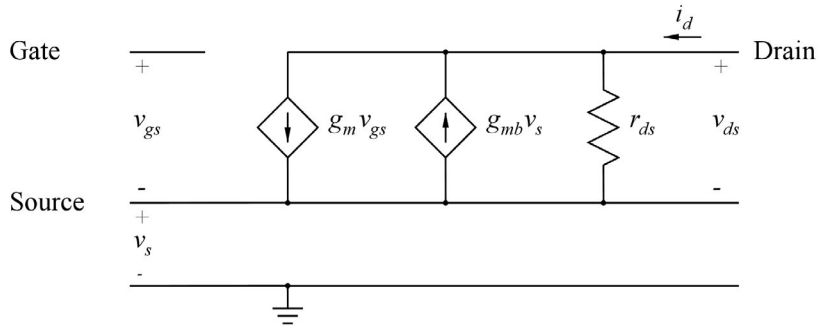


Figure 3.32: Alternative low-frequency small-signal model for an NMOS transistor with body effect and with the bulk connected to a dc voltage.

The small-signal model for a PMOS transistor. For a PMOS transistor, a low-frequency small-signal model can be derived in the same way as for the NMOS transistor but care has to be taken with the signs of voltages and the direction of current in the transistor. This is an issue which often confuses the student who is new to the field of analog CMOS design.

As shown in Fig. 3.11, for the PMOS transistor, the positive direction of the total instantaneous value of the drain current is defined as being positive *out of* the drain terminal because this leads to a positive numerical value when the transistor is in the triode region or in the active region. With this definition, we have for the active region

$$i_D = \frac{\mu_p C_{ox}}{2}\left(\frac{W}{L}\right)(v_{GS} - V_t)^2(1 - \lambda v_{DS}) \tag{3.77}$$

where the numerical values of both v_{GS}, v_{DS} and V_t are negative. The mobility μ_p and the channel-length modulation parameter λ are positive, and the resulting drain current is positive.

However, for the small-signal model, we traditionally use the same model as for the NMOS transistor, i.e., the positive direction of the small-signal value of the drain current is defined as being positive *into* the drain terminal. This implies that g_m and g_{ds} are calculated as

$$g_m = \left.\frac{\partial(-i_D)}{\partial v_{GS}}\right|_{\substack{\text{bias}\\ \text{point}}} \tag{3.78}$$

and

$$g_{ds} = \left.\frac{\partial(-i_D)}{\partial v_{DS}}\right|_{\substack{\text{bias}\\ \text{point}}} \tag{3.79}$$

In the active region, this leads to

$$\begin{aligned} g_m &= -\mu_p C_{ox}\left(\frac{W}{L}\right)(V_{GS} - V_t)(1 - \lambda V_{DS}) \\ &= \mu_p C_{ox}\left(\frac{W}{L}\right)(|V_{GS}| - |V_t|)(1 + \lambda |V_{DS}|) \end{aligned} \tag{3.80}$$

and

$$\begin{aligned} g_{ds} &= \left(-\frac{\mu_p C_{ox}}{2}\left(\frac{W}{L}\right)(V_{GS}-V_t)^2\right)(-\lambda) \\ &= \frac{\mu_p C_{ox}}{2}\left(\frac{W}{L}\right)(|V_{GS}|-|V_t|)^2\lambda \end{aligned} \tag{3.81}$$

For the triode region, we find

$$g_m = \mu_p C_{ox}\left(\frac{W}{L}\right)|V_{DS}|(1+\lambda|V_{DS}|) \tag{3.82}$$

and

$$\begin{aligned} g_{ds} &= \mu_p C_{ox}\left(\frac{W}{L}\right)\left[(|V_{GS}|-|V_t|-|V_{DS}|)(1+2\lambda|V_{DS}|)+\lambda|V_{DS}|^2/2\right] \\ &\simeq \mu_p C_{ox}\left(\frac{W}{L}\right)(|V_{GS}|-|V_t|-|V_{DS}|) \end{aligned} \tag{3.83}$$

Similarly, the bulk transconductance can be found as

$$g_{mb} = g_m\frac{\gamma}{2\sqrt{|V_{SB}|+|2\Phi_F|}} = \chi g_m \tag{3.84}$$

Thus, the PMOS small-signal model and the expressions for the small-signal parameters g_m, g_{ds} and g_{mb} are exactly the same as for the NMOS transistor, provided that absolute values are used for I_D, V_{GS}, V_{DS}, V_{SB} and V_t.

A high-frequency small-signal model. The MOS transistor inherently includes several capacitors. All the reverse-biased pn junctions contribute with their junction capacitances which are proportional to the area of the junction and also depend on the reverse-bias voltage applied to the junction. Also, the gate electrode has a capacitance to the channel and overlap capacitances to the drain, the source and the bulk. The gate-channel capacitance is the dominant capacitance, and this is proportional to the area of the gate. Normally, it is modeled as a gate-source capacitance C_{gs}, and in the active region, an approximate expression for C_{gs} is (Tsividis & McAndrew 2010)

$$C_{gs} \simeq \frac{2}{3}WLC_{ox} \tag{3.85}$$

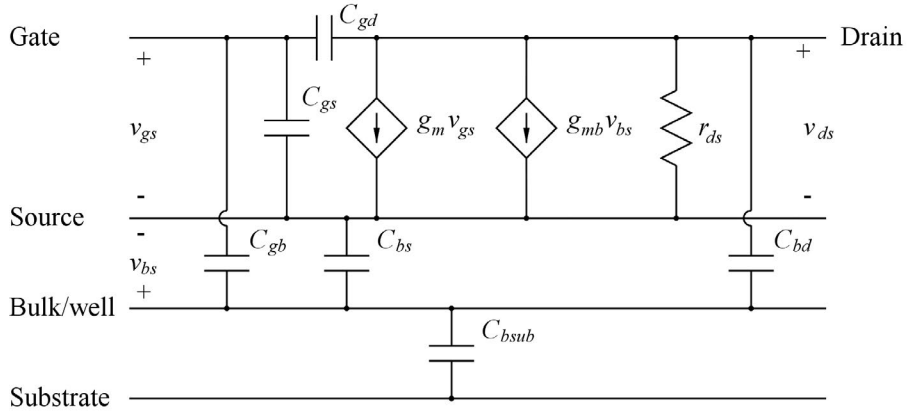


Figure 3.33: High-frequency small-signal MOS transistor model.

The overlap capacitances from gate to drain, source and bulk, $C_{gd\,\mathrm{ov}}$, $C_{gs\,\mathrm{ov}}$ and $C_{gb\,\mathrm{ov}}$ are normally much smaller than the gate-channel capacitance, so the gate-drain capacitance is normally much smaller than the gate-source capacitance. However for transistors with a very short channel length, it may be necessary to add the gate-source overlap capacitance to the value found from Eq. (3.85), in which case the gate-source capacitance is found from

$$C_{gs} \simeq \frac{2}{3} WLC_{ox} + C_{gs\,\mathrm{ov}} \tag{3.86}$$

Figure 3.33 shows the small-signal transistor model augmented with the internal capacitances of the transistor.

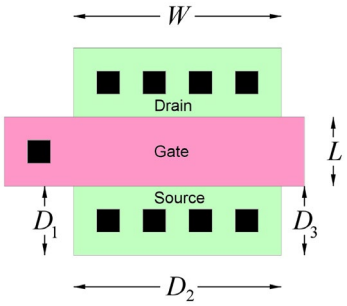


Figure 3.34: Transistor layout (top view) showing dimensions for calculating internal capacitances.

For simulating the capacitances with LTspice, the area and perimeter of the drain diffusion and the source diffusion must be specified in the specification window shown in Fig. 3.12. Figure 3.34 shows the layout of a transistor. The green areas are the drain and source diffusions. The pink area is the gate, and the black areas are contact holes for connections to a metal layer (not shown).

For the layout shown, the areas of source and drain diffusions are $\mathrm{AD} = \mathrm{AS} = D_1 \times D_2$ and the perimeters are $\mathrm{PD} = \mathrm{PS} = D_1 + D_2 + D_3$. Note that the edge facing the channel region is not included. The areas of source and drain diffusion typically have a minimum dimension of approximately W times 2.75 times the minimum length for the process. The perimeter of the drain and source diffusion typically has a minimum dimension of W plus 5.5 times the minimum length (Sedra & Smith 2016, Appendix B).

Also, the transistor model must be expanded to include parameters for junction capacitances, oxide capacitance and overlap capacitances. The parameters are defined as explained in the 'LTspiceHelp', and Fig. 3.35 shows the LTspice '.model' directives including both the parameters given in Table 3.1 and typical parameters for capacitances in a generic 0.18 µm process. In Fig. 3.35, the value of λ for $L = 1$ µm has been inserted in the models. For other values of L, the value of λ must recalculated since it is inversely proportional to L. Each of the '.model' specifications extends over more than one line with a '+' to indicate that a line is a continuation of the specification.

The capacitance from well to substrate (C_{bsub} in Fig. 3.33) is not part of the transistor model in LTspice since a well may be common to several transistors, so C_{bsub} must be inserted separately if it is needed in the circuit analysis.

```
.model NMOS-SH nmos (Kp=180u Vto=0.40 Lambda=0.10 Gamma=0.5 Phi=0.7
+TOX=4.0n CGSO=0.29n CGBO=0 CGDO=0.29n CJ=3.65m CJSW=0.79n)

.model PMOS-SH pmos (Kp=45u Vto=-0.42 Lambda=0.14 Gamma=0.5 Phi=0.7
+TOX=4.2n CGSO=0.28n CGBO=0 CGDO=0.28n CJ=1.38m CJSW=1.44n)
```

Figure 3.35: LTspice transistor models including parameters for calculating small-signal capacitances.

In hand calculations for the generic 0.18 µm CMOS model, the gate-source capacitance may be estimated from Eq. (3.85) with $C_{ox} \simeq 8.5$ fF/(µm)^2, and the parasitic junction capacitances may be estimated from the areas and perimeters of the diffusions using a junction capacitance per unit area of 3.65 fF/(µm)^2

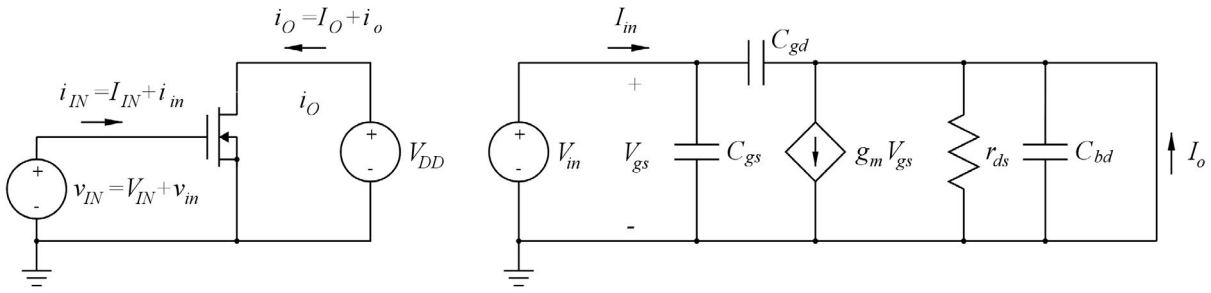


Figure 3.36: Large-signal schematic (a) and small-signal schematic (b) for finding small-signal current gain.

and a junction capacitance per unit length of the perimeter of 0.79 fF/µm for NMOS drain and source diffusions and 1.38 fF/(µm)^2 per unit area and 1.44 fF/µm per unit length of the perimeter for PMOS drain and source diffusions. The dependency on the reverse-bias voltage across the junctions may be calculated from Eq. (3.1). In parallel with the bulk-source junction capacitance is the bulk-channel capacitance. The details concerning this capacitance are quite complicated (Tsividis & McAndrew 2010) but an estimate for the bulk-channel capacitance may be

$$C_{b\,channel} \simeq \chi \frac{2}{3} WLC_{ox} \tag{3.87}$$

where χ is defined in Eq. (3.76). For the well-substrate capacitance, a junction capacitance of $C_j \simeq$ 1 fF/(µm)^2 may be used. The gate overlap capacitances towards drain and source may be estimated from the width W of the gate using an overlap capacitance of 0.3 fF/µm.

Also for the high-frequency small-signal performance, we can define a figure of merit. Traditionally, the unity-gain frequency f_T for the small-signal current gain $A_{isc} = I_o(s)/I_{in}(s)$ is used as a figure of merit. This frequency is called the transition frequency (Sedra & Smith 2016) . It is calculated from the circuit shown in Fig. 3.36 where both the large-signal circuit and the small-signal equivalent is shown. Observe that when converting the large-signal schematic to a small-signal schematic, all dc sources are reset since a dc source does not contribute to the signals in the circuit. The gate-bulk capacitance C_{gb} is not shown in Fig. 3.36. It may be included in C_{gs} but normally $C_{gb} \ll C_{gs}$, so it can be neglected.

With V_{DD} replaced by a short circuit in the small-signal diagram, a node equation at the output node gives

$$I_o = g_m V_{gs} - V_{gs}\, s C_{gd} \tag{3.88}$$

A node equation at the input node gives

$$I_{in} = V_{gs}\, s(C_{gs} + C_{gd}) \tag{3.89}$$

From Eqs. (3.88) and (3.89), we find

$$A_{\text{isc}} = \frac{I_o}{I_{in}} = \frac{g_m - sC_{gd}}{s(C_{gs} + C_{gd})} \tag{3.90}$$

For frequencies much smaller than g_m/C_{gd}, this can be approximated by

$$A_{\text{isc}} \simeq \frac{g_m}{s(C_{gs} + C_{gd})} \tag{3.91}$$

from which we find

$$f_T \simeq \frac{g_m}{2\pi\,(C_{gs} + C_{gd})} \tag{3.92}$$

Normally, $C_{gs} \gg C_{gd}$, so we find that $g_m/(C_{gs} + C_{gd})$ is indeed much smaller than g_m/C_{gd}, thus justifying the approximation used in Eq. (3.91).

With $C_{gs} \gg C_{gd}$, we find $f_T \simeq g_m/(2\pi C_{gs})$, and using Eqs. (3.70) and (3.85), we then find

$$f_T \simeq \frac{g_m}{2\pi C_{gs}} \simeq \frac{3\mu\,(V_{GS} - V_t)}{4\pi L^2} \tag{3.93}$$

Since the mobility μ is three to four times bigger for an NMOS transistor than for a PMOS transistor, NMOS transistors are generally significantly faster than PMOS transistors. Also, transistors with short channel lengths are faster than transistors with long channel lengths. For an NMOS transistor with a minimum channel length and an effective gate voltage of 400 mV, Eq. (3.93) results in a unity-gain frequency of about 60 GHz for the 0.18 μm generic CMOS process. When comparing Eqs. (3.93) and (3.74), we can see that there is a trade-off between speed and low-frequency gain. The downscaling of device geometries (L getting smaller) means faster devices but also smaller intrinsic gain.

3.6 Deriving a small-signal equivalent circuit from a large-signal schematic

With the small-signal model for the MOS transistors in place, we are ready to convert a large-signal schematic into a small-signal equivalent diagram. Then we can perform a linear circuit analysis without first having to solve nonlinear equations like we did for the inverting amplifier in Eqs. (3.37) - (3.44). We have already seen a simple example of this conversion in Fig. 3.36. In the conversion, all linear components, i.e., resistors, capacitors and inductors, remain unchanged and the nonlinear MOS transistors are replaced by linear small-signal models from Figs. 3.28, 3.31, 3.32 or 3.33.

However, in the conversion, we also need to consider voltage sources and current sources. In general, a voltage source may contain a dc term and an ac term. In a small-signal model, we consider only signal *variations*, so the dc term in a voltage source is reset which means that it is replaced by a short circuit. Thus, when drawing the small-signal circuit, supply voltage sources and dc bias voltage sources are shorted. Likewise, a current source may contain a dc term and an ac term, and the dc term is reset when drawing a small-signal equivalent circuit. Resetting a dc current source means giving the current source a value of 0, and this is the current in an open circuit. Thus, when drawing the small-signal equivalent circuit, a dc current source is simply replaced by an open circuit, i.e., it is removed.

Revisiting the inverting amplifier. Now, let us consider the inverting amplifier from Fig. 3.23 and make a small-signal equivalent circuit for the amplifier. For the small-signal equivalent, V_{DD} and V_{IN} are reset and the transistor is replaced by a small-signal model. We consider only low frequencies, and the transistor has bulk connected to source, so there is no bulk effect, meaning that we just need the simple small-signal model from Fig. 3.28. A direct replacement of the transistor symbol with the small-signal model results in the small-signal equivalent circuit shown in Fig. 3.37(a). Normally, this is redrawn with the drain resistor 'folded down' towards ground as shown in Fig. 3.37(b). This way of 'folding down' devices connected to the positive supply rail will be used extensively in small-signal equivalent circuits in this book.

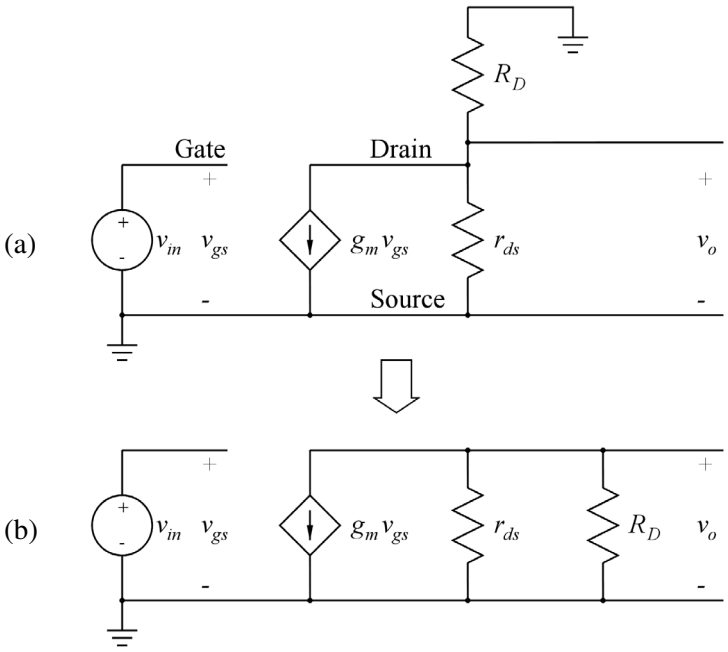


Figure 3.37: Small-signal equivalent circuit for the inverting amplifier from Fig. 3.23, first with a direct replacement of the transistor with the small-signal model from Fig. 3.28, (a), then with the drain resistor 'folded down' towards ground (b).

From this, we easily find an expression for the output voltage by using a node equation at the output.

$$\begin{aligned} \frac{v_o}{R_D} + \frac{v_o}{r_{ds}} + g_m v_{in} &= 0 \\ \Rightarrow A_v = \frac{v_o}{v_{in}} = -g_m \left(\frac{1}{R_D} + \frac{1}{r_{ds}} \right)^{-1} &= -g_m \left(R_D \parallel r_{ds} \right) \end{aligned} \tag{3.94}$$

In order to relate A_v to the component values and the bias point of the transistor, we use Eq. (3.65) to find g_m.

$$g_m = \frac{2 I_D}{V_{GS} - V_t} \tag{3.95}$$

For a bias point with $V_O = V_{DD}/2 \Rightarrow I_D = V_{DD}/(2R_D)$ and $V_{GS} = V_{IN}$, we find

$$g_m = \frac{V_{DD}}{R_D \left(V_{GS} - V_t \right)} \tag{3.96}$$

When the channel-length modulation is neglected for the transistor, we have $r_{ds} = \infty$, so $(R_D \parallel r_{ds}) = R_D$, leading to

$$A_v = -\frac{V_{DD}}{V_{IN} - V_t} \tag{3.97}$$

as also found from the previous analysis using Eqs. (3.37) - (3.44). However, in deriving Eq. (3.97), we have used only linear equations.

From the small-signal analysis, we can also find the impact of the channel-length modulation on the gain. From Eq. (3.68), we have $g_{ds} = \lambda I_D/(1+\lambda V_{DS}) = [\lambda V_{DD}/(2R_D)]/(1+\lambda V_{DD}/2)$, so

$$\begin{aligned} R_D \parallel r_{ds} &= \left(\frac{1}{R_D} + g_{ds}\right)^{-1} = \left[\frac{1}{R_D} + \left(\frac{\lambda V_{DD}}{2R_D}\right)\left(\frac{1}{1+\lambda V_{DD}/2}\right)\right]^{-1} = R_D\left(\frac{1+\lambda V_{DD}/2}{1+\lambda V_{DD}}\right) \\ \Rightarrow A_v &= -g_m(R_D \parallel r_{ds}) = -\left(\frac{V_{DD}}{V_{IN} - V_t}\right)\left(\frac{1+\lambda V_{DD}/2}{1+\lambda V_{DD}}\right) = -5.54 \text{ V/V} \end{aligned} \tag{3.98}$$

where the value $\lambda = 0.1\ \text{V}^{-1}$ is taken from Table 3.1 with $L = 1$ µm, and we use $V_{DD} = 1.8$ V and $V_{IN} = 0.7$ V as found in Eq. (3.51). As expected, the channel-length modulation causes a reduction of the absolute value of the gain which is reduced from 6.00 V/V to 5.54 V/V.

From the small-signal analysis, we can also find the impact of replacing the drain resistor R_D with an ideal dc current source. In this case, R_D is simply removed from the small-signal circuit in Fig. 3.37, and the voltage gain is

$$A_v = -g_m r_{ds} = -\left(\frac{2I_D}{V_{IN} - V_t}\right)\left(\frac{1+\lambda V_{DS}}{\lambda I_D}\right) = -\frac{2(1+\lambda V_{DD}/2)}{\lambda(V_{IN} - V_t)} = -69.7 \text{ V/V} \tag{3.99}$$

This is also the expression found for the intrinsic voltage gain A_{vi} in Eq. (3.74) with $V_{DS} = V_{DD}/2$, corresponding to the maximum achievable voltage gain from the amplifier.

When using LTspice, the small-signal transistor parameters are calculated when running a '.op' simulation. In Fig. 3.24, we saw the output file from a '.op' simulation of the inverting amplifier. Using 'Ctrl-L' to open the error log file from the simulation, we get the file shown in Fig. 3.38.

The error log file gives a warning that the channel length L for the transistor is smaller than what is recommended for a level 1 MOSFET. The level 1 MOSFET refers to the Shichman-Hodges transistor model, and the warning is an indication that this model is inaccurate for modern submicron transistors.

This message will appear whenever the Shichman-Hodges model is used for transistors with $L \leq 10$ µm. In order to get simulation results closer to the actual device behavior, more complicated transistor models must be used, see Section 3.7.

Next, we find a list of parameters for the transistor M_1. They include the bias current and voltages and the small-signal parameters. We see that the bias current and voltages are as calculated when designing the amplifier. Also, $g_m = 0.667$ mA/V which is the same value as found from Eq. (3.96). Also, $g_{ds} = 0$ as expected for $\lambda = 0$ and the values of all parasitic capacitances are 0 since no model parameters or transistor dimensions for calculating the capacitors have been specified.

Error log file

```
Instance "m1": Length shorter than recommended for a level 1 MOSFET.
Direct Newton iteration for .op point succeeded.
Semiconductor Device Operating Points:
                              --- MOSFET Transistors ---
Name:           m1
Model:       nmos-sh
Id:          1.00e-04
Vgs:         7.00e-01
Vds:         9.00e-01
Vbs:         0.00e+00
Vth:         4.00e-01
Vdsat:       3.00e-01
Gm:          6.67e-04
Gds:         0.00e+00
Gmb:         1.99e-04
Cbd:         0.00e+00
Cbs:         0.00e+00
Cgsov:       0.00e+00
Cgdov:       0.00e+00
Cgbov:       0.00e+00
Cgs:         0.00e+00
Cgd:         0.00e+00
Cgb:         0.00e+00
```

Figure 3.38: Error log file from the '.op' simulation of the inverting amplifier from Fig. 3.24.

Next, we may re-simulate the circuit with $\lambda = 0.1$ V^{-1}. The resulting output file and error log file are shown in Fig. 3.39. We see that the drain bias current is increased slightly and the bias value of the output voltage is reduced. This is as expected when the drain current is multiplied by the factor $(1 + \lambda v_{DS})$. From the error log file, we find $g_m = 0.722$ mA/V, slightly more than the previous value for $\lambda = 0$, and $g_{ds} = 0.010$ mA/V, so Eq. (3.94) gives a gain of $A_d = -5.96$ V/V. Using Eq. (3.98), we found $A_v = -5.54$ V/V. The difference between this value and the value calculated from Eq. (3.94) is caused by the larger value of the drain current, giving a larger value of g_m according to Eq. (3.95).

Output file

```
        --- Operating Point ---

V(vo):          0.825375        voltage
V(vin):         0.7             voltage
V(vdd):         1.8             voltage
Id(M1):         0.000108292     device_current
Ig(M1):         0               device_current
Ib(M1):         -8.35375e-013   device_current
Is(M1):         -0.000108292    device_current
I(Rd):          0.000108292     device_current
I(Vdd):         -0.000108292    device_current
I(Vin):         0               device_current
```

Error log file

```
                              --- MOSFET Transistors ---
Name:           m1
Model:       nmos-sh
Id:          1.08e-04
Vgs:         7.00e-01
Vds:         8.25e-01
Vbs:         0.00e+00
Vth:         4.00e-01
Vdsat:       3.00e-01
Gm:          7.22e-04
Gds:         1.00e-05
Gmb:         2.16e-04
```

Figure 3.39: Output file and error log file from the '.op' simulation of the inverting amplifier from Fig. 3.24 with λ = 0.1 V^{-1}.

LTspice can also directly find the low-frequency small-signal voltage gain by a 'DC Transfer' simulation with the simulation directive '`.tf v(VO) VIN`'. Running the '.tf' simulation (with $\lambda = 0.1$ V^{-1}) results in the output file shown in Fig. 3.40. From this, we find a gain of -5.96 V/V, corresponding to the value calculated from Eq. (3.94). We also find an output resistance of 8.256 kΩ, corresponding to $(R_D \parallel r_{ds})$.

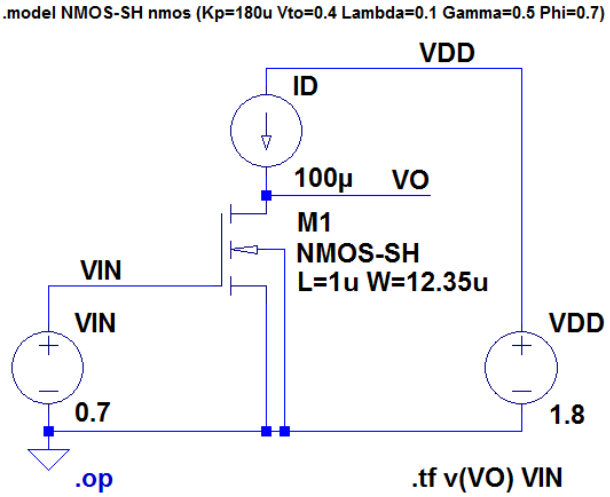
Finally, let us examine the voltage gain when R_D is replaced by a dc current source I_D as shown in Fig. 3.41. For this case, we must specify $\lambda \neq 0$ in the transistor model. Otherwise, the output node is left floating as the node between two series-connected ideal current sources. In Fig. 3.41, I_D has been specified to $I_D = 100$ µA which is the value used for our initial analysis of the circuit in Fig. 3.23.

Output file

```
        --- Transfer Function ---

Transfer_function:            -5.96084        transfer
vin#Input_impedance:          1e+020          impedance
output_impedance_at_V(vo):    8256.64         impedance
```

Figure 3.40: Output file from a '.tf' simulation of the inverting amplifier from Fig. 3.24 with λ = 0.1 V^{-1}.



Output file

```
        --- Transfer Function ---

Transfer_function:            -4.88634        transfer
vin#Input_impedance:          1e+020          impedance
output_impedance_at_V(vo):    8487.89         impedance
```

Figure 3.41: LTspice schematic and output file from a '.tf' simulation for an inverting amplifier with an ideal current source as the drain bias current.

Figure 3.41 also shows the output file from a '.tf' simulation of the circuit. We see that the gain is only -4.89 V/V, much smaller than expected from Eq. (3.99). Running a '.op' simulation reveals the reason for this. The output file from a '.op' simulation shows that the bias value of the output voltage has dropped to 0.253 V which is smaller than the saturation voltage of $V_{GS} - V_t = 0.3$ V, so the transistor is in the triode region, rather than in the active region.

In order to bias the transistor in the active region, we have different options:

- We can reduce the input bias voltage. With a dc sweep of the input voltage, we can find a bias value which gives an output bias voltage of 0.9 V. This results in $V_{IN} = 0.687$ V and from Eq. (3.99), we find $A_v = -76.0$ V/V. A '.tf' simulation gives the same value.
- We can reduce the width W of the transistor. With a '.op' simulation including a '.step' directive with W as a parameter (compare to Fig. 3.15), we can find the value of W required to get an output voltage of $V_O = 0.9$ V for an input voltage of $V_{IN} = 0.7$ V. This results in $W = 11.33$ µm and from Eq. (3.99), we find $A_v = -72.7$ V/V. A '.tf' simulation gives the same value.
- We can increase the bias current I_D. With a dc sweep of the bias current, we can find a bias value which gives an output bias voltage of 0.9 V. This results in $I_D = 109$ µA and from Eq. (3.99), we find $A_v = -72.7$ V/V. A '.tf' simulation gives the same value.

In all three cases, the simulated value of A_v is exactly as calculated from Eq. (3.99).

Notice that LTspice automatically derives a small-signal equivalent circuit for use in '.tf' simulations and '.ac' simulations on basis of the large-signal schematic and the bias point. The conversion of the large-signal schematic to a small-signal equivalent circuit is needed only for hand analysis of the circuit.

The current mirror. In Section 3.4, we analyzed a current mirror with NMOS transistors and we defined a load regulation for the output of the current mirror as the variation of the output current versus the variation of the output voltage. Ideally, the output current should remain constant and independent of the output voltage. The load regulation is a small-signal parameter which may be interpreted as the small-signal output conductance of the current mirror, so it can be analyzed using a small-signal equivalent circuit. Directly replacing the transistors in Fig. 3.16 with the small-signal model from Fig. 3.28 and the dc voltage V_{DD} with a short circuit results in the small-signal equivalent circuit shown in Fig. 3.42.

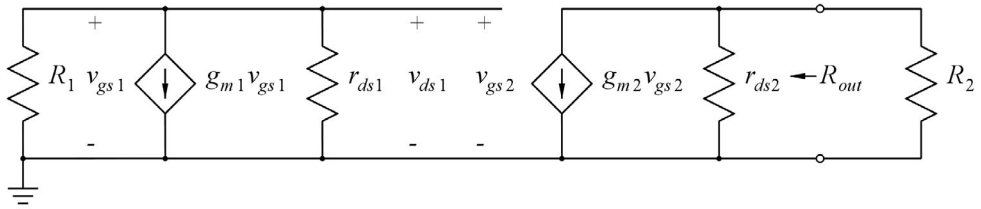


Figure 3.42: Small-signal schematic for the current mirror from Fig. 3.16.

At first glance, this may seem fairly complicated but we observe that the gate of M_1, the drain of M_1 and the gate of M_2 are connected, so $v_{gs1} = v_{ds1} = v_{gs2}$, and with this relation, a node equation at this node yields $v_{ds1}/R_1 + g_{m1}v_{gs1} + v_{ds1}/r_{ds1} = 0 \Rightarrow v_{gs1} = v_{ds1} = v_{gs2} = 0$. This implies that there is no signal variation at the inputs of the voltage-controlled current sources, and with a value of 0, they correspond to open circuits and can be deleted. The small-signal schematic of the current mirror simply reduces to the resistor r_{ds2} and the load resistor R_2 as shown in Fig. 3.43.

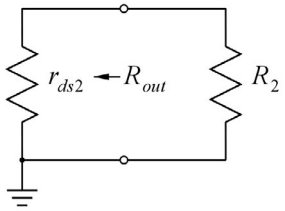


Figure 3.43: Reduced small-signal schematic for the current mirror from Fig. 3.16.

From this, we immediately find that the output conductance of the current mirror is

$$g_{ds2} = \frac{\mu_n C_{ox}}{2}\left(\frac{W_2}{L_2}\right)(V_{GS} - V_t)^2\,\lambda \tag{3.100}$$

corresponding to the result in Eq. (3.32).

We saw in this example that transistors with constant gate-source voltage reduce to resistors (and capacitors for high-frequency analysis), thus significantly simplifying the small-signal equivalent circuit. This is typically the case for transistors used to provide dc bias currents in CMOS circuits.

3.7 Advanced transistor models

The Shichman-Hodges model described in Section 3.3 was the first MOS transistor model to be widely used for the SPICE simulation program in the 1970s when transistor dimensions were typically in the µm-range (Vladimirescu 1994). However, as dimensions were scaled down, the Shichman-Hodges model turned out to be increasingly inaccurate. This is caused by a number of physical mechanisms not taken into account in the simple transistor model. Even with small voltages applied to gate and drain, the electric fields in the MOS devices become so large for submicron devices that a number of phenomena occur (Chan Carusone, Johns & Martin 2012). These include a reduction of the mobility of the free carriers in the channel and short-channel effects which change the output characteristics much more than described by the simple channel-length modulation included in Eq. (3.16). The resulting relations become too complex for simple hand calculations, and it is beyond the scope of this book to consider physical and mathematical models for these phenomena.

At very small and even negative values of the effective gate voltage, another phenomenon called weak inversion causes the transistor to follow a relation which is fundamentally different from Eq. (3.16). In weak inversion, also called the subthreshold region, the drain current is approximately described by

$$i_D = I_{D0}\left(\frac{W}{L}\right)\left(\exp\frac{v_{GS}}{nV_T}\right)\left(1-\exp\left(\frac{-v_{DS}}{V_T}\right)\right) \tag{3.101}$$

In Eq. (3.101), $V_T = kT/q$ is the thermal voltage while n and I_{D0} are parameters depending on the fabrication process and temperature (Vittoz 2004). For v_{DS} larger than approximately $3V_T$, the factor $(1-\exp(-v_{DS}/V_T))$ is close to 1 and the transistor is in the active subthreshold region. In order to model the finite output conductance of the transistor, Eq. (3.101) may be augmented with a factor $(1+\lambda v_{DS})$, similar to what was done in the Shichman-Hodges model. For a MOS transistor in a typical 0.18 µm process, Eq. (3.101) applies for a gate-source voltage smaller than the threshold voltage and Eq. (3.16) applies when the gate-source voltage is more than about 100 mV larger than the threshold voltage (the strong inversion region). The region between weak inversion and strong inversion is called the moderate inversion region, and this region of operation is often used for circuits designed to operate at very low supply voltages. Unfortunately, simple mathematical expressions for hand calculations are not available for the moderate inversion region. Alternatively, design methods based on pre-computed lookup tables may be used (Jespers & Murmann 2017).

BSIM transistor models. In order to analyze circuits with greater precision than what is obtained from the simplified equations (3.15), (3.16) and (3.101), more accurate simulation models must be used. For submicron processes, BSIM3 or BSIM4 models are often the preferred choice. BSIM is an abbreviation for 'Berkeley Short-channel IGFET Model', referring to University of California Berkeley where they have been developed (Sheu et al. 1987). Generic BSIM models are available together with several textbooks such as Chan Carusone, Johns & Martin (2012) and Baker (2010).

Models for specific CMOS processes can be obtained from foundries or from MPW (Multi-Project Wafer) service providers such as MOSIS (MPW) Integrated Circuit (IC) Fabrication Service Provider (The MOSIS Service 2022) or EUROPRACTICE (EUROPRACTICE IC Service 2022). However, most foundries require non-disclosure agreements in order to provide detailed design information.

Generic BSIM3 model for 0.18 µm CMOS process. Adapted from 'http://ptm.asu.edu/modelcard/180nm_bulk.txt '.

```
* Predictive Technology Model Beta Version
* 180nm NMOS SPICE Parametersv (normal one)
.model NMOS-BSIM NMOS
+Level=49
+Lint=4.e-08              Tox=4.e-09
+Vth0=0.3999              Rdsw=250
+Tref=27.0                version=3.1
+Xj=6.0000000E-08         Nch=5.9500000E+17
+lln=1.0000000            lwn=1.0000000          wln=0.00
+wwn=0.00                 ll=0.00
+lw=0.00                  lwl=0.00               wint=0.00
+wl=0.00                  ww=0.00                wwl=0.00
+Mobmod=1                 binunit=2
+Dwg=0.00                 Dwb=0.00
+K1=0.5613000             K2=1.0000000E-02
+K3=0.00                  Dvt0=8.0000000         Dvt1=0.7500000
+Dvt2=8.0000000E-03       Dvt0w=0.00             Dvt1w=0.00
+Dvt2w=0.00               Nlx=1.6500000E-07      W0=0.00
+K3b=0.00                 Ngate=5.0000000E+20
+Vsat=1.3800000E+05       Ua=-7.0000000E-10      Ub=3.5000000E-18
+Uc=-5.2500000E-11        Prwb=0.00
+Prwg=0.00                Wr=1.0000000           U0=3.5000000E-02
+A0=1.1000000             Keta=4.0000000E-02     A1=0.00
+A2=1.0000000             Ags=-1.0000000E-02     B0=0.00
+B1=0.00
+Voff=-0.12350000         NFactor=0.9000000      Cit=0.00
+Cdsc=0.00                Cdscb=0.00             Cdscd=0.00
+Eta0=0.2200000           Etab=0.00              Dsub=0.8000000
+Pclm=5.0000000E-02       Pdiblc1=1.2000000E-02  Pdiblc2=7.50E-03
+Pdiblcb=-1.3500000E-02   Drout=1.7999999E-02    Pscbe1=8.66E+08
+Pscbe2=1.0000000E-20     Pvag=-0.2800000        Delta=1.00E-02
+Alpha0=0.00              Beta0=30.0000000
+kt1=-0.3700000           kt2=-4.0000000E-02     At=5.5000000E+04
+Ute=-1.4800000           Ua1=9.5829000E-10      Ub1=-3.3473E-19
+Uc1=0.00                 Kt11=4.0000000E-09     Prt=0.00
+Cj=0.00365               Mj=0.54                Pb=0.982
+Cjsw=7.9E-10             Mjsw=0.31              Php=0.841
+JS=1.50E-08              JSW=2.50E-13
+N=1.0                    Xti=3.0                Cgdo=2.786E-10
+Cgso=2.786E-10           Cgbo=0.0E+00           Capmod= 2
+NQSMOD=0                 Elm=5                  Xpart=1
+Cgsl=1.6E-10             Cgdl=1.6E-10           Ckappa=2.886
+Cf=1.069e-10             Clc=0.0000001          Cle=0.6
+Dlc=4E-08                Dwc=0                  Vfbcv=-1
```

```
* Predictive Technology Model Beta Version
* 180nm PMOS SPICE Parametersv (normal one)
.model PMOS-BSIM PMOS
+Level=49
+Lint=3.e-08              Tox=4.2e-09
+Vth0=-0.42               Rdsw=450
+Tref=27.0                version=3.1
+Xj=7.0000000E-08         Nch=5.9200000E+17
+lln=1.0000000            lwn=1.0000000          wln=0.00
+wwn=0.00                 ll=0.00
+lw=0.00                  lwl=0.00               wint=0.00
+wl=0.00                  ww=0.00                wwl=0.00
+Mobmod= 1                binunit=2
+Dwg=0.00                 Dwb=0.00
+K1=0.5560000             K2=0.00
+K3=0.00                  Dvt0=11.2000000        Dvt1=0.7200000
+Dvt2=-1.0000000E-02      Dvt0w=0.00             Dvt1w=0.00
+Dvt2w=0.00               Nlx=9.5000000E-08      W0=0.00
+K3b=0.00                 Ngate=5.0000000E+20
+Vsat=1.0500000E+05       Ua=-1.2000000E-10      Ub=1.0000000E-18
+Uc=-2.9999999E-11        Prwb=0.00
+Prwg=0.00                Wr=1.0000000           U0=8.0000000E-03
+A0=2.1199999             Keta=2.9999999E-02     A1=0.00
+A2=0.4000000             Ags=-0.1000000         B0=0.00
+B1=0.00
+Voff=-6.40000000E-02     NFactor=1.4000000      Cit=0.00
+Cdsc=0.00                Cdscb=0.00             Cdscd=0.00
+Eta0=8.5000000           Etab=0.00              Dsub=2.8000000
+Pclm=2.0000000           Pdiblc1=0.1200000      Pdiblc2=8.00E-05
+Pdiblcb=0.1450000        Drout=5.0000000E-02    Pscbe1=1.00E-20
+Pscbe2=1.0000000E-20     Pvag=-6.0000000E-02    Delta=1.00E-02
+Alpha0=0.00              Beta0=30.0000000
+kt1=-0.3700000           kt2=-4.0000000E-02     At=5.5000000E+04
+Ute=-1.4800000           Ua1=9.5829000E-10      Ub1=-3.3473E-19
+Uc1=0.00                 Kt11=4.0000000E-09     Prt=0.00
+Cj=0.00138               Mj=1.05                Pb=1.24
+Cjsw=1.44E-09            Mjsw=0.43              Php=0.841
+JS=1.50E-08              JSW=2.50E-13
+N=1.0                    Xti=3.0                Cgdo=2.786E-10
+Cgso=2.786E-10           Cgbo=0.0E+00           Capmod=2
+NQSMOD=0                 Elm=5                  Xpart=1
+Cgsl=1.6E-10             Cgdl=1.6E-10           Ckappa=2.886
+Cf=1.058e-10             Clc=0.0000001          Cle=0.6
+Dlc=3E-08                Dwc=0                  Vfbcv=-1
```

Figure 3.44: Library file 'BSIM3_018.lib' with BSIM3 models for a generic 0.18 µm CMOS process, adapted from Predictive Technology Model Website (2011).

In this book, we show as an example a generic BSIM3 model for a 0.18 µm CMOS process obtained from Predictive Technology Model Website (2011) developed at Arizona State University (Cao 2011). From this website, generic models can be found for a large number of technologies, and Fig. 3.44 shows BSIM3 models for use with LTspice. The color coding used for this figure is the default color coding used by LTspice when displaying text files such as model files or netlist files.

The two models NMOS-BSIM and PMOS-BSIM are contained in a single file named BSIM3_018.lib. This makes it possible to include the models in a schematic simply by giving a reference to this file. The SPICE directive for including the library file is '`.include BSIM3_018.lib`'. The library file should be placed in the same folder as the circuit schematic file.

The file shown in Fig. 3.44 has been adapted from 'http://ptm.asu.edu/modelcard/180nm_bulk.txt ' which is found in Predictive Technology Model Website (2011). A few of the parameters in this file have been deleted since they are ignored by LTspice anyway. The parameters which are deleted will not influence the simulation results significantly for normal values of transistor geometries. Also, the transistor models are named 'NMOS-BSIM' and 'PMOS-BSIM' rather than just 'NMOS' and 'PMOS' in order to emphasize that they are BSIM models, not just the default Spice models.

Comparing the Shichman-Hodges model to the BSIM3 model. The BSIM3 models from Fig. 3.44 and the transistor parameters in Table 3.1 both refer to a generic 0.18 µm CMOS process. The BSIM3 models take both weak inversion, short-channel effects, mobility degradation and other deviations from the simple Shichman-Hodges model into account, so a direct comparison of the input characteristics (see Fig. 3.8) and the output characteristics (see Fig. 3.9) of a MOS transistor modeled by the Shichman-Hodges model and one modeled by the BSIM3 model will illustrate the accuracy which can be expected from hand calculations based on the Shichman-Hodges model.

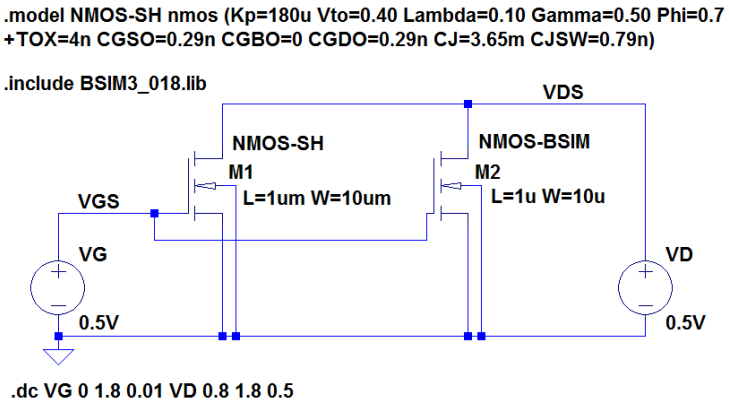


Figure 3.45: LTspice schematic for simulating input characteristics and output characteristics for both the Shichman-Hodges model and the BSIM3 model for an NMOS transistor with L=1 µm and W=10 µm.

For simulating the transistor characteristics, we use the circuit shown in Fig. 3.45 with two transistors, M_1 modeled by the Shichman-Hodges and M_2 modeled by the BSIM3 model. A dc sweep with v_G as the first source to sweep and v_D as the second source gives the input characteristics. Running the dc sweep with v_D as the first source and v_G as the second source results in the output characteristics. Figure 3.46 shows the input and output characteristics for both transistors. We observe that for small values of the gate-source voltage, $v_{GS} < 1$ V, there is a reasonable similarity between the Shichman-Hodges model and the BSIM3 model, but for $v_{GS} > 1$ V, the difference between the models is very pronounced.

Zooming in on a range from 0.3 V to 1.0 V for v_{GS}, we find the input characteristics shown in Fig. 3.47(a). The curves for M_1 (green curves) and M_2 (blue curves) look similar but we may note some differences:

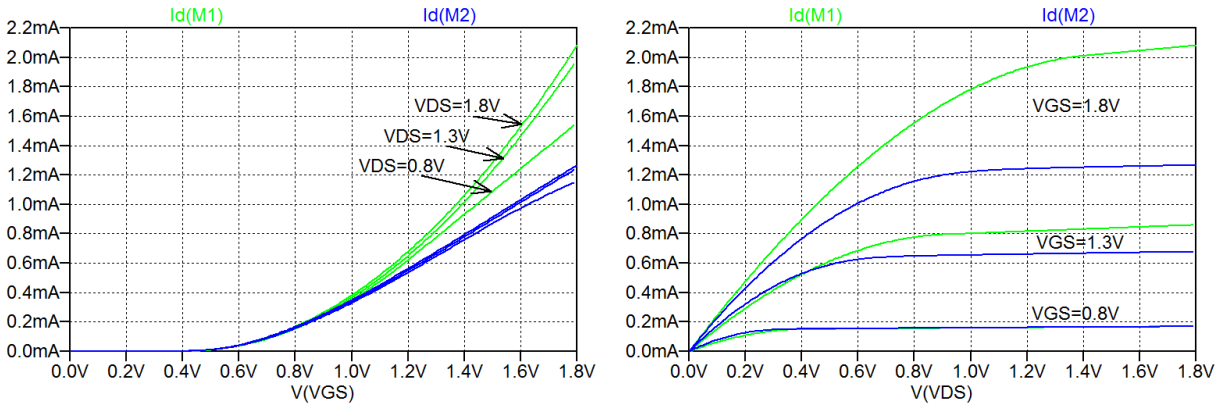


Figure 3.46: Input characteristics, i_D versus v_{GS} (a), and output characteristics, i_D versus v_{DS} (b), for the transistors from Fig. 3.45. The green curves show the Shichman-Hodges model. The blue curves show the BSIM3 model.

At voltages around the threshold voltage of 0.4 V, the BSIM3 model gives a larger current than the Shichman-Hodges model because of the subthreshold conduction in the transistor. Figure 3.47(b) shows the characteristics zoomed in around the threshold voltage, and we see that for $v_{GS} < V_t + 100$ mV, we cannot expect reasonable results from the Shichman-Hodges model.

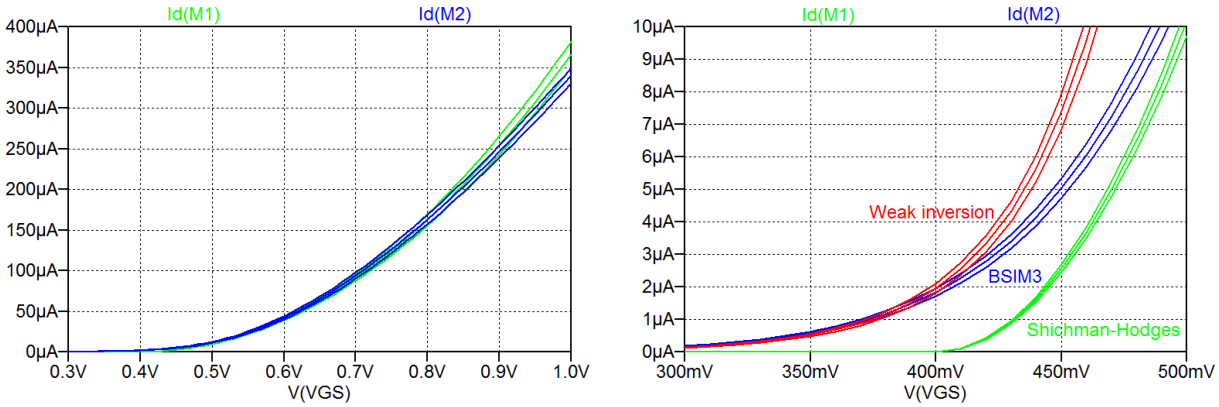


Figure 3.47: Close-up on the input characteristics for $v_{GS} < 1$ V (a) and $v_{GS} = V_t \pm 100$ mV, weak and moderate inversion region (b). The green curves show the Shichman-Hodges model. The blue curves show the BSIM3 model and the red curves shows the weak-inversion model given by Eq. (3.101), including a factor $(1 + \lambda v_{DS})$.

Also shown in Fig. 3.47(b) are red curves which are the drain current calculated from Eq. (3.101), including a factor $(1 + \lambda v_{DS})$, with $I_{D0} = 4.0$ pA, $n = 1.47$ and $\lambda = 0.18$ V^{-1} which are values fitted to match the BSIM3 model. Apparently, the weak-inversion model Eq. (3.101) provides a reasonable fit to the BSIM model for $v_{GS} < V_t$.

Also, for $v_{GS} > 0.9$ V, the difference in Fig. 3.47(a) between the BSIM3 model and the Shichman-Hodges model becomes more pronounced. In particular, the slope of the BSIM3 curves is smaller than the slope of the Shichman-Hodges curves. The slope is $\partial i_D / \partial v_{GS} = g_m$, and plotting '`d(Id(M1))`' and

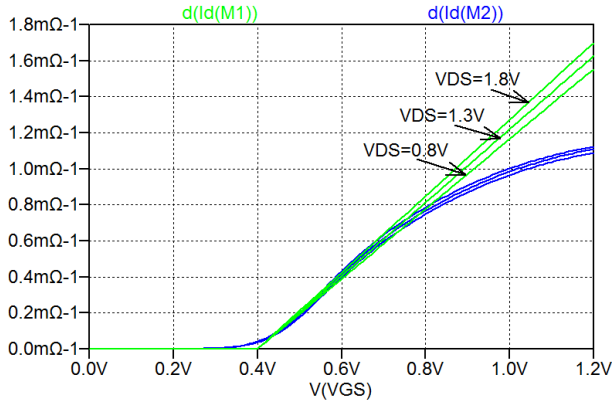


Figure 3.48: The transconductance g_m versus V_{GS}. The green curves show the Shichman-Hodges model. The blue curves show the BSIM3 model.

'`d(Id(M2))`' for $v_{GS} \leq 1.2$ V where the transistors are in the active region, we get Fig. 3.48. This clearly shows that for the BSIM3 model, g_m is not proportional to $(V_{GS} - V_t)$ as given by Eq. (3.59). Rather, it tends to saturate for high values of v_{GS}, and this is caused by the mobility degradation occurring for large electric fields (Chan Carusone, Johns & Martin 2012). However, it is beyond the scope of this book to describe the physical and mathematical modeling of the mobility degradation.

In conclusion, the Shichman-Hodges model provides reasonable results when the effective gate voltage is a few hundred mV but for larger and smaller gate voltages, simulations using a more accurate model are required. Nevertheless, a first design iteration is often performed on basis of the Shichman-Hodges model because it provides some insight into the relations between the physical and electrical design parameters and the resulting performance of the circuit, and an analysis based on the Shichman-Hodges model can be performed using reasonably simple calculations resulting in closed-form expressions for the performance parameters of a circuit.

References

Baker, RJ. 2010, *CMOS Circuit Design, Layout, and Simulation*, Third Edition, IEEE Press, Piscataway, USA.

Binkley, DM. 2008, *Tradeoffs and Optimization in Analog CMOS Design*, John Wiley & Sons, Hoboken, USA.

Brinkman, WF., Haggan, DE. & Troutman, WW. 1997, 'A History of the Invention of the Transistor and Where It Will Lead Us', *IEEE J. Solid-State Circuits*, vol. 32, no. 12, pp. 1858-1865, December 1997.

Brocard, G. 2013, *The LTspice IV Simulator – Manual, Methods and Applications*, First Edition, Swiridoff Verlag, Künzelsau, Germany.

Bruun, E. 2020, *CMOS Integrated Circuit Simulation with LTspice*, Third Edition, bookboon.
Available from: http://bookboon.com/en/cmos-integrated-circuit-simulation-with-ltspice-ebook

Cao, Y. 2011, *Predictive Technology Model for Robust Nanoelectronic Design*, Springer, New York, USA.

Chan Carusone, T., Johns, D. & Martin, K. 2012, *Analog Integrated Circuit Design*, Second Edition, International Student Version, John Wiley & Sons, Inc., Hoboken, USA.

EUROPRACTICE IC Service 2022. Retrieved from https://europractice-ic.com/

Grove, AS. 1967, *Physics and Technology of Semiconductor Devices*, John Wiley & Sons, New York, USA.

Jespers, PGA. & Murmann, B. 2017, *Systematic Design of Analog CMOS Circuits Using Pre-Computed Lookup Tables*, Cambridge University Press, Cambridge, UK.

Lilienfeld, JE. 1925, *Method and apparatus for controlling electric current*, US Patent 1745175.

Lilienfeld, JE. 1928, *Device for controlling electric current*, US Patent 1900018.

Plummer, JD., Deal, M. & Griffin, PD. 2000, *Silicon VLSI Technology: Fundamentals, Practice, and Modeling*, First Edition, Pearson Education Ltd., Harlow, UK.

Predictive Technology Model Website 2011. Retrieved from http://ptm.asu.edu/

Pujol, HL. 1975, 'Power Supply Considerations for COS/MOS Devices', RCA Application Note ICAN-6576, *1975 RCA Solid State Databook SSD-203C, COS/MOS Integrated Circuits*, pp. 562-567.

Sah, CT. 1964, 'Characteristics of the Metal-Oxide-Semiconductor Transistors', *IEEE Trans. Electron Devices*, vol. ED-11, no. 7, pp. 324-345, July 1964.

Sah, C-T. 1988, 'Evolution of the MOS Transistor – From Conception to VLSI', *Proc. IEEE*, vol. 76, no. 10, pp. 1280-1326, October 1988.

Sedra, AS. & Smith, KC. 2016, *Microelectronic Circuits*, International Seventh Edition, Oxford University Press, New York, USA.

Sheu, BJ., Scharfetter, DL., Ko, P-K. & Jeng, M-C. 1987, 'BSIM: Berkeley Short-Channel IGFET Model for MOS Transistors', *IEEE J. Solid State Circuits*, vol. SC-22, no. 4, pp. 558-566.

Shichman, H. & Hodges, DA. 1968, 'Modeling and Simulation of Insulated-Gate Field-Effect Transistor Switching Circuits', *IEEE J. Solid-State Circuit*, vol. SC-3, no. 3, pp. 285-289.

Shockley, W. 1949, 'The Theory of P-N Junctions in Semiconductors and P-N Junction Transistors', *Bell System Technical Journal*, vol. 28, no. 3, pp. 435-489.

The MOSIS Service 2022. Retrieved from https://www.themosisservice.com/

The Transistor 2014. *Nobelprize.org.* Nobel Media AB.
Retrieved from http://www.nobelprize.org/educational/physics/transistor/function/firsttransistor.html

Tsividis, Y. & McAndrew, C. 2010, *Operation and Modeling of the MOS Transistor*, Third Edition, Oxford University Press, New York, USA.

Vittoz, EA. 2004, 'Weak Inversion for Ultimate Low-Power Logic' in C. Piguet (Ed.), *Low-Power Electronics Design*, Chapter 16, CRC Press, Boca Raton, USA.

Vladimirescu, A. 1994, *The SPICE book*, First Edition, John Wiley & Sons, Hoboken, USA.

Multiple-choice test

1. Complete the following statements by selecting the appropriate continuation from the table below.

A:	The mobile electric carriers in the channel of an NMOS transistor are ...
B:	The drain and source regions of a PMOS transistor are ...
C:	The channel in an NMOS transistor is formed when the gate-source voltage is ...
D:	When the gate-source voltage for a PMOS transistor is positive, the transistor is ...
E:	For an NMOS transistor in the triode region, the drain current is ...
F:	The channel-length modulation in a MOS transistor causes the transistor to have ...
G:	The body effect in a MOS transistor occurs when ...
H:	The small-signal parameter g_m for a MOS transistor describes ...
I:	The small-signal parameter g_{ds} for a MOS transistor describes ...
J:	The Shichman-Hodges model describes ...

Continuation:

1:	a linearized relation between the drain current and the drain-source voltage.
2:	in the active region.
3:	holes.
4:	electrons.
5:	n-doped regions.
6:	a linearized relation between the drain current and the gate-source voltage.
7:	larger than the threshold voltage.
8:	a nonlinear relation between the drain current, the gate-source voltage and the drain-source voltage.
9:	p-doped regions.
10:	smaller than the threshold voltage.
11:	in the cut-off region.
12:	in the triode region.
13:	$i_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(v_{GS}-V_t)^2(1+\lambda v_{DS})$.
14:	$i_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left(2(v_{GS}-V_t)v_{DS}-v_{DS}^2\right)(1+\lambda v_{DS})$.
15:	larger than the threshold voltage.
16:	a drain current depending on the drain-source voltage.
17:	a threshold voltage depending on the source-bulk voltage.
18:	an exponential relation between drain current and gate-source voltage.
19:	drain and bulk have different voltages.
20:	source and bulk have different voltages.

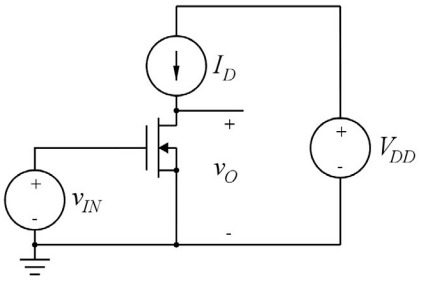
2. For an NMOS transistor with $V_t = 0.4$ V, $\mu_n C_{ox} = 180$ µA/V^2, $W/L = 10$, $\lambda = 0$, $v_{GS} = 1.2$ V and $v_{DS} = 1.0$ V, the drain current is

   A: 0 mA
   B: 0.540 mA
   C: 0.576 mA

3. For a PMOS transistor in the active region with $V_t = -0.42$ V, $\mu_p C_{ox} = 45$ µA/V^2, $W/L = 10$, $\lambda = 0$, and $i_D = 324$ µA, the gate-source voltage is

   A: -1.62 V
   B: -1.20 V
   C: $+1.62$ V

4. For the circuit shown below, assume that the transistor is in the active region. The transistor has a threshold voltage $V_t = 0.4$ V and a channel-length modulation parameter $\lambda = 0.2\ \text{V}^{-1}$. The bias current I_D is 200 μA and the bias values of v_{IN} and v_O are both 0.9 V.



The transconductance of the transistor is

A: 0.4 mA/V
B: 0.8 mA/V
C: 1.6 mA/V

5. The small-signal output resistance of the transistor in the circuit above is

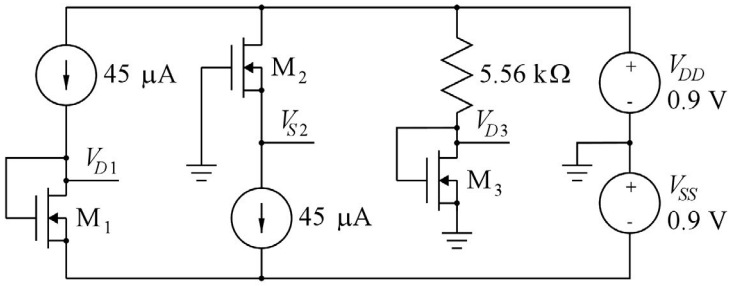
A: 15.0 kΩ
B: 29.5 kΩ
C: 59.0 kΩ

6. The small-signal voltage gain in the circuit above is

A: −23.6 V/V
B: +23.6 V/V
C: −12.0 V/V

Problems

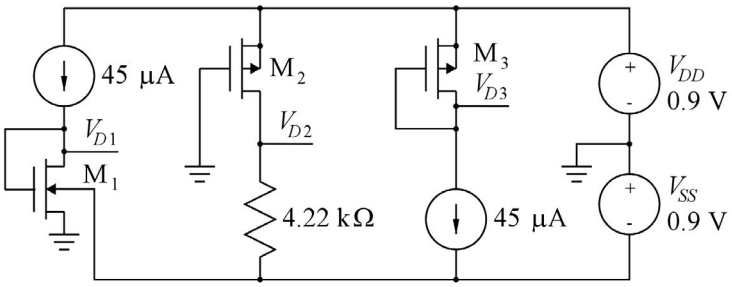
Problem 3.1



For the transistors in the circuit shown above, assume $W/L = 8$, $\mu_n C_{ox} = 180\ \mu\text{A/V}^2$, $V_t = 0.4$ V and $\lambda = 0$.

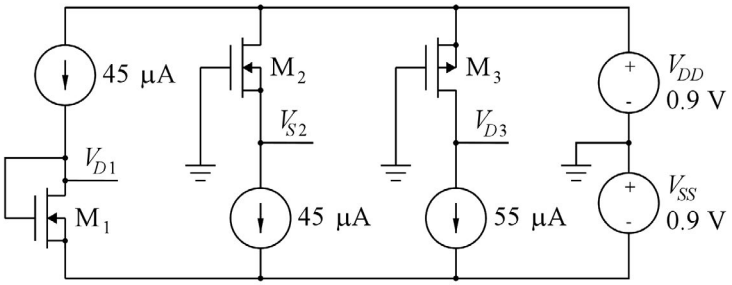
Calculate the voltages V_{D1}, V_{S2} and V_{D3}.

Problem 3.2



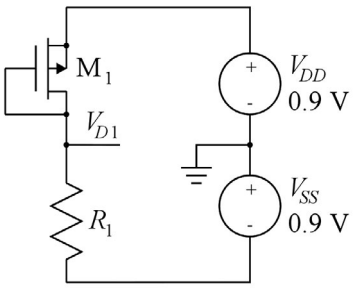
For the NMOS transistor in the circuit shown above, assume $W/L = 8$, $\mu_n C_{ox} = 180\ \mu\text{A/V}^2$, $V_{to} = 0.4$ V, $\lambda = 0$, $\gamma = 0.5\sqrt{\text{V}}$ and $|2\Phi_F| = 0.7$ V.
For the PMOS transistors in the circuit, $W/L = 32$, $\mu_p C_{ox} = 45\ \mu\text{A/V}^2$, $V_t = -0.42$ V and $\lambda = 0$.
Calculate the voltages V_{D1}, V_{D2} and V_{D3}.

Problem 3.3



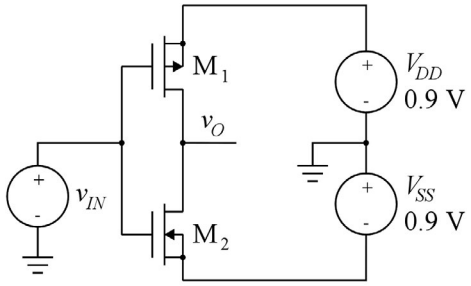
For the NMOS transistors shown in the circuit above, assume $W = 16\ \mu$m, $L = 2\ \mu$m, $\mu_n C_{ox} = 180\ \mu\text{A/V}^2$, $V_t = 0.4$ V and $\lambda L = 0.1\ \mu$m/V.
For the PMOS transistor, assume $W = 30\ \mu$m, $L = 3\ \mu$m, $\mu_p C_{ox} = 45\ \mu\text{A/V}^2$, $V_t = -0.42$ V and $\lambda L = 0.14\ \mu$m/V.
Calculate or simulate the voltages V_{D1}, V_{S2} and V_{D3}.

Problem 3.4



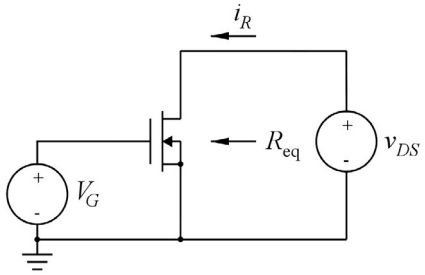
For the circuit shown above, find the resistor R_1 and the transistor width W_1 so that $V_{D1} = 0.2$ V and $I_{D1} = 10\ \mu$A. Assume $L_1 = 3\ \mu$m, $\mu_p C_{ox} = 45\ \mu\text{A/V}^2$, $V_t = -0.42$ V and $\lambda = 0$.
Repeat for $\lambda L = 0.14\ \mu$m/V.

Problem 3.5



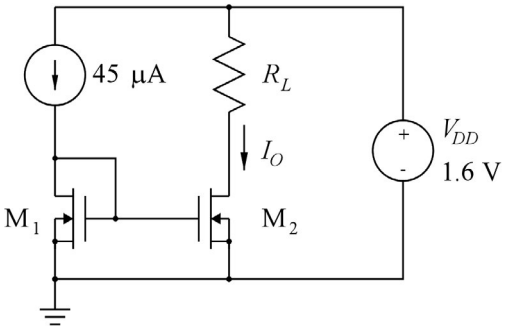
For the circuit shown above, design the width of transistor M_1 so that the output voltage v_O is 0 for an input voltage $v_{IN} = 0$. Assume $L_1 = L_2 = 0.5$ µm and $W_2 = 0.5$ µm.
For the PMOS transistor, assume $\mu_p C_{ox} = 45$ µA/V^2, $V_t = -0.42$ V and $\lambda L = 0.14$ µm/V.
For the NMOS transistor, assume $\mu_n C_{ox} = 180$ µA/V^2, $V_t = 0.4$ V and $\lambda L = 0.1$ µm$/V$.
Simulate and plot the output voltage v_O versus the input voltage for -0.9 V $\leq v_{IN} \leq 0.9$ V.
Find the output voltage for $v_{IN} = -0.2$ V and for $v_{IN} = 0.2$ V.

Problem 3.6



An NMOS transistor biased in the triode region can be operated as a voltage-controlled linear resistor R_{eq} for very small values of v_{DS} as shown above. The voltage V_G controls the value of the resistance between drain and source.
Assume the following transistor parameters: $W = 2$ µm, $L = 20$ µm, $\mu_n C_{ox} = 180$ µA/V^2, $V_t = 0.4$ V and $\lambda = 0$. What is the range of resistance values which can be obtained for 0.9 V $\leq V_G \leq$ 1.8 V?
For $V_G = 1.2$ V, what is the maximum value of v_{DS} which can be applied if the error in calculating i_R from v_{DS}/R_{eq} should be less than 5% ?

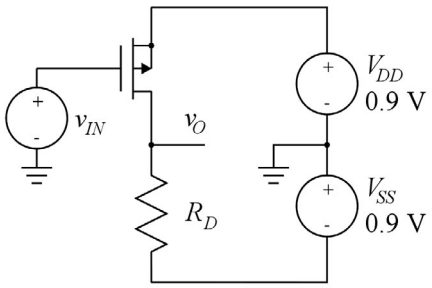
Problem 3.7



Design the current mirror shown above to provide an output current of $I_O = 90$ µA for $0 \leq R_L \leq 15$ kΩ. Assume the following transistor parameters: $L_1 = L_2 = 1$ µm, $\mu_n C_{ox} = 180$ µA/V^2, $V_t = 0.4$ V and $\lambda = 0$. Use a supply voltage $V_{DD} = 1.6$ V and use minimum values of W_1 and W_2 for which the current mirror fulfills the specifications.

Now, assume that $\lambda = 0.1$ V^{-1} for both transistors. Draw a small-signal equivalent circuit for the circuit and find the small-signal output resistance of the current mirror, assuming a bias value of 0.65 V at the output of the current mirror.

Problem 3.8

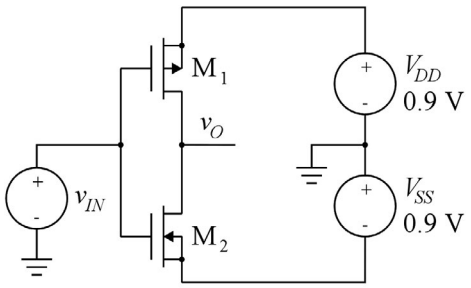


Shown above is an inverting amplifier implemented by a PMOS transistor and a drain resistor R_D. Assume the following transistor parameters: $L = 1$ µm, $\mu_p C_{ox} = 45$ µA/V^2, $V_t = -0.42$ V and $\lambda = 0$. The output voltage has a bias value of 0, and the input voltage v_{IN} is the sum of a bias voltage V_{IN} and a small-signal voltage v_{in}.

Draw a small-signal equivalent circuit for the amplifier and derive an expression for the small-signal voltage gain v_o/v_{in}. Calculate the input bias voltage so that a small-signal gain of -6 V/V is obtained.

Calculate R_D and W so that a bias current of 80 µA is obtained. Use LTspice to simulate and plot the output voltage for a sinusoidal input with an amplitude of 30 mV and a frequency of 1 kHz. Repeat the simulation with an input amplitude of 300 mV.

Problem 3.9



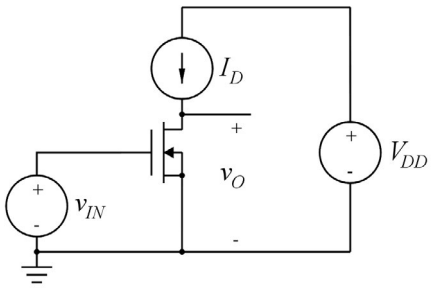
Shown above is an inverting amplifier implemented by a PMOS transistor and an NMOS transistor. Assume the following transistor geometries: $L_1 = L_2 = 0.5$ µm, $W_1 = 4$ µm and $W_2 = 1$ µm.

For the PMOS transistor, assume $\mu_p C_{ox} = 45$ µA/V^2, $V_t = -0.42$ V and $\lambda L = 0.14$ µm/V.

For the NMOS transistor, assume $\mu_n C_{ox} = 180$ µA/V^2, $V_t = 0.4$ V and $\lambda L = 0.1$ µm/V.

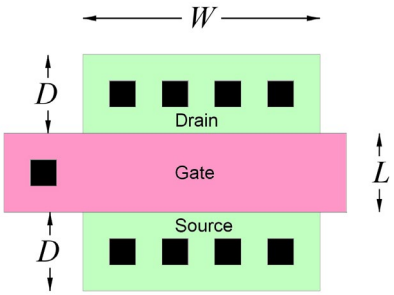
Draw a small-signal equivalent circuit for the amplifier and calculate the values of the small-signal parameters for an input bias voltage of $V_{IN} = 0$. Use LTspice to verify that both transistors are in the active region for $V_{IN} = 0$ and to check the values of the small-signal parameters.
Calculate the small-signal gain of the amplifier and verify the calculation with LTspice.

Problem 3.10



Draw a small-signal equivalent circuit for the inverting amplifier shown above, including the transistor capacitances C_{gs}, C_{gd} and C_{bd}. Find an expression for the transfer function $A_v(j\omega) = V_o(j\omega)/V_{in}(j\omega)$ and sketch the frequency response in a Bode plot, assuming the following values: $g_m = 1.0$ mA/V, $r_{ds} = 75$ kΩ, $C_{gs} = 120$ fF, $C_{gd} = 6$ fF and $C_{bd} = 60$ fF. Find the low-frequency gain and the -3 dB cutoff frequency.

Problem 3.11



Shown above is the layout of an NMOS transistor. The green areas are the drain and source diffusions and the pink area is the polysilicon layer forming the gate. The transistor has the following dimensions: $W = 1.8$ µm, $L = 0.6$ µm and $D = 0.6$ µm. Assume a gate oxide capacitance of $C_{ox} = 8.5$ fF/(µm)^2 and calculate the capacitance C_{gs}, neglecting gate overlap capacitances. Also assume a junction capacitance of 3.65 fF/(µm)^2 for the bottom and 0.79 fF/µm for the sidewall of the drain diffusion at zero reverse bias voltage and calculate C_{bd} for a drain-bulk voltage of 0.9 V.
Next, assume an overlap capacitance from gate towards drain and source of 0.3 fF/µm and calculate C_{gd} and C_{gs}, including the overlap capacitance from gate to source.
Calculate the transconductance of the transistor for an effective gate-source voltage of 0.4 V and a drain-source voltage of 0.9 V, assuming $\mu_n C_{ox} = 180$ µA/V^2 and $\lambda = 0.10$ µmV$^{-1}/L$.
Finally, calculate the transition frequency f_T.

Problem 3.12

```
.model NMOS-SH nmos (Kp=180u Vto=0.40
+lambda={0.1u/L} gamma=0.5 phi=0.7
+Tox=4n CGSO=0.29n CGBO=0
+CGDO=0.29n CJ=3.65m CJSW=0.79n)
```

Use LTspice with the transistor model shown above to simulate the transistor capacitances and the transition frequency f_T of the transistor from Problem 3.11.
Suggest a change in the layout of the transistor which will increase the transition frequency f_T by a factor of 4.

Chapter 4 – Basic Gain Stages

The MOS transistor can be configured in different ways in order to implement gain stages with different properties. In this chapter, we consider three basic single-transistor stages: The common-source stage, the common-drain stage and the common-gate stage. The transistor terminal which is termed the 'common' terminal is small-signal ground for both the input side and the output side of the gain stage.

We also consider the cascode stage and the differential pair which are gain stages using two transistors. After having studied the chapter, you should be able to

- explain the operation of the following gain stages: common source, common drain, common gate, cascode stage and differential pair.
- model each of these stages at low frequencies using a generic amplifier model characterized by an open-circuit voltage gain, an input resistance and an output resistance.
- select a suitable gain stage based on the requirements concerning gain, input resistance and output resistance.
- determine suitable bias conditions for the gain stages.
- design the stages to fulfill specification requirements concerning gain, input resistance, input voltage range, output voltage range, etc.
- analyze and simulate the small-signal frequency response of the gain stages.

We start by considering the gain stages at low frequencies where capacitors can be neglected, and subsequently, we examine the frequency characteristics of the gain stages taking transistor capacitances into consideration.

A general small-signal model for an amplifier at low frequencies is the model shown in Fig. 4.1 with the amplifier treated as a voltage-controlled voltage source with a gain $A_{v\text{oc}}$, an input resistance r_{in} and an output resistance r_{out}. The subscript for the gain indicates that it is an open-circuit voltage gain, i.e., the voltage gain when no load resistance is connected to the amplifier. When this is connected to a signal source v_s with a source resistance R_S and a load with the resistance R_L as shown in Fig. 4.1, the gain $A_v = v_o/v_{in}$ is given by

$$A_v = \frac{v_o}{v_{in}} = A_{v\text{oc}} \frac{R_L}{R_L + r_{out}} \tag{4.1}$$

and the gain $A_{vs} = v_o/v_s$ from the signal source v_s to the output is given by

$$A_{vs} = \frac{v_o}{v_s} = \left(\frac{v_o}{v_{in}}\right)\left(\frac{v_{in}}{v_s}\right) = A_v \frac{r_{in}}{R_S + r_{in}} = A_{v\text{oc}} \left(\frac{R_L}{R_L + r_{out}}\right)\left(\frac{r_{in}}{R_S + r_{in}}\right) \tag{4.2}$$

Seen from the output terminals, the amplifier is represented by a Thévenin equivalent with the Thévenin voltage $V_t = A_{v\text{oc}}\, v_{in}$ and the Thévenin resistance $R_t = r_{out}$.

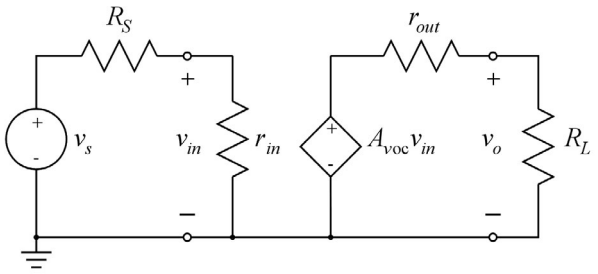


Figure 4.1: A general low-frequency small-signal amplifier model.

Alternative representations can be used with other types of controlled sources, see Fig. 2.10, but in this chapter, we use the model based on the voltage-controlled voltage source, the voltage amplifier.

4.1 The common-source stage at low frequencies

The common-source stage is also called the inverting amplifier and we have already examined this gain stage in Chapter 3, Sections 3.5 and 3.6. Here we will discuss some additional design considerations for the stage. We start by considering the gain stage with a drain resistor R_D as shown in Fig. 4.2 where both the transistor schematic and the small-signal equivalent is shown. The configuration with a drain resistor is rarely used in integrated circuit design because resistors normally take up more silicon area than transistors and also because the configuration with a drain resistor gives a smaller gain than a configuration with the drain resistor replaced by a transistor, an active load. However, it is useful for illustrating many of the fundamental properties of the common-source stage.

We can use the method described in Section 2.3 for finding the Thévenin voltage as the open-circuit output voltage and the Thévenin resistance as the resistance between the output terminals with $v_{in} = 0$. From the small-signal equivalent circuit, we find the following small-signal parameters characterizing the amplifier:

$$r_{in} = \infty \tag{4.3}$$

$$r_{out} = R_D \parallel r_{ds} \tag{4.4}$$

$$A_{voc} = -g_m (R_D \parallel r_{ds}) \tag{4.5}$$

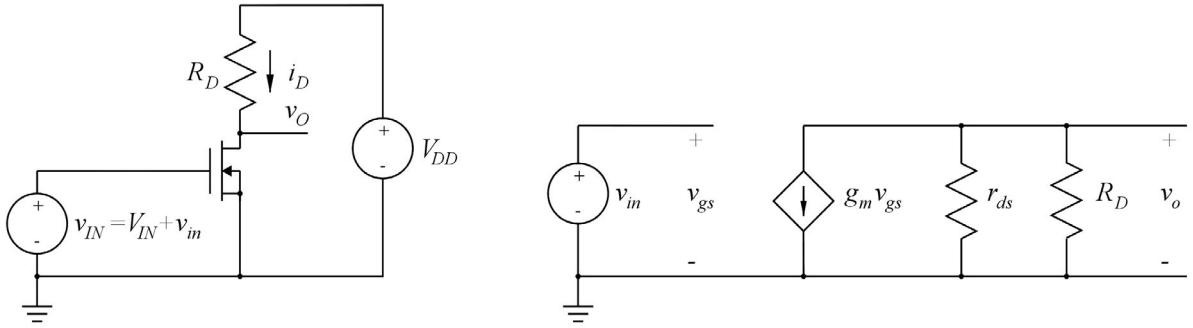


Figure 4.2: Schematic and small-signal equivalent diagram for the common-source amplifier with a drain resistor.

The small-signal parameters g_m and $r_{ds} = 1/g_{ds}$ must be found from the dc bias point using Eqs. (3.64) - (3.66) and (3.67) - (3.68) or the approximations (3.70) - (3.72) from Section 3.5.

The design criteria may comprise specifications of small-signal gain, output resistance, supply voltage, supply current, output voltage range, input voltage range and bias values of input voltage and output voltage. Also, the frequency response may be specified but we leave an investigation of the frequency response to Section 4.5.

From the example presented in Section 3.5, we recognize that the open-circuit voltage gain A_{voc} is closely related to the values of the bias voltages. Assuming $r_{ds} \gg R_D$, we find

$$A_{voc} \simeq -g_m R_D = -\left(\frac{2I_D}{V_{GS}-V_t}\right) R_D \tag{4.6}$$

$$= -\left(\frac{2I_D}{V_{IN}-V_t}\right)\left(\frac{V_{DD}-V_O}{I_D}\right) = -\frac{2(V_{DD}-V_O)}{V_{IN}-V_t} \tag{4.7}$$

From Eq. (4.7), we see that the gain is independent of both R_D and I_D and can be found merely from the dc bias voltages. With the output bias voltage selected to $V_{DD}/2$ in order to allow a large output voltage swing, we find $A_{voc} = -V_{DD}/(V_{IN}-V_t)$ as also found in Eq. (3.50). However, the output bias voltage may be selected to another value in order to match a required input bias voltage for a subsequent stage, and in this case, Eq. (4.7) applies.

While the values of the drain resistor and the transistor geometry do not enter into Eq. (4.7) for the gain, the drain current and the output resistance do relate to the drain resistor and the transistor geometry with the output resistance being approximately equal to R_D and with the drain current being dependent on the transistor geometry according to Eq. (3.13).

Often a common-source stage is used as a high-gain stage in an operational amplifier in a system with feedback, see Chapters 5, 6 and 7. However, for the configuration with a drain resistor as shown in Fig. 4.2, the gain is limited by the supply voltage and a minimum value of the overdrive voltage $V_{IN}-V_t$ at the input of the transistor. The supply voltage cannot exceed the maximum specification for the CMOS process used for the implementation. For a standard 0.18 µm process, this is typically 1.8 V. For a 0.35 µm process, it may be 3.3 V, and for a 65 nm process, it will typically be only about 1.2 V. For the transistor to work in strong inversion, the overdrive voltage at the input cannot be lower than about 100 mV. Below this, the transistor is in moderate or weak inversion, and for weak inversion, the overdrive voltage in Eq. (4.7) is replaced by nV_T which has a typical value of about 40 mV. For moderate inversion, the denominator in Eq. (4.7) assumes a value between nV_T and $V_{IN}-V_t$. Combining these limitations, the common-source stage with a drain resistor is not very well suited as a high-gain stage.

However, we saw in Section 3.6 that by replacing the resistor with a dc current source, we can separate the gain expression from the limitations caused by the supply voltage and the bias voltages. In this case, we still find $g_m = 2I_D/(V_{IN}-V_t)$ but in Eq. (4.6), R_D is replaced by $r_{ds} = (1+\lambda V_O)/(\lambda I_D)$, resulting in a gain of

$$|A_{voc}| = \frac{2(1+\lambda V_O)}{\lambda(V_{IN}-V_t)} = \frac{2(1+\lambda V_{DS})}{\lambda(V_{GS}-V_t)} \simeq \frac{2}{\lambda(V_{GS}-V_t)} \tag{4.8}$$

This is the intrinsic gain of the transistor, see Eqs. (3.73) and (3.74).

With a minimum value of $V_{IN} - V_t$ equal to about 100 mV and $\lambda = 0.1$ V^{-1} (corresponding to an NMOS transistor with $L = 1$ µm), we find a maximum intrinsic gain of about 200 V/V.

In practice, the load is not an ideal dc current source. Rather, the current source is implemented with a current mirror as shown in Fig. 3.21. This is called an active load. For the circuit shown in Fig. 4.3, the small-signal model of the active load reduces to the resistor r_{ds2} because the gate voltage of M_2 - M_3 is constant, see Figs. 3.42 and 3.43. With $r_{ds2} = (1 + \lambda_2 (V_{DD} - V_O))/(I_D \lambda_2)$ being of about the same magnitude as r_{ds1}, the resulting gain is halved compared to Eq. (4.8).

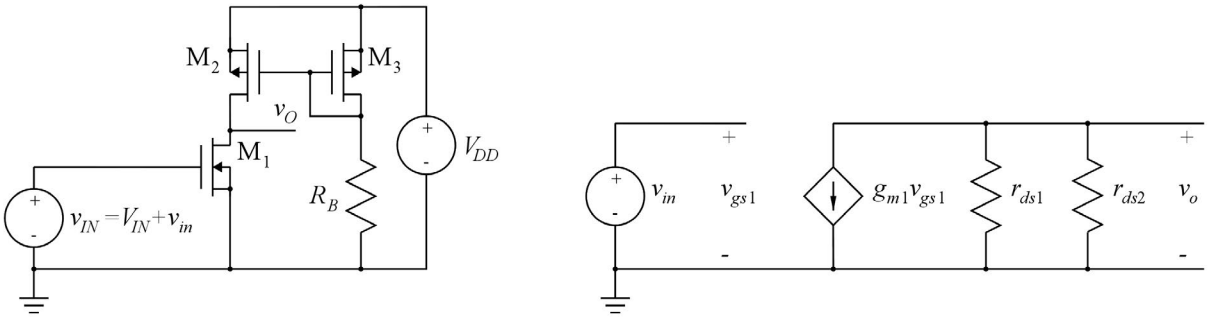


Figure 4.3: Schematic and small-signal equivalent diagram for the common-source amplifier with an active load.

From the small-signal equivalent circuit in Fig. 4.3, we also find that the output resistance of the common-source stage with an active load is $r_{out} = r_{ds1} \parallel r_{ds2}$, i.e., much higher than for the configuration with a drain resistor. This may compromise the gain if the load is resistive, see Eq. (4.1), but if the stage is connected to a subsequent stage with an infinite input resistance, the low-frequency gain is not affected.

Simulation examples. We have already seen some simulations of the common-source amplifier in Section 3.5. Figures 3.25 and 3.26 show transient simulations for the gain stage with a drain resistor, illustrating how the gain stage is overdriven if the amplitude of the input signal is too large.

We can also use simulations to illustrate the relations between gain and bias point. Using the same values for R_D and transistor geometries as in Fig. 3.24, we may run simulations in which we define the input bias voltage as a parameter '`VB`' and step '`VB`' through a number of different values ranging from 400 mV to 900 mV using a '.step param' directive as shown in Fig. 4.4.

Figure 4.5 shows the result of a transient simulation. From this, we can find the gain as the peak-to-peak output voltage divided by the peak-to-peak input voltage. This gives the small-signal gain values listed in Fig. 4.5(b). We find an increasing gain with increasing input bias voltage because of the decreasing value of V_O in Eq. (4.7). But we also see that for large values of V_{IN}, $V_{IN} = 0.8$ V and $V_{IN} = 0.9$ V,

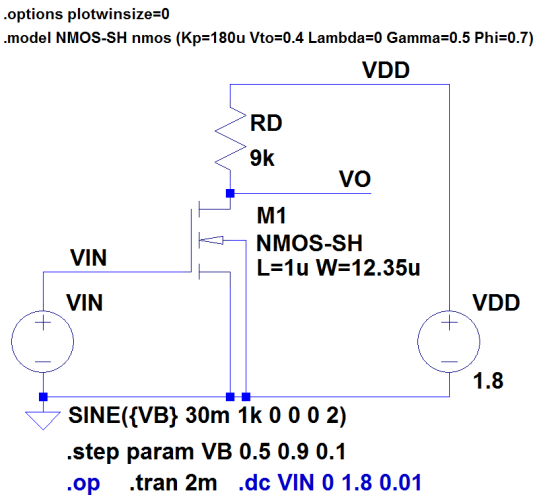
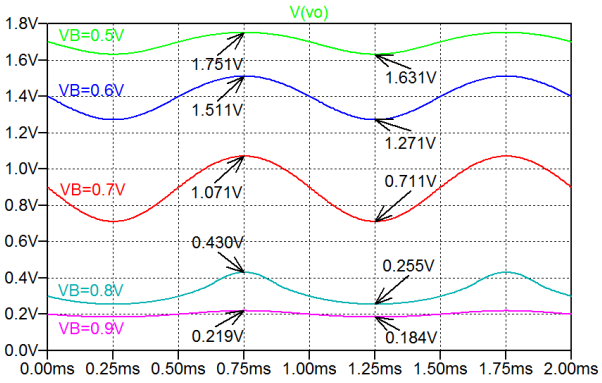


Figure 4.4: LTspice schematic for the common-source amplifier with a drain resistor.



(a)

Calculated gain

Input bias voltage [mV]	Peak-to-peak output voltage [mV]	Small-signal voltage gain [V/V]
500	120	-2.00
600	240	-4.00
700	360	-6.00
800	175	-2.92
900	035	-0.58

(b)

Figure 4.5: Transient simulation and gain calculation for the common-source stage shown in Fig. 4.4.

the gain decreases and distortion becomes visible. In particular, for $V_{IN} = 0.8$ V, the output waveform is severely distorted and the gain calculated from the peak-to-peak output voltage cannot be expected to represent the small-signal gain because the input amplitude apparently exceeds the value for which a linear approximation is reasonable. The reason for the decreasing gain and the distortion is that the transistor enters the triode region for small values of the output voltage and Eq. (4.7) no longer applies.

This can also be illustrated by a '.dc' simulation with a sweep of the input voltage. Figure 4.6 shows both the output voltage '`V(vo)`' and the derivative of the output voltage '`d(V(vo))`' versus the input voltage. Evidently, the transistor is cut off for $v_{IN} < V_t = 0.40$ V, and for $v_{IN} > 0.78$ V, the transistor is in the triode region, causing the magnitude of the gain to decrease drastically. We notice that the gain values shown in Fig. 4.6(b) match the gain values shown in Fig. 4.5(b), except for $V_{IN} = 0.8$ V where the calculated value in Fig. 4.5(b) violates the linear small-signal approximation.

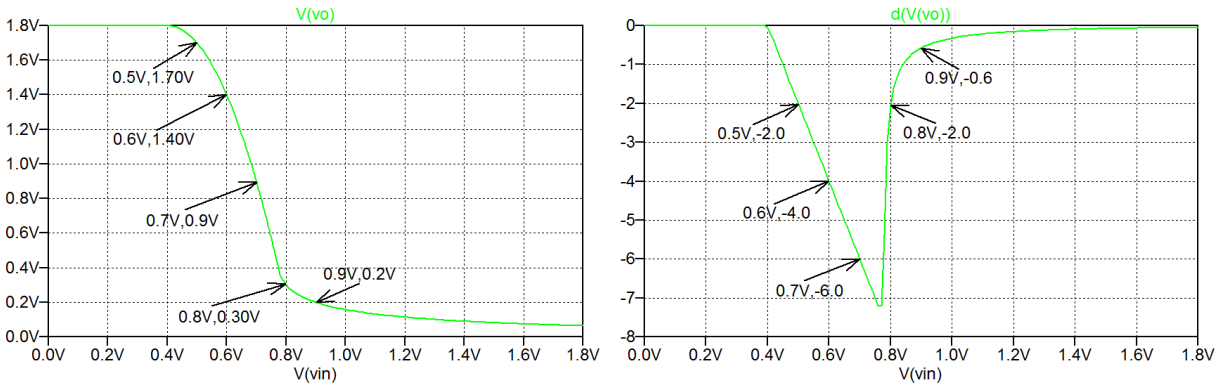


Figure 4.6: Output voltage versus input voltage (a) and small-signal gain versus input voltage (b) for the common-source stage shown in Fig. 4.4.

The simulations shown in Figs. 4.5 and 4.6 demonstrate the importance of selecting a proper input bias voltage for the circuit. For the common-source stage with an active load, it is even more important to select an input bias voltage which biases the transistor in the active region. The bias value can be estimated from Eq. (3.16) using $V_{DS} = V_{DD}/2$. However, since the gain is high, the useful range of input voltages is very small and it is necessary to run a '.dc' simulation in order to find a value for the input bias which can be used for transient simulations and simulations based on small-signal analysis.

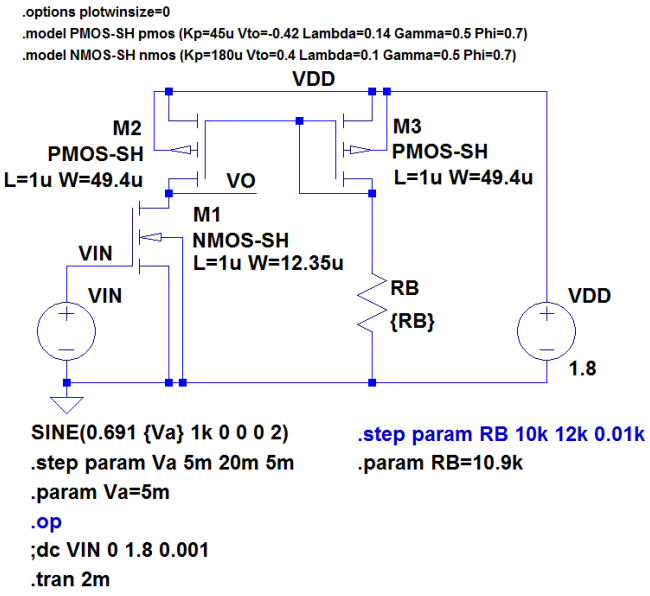


Figure 4.7: LTspice schematic of a common-source stage with an active load.

Figure 4.7 shows an LTspice schematic corresponding to the schematic from Fig. 4.3. The channel width for M_2 and M_3 has been selected to 4 times the channel width for M_1 to compensate for the difference in electron mobility and hole mobility. The bias resistor R_B has a value giving a bias current in M_3 of 100 µA and this is mirrored to M_2 to provide a bias current for M_1 of about 100 µA, slightly depending on the value of the output bias voltage. With $I_D = 100$ µA and with an output bias voltage of $V_{DD}/2 = 0.9$ V, we can calculate the output resistance to be

$$r_{ds2} = (1 + \lambda_2 V_{DD}/2)/(\lambda_2 I_D) = 80\ \text{k}\Omega \tag{4.9}$$

and

$$r_{ds1} = (1+\lambda_1 V_{DD}/2)/(\lambda_1 I_D) = 109\ \text{k}\Omega. \tag{4.10}$$

The transconductance can be calculated to be

$$g_{m1} = \sqrt{2\mu_n C_{ox}(W_1/L_1) I_D (1+\lambda_1 V_{DD}/2)} = 0.70\ \text{mA/V}, \tag{4.11}$$

implying that the gain is estimated to be

$$A_{voc} = -g_{m1}(r_{ds1} \parallel r_{ds2}) = -32\ \text{V/V}. \tag{4.12}$$

Figure 4.8(a) shows a simulation of the output voltage versus the input voltage. From the simulation, we find that an input voltage of 691 mV gives an output voltage of 0.9 V, and we also see that the useful input voltage range is from 672 mV to 711 mV, corresponding to an output range from 0.3 V to 1.5 V. Thus, even with an optimal input bias voltage, the maximum amplitude of the input signal is about 19 mV.

Figure 4.8(b) shows the output signal from a transient simulation where the input amplitude has been stepped through the values 5, 10, 15 and 20 mV. Evidently, the signal shows very little distortion for 5, 10 and 15 mV whereas the output signal for an input amplitude of 20 mV is clipped at both high and low output voltages when M_2 or M_1 enters the triode region.

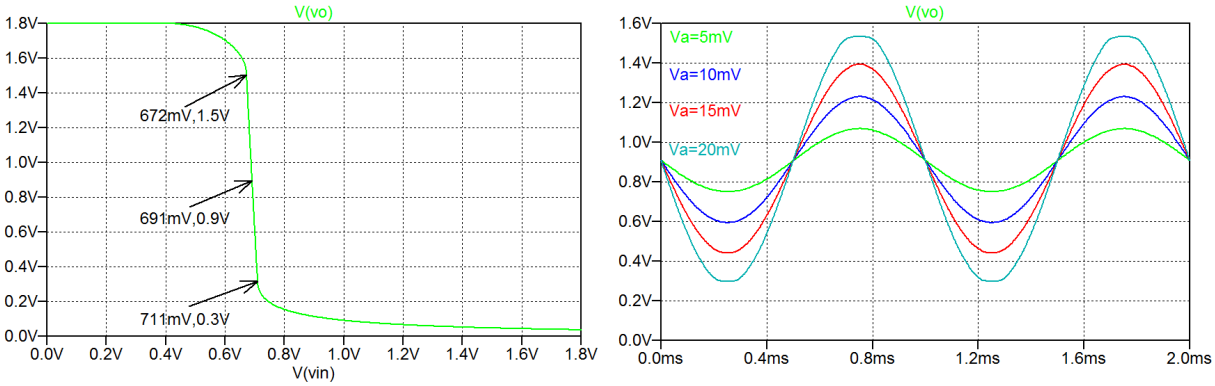


Figure 4.8: Output voltage versus input voltage (a) and output voltage versus time (b) for the common-source stage shown in Fig. 4.7.

The gain of the amplifier can be simulated directly by a '.tf' simulation with '`v(VO)`' as the output and '`VIN`' as the input, or it can be calculated from the amplitude of the sinusoids shown in Fig. 4.8(b) or found from the slope of the curve shown in Fig. 4.8(a). In all cases, we find $A_{voc} = -31.8$ V/V, confirming the hand calculations.

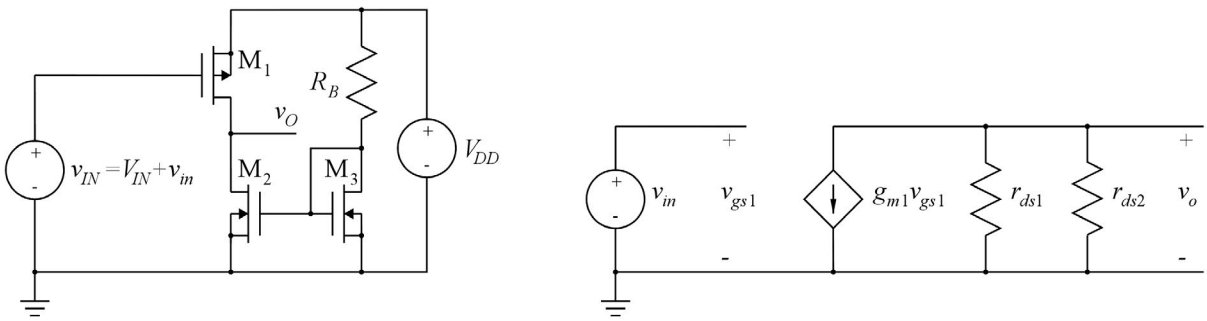


Figure 4.9: Schematic and small-signal equivalent diagram for a PMOS common-source amplifier with an active load.

The common-source stage with a PMOS transistor. The previous examples showed the common-source stage with an NMOS transistor. Of course, the gain stage can also be implemented with a PMOS transistor. Figure 4.9 shows a common-source stage with a PMOS transistor and an NMOS active load. The analysis of this stage is very similar to the analysis of the stage with an NMOS transistor, only it is important to be careful with the signs of voltages in the design equations. The use of absolute values of voltages in the device equations for the PMOS transistor is highly recommended.

The choice between a PMOS gain stage and an NMOS gain stage is often made on basis of the bias voltage available at the input. For the PMOS stage, the input bias voltage is normally close to the positive supply rail. For the NMOS stage, the input bias voltage is closer to the negative supply rail.

The detailed analysis of the PMOS common-source stage is left to the reader.

4.2 The common-drain stage at low frequencies

The common-drain stage is shown with an NMOS transistor in Fig. 4.10. The input terminal is the gate of M_1 and the output terminal is the source of M_1 while the drain of M_1 is common to the input and the output and serves as ground for signal variations. For simplicity, bulk and source of M_1 are connected so that the body effect is avoided. Transistors M_2 and M_3 form a current mirror providing a bias current for M_1. Neglecting the channel-length modulation for all transistors, M_2 provides a constant bias current

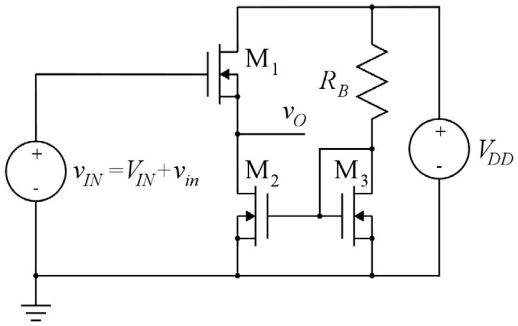


Figure 4.10: Common-drain stage with NMOS transistors.

I_{D1} for M_1, and with a constant bias current, the gate-source voltage for M_1 is also constant and given by

$$V_{GS1} = V_t + \sqrt{\frac{2I_{D1}}{\mu_n C_{ox}(W_1/L_1)}} \tag{4.13}$$

assuming that all transistors are in the active region. Thus, the output voltage is simply

$$v_O = v_{IN} - V_{GS1} \tag{4.14}$$

The common-drain stage is also called the source follower, indicating that the source voltage follows the input voltage. From Eq. (4.14), we find $A_{voc} = 1$ as V_{GS1} is a constant voltage. We also see that the small-signal input resistance is infinite.

For finding the output resistance, we need a small-signal equivalent circuit and when drawing the small-signal equivalent circuit, we will use the small-signal model from Fig. 3.28, including the output resistance, i.e., including the effect of the channel-length modulation. Just as for the common-source stage in Fig. 4.3, the small-signal model of the current mirror M_2 - M_3 reduces to the small-signal output resistance of M_2, and the complete small-signal equivalent circuit is shown in Fig. 4.11.

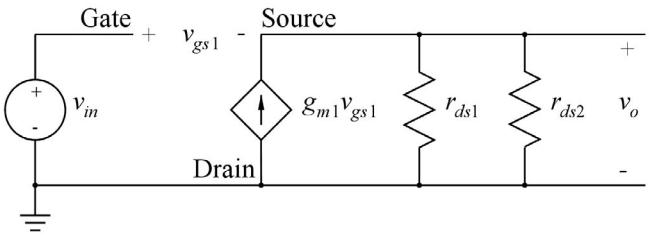


Figure 4.11: Small-signal equivalent circuit for the common-drain stage.

Notice that the voltage-controlled current source is pointing upwards as the drain of M_1 is connected to a constant supply voltage, i.e., small-signal ground, whereas the source is used as the output terminal. Also note that the controlling voltage for the current source is the input voltage v_{in} minus the output voltage v_o, not just the input voltage.

From the small-signal equivalent circuit, we can find the small-signal voltage gain taking the channel-length modulation into account. A node equation at the output gives

$$\begin{aligned} g_{m1}\,v_{gs1} = g_{m1}\,(v_{in}-v_o) &= v_o/r_{ds1}+v_o/r_{ds2} \\ \Rightarrow v_o\,(g_{m1}+1/r_{ds1}+1/r_{ds2}) &= g_{m1}\,v_{in} \\ \Rightarrow A_{voc}=\frac{v_o}{v_{in}} &= \frac{g_{m1}}{g_{m1}+1/r_{ds1}+1/r_{ds2}}=\frac{g_{m1}\,(r_{ds1}\parallel r_{ds2})}{1+g_{m1}\,(r_{ds1}\parallel r_{ds2})} \end{aligned} \tag{4.15}$$

Evidently, the open-circuit voltage gain is always smaller than 1 but with $g_{m1}\,(r_{ds1}\parallel r_{ds2}) \gg 1$ which is normally the case, we find $A_{voc} \simeq 1$.

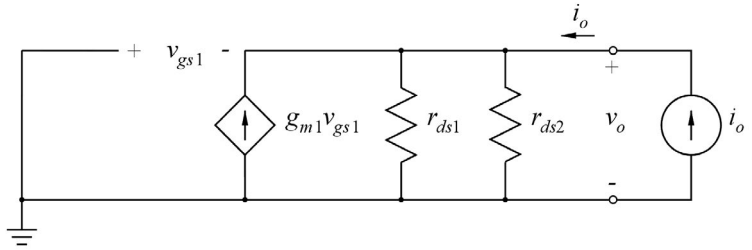


Figure 4.12: Small-signal equivalent circuit for finding the output resistance for the common-drain stage.

The small-signal output resistance is found by applying a current i_o to the output terminal as shown in Fig. 4.12 and calculating r_{out} as v_o/i_o with $v_{in}=0$, i.e., with the input shorted to ground. From a node equation at the output terminal, we find

$$\begin{aligned} i_o+g_{m1}\,v_{gs1} = i_o+g_{m1}\,(0-v_o) &= v_o/r_{ds1}+v_o/r_{ds2} \\ \Rightarrow r_{out}=\frac{v_o}{i_o} &= \frac{1}{g_{m1}+1/r_{ds1}+1/r_{ds2}}=\frac{1}{g_{m1}}\parallel r_{ds1}\parallel r_{ds2} \end{aligned} \tag{4.16}$$

Normally, $(r_{ds1} \parallel r_{ds2}) \gg 1/g_{m1}$, so $r_{out} \simeq 1/g_{m1}$.

Compared to the common-source stage, the source follower is noninverting and it has a low gain but it also has a lower output resistance, and this makes it useful as a buffer for driving a resistive load where the voltage division at the output, see Eq. (4.1), would attenuate the output signal from a common-source stage too much.

In the circuit shown in Fig. 4.10, bulk and source are connected for M_1 and therefore there is no body effect. However, this may not be possible. For a CMOS process as shown in Fig. 3.5, the bulk of the NMOS transistors is the common substrate which is connected to the negative supply rail for the circuit. With this configuration, the body effect cannot be avoided and Fig. 4.13 shows the circuit with the bulk of M_1 connected to ground.

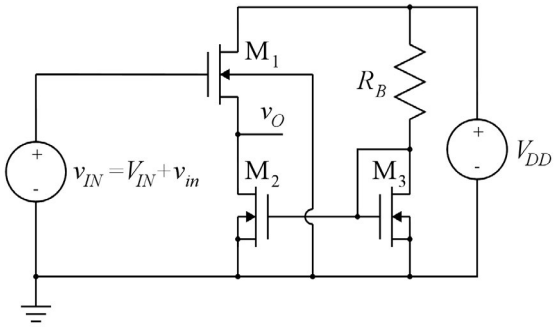


Figure 4.13: Common-drain stage with an NMOS transistor with bulk connected to ground.

The small-signal equivalent circuit is shown in Fig. 4.14(a) where the small-signal model from Fig. 3.32 has been used for M_1. In the small-signal equivalent circuit, there is now an additional voltage-controlled current source, $g_{mb1}v_{s1}$. From the small-signal equivalent circuit, we can now find A_{voc} and r_{out} using node equations just as we did in deriving Eqs. (4.15) and (4.16). But we may also apply a few simple transformations of the small-signal equivalent circuit in order to make it easier to find A_{voc} and r_{out} simply by inspection.

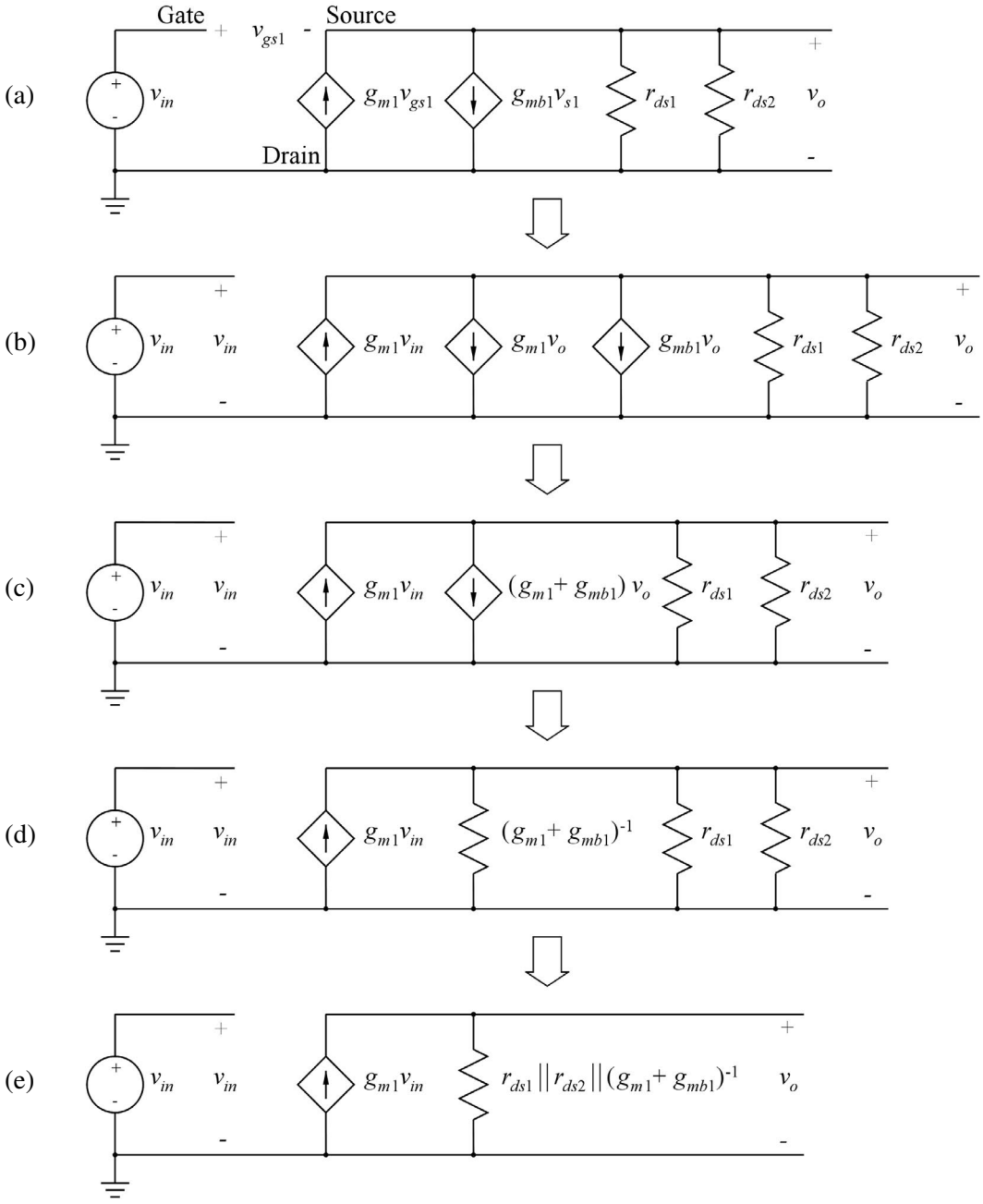


Figure 4.14: Small-signal equivalent circuit for the common-drain stage including the body effect.

We note that the current $g_{m1}\,v_{gs1} = g_{m1}\,(v_{g1} - v_{s1}) = g_{m1}\,(v_{in} - v_o)$ may be split into two current sources, one, $g_{m1}\,v_{in}$, pointing upwards, and another, $g_{m1}\,v_o$, pointing downwards as shown in Fig. 4.14(b). Combining the two current sources pointing downwards and controlled by the same voltage, the circuit can now be reduced as shown in Fig. 4.14(c).

Next, we see that the current $(g_{m1} + g_{mb1})\,v_o$ is proportional to the voltage between the two terminals of the current source, and a linear relation between current and voltage for two terminals of a device corresponds to Ohm's law for a resistor, so the current source $(g_{m1} + g_{mb1})\,v_o$ can be replaced by a resistor $1/(g_{m1} + g_{mb1})$ as shown in Fig. 4.14(d).

Finally, the three resistors r_{ds1}, r_{ds2} and $1/(g_{m1}+g_{mb1})$ are connected in parallel, so they can be replaced by an equivalent resistor $r_{ds1} \parallel r_{ds2} \parallel (g_{m1}+g_{mb1})^{-1}$ as shown in Fig. 4.14(e).

From this, we find by inspection $r_{in} = \infty$ and

$$A_{voc} = g_{m1}\left(r_{ds1} \parallel r_{ds2} \parallel (g_{m1}+g_{mb1})^{-1}\right) \simeq \frac{g_{m1}}{g_{m1}+g_{mb1}} = \frac{1}{1+\chi} \tag{4.17}$$

$$r_{out} = r_{ds1} \parallel r_{ds2} \parallel (g_{m1}+g_{mb1})^{-1} \simeq \frac{1}{g_{m1}+g_{mb1}} = \frac{1}{g_{m1}(1+\chi)} \tag{4.18}$$

where χ is defined as $\chi = g_m/g_{mb}$, see Eq. (3.76).

In Eqs. (4.17) and (4.18), the approximations are based on the assumption that the transistor output resistances are much larger than $1/g_{m1}$.

From Eqs. (4.17) and (4.18), it is apparent that the body effect reduces the gain since χ has a typical value in the range 0.1 to 0.3. Also the output resistance is reduced. Another consequence of the body effect, not seen from the small-signal analysis, is that the offset between v_{IN} and v_O is increased because of the increase of the threshold voltage V_t. This is a characteristic of the source follower: The useful voltage range at both input and output is rather limited. Especially in modern submicron processes with small supply voltages, this can be a problem. The source follower is normally used only where no other solutions to obtaining a low output resistance can be found.

Simulation examples. One of the design criteria for a common-drain stage may be the output voltage range which can be obtained for a given resistive load. As an example, we consider a common-drain stage with a supply voltage of ± 0.9 V and a load resistor of $R_L = 5$ kΩ connected from the output to ground. We also assume that the common-drain transistor is subject to body effect as shown in Fig. 4.15. The maximum value of the input voltage is equal to the positive supply voltage. Obviously, the common-drain transistor M_1 cannot drive the output all the way to the positive rail. The maximum output voltage is equal to the positive supply voltage minus the gate-source voltage for M_1. The minimum output voltage is equal to the negative supply voltage plus the drain-source saturation voltage for M_2, assuming that M_2 should be in the active region. As an example, we assume that the output voltage range is specified to be from -0.6 V to $+0.1$ V.

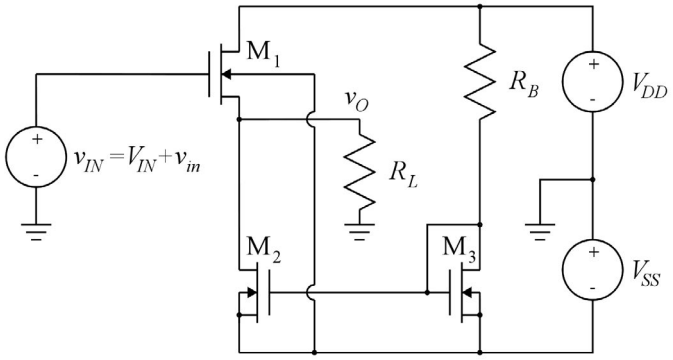


Figure 4.15: Common-drain stage with a resistive load.

In order to establish an initial design for simulation, we derive the transistor geometries by hand calculation neglecting the channel-length modulation but otherwise using the transistor parameters from Table 3.1. We use a channel length of $L = 1$ µm.

The basic design requirements arise from the output swing. In order to have a negative output swing of -0.6 V with a load resistor of 5 kΩ, we need a bias current in M_2 of at least 0.6 V/5 kΩ= 120 µA. As M_2 is configured as a current source, it should be in the active region for this bias current, implying that the overdrive voltage $V_{GS2} - V_{to}$ must be at most 0.3 V, corresponding to $V_{GS2} = 0.7$ V, in order ensure $V_{DS2} \geq V_{GS2} - V_{to}$.

Neglecting the channel-length modulation, we may calculate a minimum value of W_2/L_2 from

$$\frac{W_2}{L_2} = \frac{2\,I_{D2}}{\mu_n C_{ox}(V_{GS2} - V_{to})^2} = 14.8 \tag{4.19}$$

With $L_2 = 1$ µm, we select $W_2 = 14.8$ µm. For M_3, we may use a smaller width, scaling down the current in M_3 and R_B compared to the current in M_2 in order to reduce the current consumption of the bias circuit. Using a scale factor of 10, we select $W_3 = 1.48$ µm. The resistor R_B is calculated from $V_{GS3} = V_{GS2} = 0.7$ V, the supply voltage and the current:

$$R_B = \frac{V_{DD} + V_{SS} - V_{GS3}}{I_{D3}} = 91.7\ \text{k}\Omega \tag{4.20}$$

Finally, M_1 must be designed so that it can supply 120 µA to M_2 plus 20 µA to R_L (giving an output voltage of 0.1 V) when the gate voltage is the maximum input voltage which is assumed to be V_{DD}. This implies

$$\frac{W_1}{L_1} = \frac{2\,I_{D1}}{\mu_n C_{ox}(V_{GS1} - V_t)^2} \tag{4.21}$$

Here we need to calculate V_t from Eq. (3.17) as M_1 is subject to body effect. With $v_S = v_O = 0.1$ V and $V_B = -V_{SS} = -0.9$ V, we have $v_{SB} = 1$ V, giving $V_t = 0.63$ V and $W_1/L_1 = 53.8$, so with $L = 1$ µm, we find $W = 53.8$ µm.

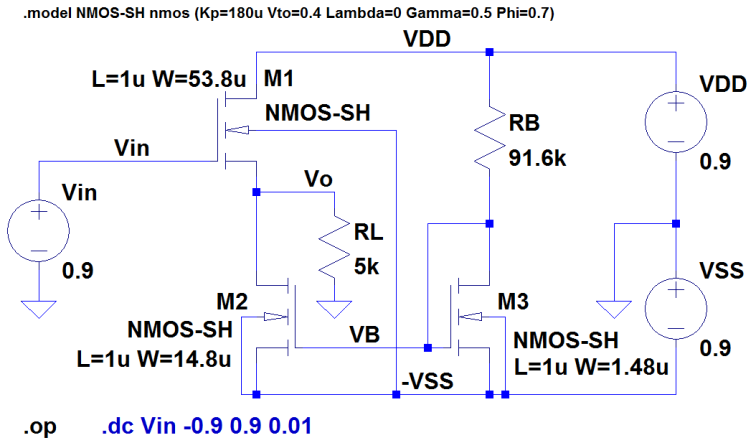


Figure 4.16: LTspice schematic for simulating the common-drain stage with a resistive load.

We now have all the specifications needed to simulate the circuit and Fig. 4.16 shows the LTspice schematic for this. In order to verify the hand calculations, we first simulate the circuit with $\lambda = 0$. For verifying the lower output-voltage limit, a '.op' simulation with the input at -0.9 V may be run, and for verifying the upper output-voltage limit, a '.op' simulation with the input at +0.9 V may be run. Figure 4.17 shows the output files from these simulations. We see that the output voltages and the device currents closely match the calculated values.

Output file with $v_{IN} = -0.9$ V.

```
        --- Operating Point ---

V(vb):          -0.199916       voltage
V(-vss):        -0.9            voltage
V(vdd):         0.9             voltage
V(vin):         -0.9            voltage
V(vo):          -0.599736       voltage
Id(M2):         0.000119947     device_current
Ig(M2):         0               device_current
Ib(M2):         -3.10264e-013   device_current
Is(M2):         -0.000119947    device_current
Id(M1):         3.30974e-012    device_current
Ig(M1):         0               device_current
Ib(M1):         -2.12026e-012   device_current
Is(M1):         -1.18947e-012   device_current
Id(M3):         1.19947e-005    device_current
Ig(M3):         0               device_current
Ib(M3):         -7.10084e-013   device_current
Is(M3):         -1.19947e-005   device_current
I(Rl):          -0.000119947    device_current
I(Rb):          1.19947e-005    device_current
I(Vin):         0               device_current
I(Vss):         -0.000131942    device_current
I(Vdd):         -1.19947e-005   device_current
```

(a)

Output file with $v_{IN} = +0.9$ V.

```
        --- Operating Point ---

V(vb):          -0.199916       voltage
V(-vss):        -0.9            voltage
V(vdd):         0.9             voltage
V(vin):         0.9             voltage
V(vo):          0.0972603       voltage
Id(M2):         0.000119947     device_current
Ig(M2):         0               device_current
Ib(M2):         -1.00726e-012   device_current
Is(M2):         -0.000119947    device_current
Id(M1):         0.000139399     device_current
Ig(M1):         0               device_current
Ib(M1):         -2.81726e-012   device_current
Is(M1):         -0.000139399    device_current
Id(M3):         1.19947e-005    device_current
Ig(M3):         0               device_current
Ib(M3):         -7.10084e-013   device_current
Is(M3):         -1.19947e-005   device_current
I(Rl):          1.94521e-005    device_current
I(Rb):          1.19947e-005    device_current
I(Vin):         0               device_current
I(Vss):         -0.000131942    device_current
I(Vdd):         -0.000151394    device_current
```

(b)

Figure 4.17: Output files from '.op' simulations of the common-drain stage with minimum input voltage (a) and maximum input voltage (b).

Next, we may repeat the simulations with $\lambda = 0.1$ V^{-1}. We do not show the results here but we find that the circuit no longer quite fulfills the requirements concerning the minimum value of the output voltage. This is hardly surprising since M_2 has a smaller drain-source voltage than M_3, so with the channel-length modulation taken into account, the scaling of the current mirror M_3 - M_2 is smaller than 10. A simple way to compensate for this is by reducing the value of R_B. A few iterations show that when R_B is reduced to 88.6 kΩ, we can achieve a minimum output voltage of -0.60 V. Similarly, a few iterations show that in order to achieve a maximum output voltage of 0.10 V, the width of M_1 should be increased to 55 µm.

With these values of W_1 and R_B, we may also illustrate the output voltage range by a '.dc' simulation with a sweep of the input voltage from -0.9 V to $+0.9$ V. Figure 4.18 shows the resulting plot of the output voltage versus the input voltage. From the plot, we find that an input voltage of 768 mV results in an output voltage of 0 V. With this input voltage, we may run a '.op' simulation and use the error log file ('Ctrl-L') to find the small-signal transistor parameters.

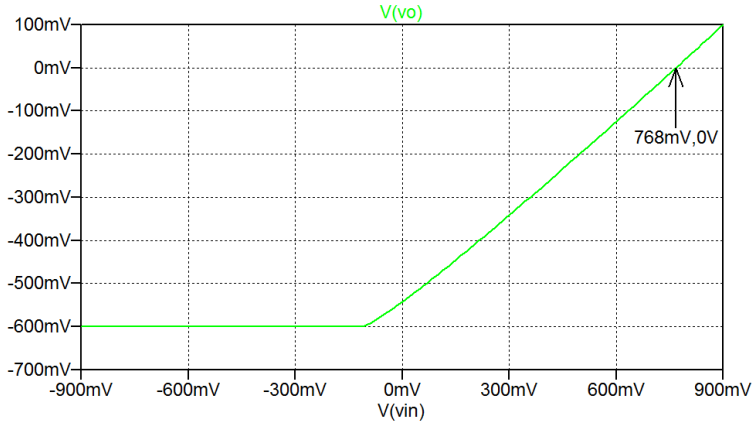


Figure 4.18: Plot of output voltage versus input voltage for the common-drain stage with channel-length modulation, $W_1 = 55$ µm and $R_B = 88.6$ kΩ.

From this, we find the error log file shown in Fig. 4.19(a). With the small-signal parameters from this file, we can calculate the open-circuit voltage gain and the output resistance using Eqs. (4.17) and (4.18) as shown in Fig. 4.19(b). These values may be verified by a '.tf' simulation with '`V(Vo)`' as the output and '`Vin`' as the input and the load resistor R_L disconnected. If R_L is not disconnected, the simulation will show $(r_{out} \parallel R_L)$ as the output resistance.

SPICE Error Log.

```
Semiconductor Device Operating Points:
                --- MOSFET Transistors ---
Name:            m3           m2           m1
Model:        nmos-sh      nmos-sh      nmos-sh
Id:          1.25e-05     1.27e-04     1.27e-04
Vgs:         6.96e-01     6.96e-01     7.68e-01
Vds:         6.96e-01     9.00e-01     9.00e-01
Vbs:         0.00e+00     0.00e+00    -9.00e-01
Vth:         4.00e-01     4.00e-01     6.14e-01
Vdsat:       2.96e-01     2.96e-01     1.53e-01
Gm:          8.43e-05     8.59e-04     1.66e-03
Gds:         1.17e-06     1.17e-05     1.17e-05
Gmb:         2.52e-05     2.57e-04     3.27e-04
```

(a)

Calculated gain and output resistance.

$$A_{voc} = \frac{g_{m1}}{g_{m1}+g_{mb1}+g_{ds1}+g_{ds2}} = 0.826$$

$$r_{out} = \frac{1}{g_{m1}+g_{mb1}+g_{ds1}+g_{ds2}} = 497\ \Omega$$

(b)

Figure 4.19: Error log file from a '.op' simulation of the common-drain stage with a bias output voltage of 0 V (a) and calculated values of open-circuit voltage gain and output resistance (b).

4.3 The common-gate stage and the cascode stage at low frequencies

The common-gate stage is shown in Fig. 4.20 with an NMOS transistor M_1 as the gain transistor. The gate of M_1 is biased by a dc voltage V_B, i.e., small-signal ground, and the gate serves as the common terminal for input and output signals. The source serves as the input, and the output signal is taken from the drain. The current mirror M_2 - M_3 provides the bias current for M_1. In the circuit shown in Fig. 4.20, the bias current flows through M_1 and also through the signal source v_{sig} with a signal source resistance R_S. The subscript 'sig' for the signal voltage is used to avoid confusion with the source voltage of M_1.

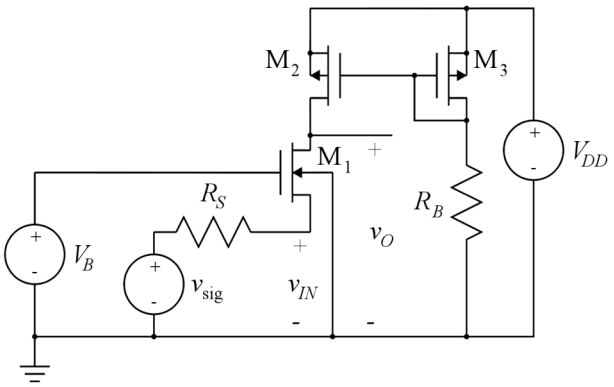


Figure 4.20: Common-gate stage with an NMOS transistor.

By applying the input voltage to the source of M_1 and having a fixed voltage on the gate of M_1, the input signal swing appears in the gate-source voltage, causing a signal swing in the drain current just as in the common-source stage and hence a signal variation in the voltage at the drain of M_1. Only, with the input voltage applied to the source instead of the gate, the output signal is not inverted with respect to the input signal as it is for the common-source configuration.

Figure 4.21 shows the small-signal equivalent circuit for the stage. The current mirror M_2 - M_3 is modeled by the resistor r_{ds2} in the same way as we did for the current mirrors used in the common-source stage and the common-drain stage. Just as for the common-drain transistor, it may not be possible to short-circuit bulk and source of M_1, so in Fig. 4.20, the bulk of M_1 is connected to the negative rail (ground), and in Fig. 4.21, the bulk effect is modeled by the current source $g_{mb1}v_{bs1} = -g_{mb1}v_{s1} = -g_{mb1}v_{in}$. This current source can be combined with the current source $g_{m1}v_{gs1} = -g_{m1}v_{s1} = -g_{m1}v_{in}$ as shown in Fig. 4.21(b). For the common-gate stage, the bulk transconductance simply adds to the gate transconductance.

From the small-signal equivalent circuit, we see that there is a signal path not only in the forward direction from input to output but also in the reverse direction. If we reset v_{sig} and apply a signal current to the output, a fraction of this current flows in r_{ds2} and a fraction flows through the transistor and through R_S, causing a variation of the input voltage v_{in}. This means that the common-gate stage is not a unilateral amplifier and a stringent treatment of the properties of the stage requires four two-port parameters, for example the inverse-hybrid parameters g_{11}, g_{12}, g_{21} and g_{22} (Sedra & Smith 2016, Appendix C).

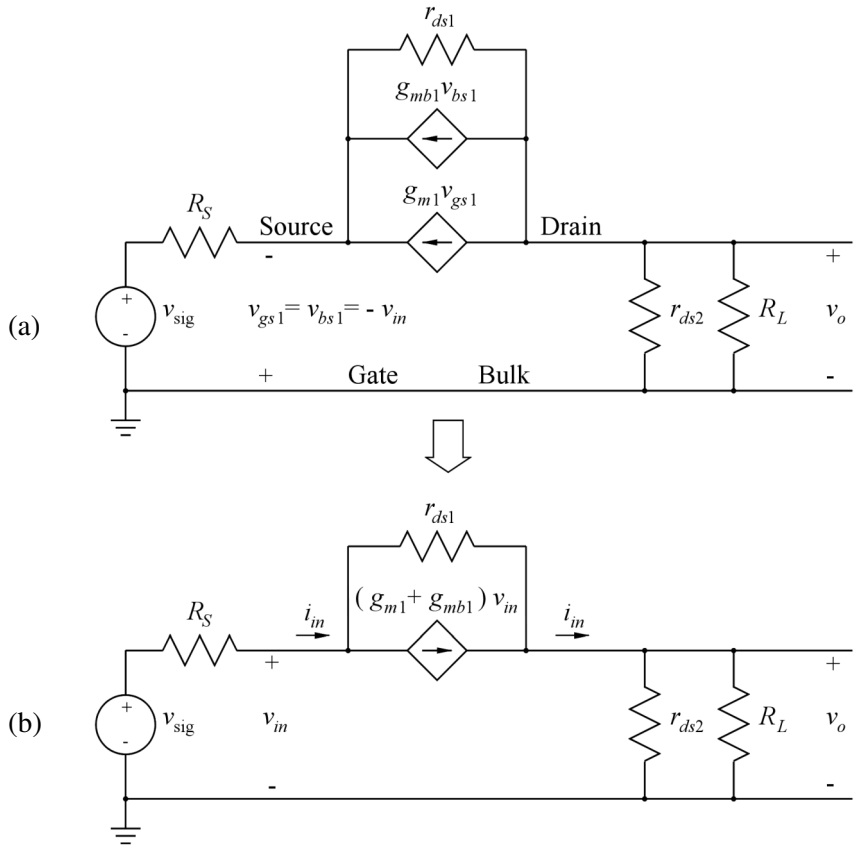


Figure 4.21: Small-signal equivalent circuit for the common-gate stage with a load resistor and a signal source resistor.

However, for practical purposes, it is easier to stay with the amplifier model shown in Fig. 4.1 but with an input resistance depending on the load resistance and an output resistance depending on the signal source resistance, and this is the reason why both R_L and R_S are included in Fig. 4.21.

From Fig. 4.21(b), we can find the open-circuit voltage gain $A_{voc} = v_o/v_{in}$ using a node equation at the output node with $R_L = \infty$:

$$\begin{aligned}(g_{m1}+g_{mb1})\,v_{in} &= \frac{v_o - v_{in}}{r_{ds1}} + \frac{v_o}{r_{ds2}} \\ \Rightarrow A_{voc} = \frac{v_o}{v_{in}} &= (g_{m1}+g_{mb1}+\frac{1}{r_{ds1}})(r_{ds1} \parallel r_{ds2}) \simeq (g_{m1}+g_{mb1})(r_{ds1} \parallel r_{ds2}) \end{aligned} \tag{4.22}$$

For finding the input resistance r_{in} with a load resistor R_L connected to the output, we calculate v_{in}/i_{in} from the small-signal equivalent circuit in Fig. 4.21(b) using a node equation at the input node, the transistor source:

$$\begin{aligned} i_{in} &= (g_{m1}+g_{mb1})v_{in} + \frac{v_{in}-v_o}{r_{ds1}} = (g_{m1}+g_{mb1})v_{in} + \frac{v_{in} - i_{in}(R_L \parallel r_{ds2})}{r_{ds1}} \\ \Rightarrow r_{in} &= \frac{v_{in}}{i_{in}} = \frac{r_{ds1}+(R_L \parallel r_{ds2})}{1+(g_{m1}+g_{mb1})r_{ds1}} \simeq \frac{1}{g_{m1}+g_{mb1}} + \frac{R_L \parallel r_{ds2}}{(g_{m1}+g_{mb1})r_{ds1}} \end{aligned} \tag{4.23}$$

We recognize the term $1/(g_{m1}+g_{mb1})$ as the output resistance from a common-drain stage, see Eq. (4.18). This is not surprising. The resistance looking into the source of M_1 in Fig. 4.13 and M_1 in Fig. 4.20 is the same if R_L in Fig. 4.20 is equal to zero, shorting the drain to small-signal ground as in the common-drain stage.

The second term shows that the input resistance increases with increasing value of R_L and with $R_L = \infty$, the increase in r_{in} is $(1/(g_{m1}+g_{mb1}))(r_{ds2}/r_{ds1})$, i.e., of a magnitude similar to the first term in Eq. (4.23), assuming that r_{ds2} is on the same order of magnitude as r_{ds1}.

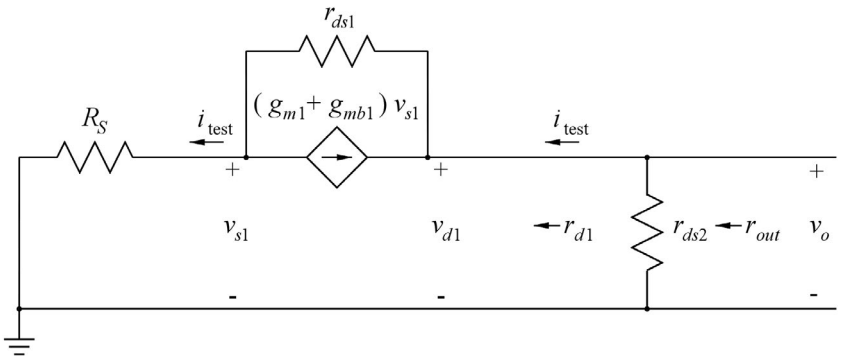


Figure 4.22: Small-signal circuit for finding the output resistance of the common-gate stage.

For finding the output resistance r_{out} with a signal source resistance R_S, we analyze the small-signal equivalent circuit shown in Fig. 4.22. We notice that r_{out} is a parallel connection of r_{ds2} and the resistance r_{d1} looking into the drain of M_1. For finding the latter, we replace r_{ds2} by a current source injecting a current i_{test} into the drain and calculate the drain voltage as

$$v_{d1} = v_{s1} + (i_{\text{test}} + (g_{m1}+g_{mb1})v_{s1})\,r_{ds1} = i_{\text{test}}R_S + (i_{\text{test}} + (g_{m1}+g_{mb1})i_{\text{test}}R_S)\,r_{ds1} \tag{4.24}$$

From Eq. (4.24), we find

$$r_{d1} = \frac{v_{d1}}{i_{\text{test}}} = R_S + (1 + (g_{m1}+g_{mb1})R_S)\,r_{ds1} \simeq r_{ds1} + (g_{m1}+g_{mb1})r_{ds1}\,R_S \tag{4.25}$$

and

$$r_{out} = r_{ds2} \parallel r_{d1} \simeq r_{ds2} \parallel (r_{ds1} + (g_{m1}+g_{mb1})r_{ds1}\,R_S) \tag{4.26}$$

We see that the resistance looking into the drain of M_1 can be much higher than just r_{ds1} because the additional term is R_S multiplied by the gain $(g_{m1}+g_{mb1})r_{ds1}$ of the transistor. With R_S being on the order of r_{ds}, we find r_{d1} on the order of $r_{ds}[(g_m+g_{mb})r_{ds}]$. This means that if r_{ds2} can be made very large, the output resistance of the common-gate stage can be made very large, a feature utilized in the cascode stage which is a combination of a common-gate stage and a common-source stage.

The cascode stage. One of the normal applications of the common-gate stage is the cascode stage where a common-source stage is followed by a common-gate stage as shown in Fig. 4.23. Transistor M_1 is the common-source transistor and M_2 is the common-gate transistor, also called the cascode transistor. We have shown the cascode stage with a current source I_B providing the bias current for both M_1 and M_2, and for our initial analysis, we assume that I_B is an ideal dc current source with an open circuit as the small-signal equivalent.

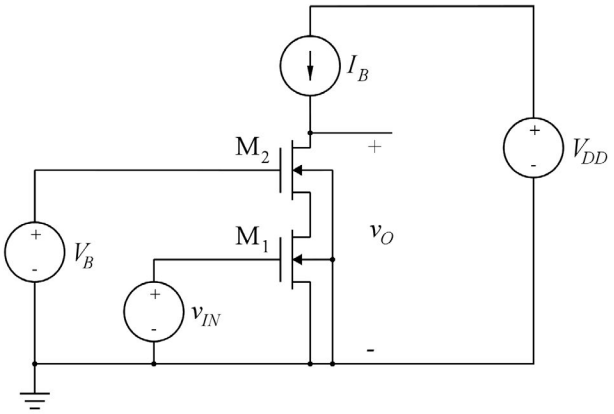


Figure 4.23: Telescopic-cascode stage with NMOS transistors.

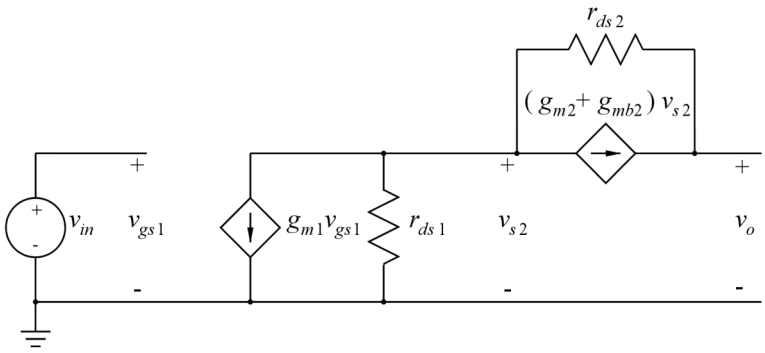


Figure 4.24: Small-signal equivalent circuit for the cascode stage.

For finding the small-signal open-circuit voltage gain, input resistance and output resistance, we consider the small-signal equivalent circuit shown in Fig. 4.24. From this, we immediately find $r_{in} = \infty$. Using Eq. (4.25) where R_S is replaced by r_{ds1} and g_{m1}, g_{mb1} and r_{ds1} are replaced by g_{m2}, g_{mb2} and r_{ds2}, we find

$$r_{out} = r_{ds1} + (1 + (g_{m2} + g_{mb2})r_{ds1})\, r_{ds2} \simeq (g_{m2} + g_{mb2})r_{ds2}\, r_{ds1} \tag{4.27}$$

For finding the open-circuit voltage gain $A_{voc} = v_o/v_{in}$, we use node equations at the input and the output of the cascode transistor M_2. A node equation at the output, i.e., the drain of M_2, gives

$$(g_{m2} + g_{mb2})v_{s2} = (v_o - v_{s2})/r_{ds2} \;\Rightarrow\; v_o = (1 + (g_{m2} + g_{mb2})r_{ds2})v_{s2} \tag{4.28}$$

Since there is no net current flowing through the parallel connection of r_{ds2} and $(g_{m2} + g_{mb2})v_{s2}$, a node equation at the source of M_2 gives

$$g_{m1}v_{in} + v_{s2}/r_{ds1} = 0 \;\Rightarrow\; v_{s2} = -g_{m1}r_{ds1}v_{in} \tag{4.29}$$

With Eq. (4.29) inserted in Eq. (4.28), we find

$$A_{voc} = v_o/v_{in} = -(1 + (g_{m2} + g_{mb2})r_{ds2})g_{m1}r_{ds1} \simeq -g_{m1}(g_{m2} + g_{mb2})r_{ds1}\, r_{ds2} \tag{4.30}$$

Rather than modeling the cascode stage as a voltage amplifier as shown in Fig. 4.1, we may model it as a transconductance amplifier as shown in Fig. 4.25. The output resistance is the same as for the voltage amplifier, i.e., it is given by Eq. (4.27). The transconductance is found from

$$g_m = -\frac{A_{voc}}{r_{out}} \simeq g_{m1} \tag{4.31}$$

When comparing Fig. 4.3 to Fig. 4.25, we may notice that the cascode stage resembles a single-transistor common-source stage, only it has a much higher output resistance, resulting in a much higher intrinsic gain $g_m r_{out}$.

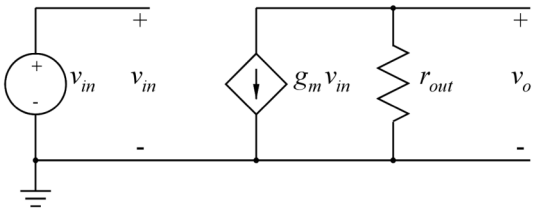


Figure 4.25: General small-signal model for an inverting transconductance amplifier.

One of the challenges with the cascode stage is the biasing of the cascode transistor and the output voltage swing. The cascode transistor M_2 needs a bias voltage for the gate ensuring that the common-source transistor M_1 is in the active region, i.e., $V_B \geq V_{DS\text{sat}1} + V_{GS2}$ where $V_{DS\text{sat}1}$ is the drain-source saturation voltage for M_1. For M_2 to stay in saturation, also V_{DS2} must be larger than the saturation voltage $V_{DS\text{sat}2}$. This implies that even with an optimal biasing of M_2, i.e., V_B as small as possible, the minimum output voltage from the cascode stage is $v_{O\min} = V_{DS\text{sat}1} + V_{DS\text{sat}2}$, i.e., twice as much as for the standard common-source stage.

Another challenge is the bias current source I_B. In the previous analysis, I_B was considered an ideal dc current source. However, in practice, I_B is implemented using transistors, so it has a finite output resistance which appears as a load at the output of the cascode stage. If I_B is implemented using a current mirror like shown in Fig. 4.3 for the common-source stage, the small-signal resistance of the current source is equal to r_{ds} for the current-mirror output resistor, and this is typically much smaller than the output resistance given by Eq. (4.27). Thus, in order to achieve a high gain, we need to increase the output resistance of the current mirror, and the previous analysis of the common-gate stage and the cascode stage showed that this can be done by placing a cascode transistor in the output of the current mirror as shown in Fig. 4.26. With the current mirror consisting of M_3 - M_5, the output resistance r_{cm} of the current mirror is derived from Eq. (4.27) as

$$r_{cm} \simeq (g_{m3} + g_{mb3}) r_{ds3}\, r_{ds4} \tag{4.32}$$

We also find that the maximum output voltage with all transistors in the active region is $v_{O\max} = V_{DD} - |V_{DS\text{sat}3}| - |V_{DS\text{sat}4}|$, assuming that the bias voltage V_{B2} is selected as high as possible while maintaining M_4 in the active region.

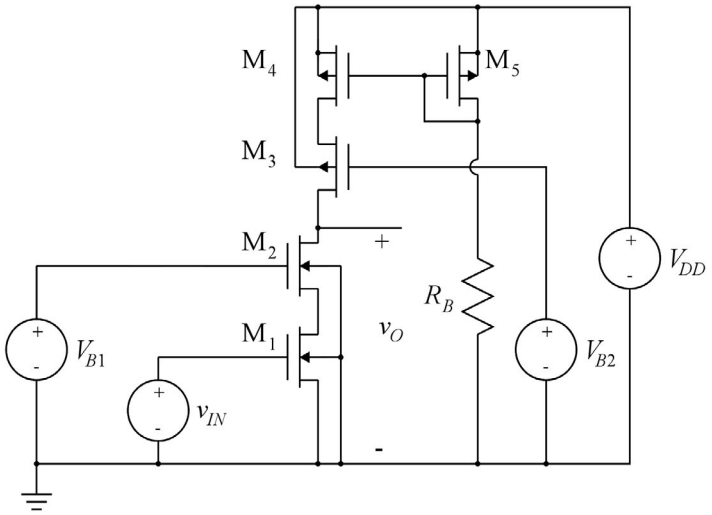


Figure 4.26: Telescopic-cascode stage with NMOS transistors and an active cascode load.

The cascode stage shown in Fig. 4.26 is called a telescopic cascode. For this, the bias voltage at the drain of the common-source transistor M_1 should be low in order to ensure a large output voltage range. However, in some circuit configurations, it is desirable to shift from a low bias voltage at the output of the common-source stage to a high bias voltage. This can be achieved by using a PMOS transistor as the cascode transistor as shown in Fig. 4.27. This is called a folded cascode. The bias current sources are shown as ideal dc current sources but in practice, they will be current sources with a finite small-signal output resistance, r_{IB1} and r_{IB2}, respectively. This leads to the small-signal equivalent circuit shown in Fig. 4.28. From the small-signal equivalent circuit, we see that r_{IB1} is in parallel with r_{ds1} so it should be at least on the same order of magnitude, i.e., the output resistance of a single transistor. Thus, the current source providing I_{B1} can be a single transistor with a dc bias voltage for the gate.

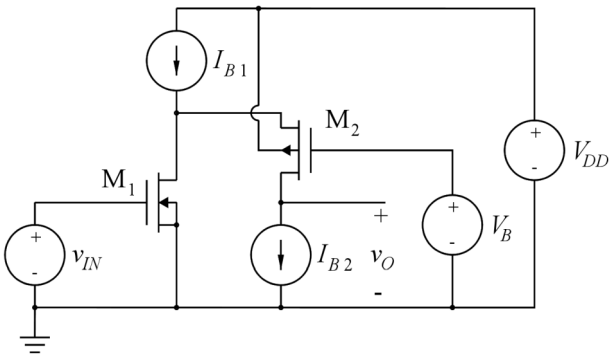


Figure 4.27: Folded-cascode stage with NMOS common-source transistor and PMOS cascode transistor.

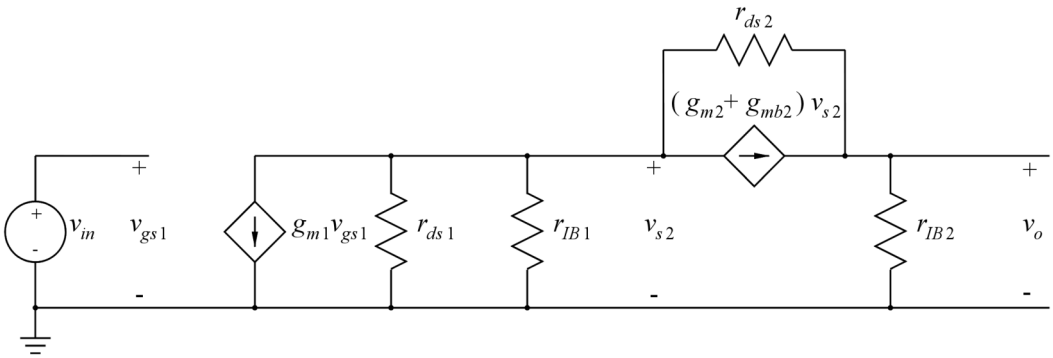


Figure 4.28: Small-signal equivalent circuit for a folded-cascode stage.

Also, r_{IB2} is connected to the output of the cascode transistor M_2 where the resistance level is very high. In order to retain the high resistance level, the current source I_{B2} should be implemented by a cascoded transistor, i.e., two series-connected transistors with appropriate dc gate bias voltages. The additional bias current source for the folded cascode makes the biasing of the folded cascode slightly more complicated than the biasing of the telescopic cascode, requiring more different dc bias voltages.

Also the folded cascode can be modeled as a transconductance amplifier as shown in Fig. 4.25 with

$$g_m \simeq g_{m1} \tag{4.33}$$

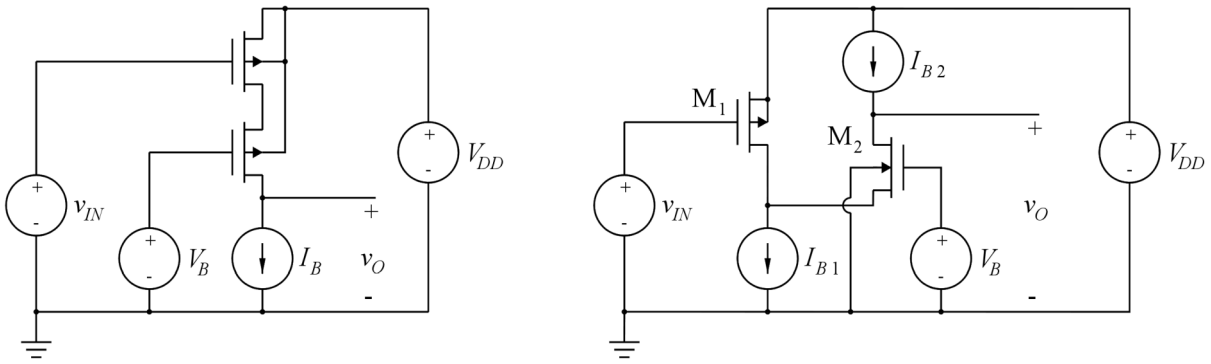
and

$$\begin{aligned} r_{out} &= r_{IB2} \parallel [r_{ds1} \parallel r_{IB1} + (1+(g_{m2}+g_{mb2})(r_{ds1} \parallel r_{IB1}))\, r_{ds2}] \\ &\simeq r_{IB2} \parallel [(g_{m2}+g_{mb2}) r_{ds2} (r_{ds1} \parallel r_{IB1})] \end{aligned} \tag{4.34}$$

For the folded-cascode stage, the output voltage range is similar to the output voltage range for the telescopic cascode but with the difference that the bias voltage at the intermediate node between the common-source transistor and the common-gate transistor is near the lower rail for the telescopic cascode and near the upper rail for the folded cascode.

The telescopic cascode may be implemented using PMOS transistors instead of NMOS transistors, and also the folded cascode may be implemented with a PMOS/NMOS transistor pair rather than an NMOS/PMOS transistor pair. These alternative versions are shown in Fig. 4.29.

Simulation examples. As an example of a simulation of a cascode stage, we consider the telescopic cascode shown in Fig. 4.23. We select the two transistors to have identical dimensions, $L = 1$ µm and $W = 12.35$ µm which were also the dimensions used for M_1 in the standard common-source stage shown in Fig. 4.7. Also, we use the transistor parameters from Table 3.1. For the supply voltage, we assume $V_{DD} = 1.8$ V, and we use the same bias current as in the common-source stage in Fig. 4.7, i.e., $I_D = I_B = 100$ µA.



(a) (b)

Figure 4.29: Telescopic cascode with PMOS transistors (a). Folded cascode with PMOS common-source transistor and NMOS cascode transistor (b).

For a simulation of the circuit, we need a value for the bias voltage V_B. We would prefer a small value for V_B in order to allow a large output voltage swing. The minimum value of V_B is

$$V_{B\min} = V_{DS\text{sat}1} + V_{GS2} \tag{4.35}$$

The saturation voltage $V_{DS\text{sat}1}$ can be found from Eq. (3.16) with $v_{GS1} - V_t = v_{DS} = V_{DS\text{sat}1}$

$$I_B = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right) V_{DS\text{sat}1}^2 (1 + \lambda V_{DS\text{sat}1}) \tag{4.36}$$

However, this is a cubic equation, so in order to have a simple analytical expression for $V_{DS\text{sat}1}$, we assume $\lambda V_{DS\text{sat}1} \ll 1$, resulting in

$$V_{DS\text{sat}1} \simeq \sqrt{\frac{2I_B}{\mu_n C_{ox}(W/L)}} = 0.30 \text{ V} \tag{4.37}$$

For V_{GS2}, we have $V_{GS2} = V_{DS\text{sat}2} + V_t$. The saturation voltage $V_{DS\text{sat}2}$ is equal to $V_{DS\text{sat}1}$ since M_1 and M_2 have the same size and the same current. The threshold voltage for M_2 must be calculated from Eq. (3.17) in order to account for the body effect. With $V_{SB2} = V_{DS\text{sat}1} = 0.30$ V, we find

$$V_t = V_{to} + \gamma(\sqrt{V_{SB2} + |2\Phi_F|} - \sqrt{|2\Phi_F|}) = 0.48 \text{ V} \tag{4.38}$$

and $V_{GS2} = V_{DS\text{sat}2} + V_t = 0.78$ V. From Eq. (4.35), we then find $V_{B\min} = 1.08$ V. In order to leave some margin, we select $V_B = 1.10$ V. With this value for V_B, we have the LTspice schematic shown in Fig. 4.30.

The first simulation to run is a dc sweep of v_{IN} in order to find a bias value V_{IN} giving a bias value of the output voltage of $V_{DD}/2 = 0.9$ V. The simulation directive shown in Fig. 4.30 steps v_{IN} from 0 to V_{DD} with a step size of 0.1 mV. From the simulation, we may plot '`V(vo)`' versus '`VIN`' as shown in Fig. 4.31. Here we observe that the y-axis extends from 0 to 60 MV which is obviously not physically reasonable for the circuit. The reason for this is the ideal dc current source I_B. This is not a realistic model for a real physical current source where the output voltage cannot exceed the supply voltage.

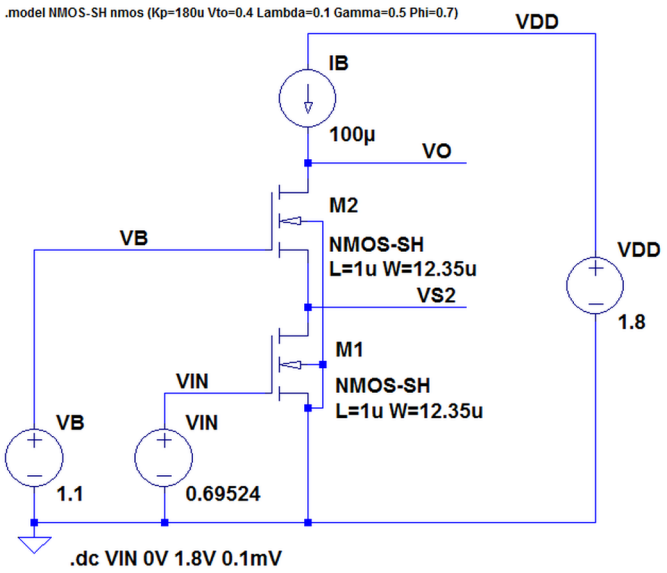


Figure 4.30: LTspice schematic for a telescopic cascode with NMOS transistors.

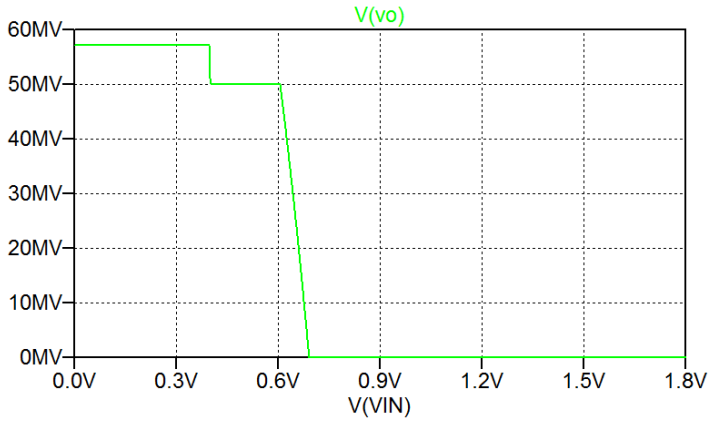


Figure 4.31: Simulation plot of v_O versus v_{IN} from the '.dc' simulation specified in Fig. 4.30.

In order to see '`V(vo)`' in a realistic range of output voltages, we set the range of the y-axis to 1.8 V. The resulting plot is shown in Fig. 4.32(a). We see that the output voltage changes abruptly for an input voltage of about 0.7 V, and in order to find the exact input voltage, we zoom in on a small part of the plot, see Fig. 4.32(b). From this, we find $V_{IN} = 695.24$ mV. Notice that this value has been specified with a resolution exceeding the step size in the '.dc' command, so it is based on an interpolation.

For a verification of the bias point, we run a '.op' simulation with a dc value of V_{IN} specified to 695.24 mV. The node voltages listed in the output file from this are shown in Fig. 4.33(a), and we see that '`V(vo)`' is sufficiently close to 0.9 V. Also shown in Fig. 4.33(b) are the small-signal transistor parameters from the error log file. We notice that the threshold voltage for M_2 is as calculated from Eq. (4.38) and also the values of the saturation voltages and the source voltage V_{S2} are as expected.

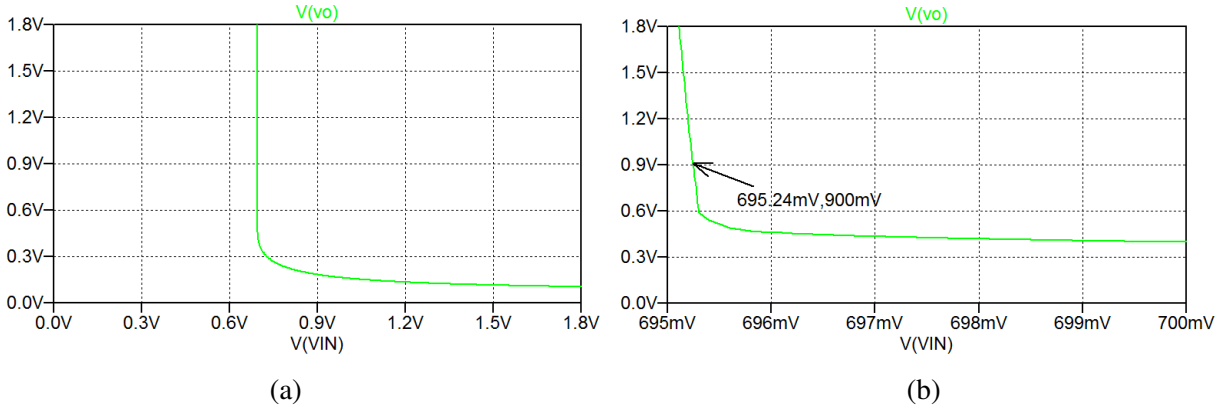


Figure 4.32: Zoom-in from Fig. 4.31 on the relevant output voltage range (a) and on a narrow input voltage range (b).

Output file

```
        --- Operating Point ---

V(vs2):         0.321436        voltage
V(vin):         0.69524         voltage
V(vo):          0.904833        voltage
V(vb):          1.1             voltage
V(vdd):         1.8             voltage
```

Error log file

```
Semiconductor Device Operating Points:
                --- MOSFET Transistors ---
Name:             m2            m1
Model:         nmos-sh       nmos-sh
Id:          1.00e-04      1.00e-04
Vgs:         7.79e-01      6.95e-01
Vds:         5.83e-01      3.21e-01
Vbs:        -3.21e-01      0.00e+00
Vth:         4.87e-01      4.00e-01
Vdsat:       2.92e-01      2.95e-01
Gm:          6.86e-04      6.77e-04
Gds:         9.45e-06      9.69e-06
Gmb:         1.70e-04      2.02e-04
```

Figure 4.33: Output file and error log file from the '.op' simulation of the telescopic cascode from Fig. 4.30.

Using the small-signal parameters from Fig. 4.33(b), we can calculate output resistance and gain from Eqs. (4.27) and (4.30): $r_{out} \simeq (g_{m2}+g_{mb2})/(g_{ds2}\,g_{ds1}) = 9.4$ MΩ and $A_{voc} \simeq -g_{m1}\,r_{out} = -6.3\times10^3$ V/V.

Gain and output resistance may also be found from a '.tf' simulation with '`v(VO)`' as the output and '`VIN`' as the input. This gives $r_{out} = 9.6$ MΩ and $A_{voc} = -6.4\times10^3$ V/V, results which match the calculated results reasonably well.

Next, we may replace the ideal current source with a cascode PMOS current mirror as shown in Fig. 4.26. In order to compensate for the difference in channel-length modulation parameter and in electron mobility and hole mobility, the length of the PMOS transistors is scaled by a factor 1.4 relative to the NMOS transistors and the W/L ratio of the PMOS transistors is scaled by a factor of 4 relative to the NMOS transistors, so $L_3 = L_4 = L_5 = 1.4$ µm and $W_3 = W_4 = W_5 = 69.16$ µm. With this scaling of the transistors, the saturation voltages for M_3 - M_5 have the same absolute value as the saturation voltages for M_1 and M_2, so we may determine a maximum value of the bias voltage V_{B2} as $V_{B2\max} = V_{DD} - |V_{DSsat4}| - |V_{GS3}| = V_{DD} - 2\,|V_{DSsat3}| - |V_{t3}|$. Using $|V_{DSsat3}| = 0.30$ V, we find $|V_{SB3}| = 0.3$ V, and from Eq. (3.17), we can calculate the threshold voltage of M_3 to be $V_{t3} = -0.5$ V, resulting in $V_{B2\max} = 0.7$ V. For the simulation, we select $V_{B2} = V_{B2\max} = 0.7$ V.

We also need a value for R_B. For this, we may use the same value as for the common-source stage in Fig. 4.7, giving a current of 100 µA in M_5 and approximately the same bias current in M_1 - M_4.

The resulting LTspice schematic is shown in Fig. 4.34. Again, we must run a '.dc' simulation to find a bias value for the input voltage. As we know from the previous simulation of the cascode in Fig. 4.30 that we need a resolution of 10 µV for 'VIN', we specify a step size of 10 µV in the '.dc' directive.

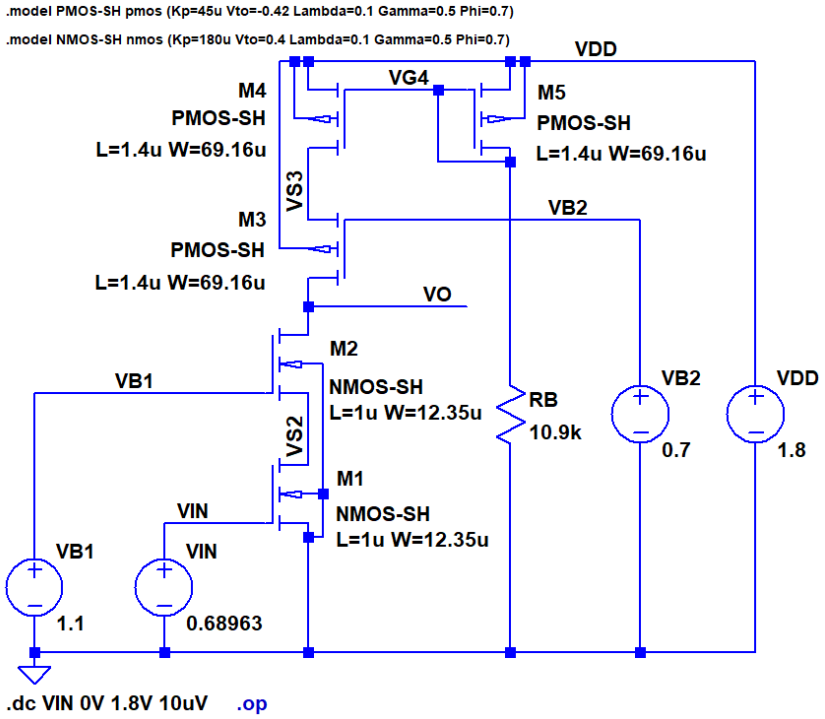
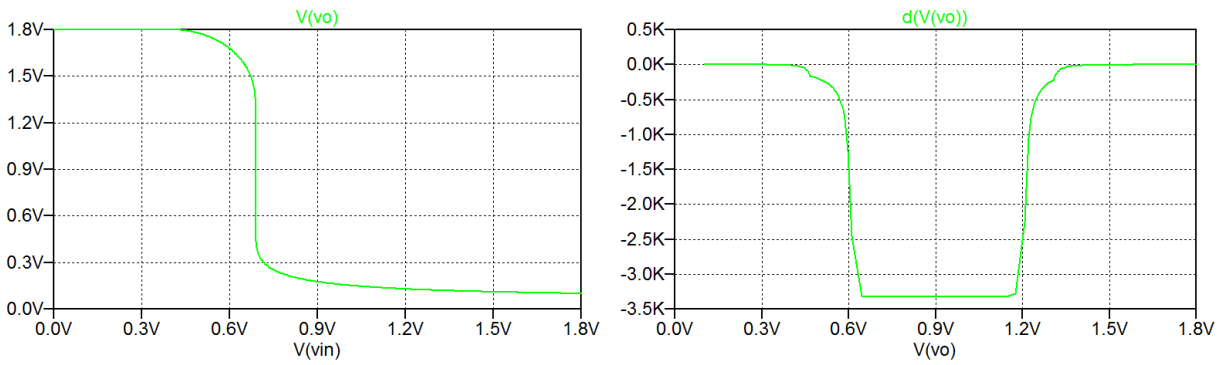


Figure 4.34: LTspice schematic for a telescopic cascode with a cascode current mirror for providing the bias current.



(a) (b)

Figure 4.35: Plot of v_O versus v_{IN} (a) and dv_O/dv_{IN} versus v_O (b) from the circuit from Fig. 4.34.

Figure 4.35(a) shows a plot of '`V(vo)`' versus '`V(vin)`'. From this, we find the bias value of the input voltage to be $V_{IN} = 689.63$ mV, and we see that the output voltage range where the stage provides a high gain is limited to a range from about 0.6 V to 1.2 V as expected from the saturation voltages. This is further illustrated in Fig. 4.35(b), showing dv_O/dv_{IN} versus v_O. From this plot, we find the gain in the useful output voltage range to be about -3.3×10^3 V/V which is about half the value found for the cascode with an ideal current source to supply the bias current. This appears reasonable. The PMOS transistors have been scaled to have the same value of the channel-length modulation parameter λ as the NMOS transistors, so from Eq. (4.32), we find an output resistance from the PMOS cascode current mirror which is about the same as the output resistance from the NMOS cascode. For the NMOS cascode, we found an output resistance of 9.6 MΩ using a '.tf' simulation, and with the same output resistance for the PMOS cascode, we can find the resulting gain as $A_v = -g_{m1} r_{out} \simeq -0.67$ mA/V $\times$ (9.6 $\|$ 9.6) MΩ $= -3.22 \times 10^3$ V/V, i.e., very close to the value found from Fig. 4.35(b).

This is about 100 times the gain found in Section 4.1 for the common-source amplifier with an NMOS transistor with the same dimensions and bias conditions as M_1 in the cascode stage. The increase in gain is caused by the increase in r_{out}. For the NMOS transistors, the output resistance increased from 109 kΩ to 9.6 MΩ, i.e., by a factor of 88, corresponding to the gain $(g_{m2} + g_{mb2}) r_{ds2}$, and for the PMOS transistors, the output resistance increased from 80 kΩ to 9.6 MΩ, i.e., by a factor of 120, partly due to the increase in channel length and partly due to the gain in M_3.

Running a '.op' simulation with $V_{IN} = 689.63$ mV, the bias values of all node voltages and the values of transistor small-signal parameters may be checked. This is an exercise left for the reader.

4.4 The differential pair at low frequencies

Often in amplifier design, it is desirable to have a gain stage with a differential input as shown in Fig. 4.36. This is needed, for instance, for the design of an operational amplifier. Ideally, the output signal from a gain stage with a differential input depends only on the differential input voltage $v_{ID} = v_{IN1} - v_{IN2}$ where v_{IN1} and v_{IN2} are the two input voltages. The common-mode input voltage $v_{ICM} = (v_{IN1} + v_{IN2})/2$ does not affect the output signal, so for an ideal differential gain stage, the small-signal output voltage is $v_o = A_d v_{id}$ where A_d is the small-signal differential gain.

In practice, the common-mode input voltage will also have some influence on the output signal. This is described by a common-mode voltage gain A_{cm}, so the small-signal output voltage is found as $v_o = A_d v_{id} + A_{cm} v_{icm}$ as shown in the general small-signal model for a differential gain stage shown in Fig. 4.36(b).

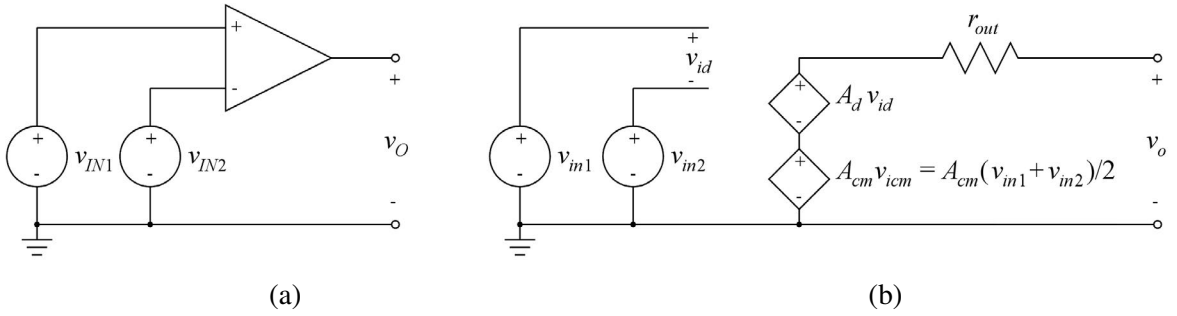


Figure 4.36: A general differential gain stage (a) and a small-signal model for the stage (b).

A frequently used figure-of-merit for a differential stage is the common-mode rejection ratio CMRR defined as

$$\text{CMRR} = \frac{|A_d|}{|A_{cm}|} \tag{4.39}$$

The common-mode rejection ratio is normally expressed in dB, i.e.,

$$\text{CMRR(db)} = 20\log\left(\frac{|A_d|}{|A_{cm}|}\right) \tag{4.40}$$

Using the definitions of the differential input voltage and the common-mode input voltage, we also see that the input voltages v_{IN1} and v_{IN2} can be decomposed into differential and common mode input voltages as shown in Fig. 4.37.

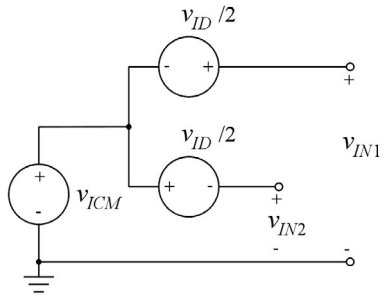


Figure 4.37: Decomposition of input voltages into common-mode voltage and differential voltage.

A gain stage with a differential input is shown in Fig. 4.38. Just as for the common-source stage, we start our analysis by considering a configuration with drain resistors as shown in Fig. 4.38 although in practice, the resistors are normally replaced by transistors providing an active load. The two resistors

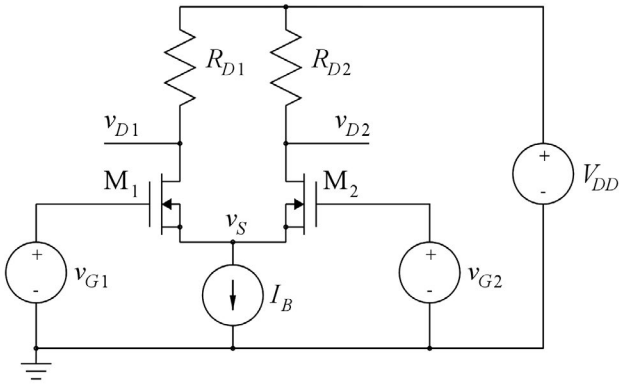


Figure 4.38: A differential gain stage with an NMOS differential pair.

R_{D1} and R_{D2} are assumed to be identical, $R_{D1} = R_{D2} = R_D$, and the two transistors M_1 and M_2 are also assumed to be identical. The bias current source I_B is called the tail current, and it is assumed to be an ideal dc current source. The gate voltages $v_{G1} = v_{IN1}$ and $v_{G2} = v_{IN2}$ are related to the differential input voltage and the common-mode input voltage as follows:

$$v_{ID} = v_{G1} - v_{G2} \tag{4.41}$$

$$v_{ICM} = (v_{G1} + v_{G2})/2 \tag{4.42}$$

$$v_{G1} = v_{ICM} + v_{ID}/2 \tag{4.43}$$

$$v_{G2} = v_{ICM} - v_{ID}/2 \tag{4.44}$$

The output voltages are $v_{O1} = v_{D1}$ and $v_{O2} = v_{D2}$ giving a differential output voltage $v_{OD} = v_{O1} - v_{O2}$. We further assume that all voltages in the circuit are kept within a range where both transistors are in the active region.

With a pure common-mode input signal, we have $v_{G1} = v_{G2}$, and as $v_{S1} = v_{S2} = v_S$ (the sources are connected), $v_{GS1} = v_{GS2}$, so $i_{D1} = i_{D2}$. From a node equation at the source of the transistors, we find $i_{D1} + i_{D2} = I_B$, so with $i_{D1} = i_{D2}$, we find $i_{D1} = i_{D2} = I_B/2$ and $v_{O1} = v_{O2} = V_{DD} - R_D I_B/2$. Thus, neither v_{O1}, nor v_{O2} are affected by a change in the common-mode input voltage, and the common-mode gain is 0.

With a pure differential-mode input signal $v_{ID} > 0$ applied to the gain stage, the gate voltage of M_1 increases while the gate voltage of M_2 decreases. This causes the drain current i_{D1} to increase and the drain current i_{D2} to decrease. Thus, the output voltage v_{O1} decreases and the output voltage v_{O2} increases. The differential output voltage $v_{OD} = v_{O1} - v_{O2}$ turns negative, so the differential gain $A_d = v_{od}/v_{id}$ is negative or inverting. In fact, the differential gain stage may be considered as two connected inverting amplifiers or common-source stages where the common terminal is not a fixed supply voltage or ground but rather the common source terminal for the two transistors M_1 and M_2. Each of these common-source stages has an input voltage of $v_{ID}/2$ and each of them delivers 50% of the differential output voltage.

We may perform a nonlinear analysis of the differential gain stage in a similar way as we did for the common-source stage in Section 3.5. In order to avoid cubic equations, we neglect the channel-length modulation, and we also neglect the body effect. For a differential input voltage, this leads to the following relations between the input voltages and the drain currents of M_1 and M_2:

$$i_{D1} = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(v_{GS1} - V_t)^2 = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(v_{ID}/2 + V_{ICM} - v_S - V_t)^2 \tag{4.45}$$

$$i_{D2} = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(v_{GS2} - V_t)^2 = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(-v_{ID}/2 + V_{ICM} - v_S - V_t)^2 \tag{4.46}$$

In these two equations, there are three unknown quantities, i_{D1}, i_{D2} and v_S, so we need a third relation. This comes from the node equation at the source of the transistors. Using this node equation and inserting Eqs. (4.45) and (4.46), we find

$$I_B = i_{D1} + i_{D2} = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(v_{ID}^2/2 + 2(V_{ICM} - v_S - V_t)^2) \tag{4.47}$$

$$\Rightarrow \quad v_S = V_{ICM} - V_t - \sqrt{\frac{I_B}{\mu_n C_{ox}(W/L)} - v_{ID}^2/4} \tag{4.48}$$

Using the Shichman-Hodges relation for the bias value V_S of the source voltage and the bias value V_{ICM} of the common-mode input voltage, we find

$$\frac{I_B}{2} = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)(V_{ICM} - V_S - V_t)^2 = \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)V_{\text{eff}}^2 \tag{4.49}$$

$$\Rightarrow \quad V_{\text{eff}}^2 = \frac{I_B}{\mu_n C_{ox}(W/L)} \tag{4.50}$$

where $V_{\text{eff}} = V_{\text{eff1}} = V_{\text{eff2}} = V_{GS} - V_t = V_{ICM} - V_S - V_t$.

Inserting Eq. (4.50) in Eq. (4.48), we can simplify this to

$$v_S = V_{ICM} - V_t - \sqrt{V_{\text{eff}}^2 - v_{ID}^2/4} = V_{ICM} - V_t - V_{\text{eff}}\sqrt{1 - \left(\frac{v_{ID}/2}{V_{\text{eff}}}\right)^2} \tag{4.51}$$

Inserting Eq. (4.51) in Eqs. (4.45) and (4.46) and also using Eq. (4.49) yields

$$\begin{aligned} i_{D1} &= \frac{\mu_n C_{ox}}{2}\left(\frac{W}{L}\right)\left(v_{ID}/2 + \sqrt{V_{\text{eff}}^2 - v_{ID}^2/4}\right)^2 \\ &= \frac{I_B}{2V_{\text{eff}}^2}\left(V_{\text{eff}}^2 + v_{ID}\sqrt{V_{\text{eff}}^2 - v_{ID}^2/4}\right) \\ &= \frac{I_B}{2} + \left(\frac{I_B}{V_{\text{eff}}}\right)\left(\frac{v_{ID}}{2}\right)\sqrt{1 - \left(\frac{v_{ID}/2}{V_{\text{eff}}}\right)^2} \qquad (4.52) \\ i_{D2} &= \frac{I_B}{2} - \left(\frac{I_B}{V_{\text{eff}}}\right)\left(\frac{v_{ID}}{2}\right)\sqrt{1 - \left(\frac{v_{ID}/2}{V_{\text{eff}}}\right)^2} \qquad (4.53) \end{aligned}$$

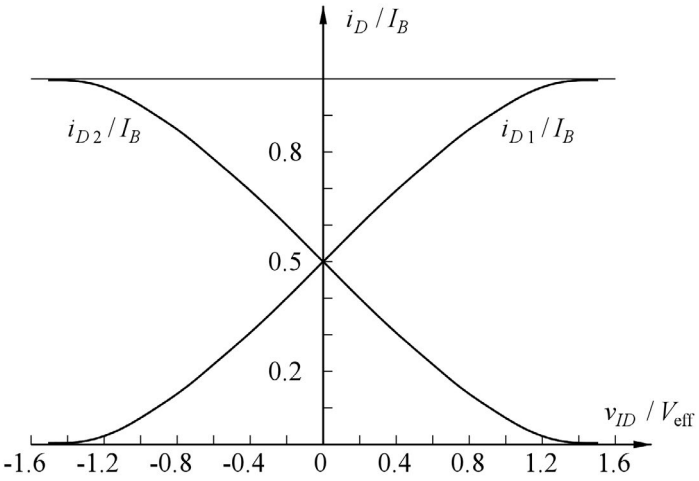


Figure 4.39: Plot of normalized drain currents for the differential pair from Fig. 4.38.

We may plot the relations (4.52) and (4.53) as shown in Fig. 4.39 where the x-axis has been normalized with respect to V_{eff} and the y-axis has been normalized with respect to I_B. We notice that for $|v_{ID}| \ll V_{\text{eff}}$, the characteristics are very linear. Also, for $|v_{ID}| \geq \sqrt{2}V_{\text{eff}}$, Eqs. (4.52) and (4.53) no longer apply since one of the transistors is cut off and the other transistor takes the entire bias current I_B.

The linearity may be further investigated by applying a Taylor expansion to Eq. (4.52) or (4.53).

Using $\sqrt{1-x} \simeq 1 - \frac{1}{2}x - \frac{1}{8}x^2 - \frac{1}{16}x^3$, we can find an approximate expression for i_{D1}:

$$\begin{aligned} i_{D1} &\simeq \frac{I_B}{2} + \left(\frac{I_B}{V_{\text{eff}}}\right)\left(\frac{v_{ID}}{2}\right)\left[1 - \left(\frac{1}{2}\right)\left(\frac{v_{ID}/2}{V_{\text{eff}}}\right)^2 - \left(\frac{1}{8}\right)\left(\frac{v_{ID}/2}{V_{\text{eff}}}\right)^4 - \left(\frac{1}{16}\right)\left(\frac{v_{ID}/2}{V_{\text{eff}}}\right)^6\right] \\ &= \frac{I_B}{2}\left[1 + \frac{v_{ID}}{V_{\text{eff}}} - \left(\frac{1}{8}\right)\left(\frac{v_{ID}}{V_{\text{eff}}}\right)^3 - \left(\frac{1}{128}\right)\left(\frac{v_{ID}}{V_{\text{eff}}}\right)^5 - \left(\frac{1}{1024}\right)\left(\frac{v_{ID}}{V_{\text{eff}}}\right)^7\right] \qquad (4.54) \end{aligned}$$

We notice that only odd powers of v_{ID} appear in Eq. (4.54). This implies that only odd harmonics will appear in the harmonic distortion when v_{ID} is a sinusoidal signal. Obviously, the higher order terms are strongly attenuated and for an estimation of the harmonic distortion, we may consider only the first few terms containing v_{ID}^3 and v_{ID}^5.

Using $\sin^3(x) = \frac{3}{4}\sin(x) - \frac{1}{4}\sin(3x)$ and $\sin^5(x) = \frac{10}{16}\sin(x) - \frac{5}{16}\sin(3x) + \frac{1}{16}\sin(5x)$, we may find the amplitude of the fundamental-frequency output and the third harmonic output when the input is a sinusoid $v_{ID} = V_{id}\sin(\omega t)$:

$$I_{d1\,\text{fund}} \simeq \left(\frac{I_B}{2}\right)\left[\left(\frac{V_{id}}{V_{\text{eff}}}\right) - \left(\frac{1}{8}\right)\left(\frac{3}{4}\right)\left(\frac{V_{id}}{V_{\text{eff}}}\right)^3 - \left(\frac{1}{128}\right)\left(\frac{10}{16}\right)\left(\frac{V_{id}}{V_{\text{eff}}}\right)^5\right] \tag{4.55}$$

$$I_{d1\,3\text{rd harm}} \simeq \left(\frac{I_B}{2}\right)\left[\left(\frac{1}{8}\right)\left(\frac{1}{4}\right)\left(\frac{V_{id}}{V_{\text{eff}}}\right)^3 + \left(\frac{1}{128}\right)\left(\frac{5}{16}\right)\left(\frac{V_{id}}{V_{\text{eff}}}\right)^5\right] \tag{4.56}$$

From these expressions, we find the harmonic distortion (third harmonic) to be

$$D = \frac{I_{d1\,3\text{rd harm}}}{I_{d1\,\text{fund}}} = \frac{64(V_{id}/V_{\text{eff}})^2 + 5(V_{id}/V_{\text{eff}})^4}{2048 - 192(V_{id}/V_{\text{eff}})^2 - 10(V_{id}/V_{\text{eff}})^4} \tag{4.57}$$

This expression is shown graphically in Fig. 4.40. We may note that the harmonic distortion is significantly smaller than for the common-source stage. When analyzing this, we found from Eq. (3.48) that an input signal amplitude of 10% of the effective gate voltage results in a distortion of 2.5%. For the differential stage, an input amplitude of 10% of V_{eff} results in only 0.031% distortion, and a distortion of 2.5% is reached for $V_{id} = 0.84\,V_{\text{eff}}$.

Simulation examples. The results found above may be illustrated and verified using LTspice. For this, we must select device parameters and bias conditions for the differential pair. For the purpose of comparison, we select the same transistor parameters as used when simulating the common-source

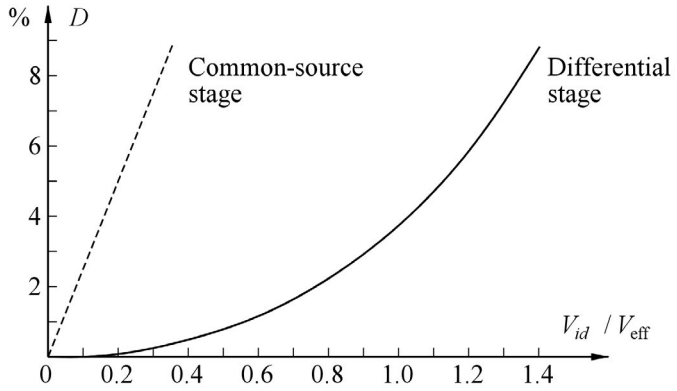


Figure 4.40: Harmonic distortion (third harmonic) in % versus normalized differential input amplitude for the differential pair from Fig. 4.38. For comparison, the dotted line shows the harmonic distortion (second harmonic) given by Eq. (3.48) in a common-source stage.

stage in Section 3.5, i.e., $W = 12.35$ µm and $L = 1$ µm. Also, we select the bias current for each of the transistors to be 100 µA, so $I_B = 200$ µA. With these bias conditions, the effective gate voltage V_{eff} of the transistors is $V_{\text{eff}} = V_{ICM} - V_S - V_t = 0.3$ V, and assuming that the voltage V_S across the bias current source should also be at least 0.3 V, the minimum value of V_{ICM} is $V_{ICM\text{min}} = V_S + V_{\text{eff}} + V_t = 1.0$ V. We select (somewhat arbitrarily) a common-mode input voltage of $V_{ICM} = 1.2$ V. With this value of V_{ICM}, we have $V_S = 0.5$ V, and as $V_{DS\,\text{sat}} = V_{\text{eff}}$, the output voltage range is from $V_S + V_{\text{eff}}$ to V_{DD}, i.e., from 0.8 V to 1.8 V. With a current swing in each of the transistors from 0 to $I_B = 200$ µA, this gives a maximum value for R_D of $R_{D\text{max}} = (1.8\text{ V} - 0.8\text{ V})/(200\text{ µA}) = 5$ kΩ. We select $R_D = 4.5$ kΩ which is half of the value used for the common-source stage in Section 3.5, and with this value of R_D, we can calculate the bias value of the output voltages to be $V_{O1} = V_{O2} = V_{DD} - R_D I_B/2 = 1.35$ V.

The LTspice schematic corresponding to this design is shown in Fig. 4.41.

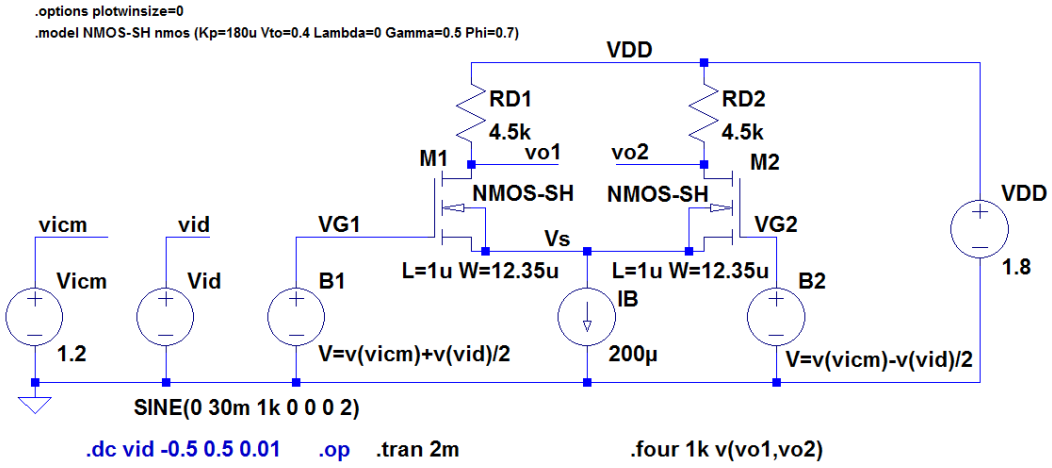


Figure 4.41: LTspice schematic for simulating a differential gain stage.

It is a good idea to verify the bias point by running a '.op' simulation before running the more advanced simulations. Figure 4.42 shows the bias voltages and currents found from the '.op' simulation.

Results from output file

```
--- Operating Point ---

V(vo1):          1.35           voltage
V(vg1):          1.2            voltage
V(vs):           0.500053       voltage
V(vdd):          1.8            voltage
V(vo2):          1.35           voltage
V(vg2):          1.2            voltage
V(vicm):         1.2            voltage
V(vid):          0              voltage
Id(M2):          0.0001         device_current
Is(M2):          -0.0001        device_current
Id(M1):          0.0001         device_current
Is(M1):          -0.0001        device_current
I(Ib):           0.0002         device_current
I(Rd2):          0.0001         device_current
I(Rd1):          0.0001         device_current
```

Results from error log file

```
Semiconductor Device Operating Points:
          --- MOSFET Transistors ---
Name:          m2          m1
Model:      nmos-sh     nmos-sh
Id:         1.00e-04    1.00e-04
Vgs:        7.00e-01    7.00e-01
Vds:        8.50e-01    8.50e-01
Vbs:        0.00e+00    0.00e+00
Vth:        4.00e-01    4.00e-01
Vdsat:      3.00e-01    3.00e-01
Gm:         6.67e-04    6.67e-04
Gds:        0.00e+00    0.00e+00
Gmb:        1.99e-04    1.99e-04
```

Figure 4.42: Results from the output file and the error log file from the '.op' simulation of the differential stage in Fig. 4.41.

Having recognized that all device currents and node voltages are as expected, we may run a '.dc' simulation to show the current in each of the transistors as a function of the differential input voltage. From the '.dc' simulation, we may also show the source voltage v_S versus the input voltage to see the variation given by Eq. (4.51). For this simulation, we sweep the differential input voltage from -0.5 V to $+0.5$ V, i.e., a voltage range extending slightly beyond $-\sqrt{2}V_{\text{eff}}$ to $+\sqrt{2}V_{\text{eff}}$.

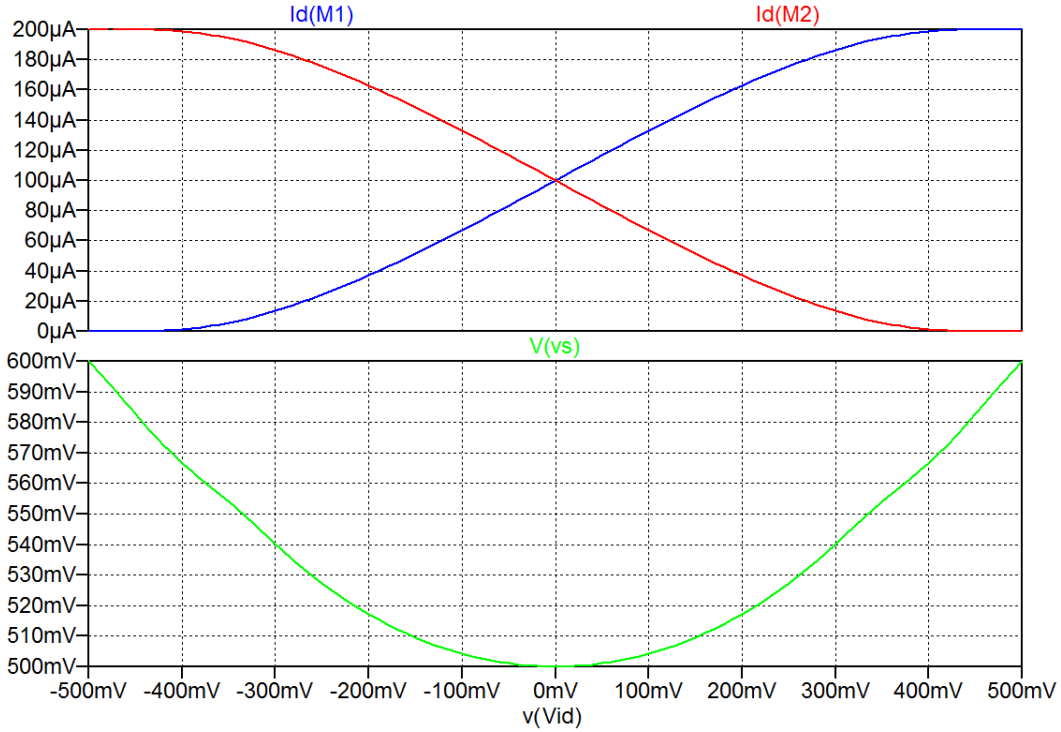


Figure 4.43: Results of a '.dc' simulation of the differential stage in Fig. 4.41.

The results of the '.dc' simulation are shown in Fig. 4.43. The resemblance between this figure and the theoretical results shown in Fig. 4.39 is apparent. Also, we see that the minimum value of v_S occurs for $v_{ID} = 0$, When $|v_{ID}|$ increases, v_S increases following the nonlinear relation Eq. (4.51).

Finally, we may run a transient simulation with a sinusoidal input as defined for v_{ID} in Fig. 4.41. Here we have selected an input amplitude of 30 mV, the same input amplitude as we used for the common-source stage in Fig. 3.24. We also include a '.four' directive in order to find the harmonic distortion.

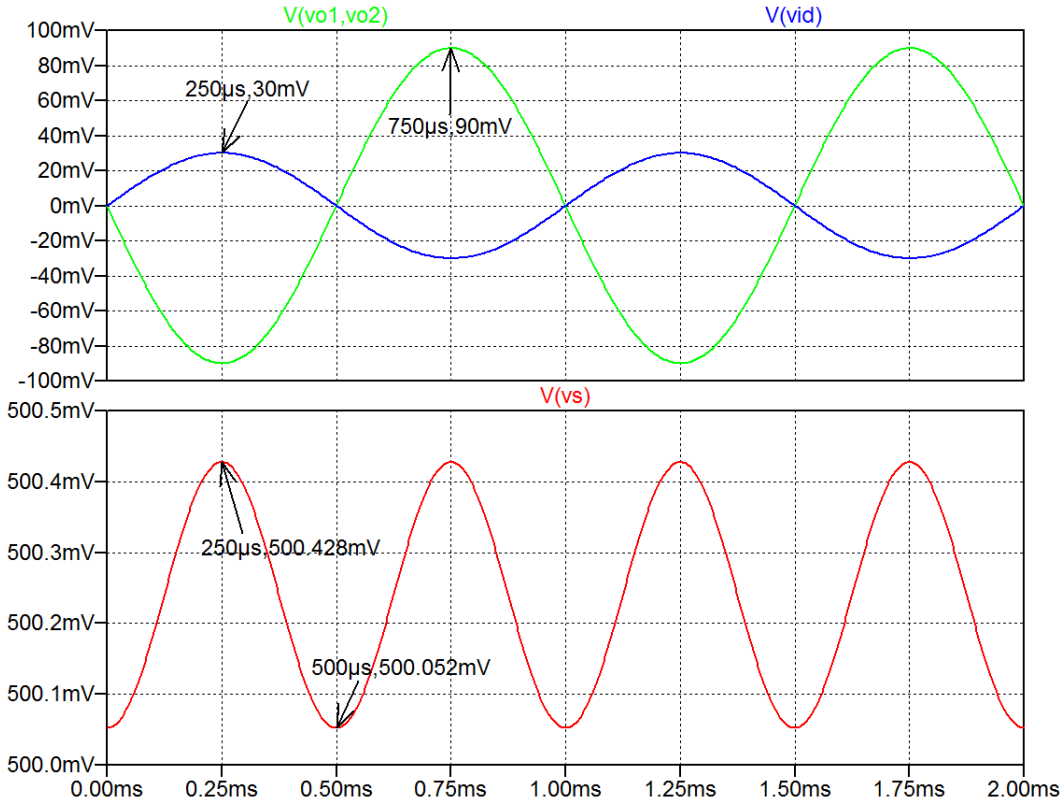


Figure 4.44: Results of a '.tran' simulation of the differential stage in Fig. 4.41.

Figure 4.44 shows resulting plots from this simulation. We see that the differential output signal amplitude is 90 mV, so with an input amplitude of 30 mV, the differential gain is 3 V/V, and it is inverting. For the common-source amplifier, we found a gain of 6 V/V (see Fig. 3.25), so the gain has dropped by a factor of 2. This is not surprising since the drain resistors have been reduced to half of the value used for the common-source amplifier, and the output signal swing is proportional to R_D.

We also note from Fig. 4.44 that v_S shows a small sinusoidal variation but with twice the frequency of the input signal. The frequency doubling comes from the fact that v_S depends on v_{IN}^2 as seen from Eq. (4.51). The small-signal variation in v_S is what cancels the second harmonic distortion in the differential gain stage compared to the inverting gain stage. With a small-signal gate-source voltage amplitude of $V_{id}/2 = 15$ mV, Eq. (3.48) gives a distortion of 1.25% for the common-source stage. The amplitude of v_s in Fig. 4.44 is 0.188 mV, corresponding to 1.25% of $V_{id}/2$, so the v_s signal cancels the distortion which would have been present with $v_s = 0$.

From the error log file, we can find the amplitude of the harmonics in the output signal and we can find the harmonic distortion. Figure 4.45 shows the results from the error log file. Clearly, the third harmonic is the dominant term in the harmonic distortion and the even harmonics are cancelled. The third harmonic is found to give a distortion of 0.031% as also calculated from Eq. (4.57).

Clearly, more simulations may be run to check the performance of the differential pair over a wider range of input amplitudes and common-mode input variations but this is an exercise left for the reader.

Error log file

Harmonic Number	Frequency [Hz]	Fourier Component	Normalized Component
1	1.000e+03	8.993e-02	1.000e+00
2	2.000e+03	2.278e-10	2.533e-09
3	3.000e+03	2.814e-05	3.129e-04
4	4.000e+03	4.540e-10	5.048e-09
5	5.000e+03	6.194e-09	6.888e-08
6	6.000e+03	6.804e-10	7.566e-09
7	7.000e+03	1.270e-08	1.412e-07
8	8.000e+03	9.067e-10	1.008e-08
9	9.000e+03	1.218e-08	1.355e-07

Total Harmonic Distortion: 0.031292%(0.031094%)

Figure 4.45: Results of distortion analysis for the differential stage in Fig. 4.41.

Small-signal analysis. The analysis and simulations above show that the differential stage exhibits a good linearity over a wide range of input signals, so a linear small-signal approximation is a very accurate model, even for fairly large signal swings for the differential pair, provided the output signal swing does not cause the transistors to enter the triode region or the cut-off region.

Using the simple transistor model without channel-length modulation and body effect, the small-signal equivalent circuit corresponding to the schematic from Fig. 4.38 is as shown in Fig. 4.46. In the small-signal equivalent circuit, the dc bias current source for M_1 and M_2 is simply replaced by an open circuit.

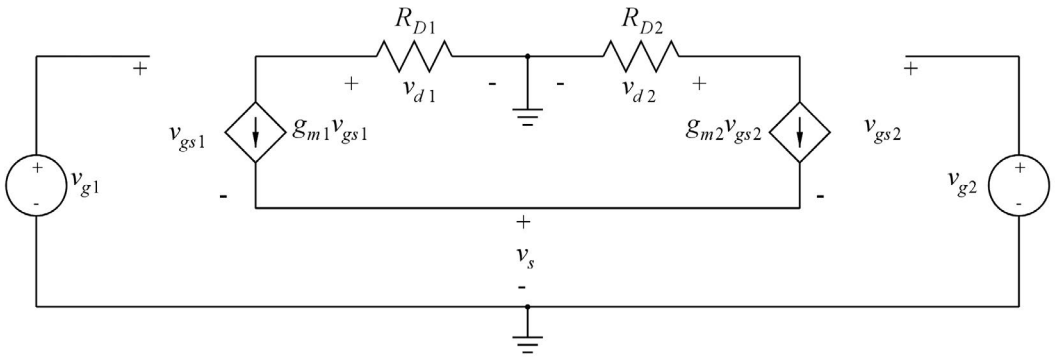


Figure 4.46: Small-signal equivalent circuit for the differential pair from Fig. 4.38.

When applying a differential input voltage to the stage, we have $v_{g1} = -v_{g2} = v_{id}/2$, and using $g_{m1} = g_{m2} = g_m$, a node equation at the source of the transistors yields

$$g_m v_{gs1} + g_m v_{gs2} = 0 \tag{4.58}$$

$$\Rightarrow\ v_{g1} - v_s + v_{g2} - v_s = 0 \ \Rightarrow\ v_s = 0 \tag{4.59}$$

We found in Eq. (4.51) that v_S actually depends on the differential input voltage but in the small-signal result given above, only linear variations are considered, and it can be seen both from Eq. (4.51) and from Fig. 4.43 that the linear term in v_S is 0.

Using $v_s = 0$ and $R_{D1} = R_{D2} = R_D$, we find the small-signal drain currents and output voltages as

$$i_{d1} = -i_{d2} = g_m v_{gs1} = g_m v_{id}/2 \tag{4.60}$$

resulting in

$$v_{d1} = -v_{d2} = -R_D g_m v_{id}/2 \tag{4.61}$$

$$\Rightarrow v_{od} = v_{o1} - v_{o2} = v_{d1} - v_{d2} = -R_D g_m v_{id} \tag{4.62}$$

so the differential small-signal gain is

$$A_d = -R_D g_m \tag{4.63}$$

If a single-ended output is needed, this is obtained by taking the output from just one of the drain terminals, e.g., the drain of M_2. In this case, the voltage gain is only $R_D g_m/2$ as seen from Eq. (4.61).

We observe that the small-signal schematic is symmetric with respect to a vertical axis, so we can analyze the circuit using just one half of the circuit, the differential half-circuit shown in Fig. 4.47. Since $v_s = 0$ for a differential input, the source of the transistor is connected to ground. If a load resistor R_L is connected between v_{d1} and v_{d2}, this may be split into two resistors, each with the value $R_L/2$ connected in series. For a differential signal, the voltage at the node connecting the two resistors is zero, so in the differential half-circuit, a resistor $R_L/2$ is inserted from the output to ground.

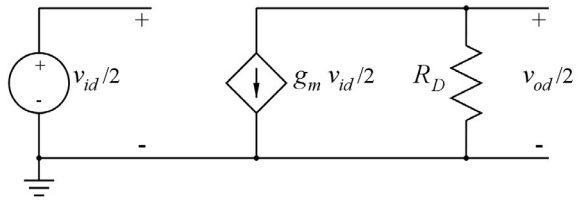


Figure 4.47: Differential half-circuit for the small-signal equivalent from Fig. 4.46.

When applying a common-mode input signal in Fig. 4.46, we have $v_{g1} = v_{g2} = v_{icm}$, so $v_{gs1} = v_{gs2}$, implying $i_{d1} = i_{d2}$. As a node equation at the source of the transistors imposes $i_{d1} + i_{d2} = 0$, we find $i_{d1} = i_{d2} = 0$ and $v_{o1} = v_{o2} = 0$, so also the differential output voltage $v_{od} = 0$. This gives a common-mode gain of $A_{cm} = 0$.

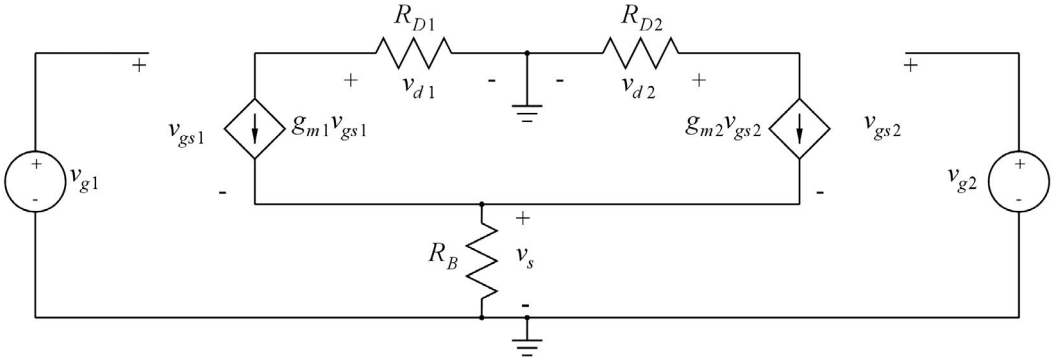


Figure 4.48: Small-signal equivalent circuit for the differential pair including a finite resistor for the bias current source.

Next, we examine the influence of a non-ideal bias current source. In a practical circuit, the bias current source I_B is implemented using a transistor in the same way as the bias current sources used for the other basic gain stages, so it has a finite small-signal output resistance. Accordingly, the small-signal equivalent circuit must be modified by connecting a resistor R_B from the source node for M_1 and M_2 to ground as shown in Fig. 4.48.

For a differential input voltage, this resistor has no influence since $v_s = 0$, so the voltage across R_B is 0. Thus, the differential half-circuit remains unchanged from Fig. 4.47.

For a common-mode input signal, a node equation at the source of M_1 and M_2 gives, assuming $g_m R_B \gg 1$

$$\begin{aligned} 2\,g_m(v_{icm} - v_s) &= \frac{v_s}{R_B} \\ \Rightarrow \quad v_s &= \frac{2\,g_m R_B}{1 + 2\,g_m R_B} v_{icm} \simeq v_{icm} \end{aligned} \tag{4.64}$$

From this, we find

$$i_{d1} = i_{d2} = \frac{v_s}{2R_B} \simeq \frac{v_{icm}}{2R_B} \tag{4.65}$$

and

$$v_{o1} = v_{o2} = -R_D\,i_{d1} = -\frac{R_D}{2R_B}\,v_s \simeq -\frac{R_D}{2R_B}\,v_{icm} \tag{4.66}$$

so with a differential output, $v_{od} = v_{o1} - v_{o2} = 0$ and $A_{cm} = 0$.

If we use only one of the outputs from the differential stage, for example v_{o2}, we find a differential small-signal gain of $v_{o2}/v_{id} = g_m R_D/2$ and a common-mode small-signal gain of $v_{ocm}/v_{icm} \simeq -R_D/(2R_B)$, so the common-mode rejection ratio is $\text{CMRR} = |A_d/A_{cm}| = g_m R_B$.

For the differential pair with a finite value of R_B, we may also define a common-mode small-signal half-circuit. Drawing the small-signal circuit from Fig. 4.48 with R_B split into two parallel-connected resistors $2R_B$ as shown in Fig. 4.49, we see that the half-circuit on each side of the vertical symmetry axis is as shown in Fig. 4.50.

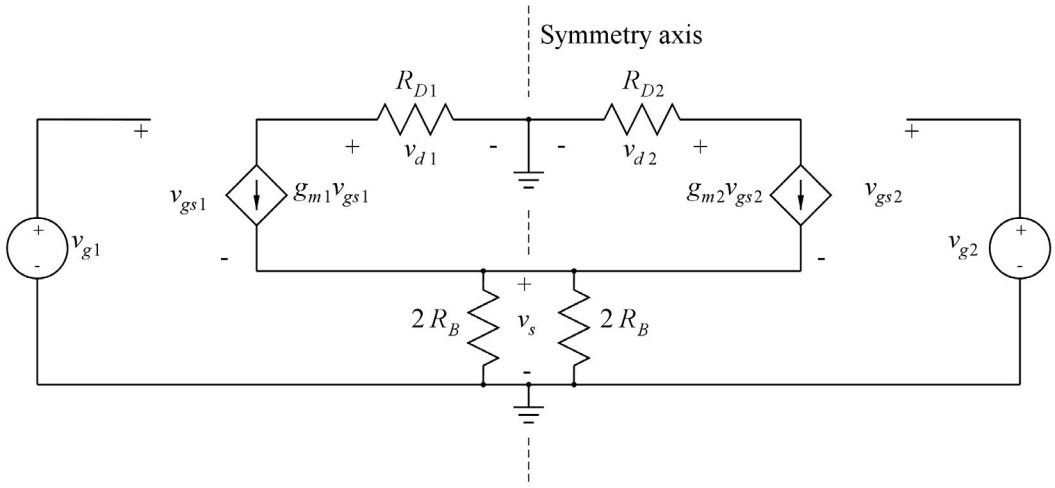


Figure 4.49: Small-signal equivalent circuit for the differential pair showing the symmetry with respect to a vertical axis.

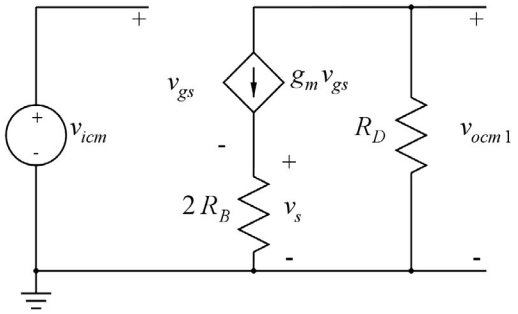


Figure 4.50: Common-mode half-circuit for the small-signal equivalent from Fig. 4.49.

Next, we include the small-signal output resistance of M_1 and M_2, i.e., we take the channel-length modulation of these transistors into account. In the small-signal equivalent circuit, this introduces the two resistors r_{ds1} and r_{ds2} as shown in Fig. 4.51, and splitting this into half-circuits, we find the differential half-circuit and the common-mode half-circuit shown in Fig. 4.52.

From the differential half-circuit, we can calculate the differential gain as

$$A_d = \frac{v_{od}}{v_{id}} = -g_m\,(R_D \parallel r_{ds}) \tag{4.67}$$

and for a single-ended output, the differential gain is only half of this value.

Again, the common-mode gain is 0 for a differential output but for a single-ended output, a detailed analysis shows that the common-mode output voltage is almost unaffected by r_{ds} if $r_{ds} \gg R_D$, so the common-mode rejection ratio turns out to be CMRR $= |A_d/A_{cm}| = g_m R_B$.

Finally, we may examine the influence of the body effect if the bulk of M_1 and M_2 is connected to the negative supply voltage rather than to the source of the transistors. This introduces a current source $g_{mb}\,v_s$ from source to drain in the small-signal half-circuits. For the differential half-circuit, it has no impact

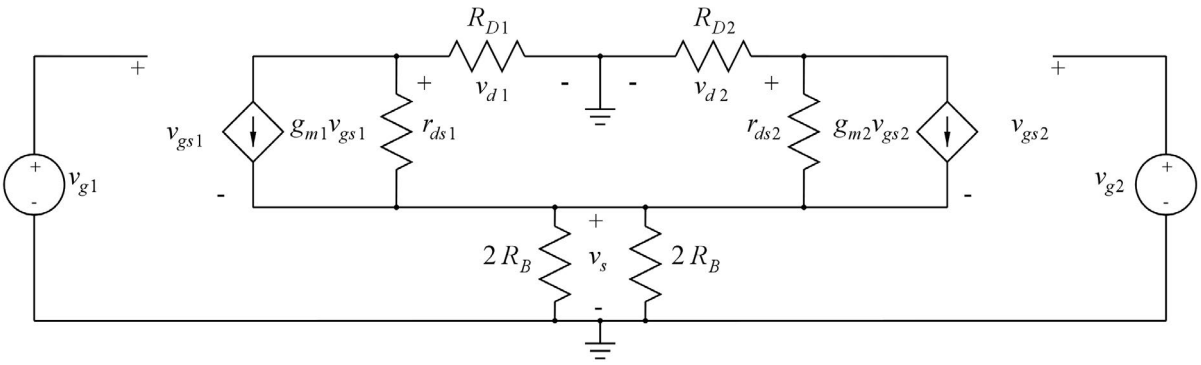


Figure 4.51: Small-signal equivalent circuit for the differential pair including small-signal output resistors for M_1 and M_2.

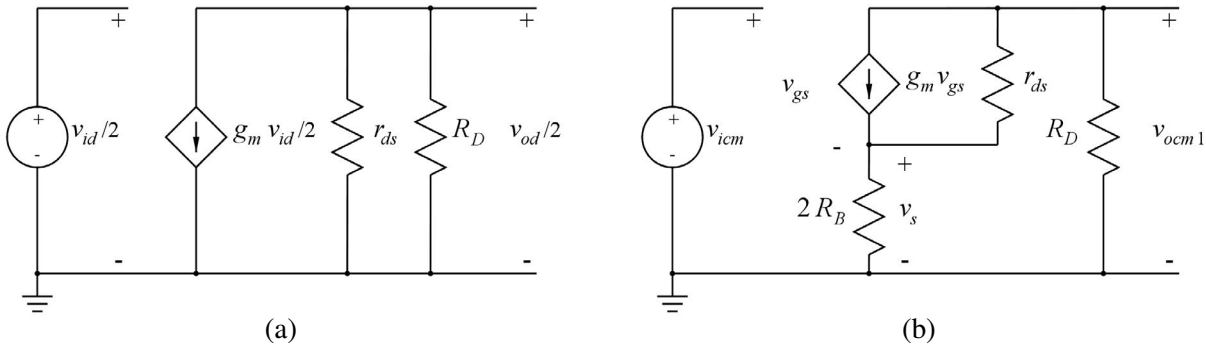


Figure 4.52: Differential half-circuit (a) and common-mode half-circuit (b) corresponding to the small-signal equivalent circuit shown in Fig. 4.51.

since $v_s = 0$, and for the common-mode half-circuit, it causes the gain from v_{icm} to v_s to be approximately

$$\frac{v_s}{v_{icm}} \simeq \frac{g_m}{g_m + g_{mb}} = \frac{1}{1+\chi} \tag{4.68}$$

compare to Eq. (4.17). From this, we find

$$v_{o1} = v_{o2} \simeq -R_D I_B \simeq -\left(\frac{R_D}{2R_B}\right)\left(\frac{v_{icm}}{1+\chi}\right) \tag{4.69}$$

and the single-ended common-mode gain is $A_{cm} = -R_D/(2R_B(1+\chi))$, giving a common-mode rejection ratio of CMRR $= g_m R_B(1+\chi)$.

Simulation examples. Also the results of the small-signal analyses may be verified using LTspice simulation. For the circuit shown in Fig. 4.41, we found the small-signal parameters listed in Fig. 4.42. From this, we can calculate the small-signal differential gain to be $A_d = -g_m R_D = -0.667$ mA/V $\times$ 4.5 kΩ $= 3.00$ V/V as also found from the transient simulation shown in Fig. 4.44.

The small-signal gain can also be found using a '.tf' simulation. Running a '.tf' simulation of the circuit from Fig. 4.41 with '`v(vo1,vo2)`' as the output and '`vid`' as the source gives the differential gain, and using '`v(vo2)`' as the output, the single-ended differential gain is found. Figure 4.53 shows the output files from these '.tf' simulations. The gain values of −3.00 V/V and 1.50 V/V are confirmed. Also, the simulations show the differential output resistance equal to $2R_D$ and the single-ended output resistance equal to R_D.

Output file, differential output, differential input

```
        --- Transfer Function ---

Transfer_function:                -3.00052     transfer
vid#Input_impedance:               1e+020      impedance
output_impedance_at_V(vo1,vo2):    9000        impedance
```

Output file, single-ended output, differential input

```
        --- Transfer Function ---

Transfer_function:                1.50026      transfer
vid#Input_impedance:              1e+020       impedance
output_impedance_at_V(vo2):       4500         impedance
```

Figure 4.53: Results from the output files from '.tf' simulations for simulating the differential small-signal gain of the differential stage in Fig. 4.41 with differential output and single-ended output.

Next, we may run the same simulations with a resistor R_B in parallel with I_B. For R_B, we use a value of 52.5 kΩ which is calculated from Eq. (3.68) as corresponding to the output resistance of an NMOS transistor with $\lambda = 0.1$ V^{-1} and the bias conditions used in Fig. 4.41, i.e., $I_D = 200$ µA and $V_{DS} = 0.5$ V. Since a current of 0.5 V/52.5 kΩ $= 9.5$ µA is flowing in R_B, the current source I_B is reduced to $I_B =$ 200 µA − 9.5 µA $= 190.5$ µA in order to ensure a bias current of 100 µA for M_1 and M_2. Running the same '.tf' simulations as before, we find that the differential gain remains unchanged.

We may also run the '.tf' simulations with '`vicm`' rather than '`vid`' as the source. This gives the output files shown in Fig. 4.54

Output file, differential output, common-mode input

```
        --- Transfer Function ---

Transfer_function:                0            transfer
vicm#Input_impedance:             1e+020       impedance
output_impedance_at_V(vo1,vo2):   9000         impedance
```

Output file, single-ended output, common-mode input

```
        --- Transfer Function ---

Transfer_function:                -0.0422537   transfer
vicm#Input_impedance:             1e+020       impedance
output_impedance_at_V(vo2):       4500         impedance
```

Figure 4.54: Results from the output files from '.tf' simulations for simulating the common-mode small-signal gain of the differential stage in Fig. 4.41 with a finite output resistance of the bias current source.

We note that the common-mode gain is 0 for a differential output, but for a single-ended output, we find a common-mode gain of -0.0423 V/V or -27.5 dB which is close to the approximate value $-R_D/(2R_B) = -0.0429$ V/V given by Eq. (4.66). The single-ended differential gain was found to be 1.5 V/V or 3.5 dB, so the common-mode rejection ration is CMRR(dB) = 31.0 dB.

In order to simulate the channel-length modulation for M_1 and M_2, we modify the transistor model to include $\lambda = 0.1\ \text{V}^{-1}$. Running the '.tf' simulations for a single-ended output and for both a differential input and a common-mode input, we find the output files shown in Fig. 4.55. We notice that the common-mode gain is almost unaffected. The output resistance is reduced slightly because R_D is now in parallel with r_{ds}. For the differential gain, we find an almost unchanged value. The parallel connection of R_D and r_{ds} reduces the gain but this reduction is compensated by a larger value of g_m, see Eq. (3.66). The values of g_m and $g_{ds} = 1/r_{ds}$ are found from a '.op' simulation to be $g_m = 0.695$ mA/V and $g_{ds} = 9.24$ µA/V.

Output file, differential gain

```
        --- Transfer Function ---

Transfer_function:            1.50043       transfer
vid#Input_impedance:          1e+020        impedance
output_impedance_at_V(vo2):   4408.95       impedance
```

Output file, common-mode gain

```
        --- Transfer Function ---

Transfer_function:            -0.0417068    transfer
vicm#Input_impedance:         1e+020        impedance
output_impedance_at_V(vo2):   4408.95       impedance
```

Figure 4.55: Results from the output files from '.tf' simulations for simulating the differential small-signal gain and the common-mode small-signal gain with single-ended output from the differential stage in Fig. 4.41 with a finite output resistance of the bias current source and channel-length modulation in the transistors.

Finally, we may modify the circuit from Fig. 4.41 by connecting the bulk of M_1 and M_2 to ground rather than to their sources. The body effect causes an increase in the threshold voltage of M_1 and M_2, so the source voltage and the voltage across the bias current source is decreased. This causes a slight decrease in the current through the resistor in parallel with the bias current source. In order to find the new bias values and the small-signal parameters, we run a '.op' simulation, the results of which are shown in Fig. 4.56. From the small-signal values, we find $\chi = g_{mb}/g_m = 0.24$.

Results from output file

```
        --- Operating Point ---

V(vo1):         1.35394         voltage
V(vg1):         1.2             voltage
V(vs):          0.406869        voltage
V(vdd):         1.8             voltage
V(vo2):         1.35394         voltage
V(vg2):         1.2             voltage
V(vicm):        1.2             voltage
V(vid):         0               voltage
Id(M2):         9.91249e-005    device_current
Is(M2):         -9.91249e-005   device_current
Id(M1):         9.91249e-005    device_current
Is(M1):         -9.91249e-005   device_current
I(Ib):          0.0001905       device_current
I(Rb):          7.74988e-006    device_current
I(Rd2):         9.91249e-005    device_current
I(Rd1):         9.91249e-005    device_current
```

Results from error log file

```
Semiconductor Device Operating Points:
          --- MOSFET Transistors ---
Name:           m2          m1
Model:       nmos-sh     nmos-sh
Id:          9.91e-05    9.91e-05
Vgs:         7.93e-01    7.93e-01
Vds:         9.47e-01    9.47e-01
Vbs:        -4.07e-01   -4.07e-01
Vth:         5.08e-01    5.08e-01
Vdsat:       2.85e-01    2.85e-01
Gm:          6.95e-04    6.95e-04
Gds:         9.05e-06    9.05e-06
Gmb:         1.65e-04    1.65e-04
```

Figure 4.56: Results from the output file and the error log file from the '.op' simulation of the differential stage in Fig. 4.41, including a resistor R_B in parallel with I_B and including channel-length modulation and body effect for M_1 and M_2.

Output file, differential gain

```
        --- Transfer Function ---

Transfer_function:            1.50163       transfer
vid#Input_impedance:          1e+020        impedance
output_impedance_at_V(vo2):   4410.91       impedance
```

Output file, common-mode gain

```
        --- Transfer Function ---

Transfer_function:            -0.033881     transfer
vicm#Input_impedance:         1e+020        impedance
output_impedance_at_V(vo2):   4410.91       impedance
```

Figure 4.57: Results from the output files from '.tf' simulations for simulating the differential small-signal gain and the common-mode small-signal gain with single-ended output from the differential stage in Fig. 4.41 with a finite output resistance of the bias current source and channel-length modulation and body effect in the transistors.

Running the '.tf' simulations to find the single-ended common-mode gain and differential gain, we find the output files shown in Fig. 4.57. Comparing these results to the results from the differential pair without body effect (Fig. 4.55), we see that the differential gain remains practically the same but the common mode gain is decreased by a factor 1.23 which is very close to $1+\chi$ as expected from Eq. (4.69).

Active load. Just as for the common-source stage, the configuration with drain resistors is not normally used. Rather, an active load is used, and for a differential stage with a single-ended output, this can be implemented as shown in Fig. 4.58.

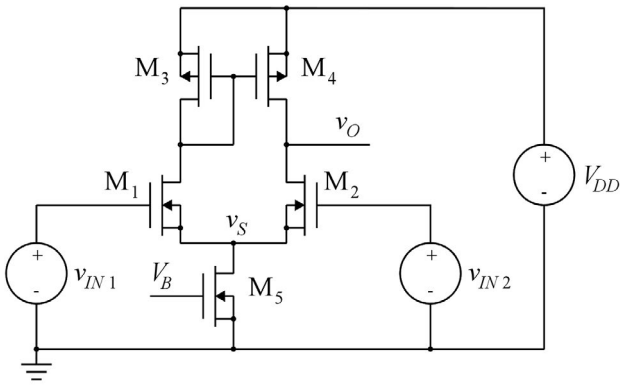


Figure 4.58: A differential gain stage with an NMOS differential pair and a PMOS current-mirror load.

The load is a current mirror which mirrors the output current signal from M_1 to the drain of M_2 so that both the output signal from M_1 and from M_2 contribute to the single-ended output. The advantage of this is that the differential gain is doubled compared to a configuration using only the output signal from M_2. With R_{D2} replaced by the small-signal output resistance of M_4, we find

$$r_{out} = r_{ds2} \parallel r_{ds4} \tag{4.70}$$

$$A_d = \frac{v_o}{v_{id}} \simeq g_{m1}(r_{ds2} \parallel r_{ds4}) \tag{4.71}$$

In Fig. 4.58, the current source I_B is also replaced by a transistor M_5 with a dc gate voltage V_B. With M_5 in the active region, it serves as a dc bias current source with a small-signal output resistance of r_{ds5}.

The current-mirror load also reduces the common-mode gain because the common-mode signal from M_2 is partly cancelled by the common-mode signal from M_1. An approximate expression for the common-mode gain, including body effect for M_1 and M_2, is (compare to Eq. (4.69))

$$A_{cm} = \frac{v_o}{v_{icm}} \simeq -\frac{1}{2\,g_{m3}\,r_{ds5}\,(1+\chi)} \tag{4.72}$$

Without the body effect, the factor $(1+\chi)$ should be omitted.

Output voltage range and input common-mode voltage range. Important parameters for a differential stage are the useful output voltage range and common-mode input voltage range. For the differential stage shown in Fig. 4.58, the bias value of the output voltage is equal to the bias value of the drain voltage of M_3 when the stage is perfectly symmetric. This voltage is given by $V_{DD}-|V_{GS3}| = V_{DD}-|V_{t3}|-|V_{DS\,\mathrm{sat}3}|$, and it serves as the input bias voltage for the next stage in a multistage amplifier design. Since it is referenced to V_{DD}, a following common-source stage should also have its input bias voltage referenced to V_{DD}, i.e., V_{DD} should be the common terminal for input and output of the common-source stage, so the gain transistor must be a PMOS transistor as shown in Fig. 4.9.

The output voltage range with all transistors in the active range extends from a minimum value

$$v_{O\,\mathrm{min}} = V_{ICM} - v_{GS2} + v_{DS\,\mathrm{sat}2} = V_{ICM} - V_{t2} \tag{4.73}$$

limited by M_2, to a maximum value

$$v_{O\,\mathrm{max}} = V_{DD} - V_{DS\,\mathrm{sat}4} \tag{4.74}$$

limited by M_4. Thus, a high common-mode input voltage reduces the output voltage range.

The minimum value of the common-mode input voltage ensuring all transistors in the active region is

$$v_{ICM\,\mathrm{min}} = V_{DS\,\mathrm{sat}5} + V_{GS1} = v_{DS\,\mathrm{sat}5} + V_{t1} + V_{DS\,\mathrm{sat}1} \tag{4.75}$$

The maximum value of the common-mode input voltage with all transistors in the active region is

$$v_{ICM\,\mathrm{max}} = V_{DD} - |V_{GS3}| - V_{DS\,\mathrm{sat}1} + V_{GS1} = V_{DD} - |V_{t3}| - |V_{DS\,\mathrm{sat}3}| + V_{t1} \tag{4.76}$$

You may notice that with $v_{ICM} = v_{ICM\,\mathrm{max}}$, the output voltage range is limited to be from the bias value of $V_{DD}-|V_{t3}|-|V_{DS\,\mathrm{sat}3}|$ to the maximum value given by Eq. (4.74). Thus, in order to operate the gain stage near its maximum common-mode input voltage limit, it is necessary to introduce an input dc offset voltage, offsetting the output bias voltage so that both a positive and a negative output voltage swing can be achieved.

While we noticed before that the body effect on M_1 and M_2 does not influence the differential gain, we see from Eqs. (4.73), (4.75) and (4.76) that it has an impact on the useful input/output voltage range. The body effect causes V_{t1} (and V_{t2}) to increase, so both the minimum common-mode input voltage and the maximum common-mode input voltage increase with M_1 and M_2 subject to body effect. It is apparent from Eq. (4.75) that the input voltage range cannot extend down to the negative supply rail, but with a suitable design of M_1 and M_3, the input voltage range may extend above the positive supply rail.

Simulation examples. We illustrate some of the properties of the gain stage from Fig. 4.58 by LTspice simulations. Figure 4.59 shows an LTspice schematic for the circuit. The input transistors M_1 and M_2 have the same size as in the previous simulations, and for the simulations shown here, the bulk terminals of the transistors are connected to ground, so M_1 and M_2 are subject to body effect.

The bias current source is replaced by a transistor M_5 with twice the channel width of M_1 and M_2 because the drain current is twice as large. The bias voltage V_B has a value giving $I_{D5} = 200$ µA.

The PMOS transistors M_3 and M_4 have been selected to have the same dimensions as the transistors used for the common-source stage in Fig. 4.7. Also the bias currents are the same, so we expect a small-signal

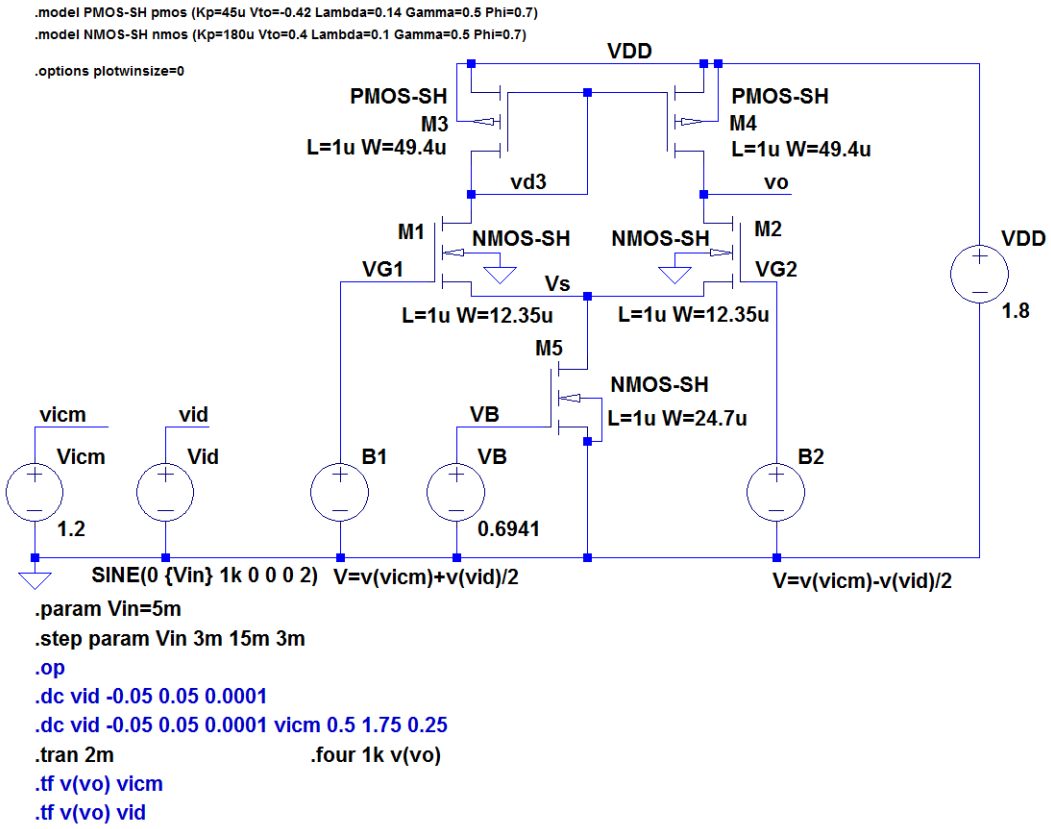


Figure 4.59: LTspice schematic for the differential gain stage with an NMOS differential pair and a PMOS current-mirror load.

differential gain which is about the same as the gain of the common-source stage, i.e., a value of about 32 V/V.

In the schematic in Fig. 4.59, several simulation commands and LTspice directives are included, some of them as comments. They are the directives used for the following simulations.

As usual, we start by running a '.op' simulation for a verification of the bias point and the small-signal parameters. Figure 4.60 shows the bias voltages and currents from the output file, and Fig. 4.61 shows the small-signal parameters from the error log file. Note that the body effect increases the threshold voltage of M_1 and M_2 by 107 mV compared to V_{to}. All voltages and currents are as expected. In particular, we recognize that the bias value of the output voltage is equal to the drain voltage of M_3. Also the small-signal parameters are as expected with values close to those found for the common-source stage.

With this in place, we may run a '.dc' simulation with v_{ID} as the input in order to examine input voltage range, output voltage range and differential gain. From the simulation shown in Fig. 4.62, we find an output voltage range from about 0.7 V to 1.5 V, corresponding to the limits given by Eqs. (4.73) and (4.74) for a common-mode input voltage of 1.2 V. Within this range, the transfer function is almost linear and a plot of dv_O/dv_{IN} shows a gain of 30.9 V/V, very close to the expected value.

Results from output file

```
        --- Operating Point ---

V(vd3):         1.09384         voltage
V(vg1):         1.2             voltage
V(vs):          0.403077        voltage
V(vdd):         1.8             voltage
V(vo):          1.09384         voltage
V(vg2):         1.2             voltage
V(vicm):        1.2             voltage
V(vid):         0               voltage
V(vb):          0.6941          voltage
Id(M5):         0.000200028     device_current
Id(M4):         -0.000100014    device_current
Id(M3):         -0.000100014    device_current
Id(M2):         0.000100014     device_current
Id(M1):         0.000100014     device_current
```

Figure 4.60: Results from the output file from a '.op' simulation of the gain stage shown in Fig. 4.59.

Results from error log file

```
Semiconductor Device Operating Points:
                        --- MOSFET Transistors ---
Name:           m4          m3          m5          m2          m1
Model:       pmos-sh     pmos-sh     nmos-sh     nmos-sh     nmos-sh
Id:         -1.00e-04   -1.00e-04    2.00e-04    1.00e-04    1.00e-04
Vgs:        -7.06e-01   -7.06e-01    6.94e-01    7.97e-01    7.97e-01
Vds:        -7.06e-01   -7.06e-01    4.03e-01    6.91e-01    6.91e-01
Vbs:         0.00e+00    0.00e+00    0.00e+00   -4.03e-01   -4.03e-01
Vth:        -4.20e-01   -4.20e-01    4.00e-01    5.07e-01    5.07e-01
Vdsat:      -2.86e-01   -2.86e-01    2.94e-01    2.90e-01    2.90e-01
Gm:          6.99e-04    6.99e-04    1.36e-03    6.89e-04    6.89e-04
Gds:         1.27e-05    1.27e-05    1.92e-05    9.36e-06    9.36e-06
Gmb:         2.09e-04    2.09e-04    4.06e-04    1.64e-04    1.64e-04
```

Figure 4.61: Results from the error log file from a '.op' simulation of the gain stage shown in Fig. 4.59.

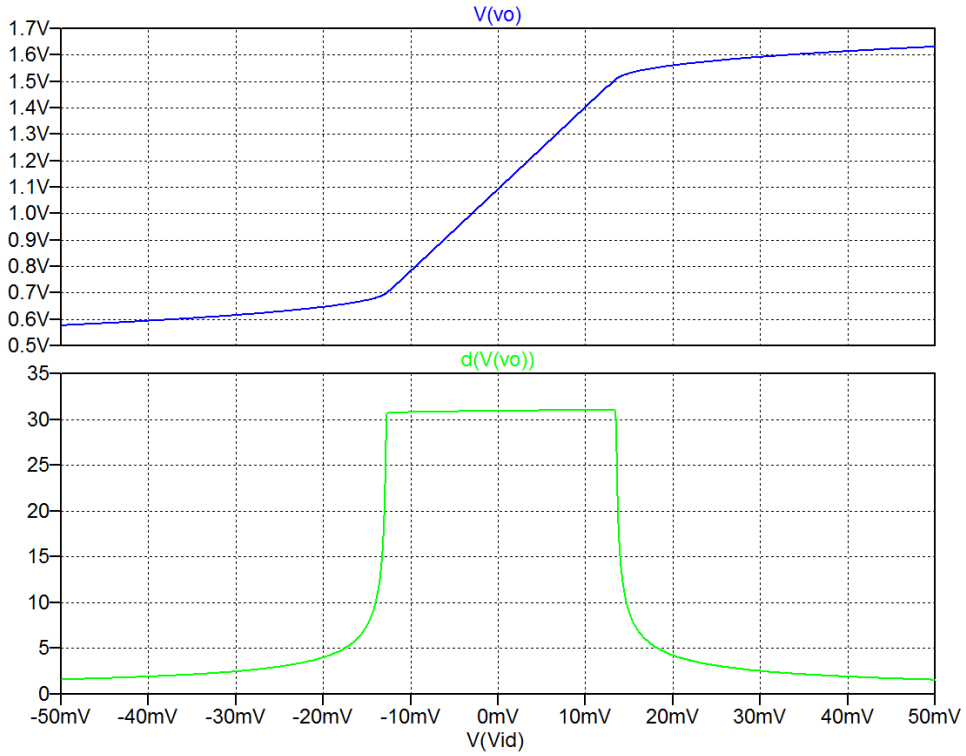


Figure 4.62: Output plot from a '.dc' simulation of the gain stage shown in Fig. 4.59.

We may also examine the common-mode input range by also stepping v_{ICM} in the '.dc' simulation. Stepping v_{ICM} from 0.50 V (well below the lower limit given by Eq. (4.73)) to 1.75 V (well above the upper limit given by Eq. (4.74)), we get the plot shown in Fig. 4.63.

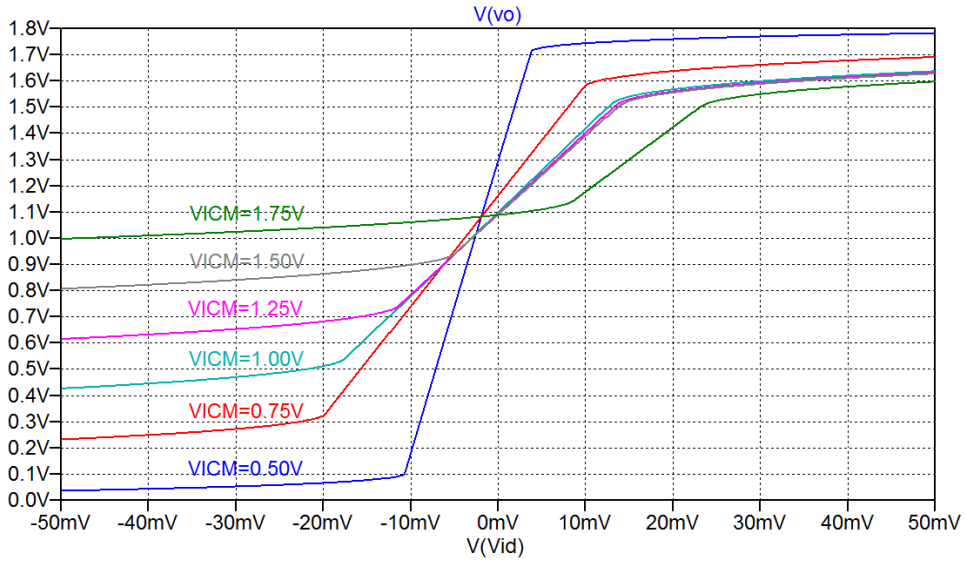


Figure 4.63: Output plot from a '.dc' simulation varying both the differential input voltage and the common-mode input voltage of the gain stage shown in Fig. 4.59.

In Fig. 4.63, we note that for the curves with 1 V $\leq V_{ICM} \leq$ 1.5 V, the transfer function shows almost the same slope for small values of $|v_{ID}|$. For $V_{ICM} = 0.50$ V and $V_{ICM} = 0.75$ V, the slope is higher, indicating a larger gain. This is caused by a reduction of the bias current in M_5 because this transistor is squeezed into the triode region. From Eqs. (3.71) and (3.72), we see that the ratio g_m/g_{ds} is proportional to $1/\sqrt{I_D}$, so the gain increases with decreasing bias current. For $V_{ICM} = 1.75$ V, the gain is much smaller as M_1 enters the triode region unless the differential input voltage is offset towards a positive value of about 15 mV.

For $V_{ICM} = 1.2$ V, we saw in Fig. 4.62 that we could apply an input signal with an amplitude of about 12 mV while maintaining a constant gain. We may run a transient simulation to verify this. With the differential input signal defined as a sinusoid with a frequency of 1 kHz and an amplitude '`Vin`' varied from 3 mV to 15 mV in steps of 3 mV, a '.tran' simulation gives the output shown in Fig. 4.64. We notice almost perfect sinusoidal outputs, except for an input amplitude of 15 mV where there is a visible clipping of both the positive and negative signal excursions.

The linearity may also be verified from the distortion analysis defined by the '.four' directive. The results are given in the error log file. The total harmonic distortion is smaller than 0.1% for all signals, except for the input amplitude of 15 mV which gives a total harmonic distortion of 3.2%. From the error log file, partly shown in Fig. 4.65, we also see that for small signal amplitudes, the second harmonic is the dominant distortion component, and it is proportional to the signal amplitude. This is the distortion

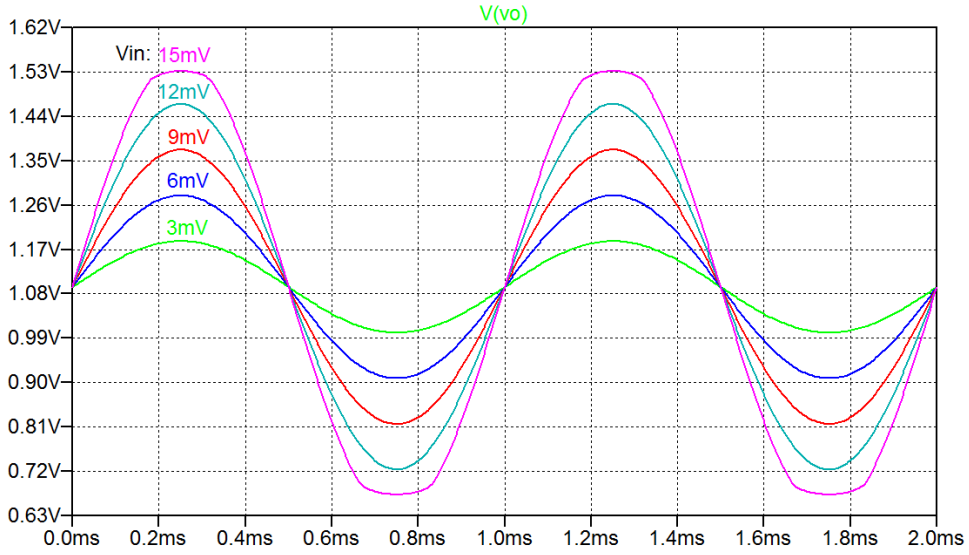


Figure 4.64: Output plot from a '.tran' simulation of the gain stage shown in Fig. 4.59.

characteristics found for a common-source stage, so it indicates that the major source of the distortion is the current mirror M_3 - M_4 and not the differential pair. When the output signal is clipped, several distortion components appear in the output as shown in Fig. 4.65.

Results from error log file

Harmonic Number	Frequency [Hz]	Fourier Component	Normalized Component
.step vin=0.003			
1	1.000e+03	9.275e-02	1.000e+00
2	2.000e+03	2.263e-05	2.440e-04
3	3.000e+03	9.253e-07	9.976e-06
.step vin=0.006			
1	1.000e+03	1.855e-01	1.000e+00
2	2.000e+03	9.051e-05	4.880e-04
3	3.000e+03	7.408e-06	3.994e-05
.step vin=0.009			
1	1.000e+03	2.782e-01	1.000e+00
2	2.000e+03	2.035e-04	7.316e-04
3	3.000e+03	2.500e-05	8.985e-05
.step vin=0.012			
1	1.000e+03	3.708e-01	1.000e+00
2	2.000e+03	3.616e-04	9.749e-04
3	3.000e+03	5.921e-05	1.597e-04
.step vin=0.015			
1	1.000e+03	4.501e-01	1.000e+00
2	2.000e+03	4.142e-03	9.202e-03
3	3.000e+03	1.106e-02	2.457e-02
4	4.000e+03	2.147e-03	4.771e-03
5	5.000e+03	7.143e-03	1.587e-02
6	6.000e+03	5.727e-04	1.272e-03
7	7.000e+03	3.305e-03	7.342e-03
8	8.000e+03	4.673e-04	1.038e-03
9	9.000e+03	5.606e-04	1.245e-03

Total Harmonic Distortion: 3.195163%(3.210823%)

Figure 4.65: Results from the distortion analysis of the gain stage shown in Fig. 4.59.

Finally, we show the results of '.tf' simulations to find the small-signal differential gain and common-mode gain. Figure 4.66 shows the output files for both of these simulations. We find a differential gain of 30.9 V/V $\sim$ 29.8 dB and a common-mode gain of 0.011 V/V ~ -39.2 dB, so the common-mode rejection ratio is 69 dB. From Eqs. (4.71) and (4.72), we find $A_d = 31.2$ V/V and $A_{cm} = 0.011$ V/V, closely matching the simulated values.

Output file, differential gain

```
        --- Transfer Function ---

Transfer_function:            30.9201       transfer
vid#Input_impedance:          1e+020        impedance
output_impedance_at_V(vo):    45548.1       impedance
```

Output file, common-mode gain

```
        --- Transfer Function ---

Transfer_function:            -0.0106715    transfer
vicm#Input_impedance:         1e+020        impedance
output_impedance_at_V(vo):    45548.1       impedance
```

Figure 4.66: Results from the output files from '.tf' simulations for simulating the differential small-signal gain and the common-mode small-signal gain of the gain stage shown in Fig. 4.59.

A PMOS differential pair. The differential gain stage can of course also be implemented using a PMOS differential pair. Figure 4.67 shows the schematic for a PMOS differential stage. The analysis of such a stage is almost identical to the analysis of the NMOS differential pair, and the small-signal diagram for the PMOS differential stage is the same as for the NMOS differential stage. An important difference is the input voltage range and the output bias voltage. For the PMOS differential pair, the input voltage range can extend below the negative supply rail but it cannot reach the positive supply rail, and the output bias voltage is referenced to the negative supply rail, so a subsequent common-source stage should use an NMOS transistor for the gain transistor.

By combining a PMOS differential pair and an NMOS differential pair, it is possible to design a differential gain stage with a rail-to-rail input voltage range (Babanezhad 1988) but this is beyond the scope of this book.

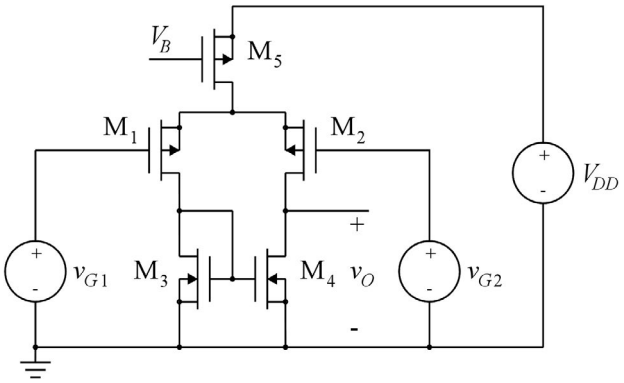


Figure 4.67: A differential gain stage with a PMOS differential pair and an NMOS current-mirror load.

4.5 Frequency response of the basic gain stages

In this section, we will present a simplified analysis of the frequency response of the basic gain stages presented in the previous sections. We saw in Chapter 3 that the MOS transistor model includes small-signal capacitances as shown in Fig. 3.33. These capacitances will cause the gain to drop at high fre-

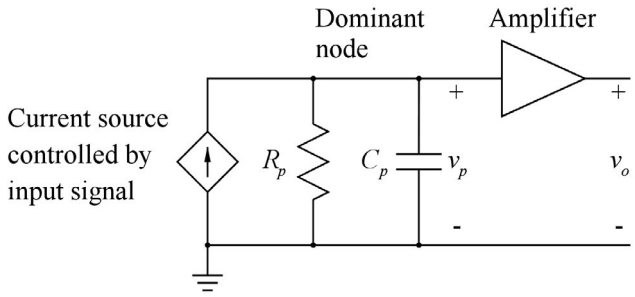


Figure 4.68: Generic gain stage model showing a dominant node.

quencies compared to the low-frequency small-signal gains found in the previous sections. Also, extra capacitances may be inserted to deliberately modify the frequency response.

In many cases, the frequency response may be characterized by a dominant pole, i.e., a pole located at a much lower frequency than any other poles and zeros in the transfer function. If this is the case, the frequency response will fall off with 20 dB per decade of frequency for frequencies above the frequency of the dominant pole. At frequencies well above the dominant pole frequency, the phase is shifted $-90°$ with respect to the low-frequency phase, compare to Fig. 2.23.

At even higher frequencies, inevitably, more poles and zeros will contribute to the gain and phase response of a gain stage, and as we will see in Chapter 6 on feedback, they have a profound effect on the overall system performance but for now, we will concentrate on investigating the frequency response of a basic gain stage when a dominant pole can be identified.

A well-known technique for investigating the high-frequency response of a gain stage is the open-circuit time constant method developed at Massachusetts Institute of Technology in the 1960s (Gray & Searle 1969; Gray, Hurst, Lewis & Meyer 2009). This involves calculating a time constant for each capacitor in the circuit, including floating capacitances, and it provides an estimation of the -3 dB bandwidth of the amplifier. Here we will use a slightly different approach where we calculate a time constant for each node in the signal path of the gain circuit. We do not provide a proof that the method leads to correct results and it should be used with some caution. The justification of the approach is illustrated in Fig. 4.68. A system with only a single, dominant pole will include a node in the signal path characterized by an equivalent capacitance C_p and an equivalent resistance R_p. Figure 4.68 shows a Norton equivalent for the node. The time constant associated with the node is $\tau_p = R_p C_p$, corresponding to the dominant pole frequency $\omega_p = 1/\tau_p$. The dominant node is driven by a frequency-independent controlled current source, representing the circuitry preceding the dominant node, and the dominant node is driving an amplifier with a flat frequency response for the range of frequencies considered.

When using the model shown in Fig. 4.68, we examine each node in the signal path of the circuit and find the equivalent resistance towards ground and capacitance towards ground and calculate the corresponding time constant. If there is a single node with a time constant much larger than all other time constants, we consider this as the dominant node.

Clearly, not all gain circuits can in a reasonable way be represented by the model shown in Fig. 4.68 and it is also important that only the dominant pole can be determined from the simple equivalent. Higher order poles and zeros cannot be determined just by examining the nodes one by one. However, as we will see in the following, an insight into the frequency response of several of the basic gain stages can be obtained by using the approach.

The common-source stage. The low-frequency small-signal model for the common-source stage with a drain resistor is shown in Fig. 4.2. With an active load, the resistor R_D is replaced by the small-signal output resistance of the active load transistor as shown in Fig. 4.3, and if a load consisting of a resistor R_L in parallel with a capacitor C_L is connected to the output, this load appears in parallel with R_D. Also, the input voltage source may have a resistor R_S in series with the voltage source. Also introducing the capacitances from the high-frequency MOS transistor small-signal model shown in Fig. 3.33, we arrive at the small-signal equivalent circuit shown in Fig. 4.69. In this figure, C_{gb} has been included in C_{gs}.

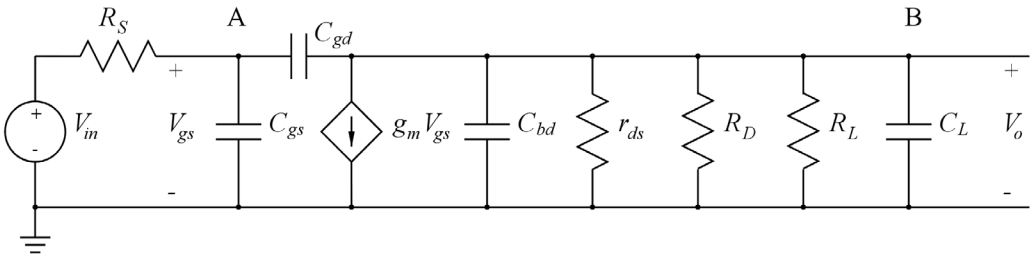


Figure 4.69: High-frequency small-signal equivalent circuit for a common-source stage.

The circuit has three nodes in addition to the common ground. However, the input node V_{in} is directly controlled by the input voltage source, so this cannot provide a limitation to the frequency response. For the other two nodes, the gate of the common-source transistor (node A) and the output node (node B), we need to find the equivalent time constant.

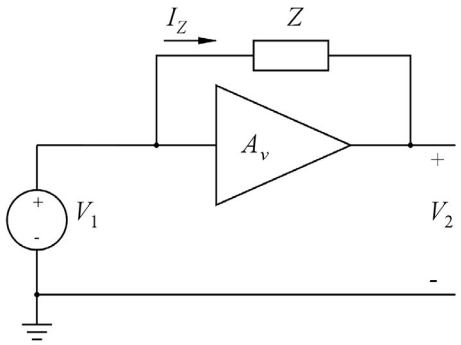


Figure 4.70: Amplifier and impedance for illustrating the Miller theorem.

For the gate of the transistor, we find the resistance R_{pA} to ground with V_{in} reset directly by inspection: $R_{pA} = R_S$. The capacitance looking into node A includes C_{gs} which is directly connected to ground but it also includes an equivalent capacitance to ground from the capacitance C_{gd}. This equivalent capacitance can be found using the Miller theorem published in 1919 by John M. Miller (1882-1962), an American electrical engineer (Miller 1919). The theorem describes how to derive an equivalent circuit for an amplifier with an impedance connected between input and output (Chan Carusone, Johns & Martin 2012). This is illustrated in Fig. 4.70, showing a voltage amplifier with a gain A_v and an impedance Z connected between input and output. In order to find the input impedance, a voltage V_1 is applied to the input. This generates a voltage $V_2 = A_v V_1$ at the output, so the current flowing into Z from the voltage source V_1 is

$$I_Z = \frac{V_1 - V_2}{Z} = \frac{V_1 - A_v V_1}{Z} = \frac{(1 - A_v) V_1}{Z} \tag{4.77}$$

Thus, the equivalent input impedance Z_M resulting from Z is

$$Z_M = \frac{V_1}{I_Z} = \frac{Z}{1 - A_v} \tag{4.78}$$

For the small-signal equivalent circuit shown in Fig. 4.69 where $A_v = -g_m (r_{ds} \parallel R_D \parallel R_L)$ and $Z = 1/(sC_{gd})$, the equivalent input impedance resulting from C_{gd} is $Z_M = 1/(sC_{gd}(1 - A_v))$, corresponding to an equivalent input capacitance

$$C_M = (1 - A_v) C_{gd} = (1 + g_m (r_{ds} \parallel R_D \parallel R_L)) C_{gd} \tag{4.79}$$

The capacitance C_M derived using the Miller theorem is called the Miller capacitance. Thus, the capacitance towards ground for node A is

$$C_{pA} = C_{gs} + C_M = C_{gs} + (1 - A_v) C_{gd} = C_{gs} + (1 + g_m (r_{ds} \parallel R_D \parallel R_L)) C_{gd} \tag{4.80}$$

and with $R_{pA} = R_S$, we find the time constant

$$\tau_{pA} = R_S (C_{gs} + (1 + g_m (r_{ds} \parallel R_D \parallel R_L)) C_{gd}) = R_S (C_{gs} + C_M) \tag{4.81}$$

For the output node, we have $(r_{ds} \parallel R_D \parallel R_L)$ and $C_L + C_{bd}$ directly connected to ground, but we also have C_{gd} connected between input and output. When a test voltage is applied to the output node (with V_{in} reset) in order to find the impedance in this node, a current flows through C_{gd} and the parallel connection of R_S and C_{gs}, and this generates a voltage V_{gs}, in turn adding a current $g_m V_{gs}$ to the current from the test voltage source.

The analysis of this is fairly complicated, so in order to limit the calculations, we first consider the case that R_S is small, $R_S \ll 1/(\omega C_{gd})$, so that the voltage division between $R_S \parallel (1/(sC_{gs}))$ and $1/(\omega C_{gd})$ results in $V_{gs} \simeq 0$ when $V_{in} = 0$, and C_{gd} appears as a capacitance from the output node to ground.

For this case, we then find the time constant of the output node to be

$$\tau_{pB} \simeq (r_{ds} \parallel R_D \parallel R_L) (C_{gd} + C_{bd} + C_L) \tag{4.82}$$

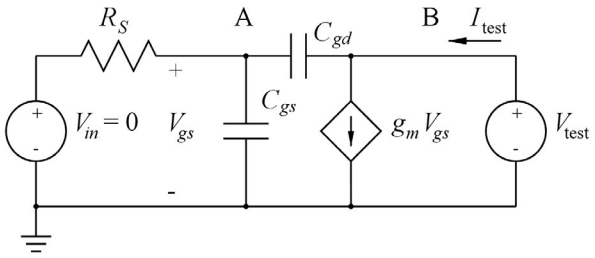


Figure 4.71: Equivalent circuit for finding the impact of C_{gd}.

Obviously, with R_S small, the time constant from the output node is the dominant time constant and we find a dominant pole at the frequency

$$\omega_{pB} \simeq \frac{1}{(r_{ds} \parallel R_D \parallel R_L)(C_{gd}+C_{bd}+C_L)} \tag{4.83}$$

Next, we consider the case that R_S is large. This is the likely situation if the common-source stage is driven from another common-source stage or from a differential stage where the output resistance is on the order of a transistor output resistance r_{ds}. For this case, we expect a large time constant from the gate node because both R_S and the Miller capacitance C_M can be expected to be fairly large compared to the capacitances and resistances determining the time constant from the output node but we need to verify this through a calculation of the time constant from the output node.

For finding the impact of C_{gd} on the impedance from node B to ground, we examine the circuit shown in Fig. 4.71. We apply a test voltage V_{test} to node B and find the current I_{test}. Using the voltage divider rule Eq. (2.29), we find

$$V_{gs} = \frac{R_S \parallel (1/(sC_{gs}))}{R_S \parallel (1/(sC_{gs}))+1/(sC_{gd})} V_{\text{test}} = \frac{C_{gd}}{C_{gd}+C_{gs}+1/(sR_S)} V_{\text{test}} \tag{4.84}$$

For $R_S \gg 1/(\omega(C_{gs}+C_{gd}))$, we may use the approximation

$$V_{gs} \simeq \left(\frac{C_{gd}}{C_{gs}+C_{gd}}\right) V_{\text{test}} \tag{4.85}$$

$$I_{\text{test}} = g_m V_{gs} + V_{\text{test}} \left(\frac{1}{R_S \parallel (1/(sC_{gs}))+1/(sC_{gd})}\right) \simeq \left(g_m \frac{C_{gd}}{C_{gs}+C_{gd}} + s\frac{C_{gs}C_{gd}}{C_{gs}+C_{gd}}\right) V_{\text{test}} \tag{4.86}$$

From Eq. (4.86), we see that C_{gd} corresponds to a resistor $R=(C_{gs}+C_{gd})/(C_{gd}g_m)$ to ground in parallel with a capacitor $C=C_{gs}C_{gd}/(C_{gs}+C_{gd})$ to ground. Combining this with $(r_{ds} \parallel R_D \parallel R_L)$ and C_L+C_{bd} also connected to ground, we find

$$\tau_{pB} \simeq \left[r_{ds} \parallel R_D \parallel R_L \parallel ((1+C_{gs}/C_{gd})/g_m)\right] \left[C_L+C_{bd}+C_{gs}C_{gd}/(C_{gs}+C_{gd})\right] \tag{4.87}$$

From this expression, we find that with a finite value of $(r_{ds} \parallel R_D \parallel R_L)$

$$\tau_{pB} < \left(1+C_{gs}/C_{gd}\right)\left(\frac{1}{g_m}\right)\left(C_L+C_{bd}+\frac{C_{gd}C_{gs}}{C_{gs}+C_{gd}}\right) = \frac{1}{g_m}\left((1+C_{gs}/C_{gd})(C_L+C_{bd})+C_{gs}\right) \tag{4.88}$$

When comparing Eqs. (4.81) and (4.88), we notice that with R_S on the order of r_{ds} and $r_{ds} \gg 1/g_m$, it is likely that τ_{pA} is larger than τ_{pB} unless C_L is very large. This implies that the dominant pole normally comes from the gate node for a common-source stage driven from a stage with a high output resistance. It also implies that the Miller capacitance C_M plays an important role when designing the frequency response. By inserting an extra capacitor between gate and drain, the dominant pole can be moved down in frequency. At the same time, the time constant τ_{pB} is reduced, moving the pole from the output node to a higher frequency so the extra capacitor gives rise to a pole splitting.

Simulation examples. For showing the frequency response of a common-source stage we revisit the circuit shown in Fig. 4.7. This is a common-source stage with an active load. For simulating the frequency response, the transistor models must be augmented with the high-frequency model parameters shown in Fig. 3.35. Also, device dimensions for source and drain diffusions must be specified as explained in Section 3.5. The schematic shown in Fig. 4.72 includes this additional information.

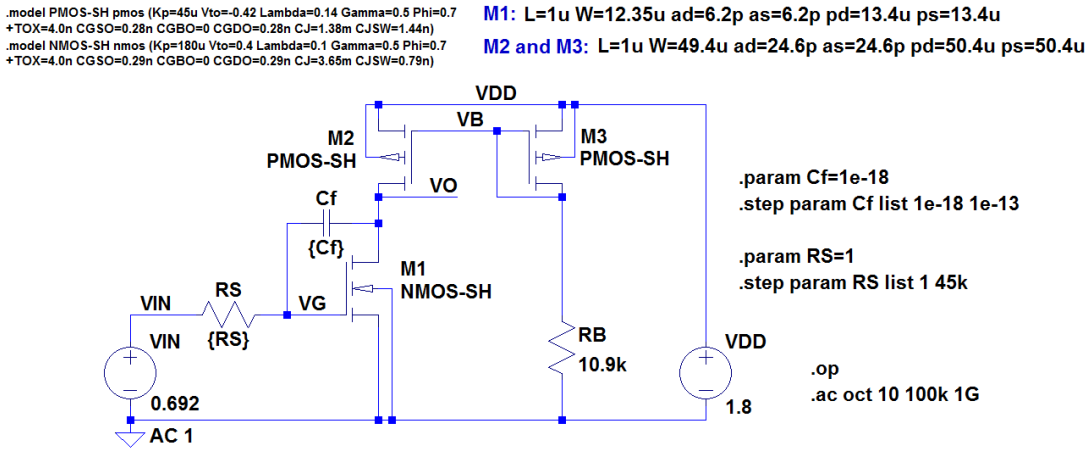


Figure 4.72: LTspice schematic for simulating the frequency response of a common-source stage with an active load.

In order to simulate the influence of a series resistance at the input and of an extra capacitance from gate to drain of the gain transistor, a resistor R_S and a capacitor C_f are inserted in the schematic. An LTspice '.step param' directive for changing the value of R_S between 0 (or 1 Ω since LTspice does not accept a value of 0) and 45 kΩ has been specified as has a directive for changing C_f between 0 and 0.1 pF.

The first simulation to run is a '.op' simulation in order to verify the bias point and to find the small-signal parameters of the transistors. When running the '.op' simulation, the '.step param' directives are turned into comments and '.param' directives are used for specifying values of 0 for R_S and C_f. The error log file from this simulation is shown in Fig. 4.73.

Error log file

```
Semiconductor Device Operating Points:
                  --- MOSFET Transistors ---
Name:           m3          m2          m1
Model:       pmos-sh     pmos-sh     nmos-sh
Id:        -1.00e-04   -1.03e-04    1.03e-04
Vgs:       -7.07e-01   -7.07e-01    6.92e-01
Vds:       -7.07e-01   -9.23e-01    8.77e-01
Vbs:        0.00e+00    0.00e+00    0.00e+00
Vth:       -4.20e-01   -4.20e-01    4.00e-01
Vdsat:     -2.87e-01   -2.87e-01    2.92e-01
Gm:         7.00e-04    7.19e-04    7.06e-04
Gds:        1.28e-05    1.28e-05    9.48e-06
Gmb:        2.09e-04    2.15e-04    2.11e-04
Cbd:        7.77e-14    7.27e-14    2.29e-14
Cbs:        1.07e-13    1.07e-13    3.32e-14
Cgsov:      1.38e-14    1.38e-14    3.58e-15
Cgdov:      1.38e-14    1.38e-14    3.58e-15
Cgbov:      0.00e+00    0.00e+00    0.00e+00
Cgs:        2.84e-13    2.84e-13    7.11e-14
Cgd:        0.00e+00    0.00e+00    0.00e+00
Cgb:        0.00e+00    0.00e+00    0.00e+00
```

Figure 4.73: Results from the error log file from a '.op' simulation of the gain stage shown in Fig. 4.72.

From this, we can find the device parameters corresponding to the small-signal equivalent circuit shown in Fig. 4.69:

$$\begin{aligned} g_m &= \texttt{Gm1} = 0.706 \text{ mA/V} \\ r_{ds} &= 1/\texttt{Gds1} = 105 \text{ k}\Omega \\ R_D &= 1/\texttt{Gds2} = 78 \text{ k}\Omega \\ C_{gs} &= \texttt{Cgs1} + \texttt{Cgsov1} + \texttt{Cgb1} = 74.68 \text{ fF} \\ C_{gd} &= \texttt{Cgd1} + \texttt{Cgdov1} = 3.58 \text{ fF} \\ C_{bd} &= \texttt{Cbd1} = 22.9 \text{ fF} \\ C_L &= \texttt{Cbd2} + \texttt{Cgd2} + \texttt{Cgdov2} = 86.5 \text{ fF} \end{aligned}$$

Here, C_L is just the parasitic capacitances from M_2. No extra load capacitance is assumed.

It should be noted that LTspice uses a more complex modeling of the capacitive effects in the transistor than defined by the small-signal model shown in Fig. 3.33. In particular, the thin gate oxide capacitance is not just assigned as a gate-source capacitance defined by Eq. (3.85) but distributed among the gate, source, drain, and bulk regions, see the 'LTspiceHelp'. This implies that the transistor capacitances found in the error log file, Fig. 4.73, should only be taken as approximate values when used in hand calculations. For finding exact values of, for example, input capacitances and output capacitances from an amplifier circuit, simulations directly showing the capacitance values may be used (Bruun 2020, Tutorial 2.6). An example of this is given in Problem 4.17.

With $R_S = 0$, we expect the dominant pole to come from the output node and we can find the pole frequency using Eq. (4.83). In Eq. (4.83), C_{gd} should be replaced by $C_{gd} + C_f$.
With $C_f \simeq 0$, this gives $f_{pB} = \omega_{pB}/(2\pi) = 31.5$ MHz.
With $C_f = 0.1$ pF, this gives $f_{pB} = \omega_{pB}/(2\pi) = 16.7$ MHz. We notice that C_f lowers f_{pB} in the same way as an extra load capacitance C_L would do.

With $R_S = 45$ kΩ, corresponding to the output resistance of a stage similar to the stage shown in Fig. 4.72, we expect the dominant pole to come from the input node and we can find the pole frequency using Eq. (4.81), again with C_{gd} replaced by $C_{gd} + C_f$.
With $C_f \simeq 0$, this gives $f_{pA} = 1/(2\pi\tau_{pA}) = 18.5$ MHz.
With $C_f = 0.1$ pF, this gives $f_{pA} = 1/(2\pi\tau_{pA}) = 1.02$ MHz.

Using a '.ac' simulation in LTspice, we can plot the frequency response, and the dominant pole can be found as the -3 dB frequency. Figure 4.74 shows the ac response for the four cases listed above and the -3 dB frequencies are indicated on the magnitude plots.

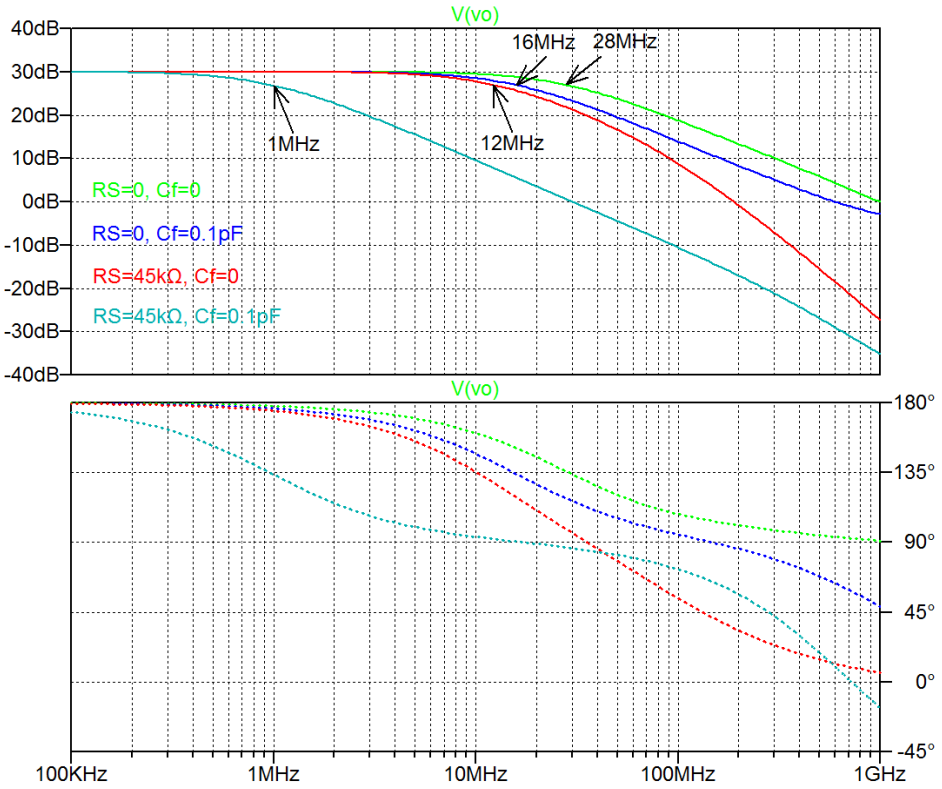


Figure 4.74: Output plot from a '.ac' simulation of the gain stage shown in Fig. 4.72.

We see that except for the case with $R_S = 45$ kΩ and $C_f = 0$, the simulated -3 dB bandwidths are fairly close to the calculated dominant pole frequencies, although tending to be somewhat lower. This is partly caused by the difference in capacitance modeling mentioned above, partly by poles and zeros at higher frequencies. From the phase response, we notice that the phase shift at high frequencies is more than $-90°$, and for $R_S = 45$ kΩ, it exceeds $-180°$. For the circuit with $R_S = 0$, the additional phase shift comes from a zero in the transfer function, see Problem 4.10.

For the circuit with $R_S = 45$ kΩ, we can also expect a pole as indicated by the non-dominant time constant τ_{pB}. From Eq. (4.87), we can estimate this time constant and the corresponding frequency.

With $C_f = 0.1$ pF, we find $f_{pB} = 1/(2\pi\tau_{pB}) \simeq 450$ MHz which is more than $100 \times f_{pA}$, so the assumption of a dominant pole at the input is clearly fulfilled as is also evident from Fig. 4.74. However, we do find a phase shift of more than $-180°$ at very high frequencies, indicating the presence of an additional pole or zero. This is an issue which will be investigated in more detail in Chapter 7.

With $C_f = 0$, we find $f_{pB} = 1/(2\pi\tau_{pB}) \simeq 77$ MHz which is only about $4.1 \times f_{pA}$, so the assumption of a dominant pole is not fulfilled, implying that the -3 dB frequency cannot be found from Eq. (4.81). Also, the assumption of $R_S \gg 1/(\omega(C_{gs}+C_{gd}))$ is not fulfilled for $\omega = 2\pi \times 77$ MHz, so Eq. (4.87) does not yield an accurate estimate for τ_B. This is also apparent from the gain and phase plots (red curves in Fig. 4.74), showing a roll-off of the gain of about 40 dB per decade for frequencies above 100 MHz and a phase shift continuing smoothly towards 0° without a flat level after $-90°$ phase shift.

The common-drain stage. The low-frequency small-signal model for the common-drain stage (or source follower) is shown in Fig. 4.14. For an analysis of the high-frequency properties of the stage, this figure must be modified to include the parasitic transistor capacitances. Also, a load consisting of a resistor R_L in parallel with a capacitor C_L should be connected to the output, and the input voltage source may have a resistor R_S in series with the voltage source. The resulting small-signal equivalent circuit is shown in Fig. 4.75 where the resistor r_{out} is given by Eq. (4.18). In this figure, C_{gb} has been included in C_{gd} and the parasitic capacitances C_{bd} and C_{gd} from the bias transistor M_2 are included in C_L.

The topology of the small-signal circuit shown in Fig. 4.75 is very similar to the topology of the circuit shown in Fig. 4.69, so we can use similar approaches when analyzing the time constants of the nodes in the circuit.

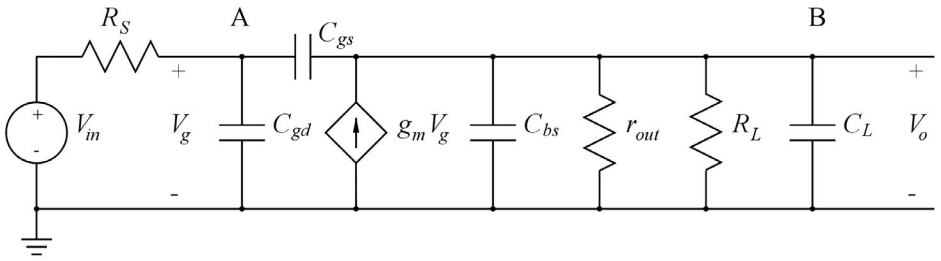


Figure 4.75: High-frequency small-signal equivalent circuit for a common-drain stage.

For the gate of the common-drain transistor, node A, we find the resistance to ground as $R_{pA} = R_S$. For the capacitance, we use Miller's theorem to find

$$C_{pA} = C_{gd} + C_{gs}(1 - V_o/V_g) = C_{gd} + C_{gs}(1 - g_m(r_{out} \parallel R_L)) \tag{4.89}$$

Using Eq. (4.18), we find

$$g_m(r_{out} \parallel R_L) = \frac{g_m}{1/r_{out} + 1/R_L} \simeq \frac{g_m R_L}{1 + (g_m + g_{mb})R_L} \tag{4.90}$$

so

$$\tau_{pA} \simeq R_S\left(C_{gd} + C_{gs}\left(\frac{1 + g_{mb}R_L}{1 + (g_m + g_{mb})R_L}\right)\right) \tag{4.91}$$

For the output node (node B), we use the same approach as for the common-source stage, i.e., we consider the case that R_S is very small and the case that R_S is very large. For R_S very small, we assume $V_g = 0$ when $V_{in} = 0$, so C_{gs} appears as a capacitor to ground. For this case, the time constant for the output node is

$$\tau_{pB} = (r_{out} \parallel R_L)(C_L + C_{bs} + C_{gs}) \simeq \left(\frac{R_L}{1 + (g_m + g_{mb})R_L}\right)(C_L + C_{bs} + C_{gs}) \tag{4.92}$$

For R_S very large, we use the same approach as used when deriving Eqs. (4.85) - (4.87). We find

$$V_g = \frac{C_{gs}}{C_{gs}+C_{gd}} V_{\text{test}} \tag{4.93}$$

$$I_{\text{test}} = -g_m V_g + V_{\text{test}}\, s\, \frac{C_{gs}C_{gd}}{C_{gs}+C_{gd}} = \left(-g_m \frac{C_{gs}}{C_{gs}+C_{gd}} + s\, \frac{C_{gs}C_{gd}}{C_{gs}+C_{gd}} \right) V_{\text{test}} \tag{4.94}$$

$$\tau_{pB} = \left[r_{out} \parallel R_L \parallel (-(1+C_{gd}/C_{gs})/g_m) \right] \left[C_L + C_{bs} + C_{gs}C_{gd}/(C_{gs}+C_{gd}) \right] \tag{4.95}$$

Notice that since the controlled current source $g_m V_g$ has the opposite direction compared to Fig. 4.71, a negative resistance appears in parallel with $r_{out} \parallel R_L$. Using Eq. (4.18) and $C_{gs} \gg C_{gd}$, we find

$$r_{out} \parallel R_L \parallel (-(1+C_{gd}/C_{gs})/g_m) \simeq \frac{1}{g_m + g_{mb} + 1/R_L - g_m} = \frac{R_L}{1+g_{mb}R_L} \tag{4.96}$$

$$\tau_{pB} \simeq \left(\frac{R_L}{1+g_{mb}R_L} \right) \left(C_L + C_{bs} + \frac{C_{gd}C_{gs}}{C_{gs}+C_{gd}} \right) \simeq \left(\frac{R_L}{1+g_{mb}R_L} \right) \left(C_L + C_{bs} + C_{gd} \right) \tag{4.97}$$

We have considered the common-drain stage including the body effect for M_1. If M_1 has source and bulk connected, there is no body effect and g_{mb} should be reset ($g_{mb} = 0$) in the equations above.

The common-drain stage is normally used as a buffer amplifier for driving a load which cannot be directly driven by a common-source stage. Therefore, with R_S representing the output resistance of a common-source stage, it is likely that $R_S \gg R_L/(1+g_{mb}R_L)$, so unless C_L is very large, the input time constant τ_{pA} is larger than the output time constant τ_{pB}. However, whether the pole from the input node can be considered as a dominant pole depends on the actual values of the transistor parameters and the load impedance involved. A general, detailed analysis is beyond the scope of the present book but can be found in Chan Carusone, Johns & Martin (2012). Here, we illustrate the frequency performance of the common-drain stage by LTspice simulations of the same common-drain stage as used for the low-frequency simulations discussed in Section 4.2.

Simulation examples. Figure 4.76 shows the LTspice schematic for simulating the frequency response of the common-drain stage previously simulated at low frequencies. The transistor model includes the high-frequency model parameters and also parameters for both channel-length modulation and body effect. The transistor specifications include dimensions for source and drain diffusions.

Directives for stepping the resistor R_S between 0 and 45 kΩ and the load capacitor C_L between 0 and 3.5 pF are also included. First, we run a '.op' simulation to verify the bias point and to find the small-signal parameters. The error log file from this simulation is shown in Fig. 4.77.

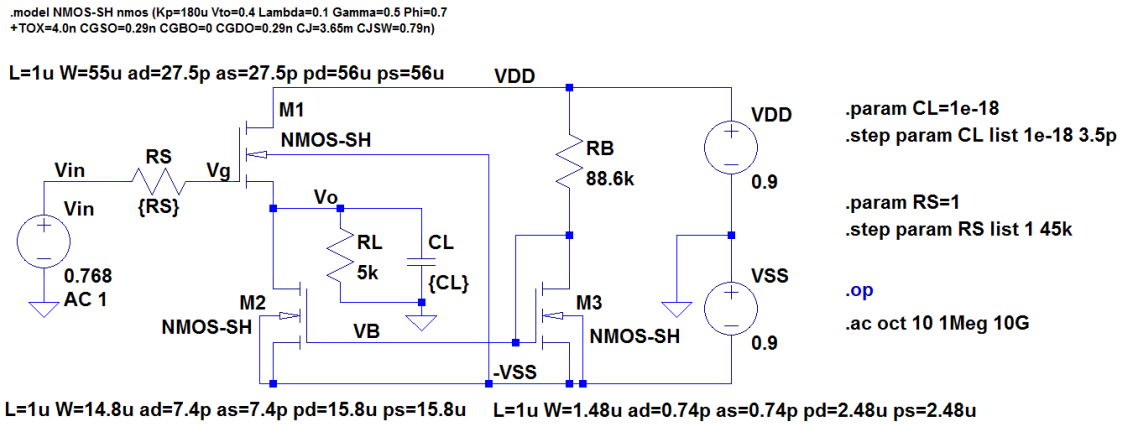


Figure 4.76: LTspice schematic for simulating the frequency response of a common-drain stage.

Error log file

```
Semiconductor Device Operating Points:
                    --- MOSFET Transistors ---
Name:           m3          m2          m1
Model:       nmos-sh     nmos-sh     nmos-sh
Id:         1.25e-05    1.27e-04    1.27e-04
Vgs:        6.96e-01    6.96e-01    7.68e-01
Vds:        6.96e-01    9.00e-01    9.00e-01
Vbs:        0.00e+00    0.00e+00   -9.00e-01
Vth:        4.00e-01    4.00e-01    6.14e-01
Vdsat:      2.96e-01    2.96e-01    1.53e-01
Gm:         8.43e-05    8.59e-04    1.66e-03
Gds:        1.17e-06    1.17e-05    1.17e-05
Gmb:        2.52e-05    2.57e-04    3.27e-04
Cbd:        3.41e-15    2.71e-14    8.02e-14
Cbs:        4.66e-15    3.95e-14    9.92e-14
Cgsov:      4.29e-16    4.29e-15    1.59e-14
Cgdov:      4.29e-16    4.29e-15    1.59e-14
Cgbov:      0.00e+00    0.00e+00    0.00e+00
Cgs:        8.52e-15    8.52e-14    3.17e-13
Cgd:        0.00e+00    0.00e+00    0.00e+00
Cgb:        0.00e+00    0.00e+00    0.00e+00
```

Figure 4.77: Results from the error log file from a '.op' simulation of the gain stage shown in Fig. 4.76.

From this, we can find the device parameters corresponding to the small-signal equivalent circuit shown in Fig. 4.75:

$$
\begin{aligned}
g_m &= \texttt{Gm1} = 1.66\ \text{mA/V} \\
g_{mb} &= \texttt{Gmb1} = 0.327\ \text{mA/V} \\
r_{out} &= 1/(\texttt{Gds1}+\texttt{Gds2}+\texttt{Gm1}+\texttt{Gmb1}) = 498\ \Omega \\
C_{gs} &= \texttt{Cgs1}+\texttt{Cgsov1} = 0.333\ \text{pF} \\
C_{gd} &= \texttt{Cgd1}+\texttt{Cgdov1}+\texttt{Cgb1}+\texttt{Cgbov1} = 0.016\ \text{pF} \\
C_{bs} &= \texttt{Cbs1} = 0.099\ \text{pF} \\
C_L &= \texttt{CL}+\texttt{Cbd2}+\texttt{Cgd2}+\texttt{Cgdov2} = \texttt{CL}+0.031\ \text{pF}
\end{aligned}
$$

Next, we run a '.ac' simulation. This results in the frequency response shown in Fig. 4.78.

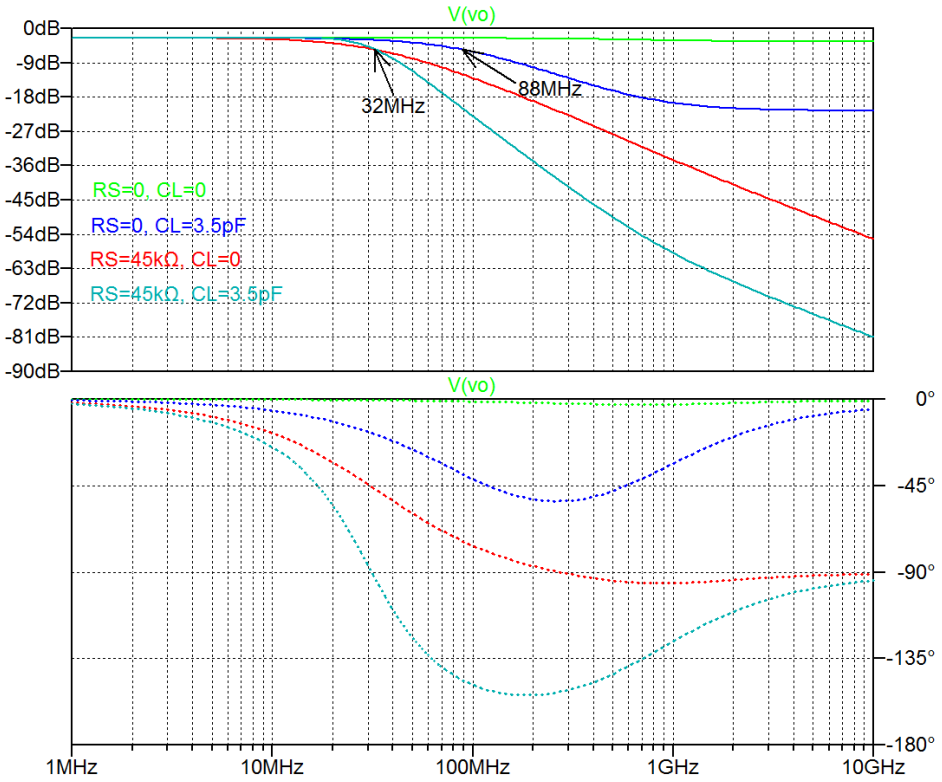


Figure 4.78: Output plot from a '.ac' simulation of the gain stage shown in Fig. 4.76.

With $R_S = 0$ and $C_L = C_{bd2} + C_{gd2}$, we notice an almost flat gain response and phase response. From Eq. (4.92), we can calculate a time constant and a corresponding frequency $f_{pB} = 1/\tau_{pB} \simeq 750$ MHz. Apparently, the pole caused by this is cancelled by a zero caused by C_{gs} and g_m. The verification of this is an exercise left for the reader as Problem 4.14.

With $R_S = 0$ and $C_L = 3.5$ pF, the simulation shows a -3 dB frequency of about 88 MHz. Equation (4.92) also results in $f_{pB} = 88$ MHz, so there is good agreement between simulation and calculation. However, we see from the output plot that the gain response is flat for higher frequencies and the phase response returns to 0°, again an indication of the presence of a zero.

With $R_S = 45$ kΩ and $C_L = C_{bd2} + C_{gd2}$, we use Eq. (4.97) to find $f_{pB} \simeq 570$ MHz, and from Eq. (4.91), we find $f_{pA} \simeq 37$ MHz. This indicates a dominant pole from the input node, and the simulation confirms this. The second pole reduces the -3 dB frequency only slightly compared to the value estimated from the dominant pole.

With $R_S = 45$ kΩ and $C_L = 3.5$ pF, the time constant from the output node is increased according to Eq. (4.97) which now gives $f_{pB} \simeq 23$ MHz and as the time constant from the input remains unchanged, we clearly do not have a dominant pole. The simulated response also shows a -3 dB frequency of about 32 MHz and a fall-off of the gain response with 40 dB/decade, indicating the presence of two poles. Again, both gain response and phase response indicate the presence of a zero as well.

The common-gate stage and the cascode stage. The low-frequency small-signal model for the common-gate stage is shown in Fig. 4.21. Inserting a load capacitance C_L and the transistor capacitances for the common-gate transistor, we arrive at the small-signal equivalent circuit shown in Fig. 4.79. For simplicity, the small-signal parameters for the bias transistor M_2 shown in Fig. 4.20 are included in R_L and C_L, i.e., R_L is a parallel connection of r_{ds2} and an extra load resistor, and C_L is a parallel connection of C_{gd2}, C_{bd2} and an extra load capacitor.

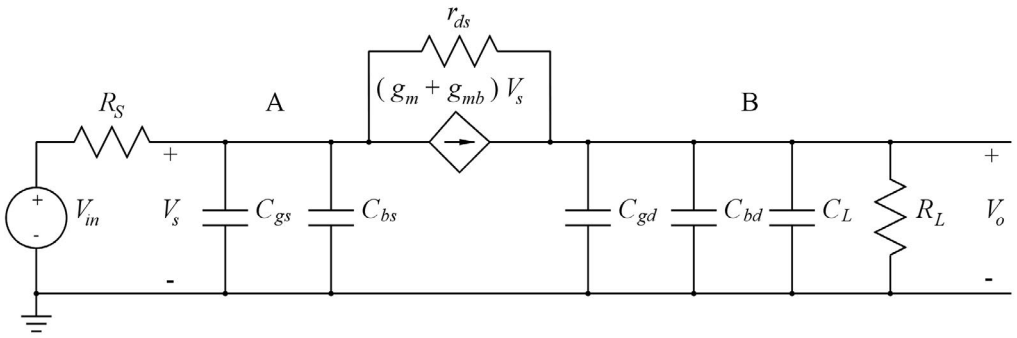


Figure 4.79: High-frequency small-signal equivalent circuit for a common-gate stage.

From the analysis in Section 4.3, we found that the input resistance is much smaller than the output resistance, so a dominant pole in a common-gate stage comes from the output node, node B in Fig. 4.79, unless R_L is small, i.e., on the order of $1/g_m$ or smaller. Using Eq. (4.26) with r_{ds2} replaced by R_L, we find

$$\tau_{pB} \simeq (R_L \parallel [r_{ds} + (g_m + g_{mb})\, r_{ds} R_S])(C_L + C_{gd} + C_{bd}) \tag{4.98}$$

For the input node, node A in Fig. 4.79, we find, using Eq. (4.23) with $(R_L \parallel r_{ds2})$ replaced by R_L

$$\tau_{pA} \simeq \left(\frac{C_{gs} + C_{bs}}{g_m + g_{mb}}\right)\left(1 + \frac{R_L}{r_{ds}}\right) \tag{4.99}$$

We will consider in more detail the common-gate stage used as a cascode as shown in Figs. 4.23 and 4.24. The small-signal equivalent circuit for the cascode is redrawn in Fig. 4.80, augmented with the signal source resistance R_S, the load impedance R_L in parallel with C_L and the transistor capacitances combined into the capacitors C_1, C_2 and C_3 where $C_1 = C_{gs1} + C_{gb1}$, $C_2 = C_{gd1}$, $C_3 = C_{bd1} + C_{bs2} + C_{gs2}$ and C_{gd2} and C_{bd2} are included in C_L. We notice that in addition to the input node V_{in}, there are three nodes in the stage, labeled A, B and C in Fig. 4.80.

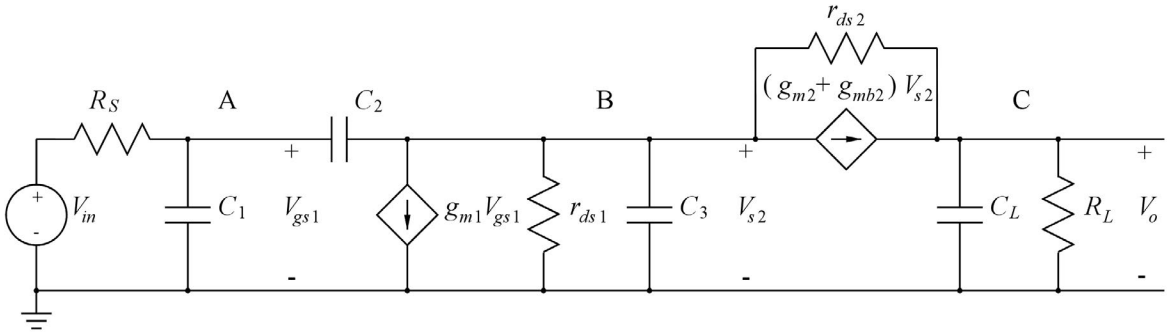


Figure 4.80: High-frequency small-signal equivalent circuit for a cascode stage.

The time constants of the three nodes are estimated as follows:

Node A: This is the input node to the common-source stage, so the time constant τ_{pA} can be estimated using Eqs. (4.80) and (4.81).

$$\tau_{pA} \simeq R_S \left[C_1 + C_2(1 - A_{v1})\right] \tag{4.100}$$

where A_{v1} is the gain of the common-source transistor M_1, i.e., $A_{v1} = -g_{m1}\, r_B$ where r_B is the small-signal resistance to ground from node B with $V_{in} = 0$. Using Eq. (4.23), we find

$$r_B \simeq r_{ds1} \parallel \left(\frac{1}{g_{m2} + g_{mb2}} + \frac{R_L}{(g_{m2} + g_{mb2}) r_{ds2}} \right) = r_{ds1} \parallel \left(\frac{1 + R_L / r_{ds2}}{g_{m2} + g_{mb2}} \right) \tag{4.101}$$

Node B: This is the output node of the common-source stage. With R_S small, the time constant τ_{pB} can be estimated using Eq. (4.82) with $(r_{ds} \parallel R_D \parallel R_L) = r_B$ and $C_{gd} + C_{bd} + C_L = C_2 + C_3$. With R_S large, we can use Eq. (4.87) with $g_m = g_{m1}$, $(r_{ds} \parallel R_D \parallel R_L) = r_B$, $C_{gs} = C_1$, $C_{gd} = C_2$ and $C_L = C_3$ but the details of this calculation are left to the reader.

Node C: This is the output node of the common-gate stage. We find the time constant τ_{pC} as

$$\tau_{pC} = C_L \left(R_L \parallel r_{out}\right) \tag{4.102}$$

where r_{out} is found from

$$r_{out} \simeq (g_{m2} + g_{mb2})\, r_{ds2}\, r_{ds1} \tag{4.103}$$

Let us consider two different cases. For the first case, R_L is very large, i.e., R_L is on the same order of magnitude as r_{out} in Eq. (4.102). In this situation, we have a very high low-frequency gain. A typical example is a cascode gain stage biased by a cascode current mirror as shown in Fig. 4.26.

Since the output resistance is very high, it is likely that a dominant pole comes from the output node, giving a pole at the frequency

$$f_{pC} = \frac{1}{2\pi\tau_{pC}} = \frac{1}{2\pi C_L \left(r_{cm} \parallel r_{out}\right)} = \frac{1}{2\pi C_L \left(r_{cm} \parallel ((g_{m2} + g_{mb2}) r_{ds2}\, r_{ds1})\right)} \tag{4.104}$$

where r_{cm} is found from Eq. (4.32).

We may note that the product GBW (gain-bandwidth product) of the dominant pole frequency and the low-frequency gain is given by $\text{GBW} = g_{m1}/(2\pi C_L)$.

The resistance r_B in the intermediate node between the common-source stage and the common-gate stage is found from Eq. (4.101) to be r_{ds1} in parallel with r_{cm} divided by the gain $(g_{m2} + g_{mb2}) r_{ds2}$. With r_{cm} given by Eq. (4.32), this is a high resistance, so the gain A_{v1} in the common-source stage is high. Thus, the Miller effect may cause a large input capacitance, resulting in a pole in combination with R_S.

For the second case, R_L is much smaller. We assume that R_L is on the same order of magnitude as r_{ds2} or smaller. A typical example is a cascode gain stage where the bias current source is implemented as a single transistor without a cascode transistor. In this case, we find the low-frequency gain as

$$A_v = -g_{m1} \left(r_{out} \parallel R_L\right) \simeq -g_{m1} \left(\left[(g_{m2} + g_{mb2})\, r_{ds2}\, r_{ds1} \right] \parallel R_L \right) \simeq -g_{m1} R_L \tag{4.105}$$

This is the same low-frequency gain as from a standard common-source stage when the load resistor is significantly smaller than the output resistance.

With R_L on the same order of magnitude as r_{ds2}, the factor $(1+R_L/r_{ds2})$ in Eq. (4.101) has a value of about two. In this case, Eq. (4.101) gives a resistance r_B of about $2/(g_{m2}+g_{mb2})$, so the gain in the common-source stage is $A_{v1} \simeq -2g_{m1}/(g_{m2}+g_{mb2})$ which is a small value, on the order of one. Thus, the Miller effect is limited and the input capacitance is smaller than for the cascode stage with a very high value of R_L.

Assuming that the dominant pole comes from the output, the dominant pole frequency is

$$f_{pC} = \frac{1}{2\,\pi\tau_{pC}} \simeq \frac{1}{2\,\pi C_L R_L} \tag{4.106}$$

We may note that the product GBW (gain-bandwidth product) of the dominant pole frequency and the low-frequency gain is given by GBW $= g_{m1}/(2\pi C_L)$, i.e., we find the same gain-bandwidth product as for the cascode with a high low-frequency gain. However, with a small R_L, the cascode transistor provides the advantage of reducing the input capacitance.

Simulation examples. For illustrating the high-frequency properties of the cascode, we revisit the cascode stage shown in Fig. 4.34. This cascode stage has a bias current source implemented as a cascoded current mirror. For comparison, we also investigate the stage with just a single transistor for the bias current. Both versions of the stage are shown in Fig. 4.81 with the transistor models modified to include the capacitance parameters. Transistors M_1, M_2, M_3 and M_4 constitute a cascode stage with a cascode bias current source while transistors M_6, M_7 and M_8 constitute a cascode stage with a single transistor for the bias current source. All transistors are specified with dimensions for drain and source diffusions, a resistor R_S has been included in series with the input voltage and a load capacitor $C_L = 3$ pF (much larger than the transistor capacitances) has been connected to each of the outputs.

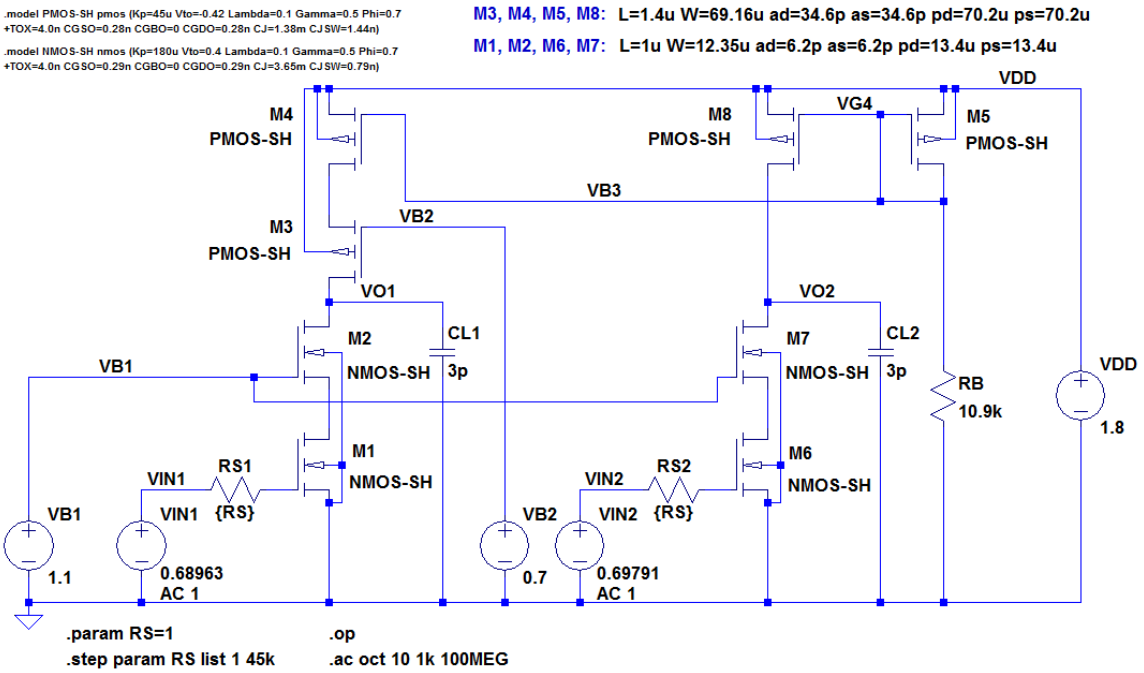


Figure 4.81: LTspice schematic for simulating the frequency response of cascode stages with different bias current sources.

A directive for stepping the resistor R_S between 0 and 45 kΩ is also included. First, we run '.dc' simulations to find bias values for the input voltages resulting in bias values of the output voltages of 0.9 V.

Next, we may run a '.op' simulation to verify the bias point and to find the small-signal parameters. We do not show all the results from the '.op' simulation but just repeat that g_{m1} and g_{m6} has values of about 0.67 mA/V as also found for the cascode stage from Fig. 4.30. With this value of g_m and a load capacitance of 3 pF, we expect a gain-bandwidth product of GBW $= g_m/(2\pi C_L) = 35$ MHz.

Next, we run a '.ac' simulation. The output from this is shown in Fig. 4.82. The green traces correspond to the cascode with a cascode bias current source (very large R_L) and the red traces are for the cascode with a single-transistor bias current source (smaller R_L). With $R_S = 0$, the gain falls off with 20 dB/decade over a wide frequency range and the low-frequency gain is much higher with the cascode bias current source than with the single transistor bias current source.

For frequencies well above the dominant pole, the two stages have almost identical characteristics and the simulated gain-bandwidth product matches the calculated value.

With $R_S = 45$ kΩ, we see that the pole from the input side modifies both the gain response and the phase response and also in this case, the two circuits have almost identical responses for frequencies well above the dominant pole.

These characteristics closely match the analysis based on Fig. 4.80.

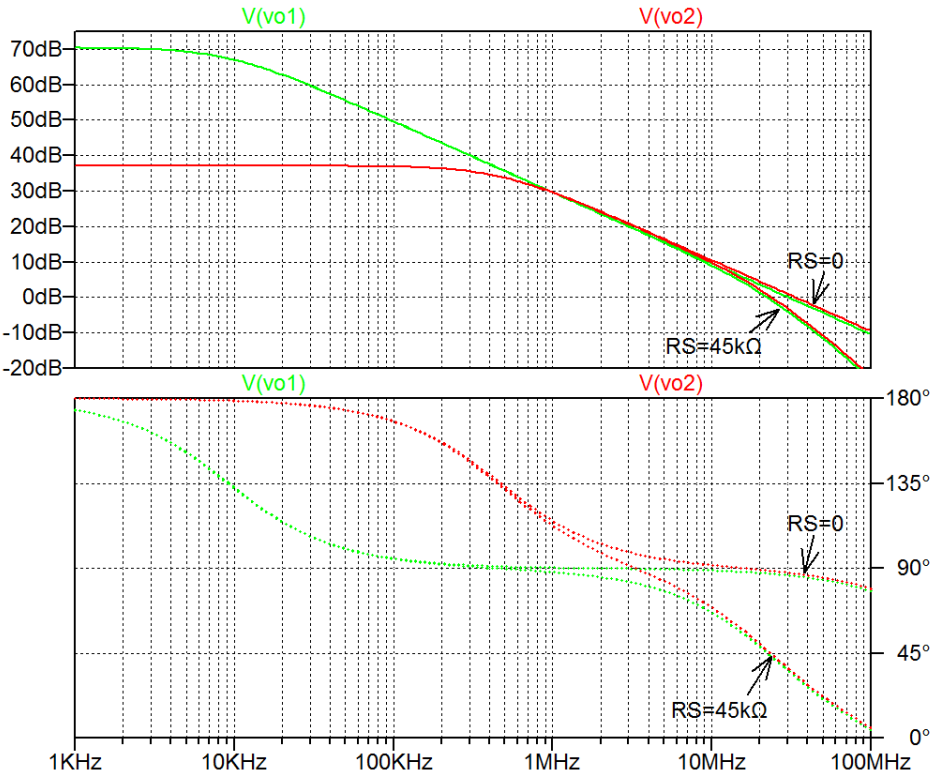


Figure 4.82: Output plot from a '.ac' simulation of the cascode stages shown in Fig. 4.81.

The differential pair. As explained in Section 4.4, the differential pair in many respects resembles a common-source stage when a differential input voltage is applied. The differential half-circuit shown in Fig. 4.52 may be augmented with the transistor capacitances, a load capacitor and a series resistor for the signal source as shown in Fig. 4.83. In the half-circuit, a load capacitance of $2C_L$ must be inserted: The load capacitor C_L connected between v_{d1} and v_{d2} in Fig. 4.51 may be split into a series connection of two capacitors, each with a value of $2C_L$, i.e., one for each half of the circuit.

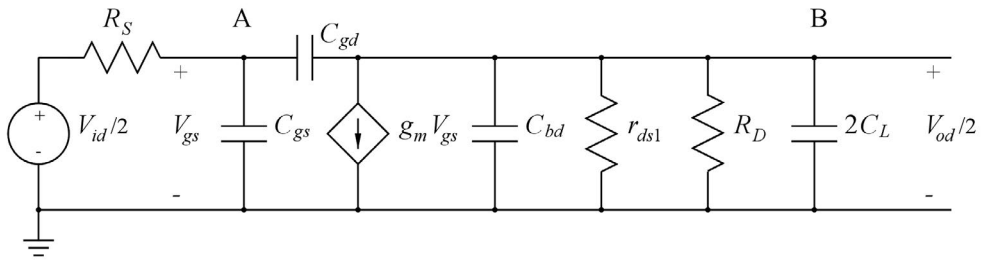


Figure 4.83: High-frequency small-signal equivalent circuit for the differential half-circuit of a differential pair with drain resistors.

This is exactly the same circuit topology as shown in Fig. 4.69 for the common-source stage, so the results found for this stage are also applicable to the differential gain in the differential pair.

For the common-mode gain, we found that the common-mode gain is zero when using a differential output, and this also applies at high frequencies for a fully symmetric circuit. When using a single-ended output, the common-mode gain is not zero as shown in Eq. (4.69). An analytical solution to the high-frequency response of the common-mode single-ended output is outside the scope of this book. For details, see for example Gray, Hurst, Lewis & Meyer (2009).

For the differential pair with a current-mirror load and a single-ended output as shown in Fig. 4.58, we often find a dominant time constant at the output since the resistance in this node may be high, see Eq. (4.70). Often the differential pair is followed by a common-source stage providing a capacitive load C_L. In this case, the time constant from the output node is

$$\tau_{out} = r_{out} C_L = (r_{ds2} \parallel r_{ds4}) C_L \tag{4.107}$$

where the transistor numbering refers to Fig. 4.58. In Eq. (4.107), the transistor capacitances C_{gd2}, C_{bd2}, C_{gd4} and C_{bd4} are included in C_L, and it is assumed that the signal source resistance R_S is small.

The node connecting drain of M_1, drain and gate of M_3 and gate of M_4 has a small-signal resistance to ground of approximately $1/g_{m3}$ since M_3 has gate and drain connected, so the time constant τ from this node is approximately

$$\tau = (1/g_{m3})(C_{gs3} + C_{gs4} + C_{bd3} + C_{bd1} + C_{gd4}(1 - A_{v4})) \tag{4.108}$$

where A_{v4} is the small-signal voltage gain from gate to drain of M_4 at frequencies of about $1/(2\pi\tau_A)$. With $1/g_{m3} \ll r_{out}$ and the capacitances in Eq. (4.108) also often being smaller than the input capacitance of a following common-source stage, we have a dominant pole from the output node. However, the current mirror does provide an addition pole, and it also provides a zero. For an analysis of this, the reader is referred to Gray, Hurst, Lewis & Meyer (2009).

In the discussion above, it was assumed that the signal source resistance of the differential input voltage was small enough that it did not cause a large time constant from the input terminals to the differential pair. The input capacitance for each of the two inputs can be determined in the same way as the input capacitance for a common-source stage, i.e., it is the sum of the gate-source capacitance of the differential input transistor and the Miller capacitance caused by the gate-drain capacitance and the gain in the input transistor. Note that the gain in transistor M_1 in Fig. 4.58 is fairly small, about $-g_{m1}/g_{m3}$ since the load resistance at the drain of M_1 is the rather small input resistance of the current mirror M_3 - M_4. The gain in M_2 is much larger since the load resistance of M_2 is large.

Simulation example. We illustrate the frequency response of the differential stage by a simulation of the stage with a PMOS current-mirror load shown in Fig. 4.59. In Fig. 4.84, this stage is redrawn with the transistor capacitance model parameters included and with an added load capacitance $C_L = 3$ pF which is much larger than the internal transistor capacitances.

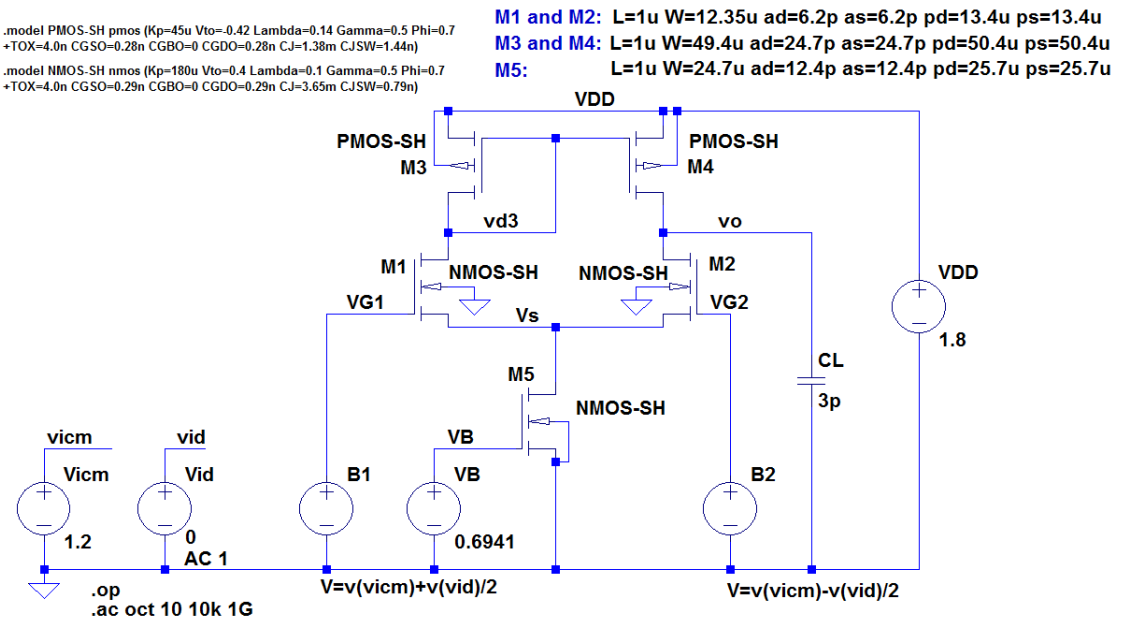


Figure 4.84: LTspice schematic for the differential gain stage with an NMOS differential pair, a PMOS current-mirror load and a load capacitance of 3 pF.

From the simulation results shown in Fig. 4.61, we know that $g_{m1} = 0.689$ mA/V, so with $C_L = 3$ pF, we expect a gain-bandwidth product of GBW $= g_{m1}/(2\pi C_L) = 36.6$ MHz, assuming that the dominant pole is caused by C_L and the output resistance of the gain stage.

Figure 4.85 shows the plot of the output voltage resulting from a '.ac' simulation. From the plot, we find a low-frequency gain of 29.8 dB $\sim$ 30.9 V/V and a -3 dB bandwidth of 1.1 MHz, giving a gain-

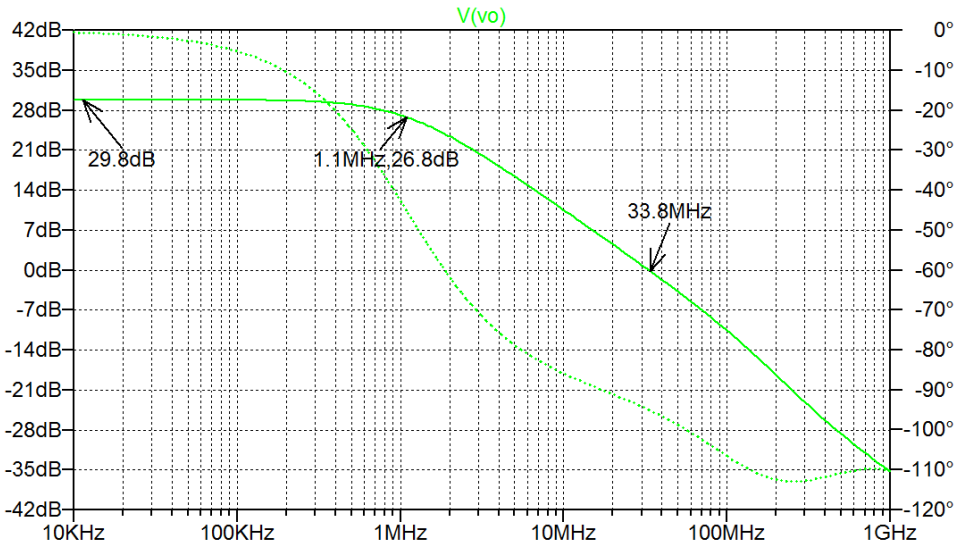


Figure 4.85: Output plot from a '.ac' simulation of the cascode stages shown in Fig. 4.84.

bandwidth product of 34 MHz. We can also read GBW as the unity-gain frequency to be 33.8 MHz. These values match the calculated value of 36.6 MHz, considering that transistor capacitances on the order of 100 fF have been neglected.

From Fig. 4.85, we also see that the gain response falls off by 20 dB/decade for frequencies below GBW but for frequencies above GBW, the slope increases over a range of frequencies. Also the phase response shows a phase shift in excess of 90° for frequencies above 20 MHz, and for frequencies above approximately 200 MHz, the phase shift is reduced again. This is an indication that a pole-zero pair from the current-mirror active load influences the frequency response at high frequencies (Gray, Hurst, Lewis & Meyer 2009).

References

Babanezhad, JN. 1988, 'A Rail-to-Rail CMOS Op Amp', *IEEE J. Solid-State Circuits*, vol. 23, no. 6, pp. 1414-1417, December 1988.

Bruun, E. 2020, *CMOS Integrated Circuit Simulation with LTspice*, Third Edition, bookboon.
Available from: http://bookboon.com/en/cmos-integrated-circuit-simulation-with-ltspice-ebook

Chan Carusone, T., Johns, D. & Martin, K. 2012, *Analog Integrated Circuit Design*, Second Edition, John Wiley & Sons, Inc., Hoboken, USA.

Gray, PE. & Searle, CL. 1969, *Electronic Principles, Physics, Models, and Circuits*, John Wiley & Sons, New York, USA.

Gray, PR., Hurst, PJ., Lewis, SH. & Meyer, RG. 2009, *Analysis and Design of Analog Integrated Circuits*, Fifth Edition, John Wiley & Sons, New York, USA.

Miller, JM. 1919, 'Dependence of the input impedance of a Three-Electrode Vacuum Tube upon the Load in the Plate Circuit', *Scientific Papers of the Bureau of Standards*, vol. 15, paper no. 351, pp. 367-385, November 1919.

Sedra, AS. & Smith, KC. 2016, *Microelectronic Circuits*, International Seventh Edition, Oxford University Press, New York, USA.

Multiple-choice test

1. Complete the following statements by selecting the appropriate continuation from the table below.

A:	The low-frequency voltage gain in a common-source stage is ...
B:	The magnitude of the low-frequency voltage gain in a common-drain stage is ...
C:	The low-frequency voltage gain in a common-gate stage is ...
D:	The maximum output voltage from an NMOS common-source stage with an active load is ...
E:	The output resistance of a common-drain stage where the gain transistor has source and bulk connected is ...
F:	The output resistance of a common-source stage biased by an active load is ...
G:	The maximum output voltage from an NMOS common-drain stage is ...
H:	A source follower is the same as ...
I:	A folded cascode is a combination of ...
J:	The output resistance for a telescopic cascode biased by an ideal current source is ...

Continuation:

1:	less than the positive supply voltage by approximately $\lvert V_{\text{eff}}\rvert$.
2:	on the order of $1/g_m$.
3:	a common-source stage followed by another common-source stage.
4:	half of the supply voltage.
5:	always positive.
6:	on the order of $g_m r_{ds}^2$.
7:	always negative.
8:	smaller than one.
9:	larger than one.
10:	less than the positive supply voltage by approximately $\lvert V_{GS}\rvert$.
11:	on the order of r_{ds}.
12:	two common-source stages.
13:	a common-source stage followed by a common-drain stage.
14:	infinite.
15:	a common-drain stage.
16:	a common-source stage and a common-gate stage using identical transistor types (NMOS/NMOS or PMOS/PMOS).
17:	a common-source stage and a common-gate stage using different transistor types (NMOS/PMOS or PMOS/NMOS).
18:	an inverting amplifier.
19:	the negative supply voltage.
20:	zero.

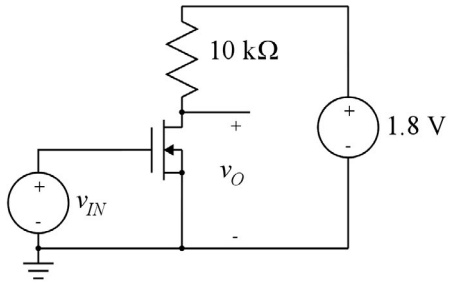
2. Complete the following statements by selecting the appropriate continuation from the table below.

A:	The common-mode input voltage to a differential pair is ...
B:	The differential input voltage to a differential pair is ...
C:	A differential pair with a current mirror as an active load provides ...
D:	For an ideal differential stage, the output voltage depends ...
E:	The dominant pole for a common-source stage driven by an ideal voltage source comes from ...
F:	For a common-source stage, the Miller effect causes ...
G:	For a source follower, the Miller effect causes ...
H:	For a common-source stage, the dominant pole may come from the input node when ...
I:	The gain-bandwidth product of a cascode stage with a transconductance of g_{m1} for the common-source stage and a load of $R_L\|\|(1/(sC_L))$ is ...
J:	The input capacitance of a cascode stage with a load of $R_L\|\|(1/(sC_L))$ may be reduced by ...

Continuation:

1:	the difference between the input voltages.
2:	only on the differential input voltage.
3:	the output node.
4:	an increased influence of C_{gs}.
5:	the sum of the input voltages.
6:	a reduced influence of C_{gs}.
7:	a single-ended output.
8:	the average of the input voltages.
9:	the input node.
10:	an increased influence of C_{gd}.
11:	decreasing R_L.
12:	a differential output.
13:	only on the common-mode input voltage.
14:	decreasing C_L.
15:	on both the differential input voltage and the common-mode input voltage.
16:	the common-source stage is driven from a voltage source with a large series resistance.
17:	$g_{m1}/(2\pi C_L)$.
18:	an inverting amplifier.
19:	$1/(2\pi C_L R_L)$.
20:	when the load capacitance is larger than C_{gs}.

3. For the circuit shown below, assume that the transistor has a bias overdrive voltage of 0.3 V and a bias current of 69 μA. Also assume that the channel-length modulation for the transistor can be neglected.



The bias value of the output voltage is

A: 0.69 V
B: 0.90 V
C: 1.11 V

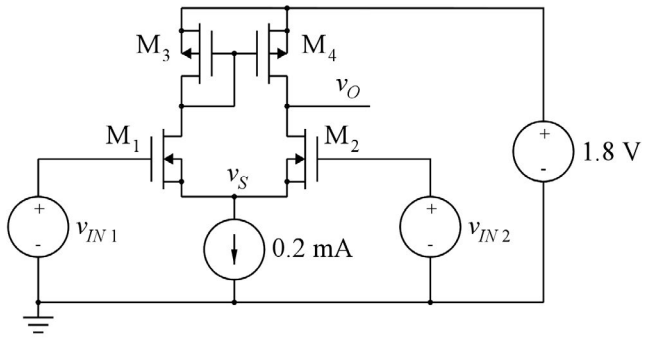
4. The small-signal gain v_o/v_{in} for the circuit above is

A: +2.3 V/V
B: −2.3 V/V
C: −4.6 V/V

5. The minimum output voltage obtainable from the circuit above with the transistor in the active region is approximately

A: 0.30 V
B: 0.42 V
C: 0.69 V

6. For the differential gain stage shown below, assume that all transistors are in the active region and have a channel-length modulation parameter $\lambda = 0.1\ \text{V}^{-1}$.



The small-signal output resistance of the differential gain stage is approximately

A: 50 kΩ
B: 100 kΩ
C: 200 kΩ

7. With all transistors in the differential gain stage shown before having an effective gate voltage of $|V_{GS} - V_t| = 0.25$ V, the small-signal differential gain is approximately

   A: 20 V/V
   B: 40 V/V
   C: 80 V/V

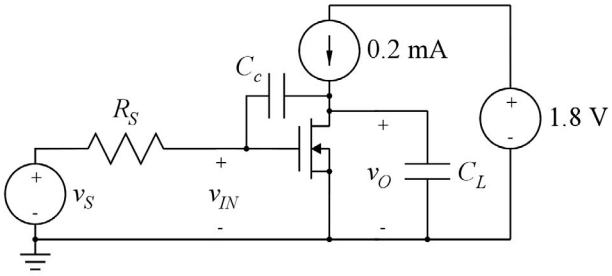
8. For the differential gain stage shown before, assume that the threshold voltage for the PMOS transistors is -0.42 V. With a bias voltage of 0.9 V for both input voltages, the bias value of the gate voltage for the PMOS transistors is

   A: 0.67 V
   B: 1.13 V
   C: 2.22 V

9. If the input voltages for the differential gain stage shown before are changed so that all of the bias current flows in M_1 - M_3, the gate voltage of the PMOS transistors is changed to approximately

   A: 0.63 V
   B: 0.77 V
   C: 1.03 V

10. For the circuit shown below, assume that the transistor is in the active region with $r_{ds} = 40$ kΩ and $g_m = 0.5$ mA/V. Also assume that all transistor capacitances are negligible compared to C_L and C_c.



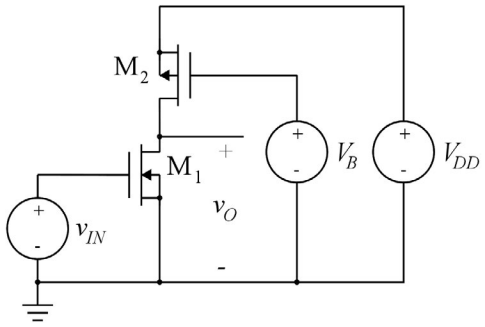
   With $C_c = 0.1$ pF, $C_L = 0.1$ pF and $R_S = 40$ kΩ, the dominant time constant
   A: originates from the output and is about 4 ns.
   B: originates from the input and is about 4 ns.
   C: originates from the input and is about 80 ns.

11. With $C_c = 0.01$ pF, $C_L = 2$ pF and $R_S = 40$ kΩ, the dominant time constant

   A: originates from the output and is about 80 ns.
   B: originates from the input and is about 0.4 ns.
   C: originates from the input and is about 80 ns.

Problems

Problem 4.1



For both transistors M_1 and M_2 in the common-source stage shown above, assume $W/L = 3$. Further, assume that the channel-length modulation can be neglected for both transistors and that $\mu_n C_{ox} = 180\ \mu A/V^2$, $\mu_p C_{ox} = 45\ \mu A/V^2$, $V_{tn} = 0.4$ V and $V_{tp} = -0.42$ V. The supply voltage is $V_{DD} = 1.8$ V and the bias voltage V_B is $V_B = 980$ mV.

Calculate the bias current in M_1 and M_2, assuming that they are both in the active region and find the bias value of the input voltage for which the transistors are in the active region.

Find the output voltage range for which the transistors are in the active region.

Problem 4.2 is a follow-up to this problem.

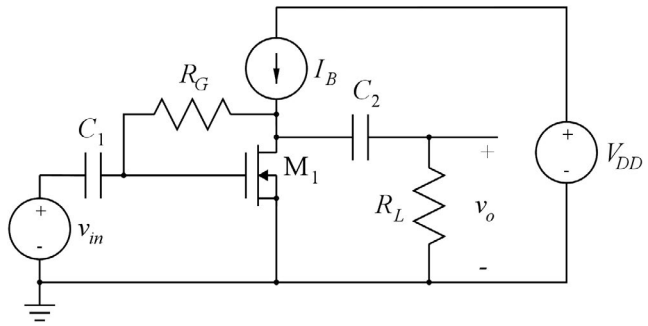
Problem 4.2

For the common-source stage from Problem 4.1, assume now that the channel-length modulation is defined by the parameters specified in Table 3.1. The channel length is $L = 1.5\ \mu m$ for both transistors.

Calculate the small-signal voltage gain $A_v = v_o/v_{in}$. You may use reasonable approximations when calculating the small-signal parameters.

Also use LTspice to find the value of $A_v = v_o/v_{in}$.

Problem 4.3



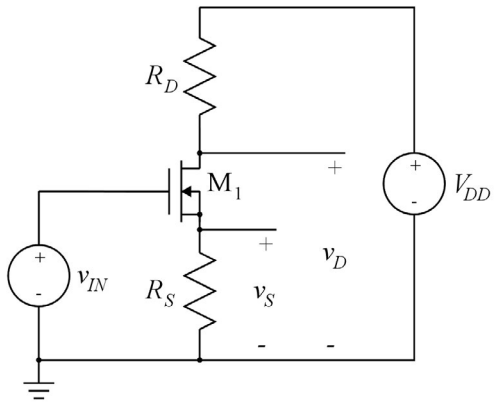
For the common-source stage shown above, we assume $R_G = 1\ \text{M}\Omega$, $I_B = 20\ \mu\text{A}$ and $R_L = 100\ \text{k}\Omega$. The capacitors C_1 and C_2 are very large and can be treated as short circuits for ac currents but open circuits for dc currents. Transistor M_1 has $V_t = 0.4\ \text{V}$ and $\mu_n C_{ox}(W/L) = 0.6\ \text{mA/V}^2$. The channel-length modulation can be neglected.

Find the bias voltages V_{GS1} and V_{DS1}.

Find an expression for the small-signal gain v_o/v_{in} for frequencies where C_1 and C_2 are ac short circuits and the transistor capacitances can be neglected.

Calculate the numerical value of the small-signal gain.

Problem 4.4



Shown above is a gain stage with two outputs, v_D and v_S. The resistors R_D and R_S have the values $R_D = 10\ \text{k}\Omega$ and $R_S = 5\ \text{k}\Omega$. The supply voltage is $V_{DD} = 1.8\ \text{V}$. The transistor has $V_t = 0.4\ \text{V}$ and $\mu_n C_{ox} = 180\ \mu\text{A/V}^2$. The channel-length modulation and the body effect can be neglected.

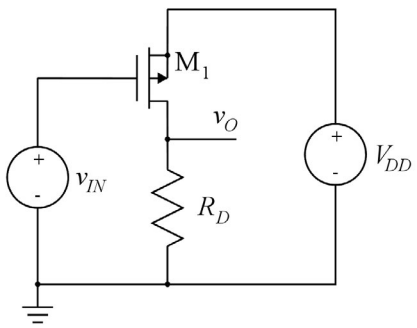
Design the transistor to have a transconductance of 0.6 mA/V and to give a bias value of 1.2 V for V_D. Use a channel length of $L = 3\ \mu\text{m}$.

Calculate the bias values of v_S and v_{IN}.

Draw a small-signal equivalent circuit and find expressions for the small-signal gains v_d/v_{in} and v_s/v_{in}.

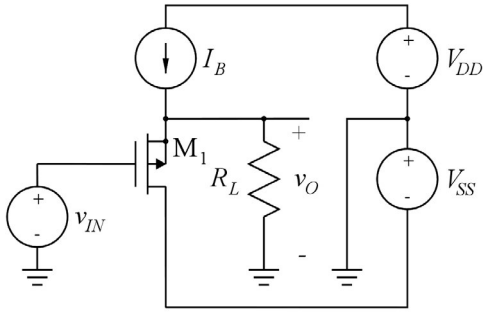
Find numerical values for the small-signal gains.

Problem 4.5



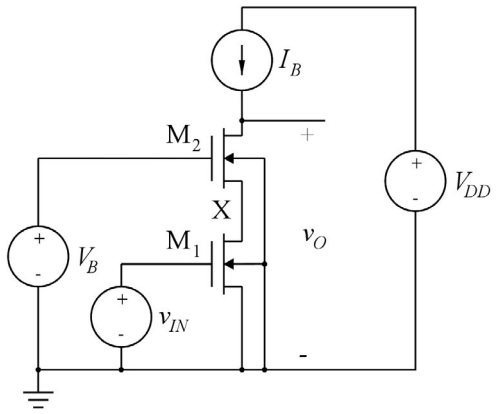
Design the common-source stage shown above to have a low-frequency small-signal gain of 17 dB. For the transistor, use $V_t = -0.42$ V, $\mu_p C_{ox} = 45$ μA/V^2 and $\lambda = 0$. Use a channel length of 2 μm. The supply voltage is $V_{DD} = 1.8$ V, and you should design for a bias value of the output voltage of $V_{DD}/2$ and a bias current in M_1 of 50 μA.
Calculate the bias value of the input voltage for your design.

Problem 4.6



Shown above is a common-drain stage with a PMOS transistor. Assume the following transistor parameters: $L = 1$ μm, $W = 80$ μm, $\mu_p C_{ox} = 45$ μA/V^2, $V_t = -0.42$ V and $\lambda = 0.14$ V^{-1}. The bias current I_B has a value of 125 μA and the supply voltages are $V_{DD} = V_{SS} = 0.9$ V.
Find the bias value of v_{IN} which gives a bias value of 0 V for v_O.
Find the small-signal open-circuit voltage gain $A_{v\text{oc}}$ and output resistance r_{out}.
Find the small-signal voltage gain $A_v = v_o/v_{in}$ with $R_L = 5$ kΩ.
Use LTspice to find the output voltage range for $R_L = 5$ kΩ, assuming that the voltage across the current source I_B must be at least 0.3 V.

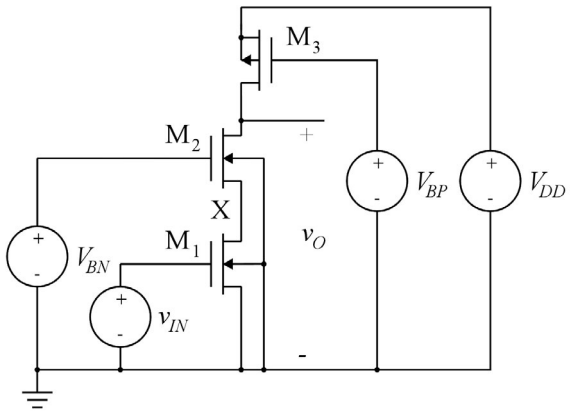
Problem 4.7



For the telescopic cascode shown above, use LTspice to find the bias value V_{IN} of the input voltage required to give an output bias voltage of 0.9 V. Assume $V_{DD} = 1.8$ V, $V_B = 1.0$ V, $I_B = 20$ μA, $L_1 = L_2 = 0.7$ μm, $W_1 = W_2 = 7$ μm and use transistor parameters as specified in Table 3.1.
Also find the small-signal gain $A_{v\text{oc}}$ and output resistance r_{out}.

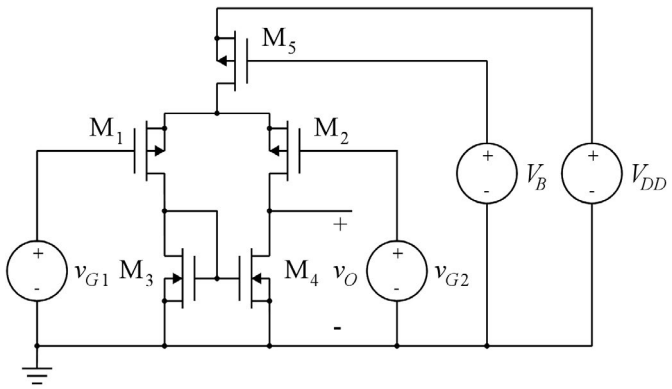
Find the small-signal resistance r_x to ground from the node X between the source of M_2 and the drain of M_1 and find the small-signal voltage gain from the input to node X.

Problem 4.8



For the telescopic cascode shown above, use LTspice to find the bias value V_{IN} of the input voltage required to give an output bias voltage of 0.9 V. Assume $V_{DD} = 1.8$ V, $V_{BN} = 1.0$ V, $V_{BP} = 1.25$ V, $L_1 = L_2 = L_3 = 0.7$ µm, $W_1 = W_2 = 7$ µm, $W_3 = 28$ µm and use transistor parameters as specified in Table 3.1. Also find the small-signal gain A_{voc} and output resistance r_{out}.
Find the small-signal resistance r_x to ground from the node X between the source of M_2 and the drain of M_1 and find the small-signal voltage gain from the input to node X.

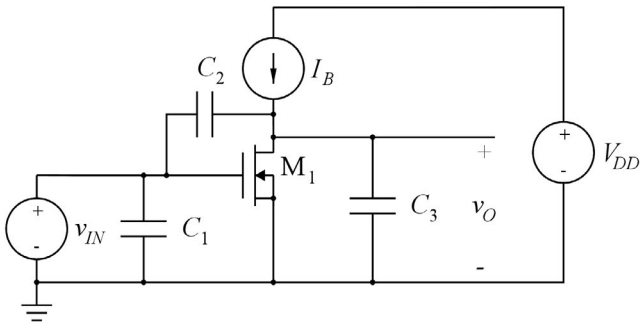
Problem 4.9



The figure above shows a PMOS differential pair. For all transistors, assume $L = 1$ µm and $W = 20$ µm and transistor parameters as specified in Table 3.1. The supply voltage is $V_{DD} = 1.8$ V and the bias voltage V_B is $V_B = 1.1$ V.
Neglecting the channel-length modulation in M_5, calculate the bias current in each transistor, assuming all transistors to be in the active region.
Calculate the small-signal parameters g_m and r_{ds} for each transistor, using reasonable approximations and assuming all transistors to be in the active region.

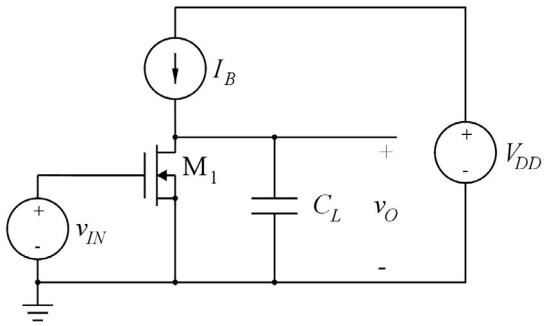
Calculate the differential small-signal gain $A_d = v_o/(v_{g1} - v_{g2})$.
Check your results with LTspice and explain differences between simulated and calculated results.
Use LTspice to find the input common-mode voltage range where all transistors are in the active region.

Problem 4.10



For the common-source stage shown above, we assume that the transistor has the small-signal parameters g_m and r_{ds} and that all transistor capacitances can be neglected compared to the capacitors C_1, C_2 and C_3. Draw a small-signal equivalent circuit for the stage and use a node equation to find the small-signal transfer function V_o/V_{in}. Find expressions for poles and zeros in the transfer function.

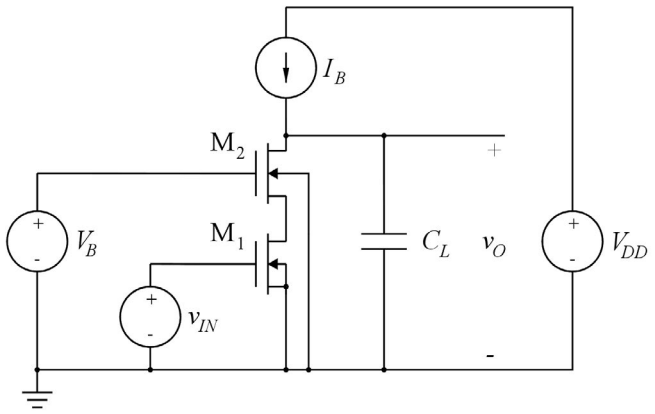
Problem 4.11



The figure above shows a common-source stage with a capacitive load $C_L = 5$ pF which is much larger than the transistor capacitances. The transistor is biased in the active region and it has the small-signal parameters $g_{m1} = 0.20$ mA/V and $r_{ds1} = 250$ kΩ.
Calculate the output resistance, the small-signal voltage gain at low frequencies, the -3 dB bandwidth and the gain-bandwidth product of the amplifier stage.

Problems 4.12 and 4.13 are follow-ups to this problem.

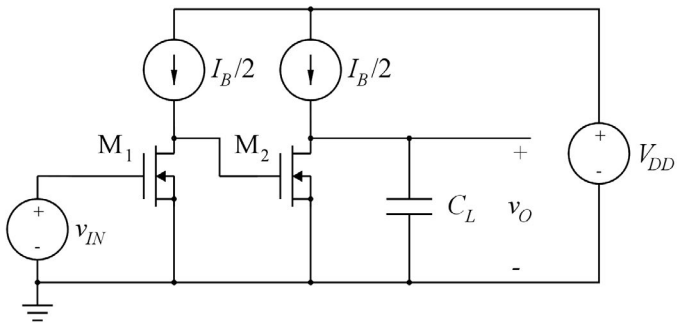
Problem 4.12



In order to increase the gain of the amplifier from Problem 4.11, a cascode transistor M_2 is inserted as shown above. The cascode transistor is assumed to be identical to M_1 and it has a bulk transconductance of $g_{mb2} = 0.04$ mA/V. Both M_1 and M_2 are biased in the active region with an unchanged bias current I_B.

Calculate the output resistance, the small-signal voltage gain at low frequencies, the -3 dB bandwidth and the gain-bandwidth product of the cascode amplifier.

Problem 4.13

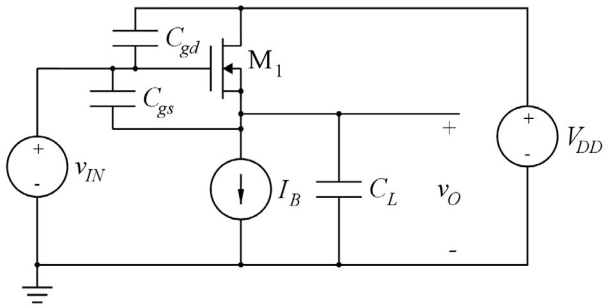


An alternative to the cascode amplifier from Problem 4.12 is a two-stage amplifier using common-source stages as shown above. The two transistors are assumed to be identical and identical to the transistors in Problems 4.11 and 4.12 and they are both biased in the active region, however with a bias current which is half of the bias current used in Problems 4.11 and 4.12. The reduction in bias current affects the small-signal parameters of the transistors.

Calculate the new values of g_m and r_{ds} using the Shichman-Hodges model with reasonable approximations.

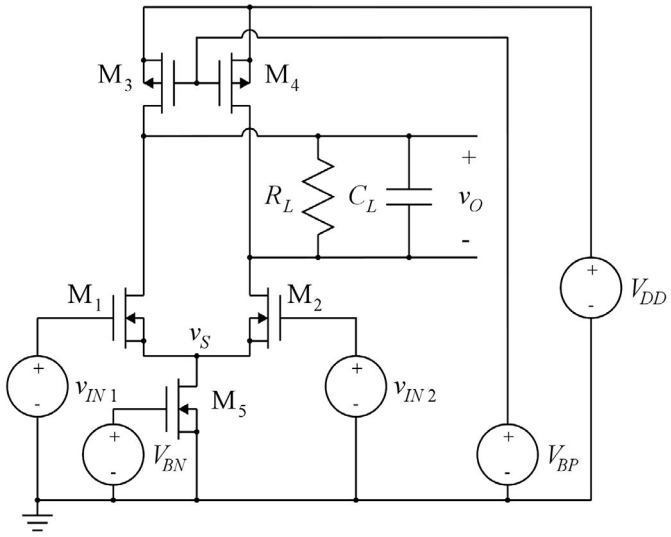
Calculate the output resistance, the small-signal voltage gain at low frequencies, the -3 dB bandwidth and the gain-bandwidth product of the two-stage common-source amplifier.

Problem 4.14



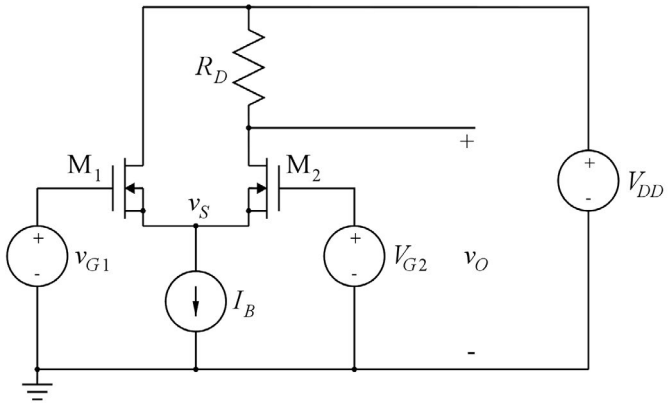
For the common-drain stage shown above, we assume that the transistor has the transconductance g_m and that the channel-length modulation and body effect can be neglected. Also, only the transistor capacitances shown in the figure above and the load capacitance C_L need to be considered.
Draw a small-signal equivalent circuit for the stage and use a node equation to find the small-signal transfer function V_o/V_{in}. Find expressions for poles and zeros in the transfer function.

Problem 4.15



The figure above shows a differential pair with a differential output v_O. The load resistor R_L and capacitor C_L and the supply voltages have the following values: $R_L = 20$ kΩ, $C_L = 2$ pF, $V_{DD} = 1.8$ V, $V_{BN} = 0.7$ V and $V_{BP} = 1.08$ V.
All transistors have a channel length of 1 μm, are in the active region, and have transistor parameters as specified in Table 3.1, except the channel-length modulation can be neglected.
Draw a small-signal differential half-circuit for the differential pair and find the small-signal parameters such that a gain-bandwidth product of 25 MHz for the differential gain $V_o/(V_{in1} - V_{in2})$ is obtained. Calculate the low-frequency small-signal gain.
Find the required bias current for all transistors using an effective gate voltage of 0.3 V. Find the channel widths of all transistors.

Problem 4.16



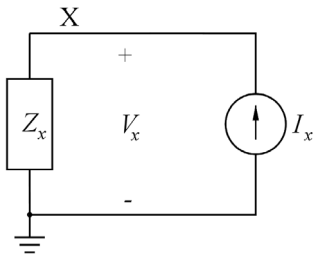
The circuit shown above with the gate of M_2 connected to a dc bias voltage V_{G2} can be used as an amplifier with a single-ended input $v_{IN} = v_{G1}$ and a single-ended output v_O.

Neglect the channel-length modulation and assume that both transistors are in the active region and have a transconductance of g_m.

Find an expression for the low-frequency small-signal gain v_o/v_{in} and the low-frequency small-signal gain v_s/v_{in} where v_S is the source voltage of M_1.

Find an expression for the input capacitance of the amplifier, assuming that the only capacitances to be considered are C_{gs} and C_{gd} for the transistors.

Problem 4.17



The small-signal impedance to ground from a node X can be found by applying an ac current I_x to the node and measuring the voltage V_x while all other independent sources in the circuit are reset, see the figure above. If Z_x can be approximated by a parallel combination of a resistor R_x and a capacitor C_x, R_x and C_x can be found from $R_x = |V_x|^2/(I_x \mathrm{Re}(V_x))$ and $C_x = -(I_x \mathrm{Im}(V_x))/(2\pi f |V_x|^2)$ where f is the frequency (Bruun 2020, Tutorial 2.6).

With the ac value of I_x defined to be 1 and the dc value defined to be 0 in LTspice, we can find R_x and C_x from a plot of the expressions `Rx=Abs(v(Vx))**2/Re(v(Vx))` and `Cx=-Im(v(Vx))/(2*pi*frequency*Abs(v(Vx)**2))` after having run a '.ac' simulation over a suitable frequency range.

Use this method to find the output resistance and output capacitance of the common-source stage shown in Fig. 4.72 with $C_f = 0$ and $R_S = 0$. Insert a decoupling capacitor in parallel with R_B to ensure that V_B can be treated as a dc voltage for the frequencies of interest. Compare to the values found using the small-signal parameters from a '.op' simulation.

Repeat for the cascode stages shown in Fig. 4.81 with $R_S = 0$.

Chapter 5 – Multistage Amplifiers

A CMOS amplifier with only one stage to provide gain cannot provide a high gain. In order to achieve a high gain, more gain stages must be applied. In this chapter, we consider multistage amplifiers capable of giving a large voltage gain and we focus on configurations which are useful for operational amplifiers (opamps) with a differential input and a single-ended output.

After having studied the chapter, you should be able to

- explain the operation of the folded-cascode opamp and calculate the small-signal low-frequency gain.
- find the frequency of the dominant pole in a folded-cascode opamp.
- explain the operation of the two-stage opamp and calculate the small-signal low-frequency gain.
- design the two-stage opamp to have bias conditions ensuring symmetry in the differential input stage.
- find the frequency of the dominant pole in a two-stage opamp with the dominant pole originating from the input to the second stage.
- simulate the frequency response and the transient response of the two-stage opamp in a feedback configuration.
- use transient simulation to illustrate stability problems in an opamp configuration with feedback.

In Chapter 4, we examined the basic gain stages which can be configured from MOS transistors. We found that the low-frequency small-signal gain in a single-transistor stage is often calculated as the product of a transistor transconductance and small-signal output resistance. While the transconductance is a parameter which is fairly well controlled in the manufacturing process, the output resistance is less well controlled, so the gain in a single-transistor stage will show a large variation over temperature, supply voltage and process variations. Often in order to achieve precision, an amplifier with a specific gain is built from an operational amplifier and a feedback network. This is a concept which you have already seen if you have followed an introductory course in electronic circuit design.

Figure 5.1 shows a noninverting amplifier circuit built from an ideal operational amplifier and a feedback network consisting of two resistors R_1 and R_2. The operational amplifier is an amplifier with a differential input and a single-ended output. For an ideal opamp, the differential gain v_o/v_{id} is infinite, implying that with a finite value of v_o, the differential input voltage is 0. For the ideal opamp, also the input currents are 0 and the output resistance of the opamp is 0.

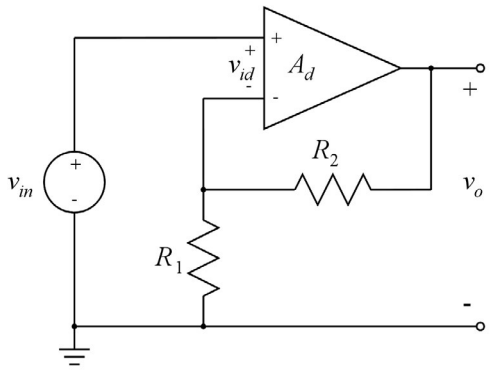


Figure 5.1: Noninverting amplifier circuit using an operational amplifier.

With $v_{id} = 0$, the voltage at the inverting input of the opamp is v_{in}, and using the voltage divider rule given by Eq. (2.24), we also find the voltage at the inverting input as $v_o R_1/(R_1 + R_2)$. This leads to the well-known expression for the gain in a noninverting amplifier built from an ideal operational amplifier:

$$v_{in} = v_o \frac{R_1}{R_1 + R_2} \Rightarrow A_v = \frac{v_o}{v_{in}} = 1 + \frac{R_2}{R_1} \tag{5.1}$$

This equation shows that with a very high gain, ideally infinite gain, in the opamp, the resulting gain is depending only on the ratio between the resistors R_1 and R_2 and this is a quantity which can be very precisely controlled over temperature variations and process variations.

If the differential gain of the opamp is not infinite but has a finite value, $A_d = v_o/v_{id}$, Eq. (5.1) must be modified to

$$v_{in} - v_{id} = v_{in} - \frac{v_o}{A_d} = v_o \frac{R_1}{R_1 + R_2} \Rightarrow A_v = \left(1 + \frac{R_2}{R_1}\right)\left(\frac{1}{1 + (1 + R_2/R_1)/A_d}\right) \tag{5.2}$$

From Eq. (5.2), we see that if A_d is much larger than the desired gain of $(1 + R_2/R_1)$, the error made in the gain calculation by assuming the opamp to be ideal is small. Hence, an amplifier with a differential input and a large gain is a very useful building block. We saw in Chapter 4 that a differential pair with a single-ended output as shown in Fig. 4.58 has a gain of $g_{m1}\,(r_{ds2} \parallel r_{ds4})$, see Eq. (4.71). This is about half of the intrinsic gain in a transistor, and for most applications, this is not sufficient for an amplifier to be used as an approximation to an ideal opamp.

Another problem with the differential pair shown in Fig. 4.58 is the limited range of voltage swing at the output. With a common-mode input voltage of V_{ICM}, the output voltage cannot go below $V_{ICM} - V_{GS2} + V_{DS\text{sat}2} = V_{ICM} - V_{t2}$ which may well be significantly higher than the negative supply rail.

Therefore, an amplifier to be used as an ideal opamp must have more than one high-gain stage. We have already seen examples of two-stage amplifiers in Chapter 4. One example investigated in some detail is the cascode stage which is a common-source stage followed by a common-gate stage. This can yield a high low-frequency gain as analyzed in Chapter 4.3. However, in the configuration shown in Fig. 4.26, it has a single-ended input, so the common-source stage must be replaced by a differential pair.

5.1 Cascode opamps

Using cascoding to obtain a higher gain in a differential pair, we may devise the circuit shown in Fig. 5.2. Transistor M_6 is a cascode transistor for M_2 and transistor M_7 is a cascode transistor for M_4. Thus, the small-signal output resistance is very high. Using Eq. (4.27) while neglecting the bulk effect, we can find the output resistance as

$$r_{out} \simeq (g_{m6}\, r_{ds6}\, r_{ds2}) \parallel (g_{m7}\, r_{ds7}\, r_{ds4}) \tag{5.3}$$

and the differential gain $A_d = v_o/(v_{in1} - v_{in2}) \simeq g_{m1}\, r_{out}$ is very high.

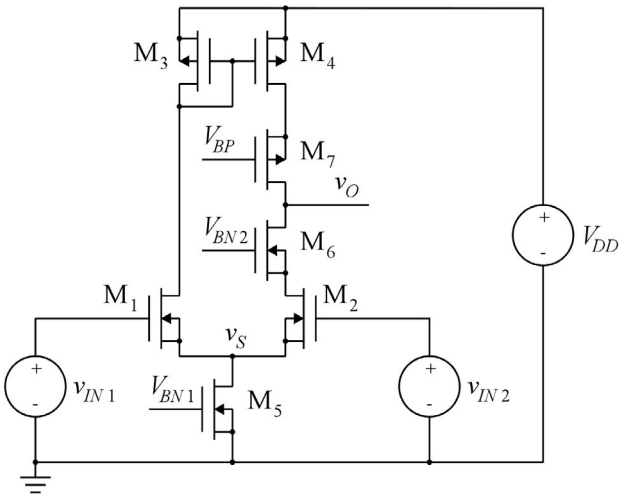


Figure 5.2: A differential gain stage with cascode transistors to increase the gain.

However, there are some serious problems with the circuit shown in Fig. 5.2. First, it is not symmetric. This could be fixed by inserting additional cascode transistors for M_1 and M_3 in the same way as for M_2 and M_4. Second, the output voltage range is limited even more than for the differential pair without cascode transistors. The output voltage cannot go below $V_{BN2} - V_{t6}$ and V_{BN2} must be selected large to allow a large common-mode input voltage but small to allow a large output voltage swing. A solution could be to let V_{BN2} depend on the common-mode input voltage but this complicates the design to a considerable extent (Allen & Holberg 2012). Another much more common solution is to use PMOS transistors for the cascoding of M_1 and M_2. In this way, the cascode outputs are folded down towards the negative supply rail in the same way as shown for the folded cascode in Fig. 4.27. Using this approach and placing the current-mirror active load at the output of the cascode transistors, we arrive at the folded-cascode opamp shown in Fig. 5.3.

For this, we find the small-signal output resistance to be

$$r_{out} \simeq (g_{m9}\, r_{ds9}\, r_{ds11}) \parallel (g_{m7}\, r_{ds7}\, (r_{ds1} \parallel r_{ds3})) \tag{5.4}$$

and the differential gain $A_d = v_o/(v_{in2} - v_{in1}) \simeq g_{m1}\, r_{out}$ is again very high.

With a very high output resistance, the dominant pole in the frequency response of the folded-cascode opamp normally originates from the output node, giving rise to a -3 dB bandwidth of $1/(2\pi\, r_{out}\, C_L)$ where C_L is the load capacitance at the output, including the parasitic transistor capacitances.

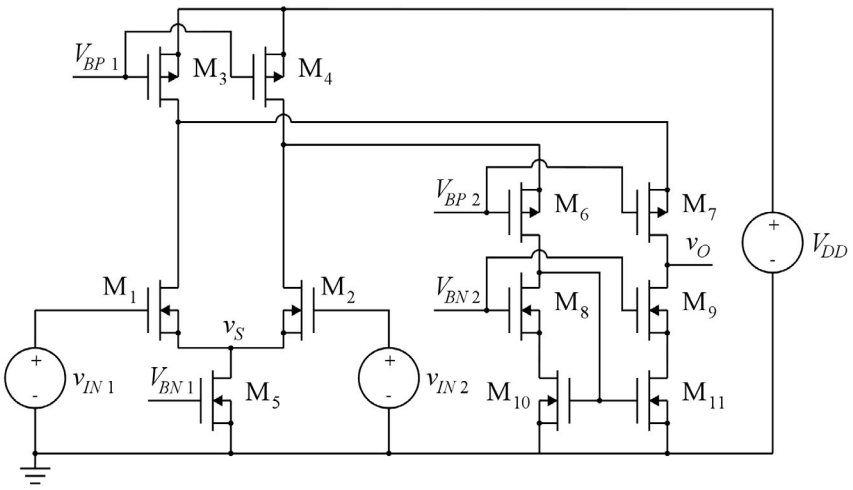


Figure 5.3: A folded-cascode opamp.

The folded-cascode opamp is one of the frequently used standard configurations for CMOS opamps. However, a detailed analysis of the folded-cascode opamp is beyond the scope of this book. The reader is referred to Chan Carusone, Johns & Martin (2012) or Allen & Holberg (2012).

5.2 The two-stage opamp

An alternative to the use of a common-source stage or a differential stage followed by a common-gate stage is the use of a cascade of common-source stages. In Problem 4.13, we saw an example of a cascade of two common-source stages. However, the configuration shown in Problem 4.13 has a single-ended input, so for an opamp requiring a differential input, we must replace the first common-source stage by a differential pair as shown in Fig. 5.4. Here the second stage is a PMOS common-source stage because the input voltage range for this matches the output voltage range from the differential pair, also when the value of the supply voltage V_{DD} is changed.

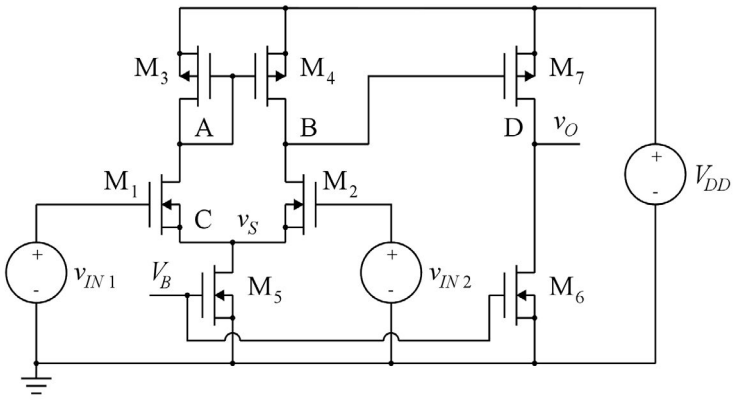


Figure 5.4: A two-stage opamp.

Scaling of bias currents and transistor geometries. In order to ensure symmetry in the input stage, the bias voltage of node B in Fig. 5.4 should be the same as the bias voltage of node A. This implies that M_3, M_4 and M_7 have identical gate-source bias voltages and full symmetry is obtained with an output bias voltage (node D) which is equal to the bias voltage at node B. This is not necessarily mid-rail, so in order to achieve full symmetry with no systematic offset in the input stage, we must accept an output bias voltage different from the mid-rail voltage. With identical drain-source bias voltages of M_3, M_4 and M_7 and assuming that these transistors have the same value of the channel-length modulation parameter λ, the Shichman-Hodges relation (3.13) for the drain current shows that the bias currents in M_3, M_4 and M_7 must have the same current density $I_D/(W/L)$, i.e.,

$$\frac{I_{D3}}{(W/L)_3} = \frac{I_{D4}}{(W/L)_4} = \frac{I_{D7}}{(W/L)_7} \tag{5.5}$$

From Fig. 5.4, it can also be seen that $I_{D5} = 2I_{D3}$ and $I_{D6} = I_{D7}$. With the same gate bias voltage for M_5 and M_6, we find (approximately)

$$\frac{I_{D5}}{(W/L)_5} = \frac{I_{D6}}{(W/L)_6} \Rightarrow \frac{(W/L)_6}{(W/L)_5} = \frac{I_{D6}}{I_{D5}} = \frac{I_{D7}}{2I_{D3}} = \frac{(W/L)_7}{2(W/L)_3} \tag{5.6}$$

In Eq. (5.6), we have neglected the difference in channel-length modulation between M_5 and M_6. This will cause a slightly larger bias current in M_6 than found from Eq. (5.6), pulling the bias value of the output voltage towards the negative supply rail. Nevertheless, Eqs. (5.5) and (5.6) are the equations to use for scaling the transistor dimensions once the bias currents of the transistors have been designed.

Small-signal gain, output resistance and frequency response. The low-frequency small-signal gain is found as the product of the gain in the differential stage and the gain in the common-source stage. Using the results from Chapter 4, we find the low-frequency small-signal gain to be

$$A_d = \frac{v_o}{v_{id}} = \frac{v_o}{v_{g2} - v_{g1}} \simeq g_{m1}\,(r_{ds2} \parallel r_{ds4})\,g_{m7}\,(r_{ds6} \parallel r_{ds7}) \tag{5.7}$$

We also see that the output resistance is equal to the output resistance of the common-source stage, i.e.,

$$r_{out} = r_{ds6} \parallel r_{ds7} \tag{5.8}$$

Examining the nodes in Fig. 5.4, we notice that the high-resistance nodes are node B and node D. The capacitance at node B is often fairly large because of the Miller effect from transistor M_7, causing the gate-drain capacitance to be multiplied by the gain in the common-source stage. So unless a large load capacitance is connected to the output, the dominant pole comes from node B. The capacitance at node B can be found using Eq. (4.80) with $A_v = -g_{m7}\,(r_{ds6} \parallel r_{ds7})$ and C_{gs} replaced by a capacitance C_1 which is the sum of C_{gs7} and the parasitic capacitances to ground from M_2 and and M_4. This results in a dominant pole at the frequency

$$f_{p1} \simeq \frac{1}{2\pi\,(r_{ds2} \parallel r_{ds4})(C_1 + C_{gd7}(1 + g_{m7}\,(r_{ds6} \parallel r_{ds7})))} \tag{5.9}$$

Often, zeros and non-dominant poles in the opamp are located at frequencies which are not very much higher than f_{p1} and as we will see in the following, this may cause problems when using the opamp in a feedback configuration like shown in Fig. 5.1. The dominant pole may be separated from the non-dominant poles by deliberately inserting an extra capacitance C_c between gate and drain of M_7. This creates a large Miller capacitance $C_M \gg C_1$ and causes a pole splitting as explained in Chapter 4.5. If $C_c \gg C_{gd7}$ and $g_{m7}\,(r_{ds6} \parallel r_{ds7}) \gg 1$, we can simplify Eq. (5.9) to

$$f_{p1} \simeq \frac{1}{2\pi\,(r_{ds2} \parallel r_{ds4})C_c\,(g_{m7}\,(r_{ds6} \parallel r_{ds7}))} \tag{5.10}$$

and combining this with Eq. (5.7), we find the gain-bandwidth product GBW $\simeq g_{m1}/(2\pi C_c)$.

Simulation example. As an example of a two-stage opamp, we may combine the differential pair shown in Figs. 4.59 and 4.84 with a common-source stage using a PMOS common-source transistor. The resulting LTspice schematic is shown in Fig. 5.5. For M_1 - M_5, we are using the same dimensions as in Figs. 4.59 and 4.84. For M_7 and M_6, we use $L = 1$ µm as for the other transistors, and we select (somewhat arbitrarily) $W_6 = 2W_5$ and $W_7 = 4W_3$. This ensures that Eqs. (5.5) and (5.6) are fulfilled.

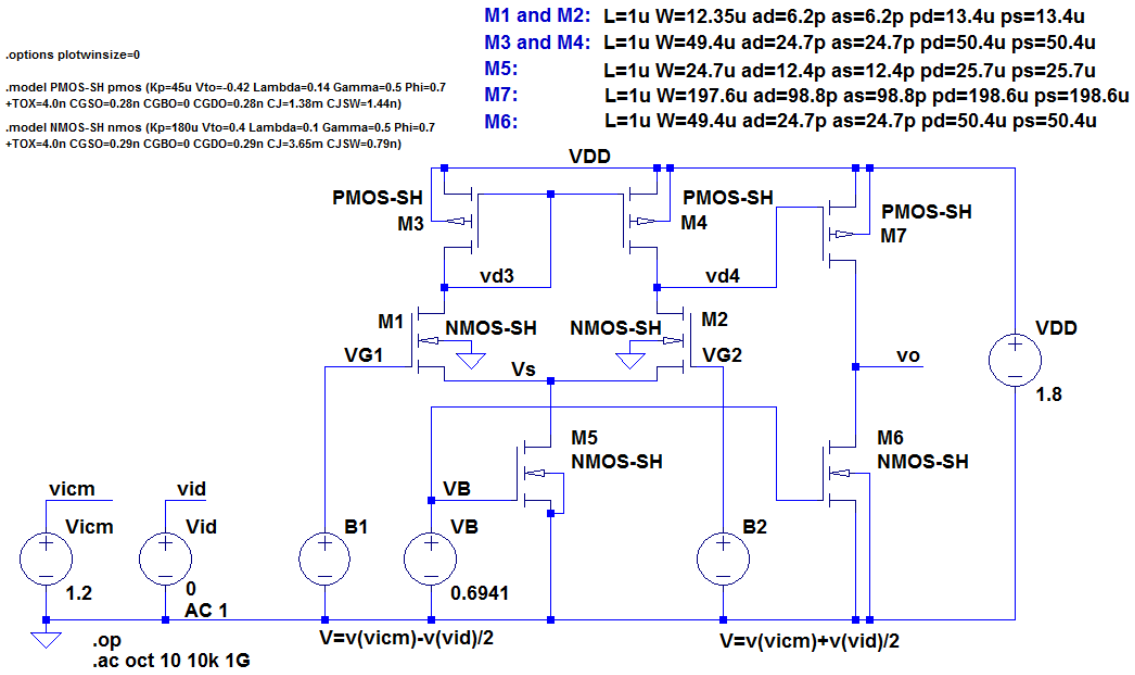


Figure 5.5: LTspice schematic for simulating a two-stage opamp.

As usual, we run a '.op' simulation to verify the bias point. From this, we find a bias value of 0.79679 V for the output voltage which is somewhat lower than the voltage of 1.09384 V at the drain of M_4 because the channel-length modulation causes the current in M_6 to be slightly higher than the current found from the simple expression (5.6). From the error log file, we can find the small-signal parameters for

Results from error log file

Semiconductor Device Operating Points:
--- MOSFET Transistors ---

Name:	m7	m6	m4	m2	m1
Model:	pmos-sh	nmos-sh	pmos-sh	nmos-sh	nmos-sh
Gm:	2.90e-03	2.82e-03	6.99e-04	6.89e-04	6.89e-04
Gds:	5.10e-05	3.85e-05	1.27e-05	9.36e-06	9.36e-06
Gmb:	8.67e-04	8.44e-04	2.09e-04	1.64e-04	1.64e-04
Cbd:	2.81e-13	9.20e-14	7.77e-14	2.16e-14	2.16e-14
Cbs:	4.22e-13	1.30e-13	1.07e-13	2.71e-14	2.71e-14
Cgsov:	5.53e-14	1.43e-14	1.38e-14	3.58e-15	3.58e-15
Cgdov:	5.53e-14	1.43e-14	1.38e-14	3.58e-15	3.58e-15
Cgbov:	0.00e+00	0.00e+00	0.00e+00	0.00e+00	0.00e+00
Cgs:	1.14e-12	2.84e-13	2.84e-13	7.11e-14	7.11e-14
Cgd:	0.00e+00	0.00e+00	0.00e+00	0.00e+00	0.00e+00
Cgb:	0.00e+00	0.00e+00	0.00e+00	0.00e+00	0.00e+00

Figure 5.6: Results from the error log file from a '.op' simulation of the opamp shown in Fig. 5.5.

finding the gain and the dominant pole. The small-signal parameters for the relevant transistors are listed in Fig. 5.6.

Using Eq. (5.7), we can calculate a low-frequency small-signal gain of $A_d \simeq 1012$ V/V ~ 60 dB. From Eq. (5.8), we find $r_{out} = 11.2\,\text{k}\Omega$, and from Eq. (5.9), we find the frequency of the dominant pole to be $f_{p1} \simeq 1.13$ MHz.

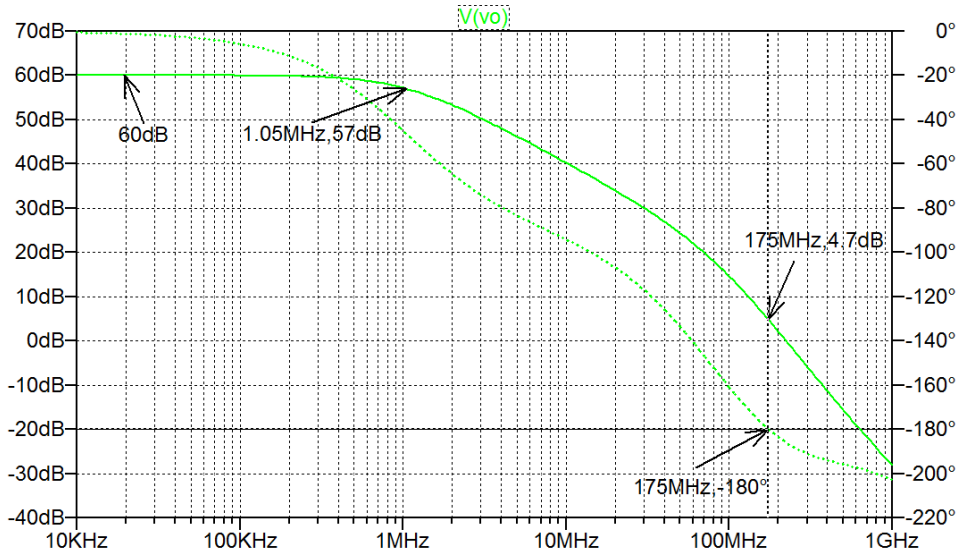


Figure 5.7: Output plot from a '.ac' simulation of the opamp shown in Fig. 5.5.

Although the gain is not impressive for an opamp, it is clearly much more than we can achieve just using a differential pair. An output resistance of 11.2 kΩ is also rather large for an opamp but it will do if the opamp is not loaded by a small resistance. If a smaller output resistance is needed, a solution can be to buffer the output by a common-drain stage but as explained in Chapter 4, this severely limits the output voltage range. The gain-bandwidth product is about 1.13 GHz but as we will see in the following, it may not be possible to use the bandwidth of the opamp to the full extent. It may be necessary to limit the bandwidth for stability reasons.

We may verify the gain and bandwidth results by running a '.ac' simulation. Figure 5.7 shows a plot of the output voltage from this simulation. We find a low-frequency gain and a -3 dB bandwidth as expected from the calculations. From the gain plot, we see that the gain drops by 40 dB per decade of frequency for frequencies above approximately 100 MHz, indicating the presence of a second pole in the system. From the phase plot, we see that the phase shifts by more than 180° for frequencies above approximately 175 MHz, so the amplifier has more than two poles and/or zeros. When the phase shift has reached 180°, the gain is still more than 0 dB.

In Chapter 4.4, we examined the common-mode input voltage range for the differential pair and found that $V_{ICM} = 1.2$ V resulted in an almost symmetric input voltage range. Of course, with an additional gain stage providing a gain of about 30 dB, the input voltage range of the opamp is smaller than that of the differential pair. With a bias value of the output voltage of 0.8 V and a saturation voltage for M_6 of about 0.3 V, we expect a maximum output voltage amplitude of about 0.5 V, corresponding to an input amplitude of 0.5 mV.

This may be verified by a transient simulation with a sinusoidal input where the input amplitude is varied up to about 0.9 mV. Figure 5.8 shows the results of a transient simulation. We notice that clipping sets in during the negative signal excursions for input amplitudes larger than approximately 0.5 mV, whereas clipping of the positive signal excursions sets in for amplitudes larger than approximately 0.7 mV. This difference is observed because the bias value of the output voltage is not mid-rail. It may be compensated for by applying a small offset voltage at the input or by adjusting the geometry of M_6 to provide an output bias voltage which is mid-rail.

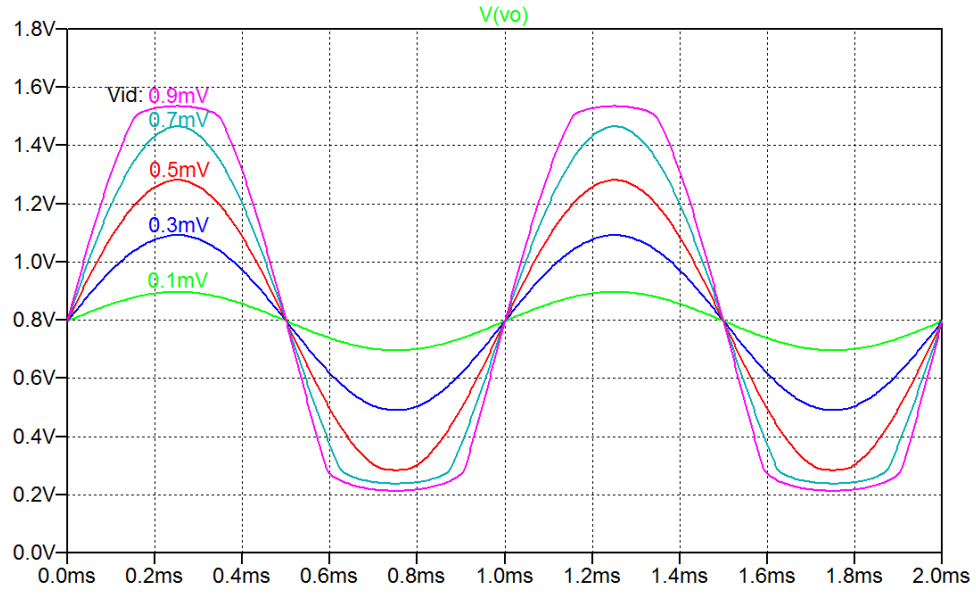


Figure 5.8: Output plot from a '.tran' simulation with a sinusoidal differential input to the opamp shown in Fig. 5.5.

5.3 The two-stage opamp with feedback

The configuration as a noninverting amplifier as shown in Fig. 5.1 is one of the standard applications for an opamp like the two-stage opamp shown in Fig. 5.4, so we use this to illustrate some of the design challenges for the two-stage opamp. The opamp shown in Fig. 5.4 uses a single-ended supply voltage V_{DD}, so the input signal and the feedback signal must be referenced to a common-mode input voltage within the supply voltage range as shown in Fig. 5.9. Also, the output voltage is the sum of an output signal voltage v_o and an output bias voltage V_O corresponding to an input signal voltage $v_{in} = 0$.

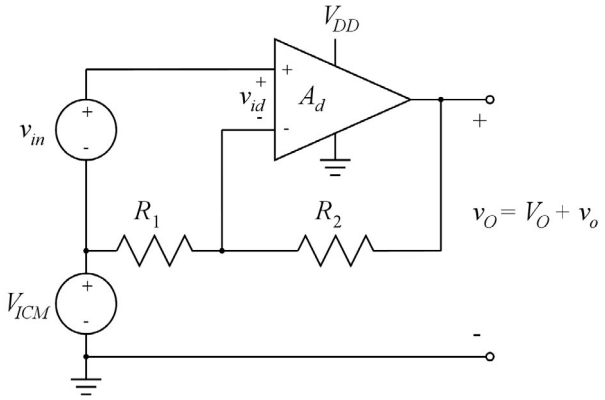


Figure 5.9: Noninverting amplifier with a single-ended supply voltage.

With a high gain A_d and the two resistors R_1 and R_2 for the feedback network, the expected gain is $A_v = v_o/v_{in} = 1 + R_2/R_1$, see Eq. (5.1).

Simulation example. As an example, we use the opamp from Fig. 5.5 and we use $R_1 = R_2 = 100\ \text{k}\Omega$. The feedback resistors have been chosen to be much larger than the output resistance of the opamp. The common-mode input voltage has been chosen to be 1.2 V, i.e., the same value as was used for the simulations of the circuit in Fig. 5.5.

Figure 5.10 shows the LTspice schematic for simulating the opamp with feedback. A '.op' simulation of the circuit shows bias values for currents and voltages as expected with a bias value $V_O = 1.19917$ V for the output voltage. Next, a '.ac' simulation is run. The result is shown in Fig. 5.11. We see that at low frequencies, the small-signal gain is 6 dB as expected but at a frequency of about 60 MHz, a very pronounced peaking of the gain is observed. This is a worrying sign, indicating a potentially unstable circuit.

We may also run a transient simulation. Figure 5.12 shows the result of a transient simulation with an input voltage v_{in} which is a pulse with a value of 0.1 V, a duration of 0.5 µs, a rise time and a fall time of 10 ns, and a period of 1 µs. Both the output signal (green trace) and the input signal (blue trace) are shown. Apparently, this input signal triggers an oscillation of the output signal where the output voltage switches between a maximum value close to the supply voltage and a minimum value close to ground.

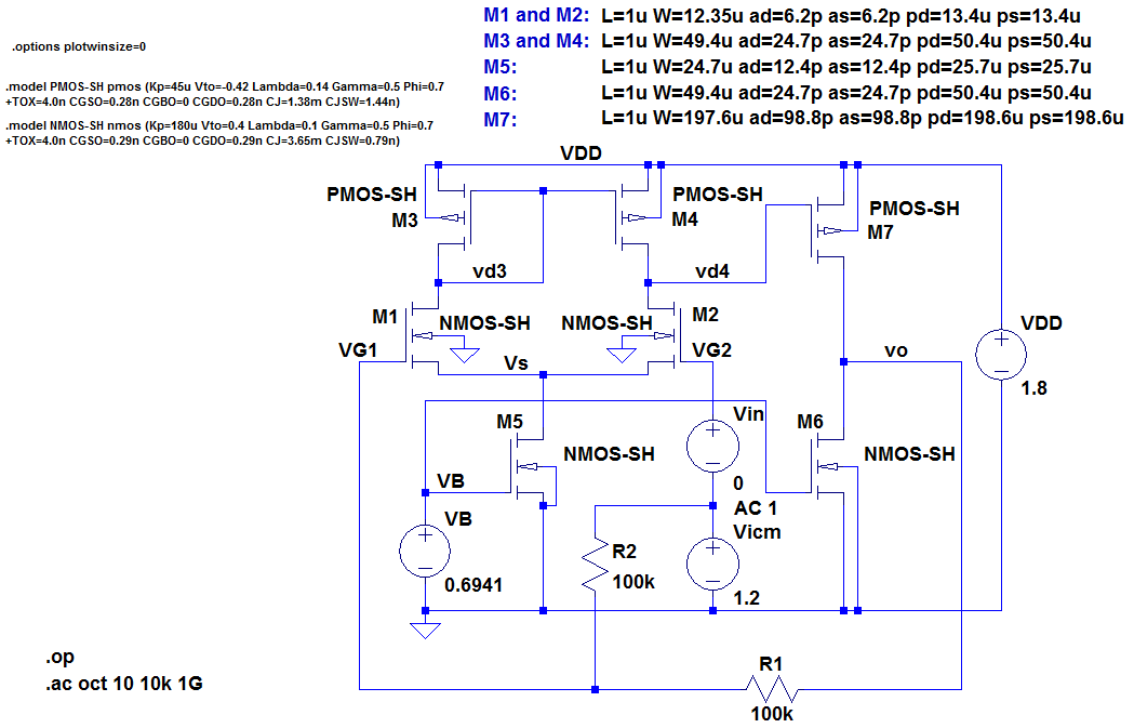


Figure 5.10: LTspice schematic for simulating the two-stage opamp with feedback.

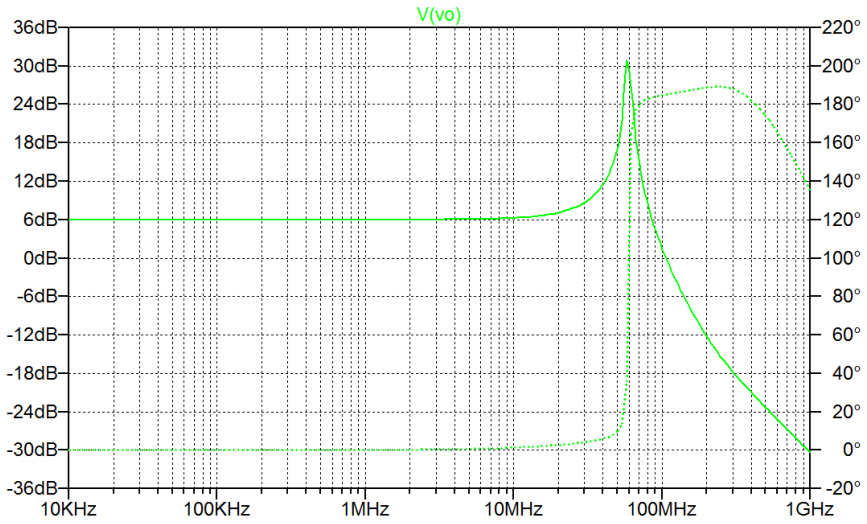


Figure 5.11: Output plot from a '.ac' simulation of the opamp with feedback shown in Fig. 5.10.

Examining the oscillations, we find an oscillation frequency of about 55 MHz, which is close to the frequency of the peak in the frequency response.

An instability like this is a major problem related to feedback systems. We will treat this in detail in Chapter 6. However, already at this point, we can give a qualitative explanation.

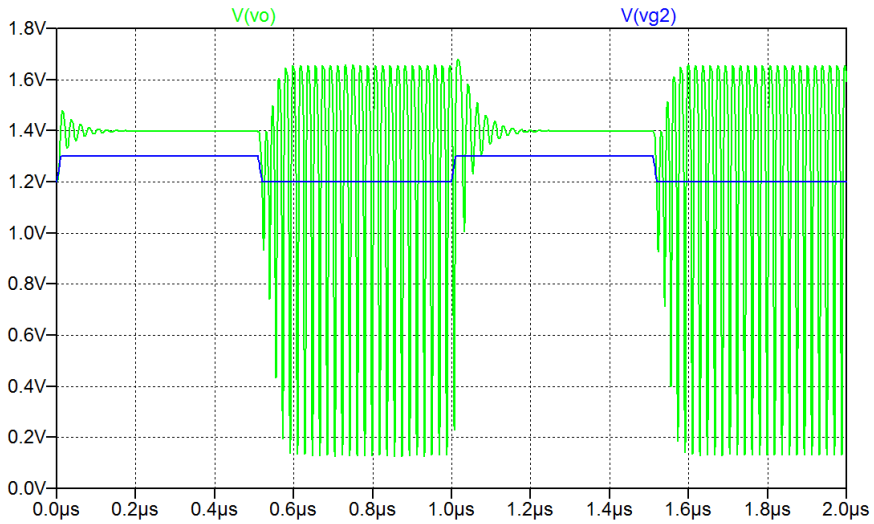


Figure 5.12: Output plot from a '.tran' simulation with a pulse input to the opamp with feedback shown in Fig. 5.10.

The stability problem. Consider the circuits shown in Fig. 5.13. The gain A_d is assumed to have a real, positive value. The circuit shown in Fig. 5.13(a) has negative feedback, i.e., an increase in the input voltage v_{in} produces a feedback voltage which is subtracted from the input voltage, thus reducing the differential input voltage v_{id} to the amplifier. This reduces the output voltage compared to a situation where the voltage v_{in} is connected directly to the input terminals of the amplifier.

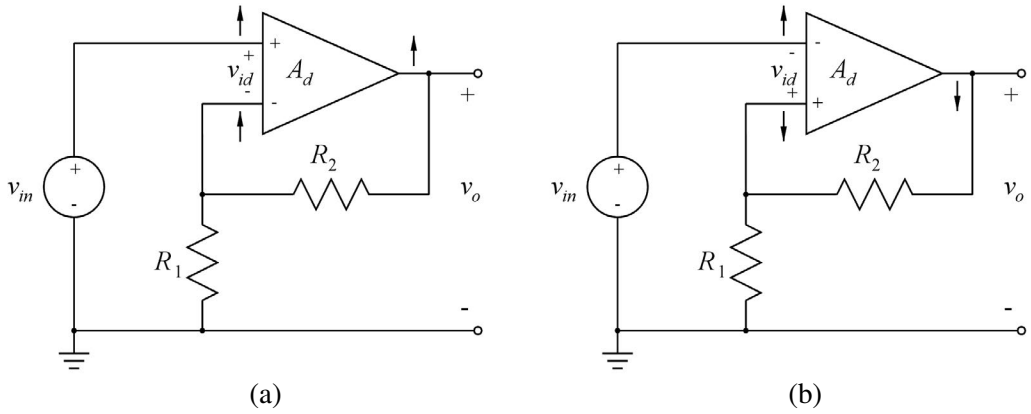


Figure 5.13: Feedback amplifiers with negative feedback (a) and positive feedback (b).

The circuit shown in Fig. 5.13(b) has positive feedback, i.e., an increase in the input voltage v_{in} produces a feedback voltage which is added to the input voltage, thus increasing the differential input voltage v_{id}. This further increases the output voltage swing compared to a situation where the voltage v_{id} is connected directly to the input terminals of the amplifier. If this increase is too large, it causes instability.

Mathematically, the output voltage for the circuit with positive feedback is given by

$$v_O = v_{IN}\left(1+\frac{R_2}{R_1}\right)\left(\frac{1}{1-(1+R_2/R_1)/A_d}\right) \tag{5.11}$$

compare to Eq. (5.2). We see that with $A_d = 1+R_2/R_1$, the denominator in Eq. (5.11) is 0, and this indicates that the output voltage is no longer controlled by the input voltage.

For an amplifier as the opamp shown in Fig. 5.5, the output voltage is in phase with the input voltage at low frequencies but at high frequencies, we observe from Fig. 5.7 that the phase of the output is shifted. If this shift is 180°, it corresponds to a sign inversion and the situation shown in Fig. 5.13(b) may occur. If the output voltage swing is too large, the positive feedback thus established will cause instability.

The trick to use in order to avoid instability is to modify the frequency response of the amplifier in such a way that the voltage gain is small when the phase shift reaches 180°. For the opamp shown in Fig. 5.5, we found that a phase shift of 180° was reached at a frequency of about 175 MHz where there is still a gain of about 5 dB. We need to reduce this gain. One way of doing this is to insert an extra capacitance between the gate and drain of the common-source transistor M_7 as explained in Chapter 4.5. This will reduce the frequency of the dominant pole so that the gain starts to roll off at a lower frequency and has reached a lower value when the phase shift is 180°. For the circuits from Fig. 5.5 and Fig. 5.10, we may insert a capacitance of 1.5 pF between the gate and source of M_7. This is about 27 times the value of C_{gd} for M_7, see Fig. 5.6, and because of the Miller effect, it will cause a substantial reduction in the frequency of the dominant pole.

Figure 5.14 shows the frequency response of the opamp from Fig. 5.5 with this extra capacitance inserted. Now the gain has dropped well below 0 dB at the frequency where the phase has reached 180°.

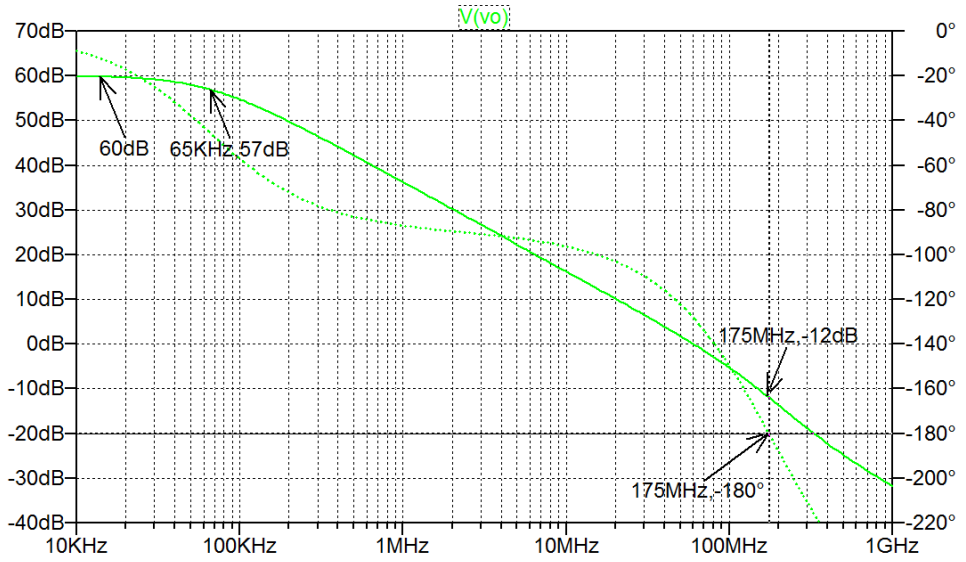


Figure 5.14: Output plot from a '.ac' simulation of the opamp shown in Fig. 5.5 with an additional capacitor of 1.5 pF between gate and drain of M_7.

For the opamp with feedback (Fig. 5.10), we repeat the '.ac' simulation and the '.tran' simulation with the extra capacitor between gate and drain of M_7. The resulting plots are shown in Figs. 5.15 and 5.16. We see that the peaking in the frequency response is now less than 3 dB, and the transient output plot shows a stable output signal although some ringing can be observed after the step in input voltage.

The lesson learned from this example is that the properties of a circuit with feedback depend in a complicated way on the detailed frequency response of the amplifier used for the feedback circuit. It is not sufficient to know the location of the dominant pole. Also higher order poles and zeros have a strong influence, so for feedback systems, we must perform a more detailed analysis of the frequency response of the amplifier to be used in the system. Before doing so, we will in the next chapter consider the properties of systems with negative feedback so that we know what to be aware of when analyzing the details of the frequency response of an amplifier.

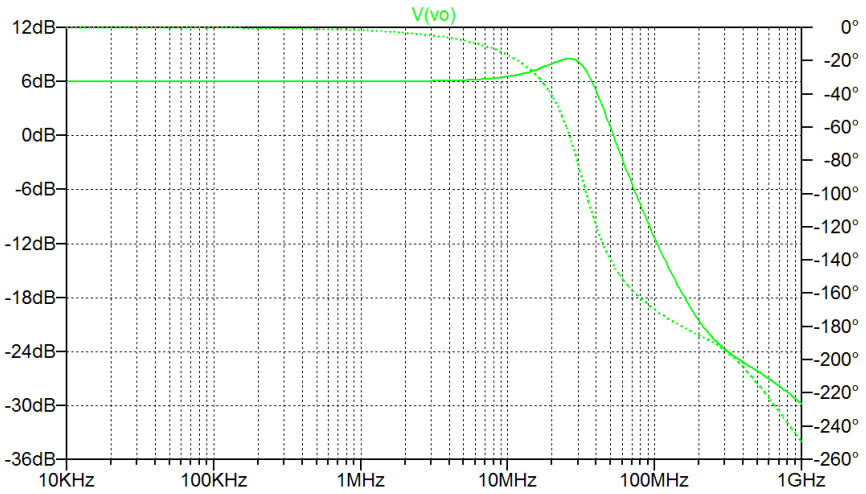


Figure 5.15: Output plot from a '.ac' simulation of the opamp with feedback shown in Fig. 5.10 with an additional capacitor of 1.5 pF between gate and drain of M_7.

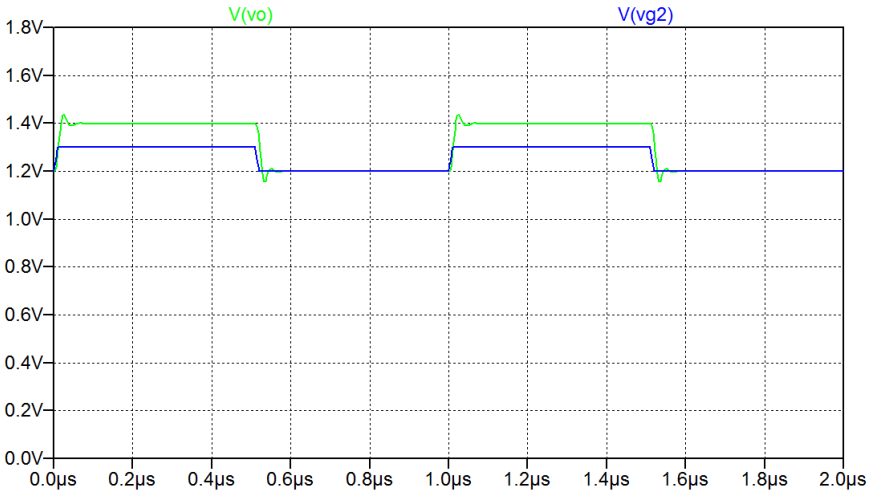


Figure 5.16: Output plot from a '.tran' simulation with a pulse input to the opamp with feedback shown in Fig. 5.10 with an additional capacitor of 1.5 pF between gate and drain of M_7.

References

Allen, PE., & Holberg, DR. 2012, *CMOS Analog Integrated Circuit Design*, Third Edition, Oxford University Press, New York, USA.

Chan Carusone, T., Johns, D. & Martin, K. 2012, *Analog Integrated Circuit Design*, Second Edition, John Wiley & Sons, Inc., Hoboken, USA.

Multiple-choice test

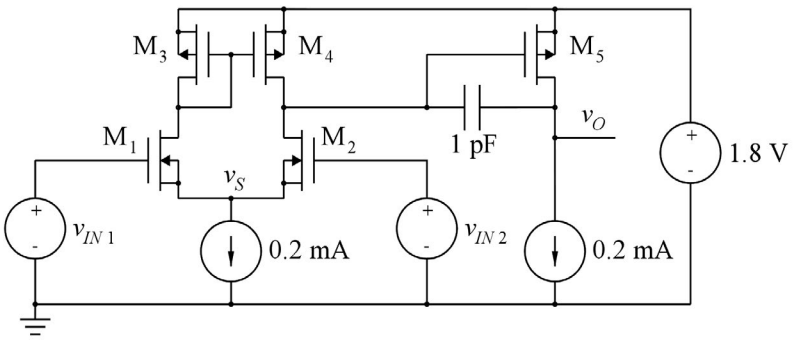
1. Complete the following statements by selecting the appropriate continuation from the table below.

A:	The input stage in a CMOS opamp is normally ...
B:	The second stage in a two-stage opamp is normally ...
C:	For a folded-cascode opamp with a differential pair using NMOS input transistors, the folded-cascode transistors are ...
D:	The dominant pole in a folded-cascode opamp normally comes from ...
E:	For a two-stage opamp with a differential pair as the input stage and a common-source stage with a PMOS transistor for providing gain, the input transistors are ...
F:	For an opamp with a differential pair using PMOS input transistors, the maximum common-mode input voltage is ...
G:	The dominant pole in a two-stage opamp normally comes from ...
H:	In order to obtain a small output resistance from a CMOS opamp, an output buffer may be added. It should be configured as ...
I:	The frequency of the dominant pole in a two-stage opamp may be reduced by inserting a capacitor between ...
J:	A circuit using an opamp with feedback may show instability if the feedback signal is ...

Continuation:

1:	less than the positive supply voltage by approximately $\|V_{\text{eff}}\|$.
2:	a common-source stage.
3:	a common-drain stage.
4:	half of the supply voltage.
5:	a differential pair.
6:	PMOS transistors.
7:	NMOS transistors.
8:	the input node.
9:	the source node for the input differential pair.
10:	less than the positive supply voltage by approximately $\|V_{GS}\|$.
11:	the output node.
12:	gate and drain of the common-source stage.
13:	input and output.
14:	less than the positive supply voltage by approximately $\|V_{GS}\| + \|V_{\text{eff}}\|$.
15:	drain and source of the common-source stage.
16:	in phase with the input signal.
17:	an active load.
18:	the input to the second stage.
19:	the positive supply voltage.
20:	inverted with respect to the input signal.

2. For the two-stage opamp shown below, assume that all transistors are in the active region and have a channel-length modulation parameter $\lambda = 0.1\ \text{V}^{-1}$ and an effective gate voltage $|V_{\text{eff}}| = 0.3$ V.



The channel-width-to-length ratio W_5/L_5 is equal to

A: $0.5 \times W_4/L_4$

B: $1.0 \times W_4/L_4$

C: $2.0 \times W_4/L_4$

3. The small-signal low-frequency output resistance of the opamp shown above is approximately

A: 50 kΩ

B: 100 kΩ

C: 200 kΩ

4. The small-signal low-frequency differential gain of the opamp shown above is approximately

A: 61 dB

B: 67 dB

C: 73 dB

5. Assuming that all transistor capacitances are much smaller than 1 pF, the frequency of the dominant pole in the opamp shown above is approximately

A: 47 kHz

B: 295 kHz

C: 3.18 MHz

6. Assuming that all zeros and non-dominant poles are at frequencies above 1 GHz, the unity-gain frequency of the opamp shown above is approximately

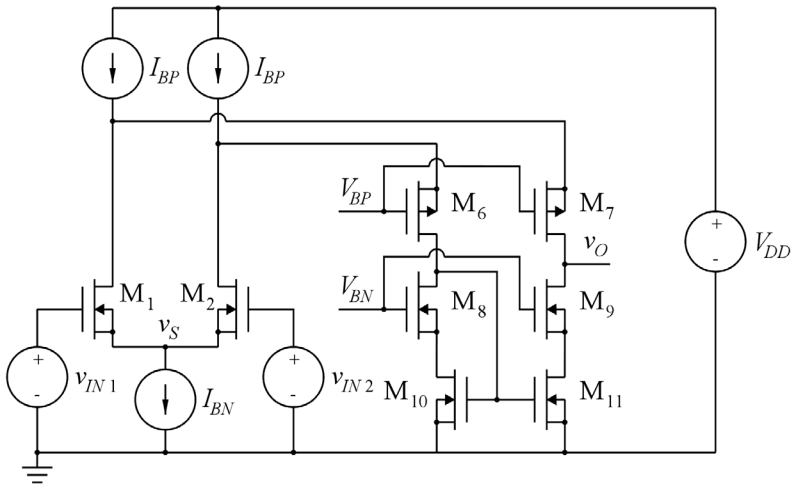
A: 31.8 MHz

B: 104 MHz

C: 654 MHz

Problems

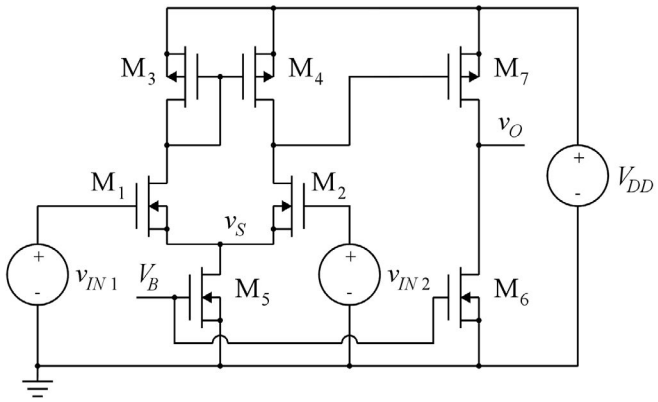
Problem 5.1



For the folded-cascode opamp shown above, assume that all transistors have a channel length $L = 1$ µm and a channel width $W = 20$ µm. Assume transistor parameters as specified in Table 3.1. The bias currents are $I_{BP} = 0.3$ mA and $I_{BN} = 0.4$ mA. Assume that the bias voltages V_{BP} and V_{BN} and the common-mode input bias voltage have values ensuring that all transistors are in the active region.

Calculate the low-frequency small-signal differential gain and the -3 dB bandwidth for a load capacitance of $C_L = 1.5$ pF which is much larger than the parasitic transistor capacitances. Use reasonable approximations when calculating small-signal parameters.

Problem 5.2



Design the two-stage opamp shown above to have a differential small-signal gain of 30 dB in the first stage and a small-signal gain of 26 dB in the second stage. Use a channel length of $L = 0.8$ µm for all transistors and design M_5 and M_6 to give bias currents of $I_{D5} = I_{D6} = 0.1$ mA with a bias voltage $V_B = 0.7$ V. Assume transistor parameters as specified in Table 3.1 and use reasonable approximations in the design equations for the transistors. Assume that all transistors are in the active region.

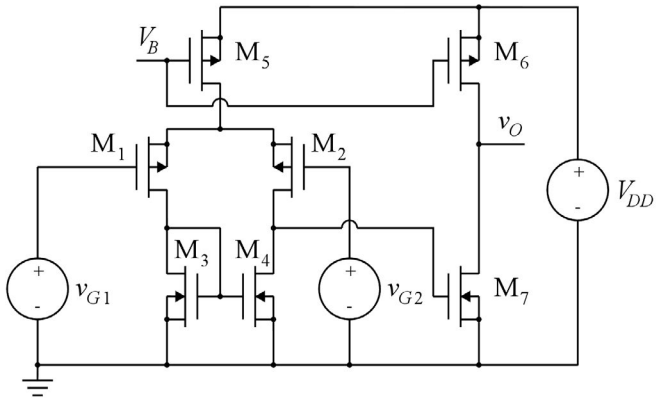
Problem 5.3

Use LTspice to simulate your design from Problem 5.2. Use $V_{DD} = 1.8$ V and find simulated values for the differential small-signal gain in the first stage and the small-signal gain in the second stage. Explain the differences between the simulated values and the values given in Problem 5.2.

Problem 5.4

Insert a capacitor $C_c = 1.5$ pF between gate and drain of M_7 in your design from Problem 5.2 and calculate the frequency of the dominant pole, assuming that C_c is much larger than the transistor capacitances. Simulate the -3 dB bandwidth and compare the simulated result to your calculated result.

Problem 5.5



The figure above shows a two-stage opamp using a PMOS differential pair for the input stage. Design the opamp using transistors with parameters as specified in Table 3.1 and use a channel length of $L = 0.9$ µm and an effective gate voltage $|V_{\text{eff}}| = |V_{GS} - V_t| = 0.3$ V for all transistors. The supply voltage is $V_{DD} = 1.8$ V.

Design the opamp to provide $g_{m1} = 0.3$ mA/V and $g_{m7} = 10\,g_{m1}$. Assume that all transistors are in the active region and use reasonable approximations in the design equations for the transistors.

Calculate the low-frequency differential small-signal gain and output resistance of the opamp.

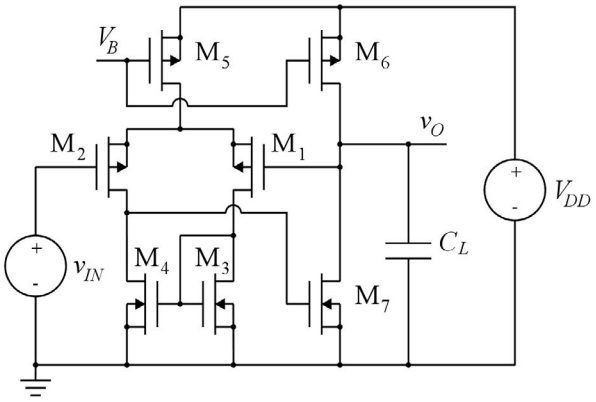
Problem 5.6

Use LTspice to simulate your design from Problem 5.5. Find simulated values for g_{m1} and g_{m7} and the differential small-signal gain and output resistance. Explain the differences between the simulated values and the values given in Problem 5.5.

Problem 5.7

Insert a capacitor C_c between gate and drain of M_7 in your design from Problem 5.5 and calculate the value of C_c required to obtain a gain-bandwidth product of 40 MHz. Assume that C_c is much larger than the transistor capacitances. You may also assume that the gain in the second stage is much larger than 1. Simulate the gain-bandwidth product and compare the simulated result to your calculated result.

Problem 5.8




```
.model NMOS-SH nmos (Kp=180u Vto=0.40
+lambda={0.1u/L} gamma=0.5 phi=0.7
+Tox=4n CGSO=0.29n CGBO=0
+CGDO=0.29n CJ=3.65m CJSW=0.79n)
.model PMOS-SH nmos (Kp=45u Vto=-0.42
+lambda={0.14u/L} gamma=0.5 phi=0.7
+Tox=4n CGSO=0.28n CGBO=0
+CGDO=0.28n CJ=1.38m CJSW=1.44n)
```

The figure above shows the opamp from Problem 5.5 with the output fed back to the inverting input and a capacitor $C_L = 1$ pF connected to the output. Find the low-frequency small-signal gain. Use the transistor models shown above including capacitances to simulate the frequency response and find the frequency of the peak in the response. Use a bias value of the input voltage ensuring that all transistors are in the active region.

Simulate the transient response with a pulse input with a value of 0.2 V, a duration of 100 ns, rise time and fall time of 1 ns and a period of 200 ns. What is the frequency of oscillation?

Insert a capacitor $C_c = 1.2$ pF between gate and drain of M_7 and repeat the simulations.

What is the -3 dB bandwidth? Find the low-frequency output resistance.

Chapter 6 – Feedback

We have already in Chapter 5 seen an example of an amplifier using negative feedback, the noninverting amplifier, and we have observed that feedback is useful for controlling properties such as the precision of the gain. We have also seen that feedback may cause problems related to stability. In this chapter, we examine the basic properties of feedback in more detail and we introduce methods to analyze the stability and modify the circuit design to achieve stability.

After having studied the chapter, you should be able to

- explain the general structure of a system with negative feedback.
- calculate the effects of feedback on gain, bandwidth, input resistance and output resistance.
- explain the effects of feedback on nonlinearity and distortion.
- explain the different feedback topologies using voltage sensing or current sensing and voltage mixing or current mixing.
- explain stability criteria for a system with feedback.
- find the loop gain and use the loop gain to find phase margin, amplitude margin and stability margin in a system with feedback.
- use methods of frequency compensation to obtain and improve stability in a system with feedback.

6.1 The basic feedback structure

The basic idea in a system with negative feedback is that a fraction of the output signal from the system is fed back to the input where it is subtracted from the external input signal. As we have already seen in Chapter 5, this causes a reduction of the overall gain but it provides a number of advantages which we will examine in the following. The concept is illustrated in Fig. 6.1, showing a signal flow diagram of a feedback system (Lathi 2009; Sedra & Smith 2016).

Block A is an amplifier with a gain $A = y_o/x_i$. The gain A is called the open-loop gain. Block β is an attenuator creating a feedback signal $x_f = \beta y_o$ where β is called the feedback factor. The input to the amplifier is the difference between an external input x_s and the feedback signal x_f. The signals x_s, x_i and x_f are of the same kind, e.g., they are all voltages or they are all currents. The signal y_o may be of a different kind than the signals x_s, x_i and x_f. In the analysis of the system, it is assumed that the feedback circuit (the β-circuit) does not load the output of the amplifier A (the A-circuit). Also, the β-circuit does not load the input of the amplifier.

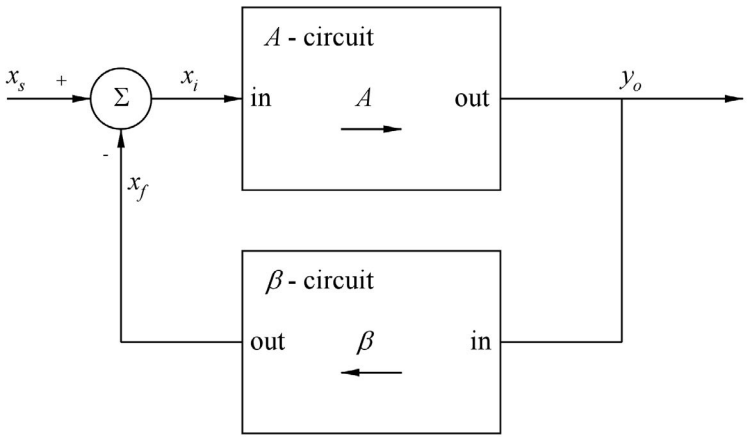


Figure 6.1: Basic structure of a system with negative feedback.

Assuming unidirectional signal paths through both the A-network and the β-network, we find

$$\begin{aligned} y_o &= A x_i = A(x_s - x_f) = A(x_s - \beta y_o) \\ \Rightarrow A_{CL} &= \frac{y_o}{x_s} = \frac{A}{1 + A\beta} \end{aligned} \tag{6.1}$$

The gain A_{CL} from the external signal source x_s to the output y_o of the system is called the closed-loop gain. The quantity $L = A\beta$ is the gain in the loop consisting of the amplifier A and the feedback network β and it is called the loop gain. The quantity $1 + L = 1 + A\beta$ is called the amount of feedback (Sedra & Smith 2016).

We see from Eq. (6.1) that the closed-loop gain A_{CL} is smaller than the open-loop gain A (assuming L positive), and if the loop gain is much larger than 1, i.e., $A\beta \gg 1$, we find an approximate value for the closed-loop gain as

$$A_{CL} \simeq \frac{1}{\beta} \tag{6.2}$$

This was also the result found in Chapter 5 for the noninverting amplifier shown in Fig. 5.1 and this is the reason why an opamp with a very large gain is such a useful device in feedback systems. The resulting closed-loop gain is determined solely by the feedback network which is normally implemented from passive devices, and often the closed-loop gain is determined by the ratio between device values. In IC technology, ratios between device values can be controlled with high precision.

6.2 Advantages of feedback

We have already seen that one of the advantages of using negative feedback is that the closed-loop gain is controlled mainly by passive components which can be selected with high precision whereas an open-loop gain in a CMOS amplifier depends on transistor parameters which are rather imprecisely defined, partly due to the spread in the manufacturing process, partly due to temperature variations and supply voltage variations. In this section, we quantify the improvements in gain control and we investigate the benefits of using negative feedback on other system parameters, including input resistance, output resistance, bandwidth and distortion.

Gain control. As an example of the gain control achieved by using negative feedback, we may revisit the noninverting opamp shown in Fig. 5.1. Using the notation from Fig. 6.1, we have $x_s = v_{in}$, $y_o = v_o$, $A = A_d$, $A_{CL} = A_v$ and $x_f = \beta y_o = v_o R_1/(R_1+R_2)$, i.e., $\beta = R_1/(R_1+R_2)$. With an ideal opamp, the closed-loop gain is $G_0 = 1/\beta = 1+R_2/R_1$ as also found in Eq. (5.1). With a finite value of A_d, we find from Eq. (6.1)

$$A_v = A_{CL} = \frac{A_d}{1+A_d R_1/(R_1+R_2)} = \left(1+\frac{R_2}{R_1}\right)\left(\frac{1}{1+(1+R_2/R_1)/A_d}\right) = G_0\,\frac{1}{1+G_0/A_d} \tag{6.3}$$

This result was also found in Eq. (5.2).

As an example, let us assume that the target gain is $G_0 = 10$ V/V ~ 20 dB. Using a basic amplifier with gains of 80 dB, 60 dB or 40 dB, we find

$$\begin{aligned} A_d \sim 80 \text{ dB:} \qquad & A_{CL} = 10\,\frac{1}{1+10/10^4} \simeq 10\times 0.999 \\ A_d \sim 60 \text{ dB:} \qquad & A_{CL} = 10\,\frac{1}{1+10/10^3} \simeq 10\times 0.990 \\ A_d \sim 40 \text{ dB:} \qquad & A_{CL} = 10\,\frac{1}{1+10/10^2} \simeq 10\times 0.909 \end{aligned}$$

Thus, even very large variations (orders of magnitude) in A_d cause only moderate variations in the closed-loop gain.

We can quantify the gain control obtained from feedback by comparing the relative variation of the closed-loop gain to the relative variation of the open-loop gain. Let us assume that the open-loop gain may vary by ΔA, i.e., the relative variation is $\Delta A/A$. This causes a variation ΔA_{CL} in the closed-loop gain. Using a linear approximation to the calculation of ΔA_{CL}, we may find $\Delta A_{CL}/\Delta A$ as

$$\begin{aligned} \frac{\Delta A_{CL}}{\Delta A} \simeq \frac{dA_{CL}}{dA} &= \frac{d}{dA}\left(\frac{A}{1+A\beta}\right) = \frac{d}{dA}\left(\frac{1}{1/A+\beta}\right) \\ &= -(1/A+\beta)^{-2}(-A^{-2}) = \frac{1}{(1+A\beta)^2} \end{aligned} \tag{6.4}$$

From Eqs. (6.4) and (6.1), we find

$$dA_{CL} = \frac{dA}{(1+A\beta)^2} \Rightarrow \frac{dA_{CL}}{A_{CL}} = \left(\frac{1}{1+A\beta}\right)\frac{dA}{A} \tag{6.5}$$

Thus, the relative variation in closed-loop gain equals the relative variation in open-loop gain divided by $(1+A\beta)$, the amount of feedback.

Linearity. The fact that feedback controls the gain such that for large values of A, the closed-loop gain is close to $1/\beta$, regardless of the exact value of A, serves to improve the linearity of the amplifier. As an example, let us consider an amplifier with a nonlinear transfer function given by

$$v_O = \frac{2\text{ V}}{1+\exp(-800\text{ V}^{-1}\cdot v_{IN})} - 1\text{ V} \tag{6.6}$$

This function is shown in Fig. 6.2. For $v_{IN} = 0$, the output voltage is 0. For large positive values of v_{IN}, the output voltage approaches 1 V and for large negative values of v_{IN}, the output voltage approaches -1V. The small-signal gain for an input bias voltage $V_{IN} = 0$ is 400 V/V.

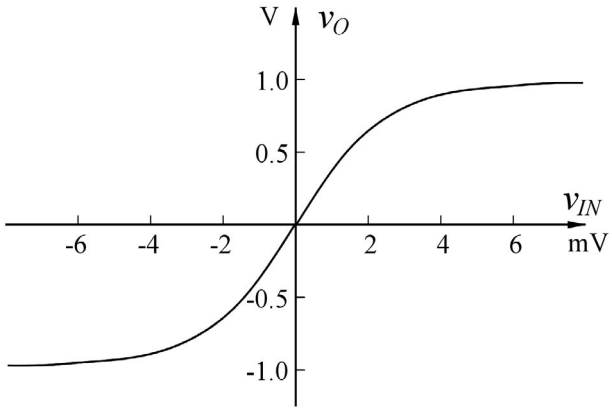


Figure 6.2: Example of a nonlinear amplifier transfer function.

We may simulate the circuit using LTspice with the amplifier modeled as an arbitrary behavioral voltage source, see the LTspice schematic in Fig. 6.3. Running a '.dc' simulation with v_{IN} varying from -2 mV to $+2$ mV results in the output plots shown in Fig. 6.4.

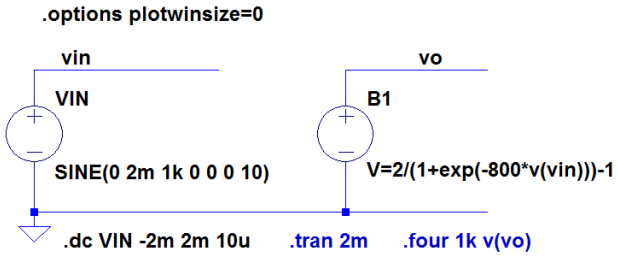


Figure 6.3: LTspice schematic showing an amplifier with the nonlinear transfer function shown in Fig. 6.2.

The top trace shows v_O versus v_{IN} and the bottom trace shows the small-signal gain dv_O/dv_{IN} versus v_{IN}. We see that the transfer function is fairly linear for small values of v_{IN}, but for larger values, the output voltage approaches the saturation limits and the plot of dv_O/dv_{IN} also shows that the small-signal gain drops to about 220 V/V for $|v_{IN}| = 2$ mV.

Now, let us apply feedback to the amplifier. Suppose we wish to obtain a closed-loop gain of $A_{CL} =$ 20 V/V. We can find the required feedback factor β using Eq. (6.1):

$$A_{CL} = \frac{y_o}{x_s} = \frac{A}{1+A\beta} \Rightarrow \beta = \frac{1}{A}\left(\frac{A}{A_{CL}} - 1\right) = \frac{1}{A_{CL}} - \frac{1}{A} \tag{6.7}$$

With $A_{CL} = 20$ V/V and $A = 400$ V/V, we find $\beta = 0.0475$. Figure 6.5 shows the LTspice schematic for the amplifier with feedback. As the feedback reduces the gain by a factor $1 + \beta A = 20$, we must use an

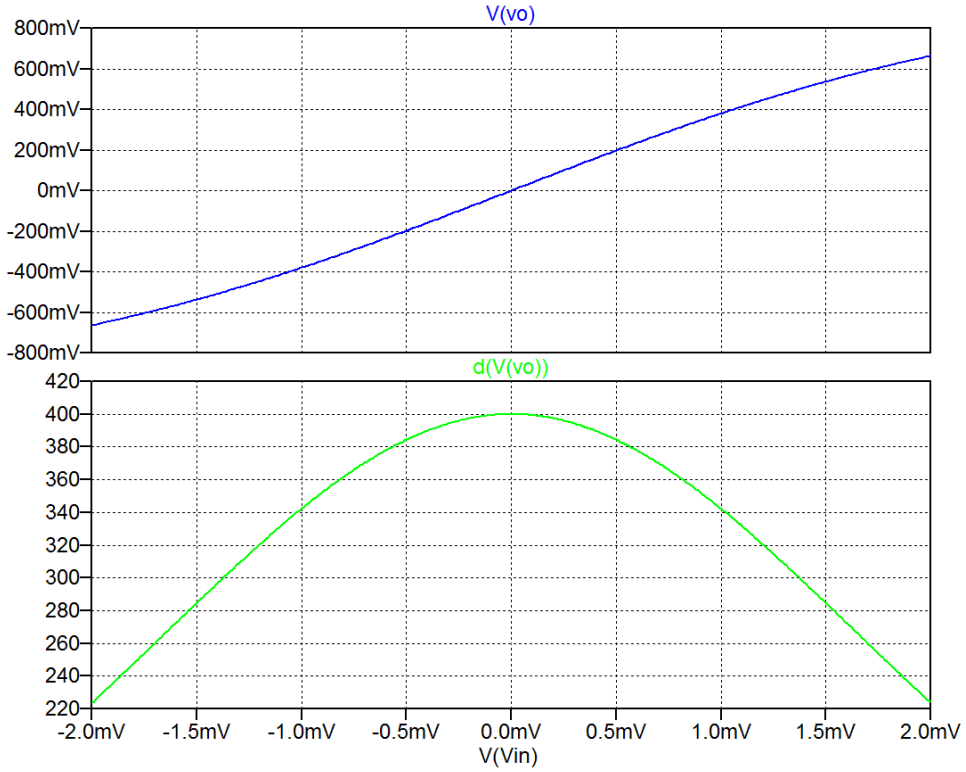


Figure 6.4: Output plots from a '.dc' simulation of the amplifier shown in Fig. 6.3.

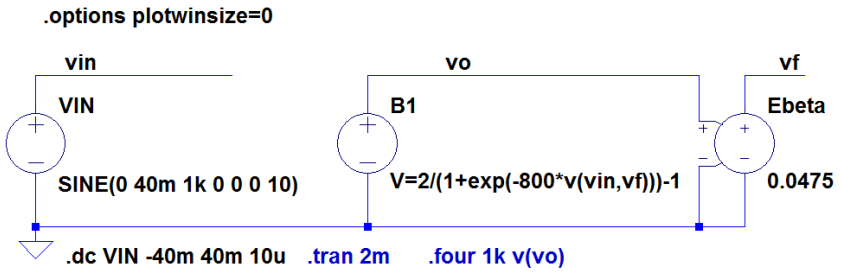


Figure 6.5: LTspice schematic showing the amplifier from Fig. 6.3 with feedback using $\beta = 0.0475$.

input voltage range from -40 mV to $+40$ mV in order to get an output voltage range comparable to the range for the amplifier without feedback.

Figure 6.6 shows a simulation of the output voltage v_O versus v_{IN} and of the small-signal gain dv_O/dv_{IN} versus v_{IN}. Compared to Fig. 6.4, it is apparent that the feedback has improved the linearity. The reduction in small-signal gain at the limits of the input range is now only about 7.5% compared to 45% for the amplifier without feedback.

The linearity improvement may also be illustrated from the distortion. Applying a sinusoidal input with an amplitude of 2 mV to the amplifier without feedback, we find (using a '.tran' simulation with a '.four' directive) a harmonic distortion of 4.6%. For the amplifier with feedback, a sinusoidal input with an

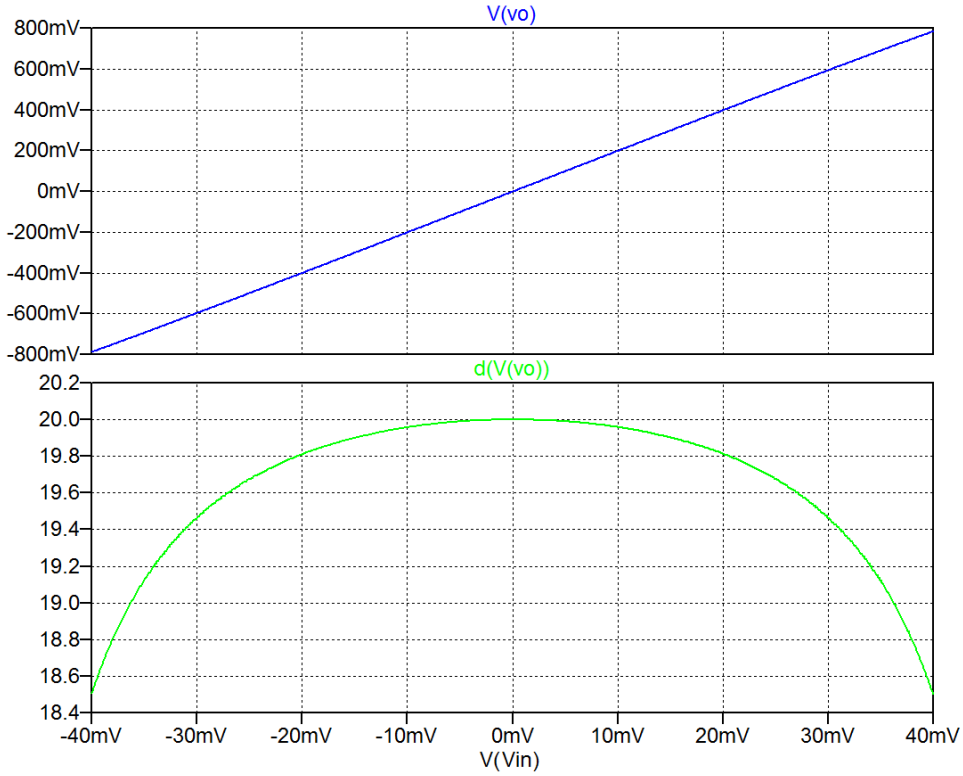


Figure 6.6: Output plots from a '.dc' simulation of the amplifier shown in Fig. 6.5.

amplitude of 40 mV results in a distortion of 0.48%. Thus, we achieve a substantial improvement in linearity at the expense of a reduction in gain. However, by cascading two amplifier stages of the kind shown in Fig. 6.5, we achieve an overall gain of 400 V/V, i.e., the same as for a single-stage amplifier without feedback. Figure 6.7 shows the LTspice schematic for such a two-stage amplifier and Fig. 6.8

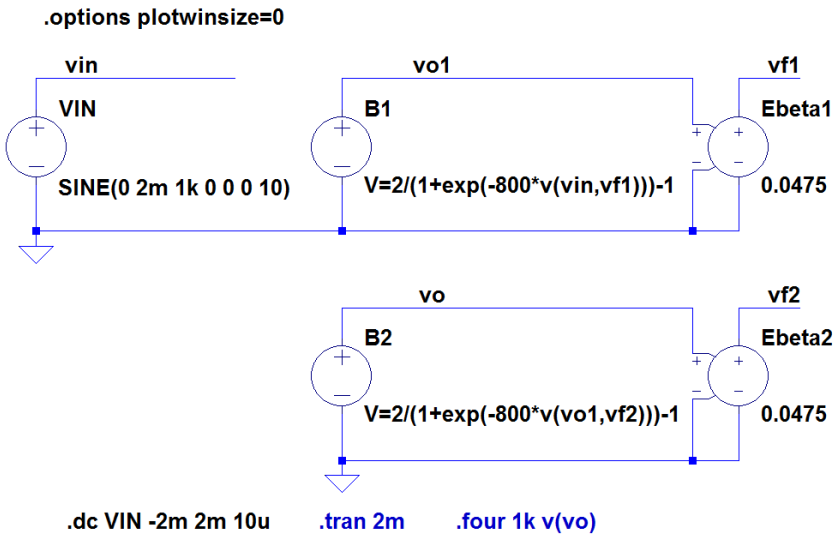


Figure 6.7: LTspice schematic showing a two-stage amplifier using the gain stages from Fig. 6.5.

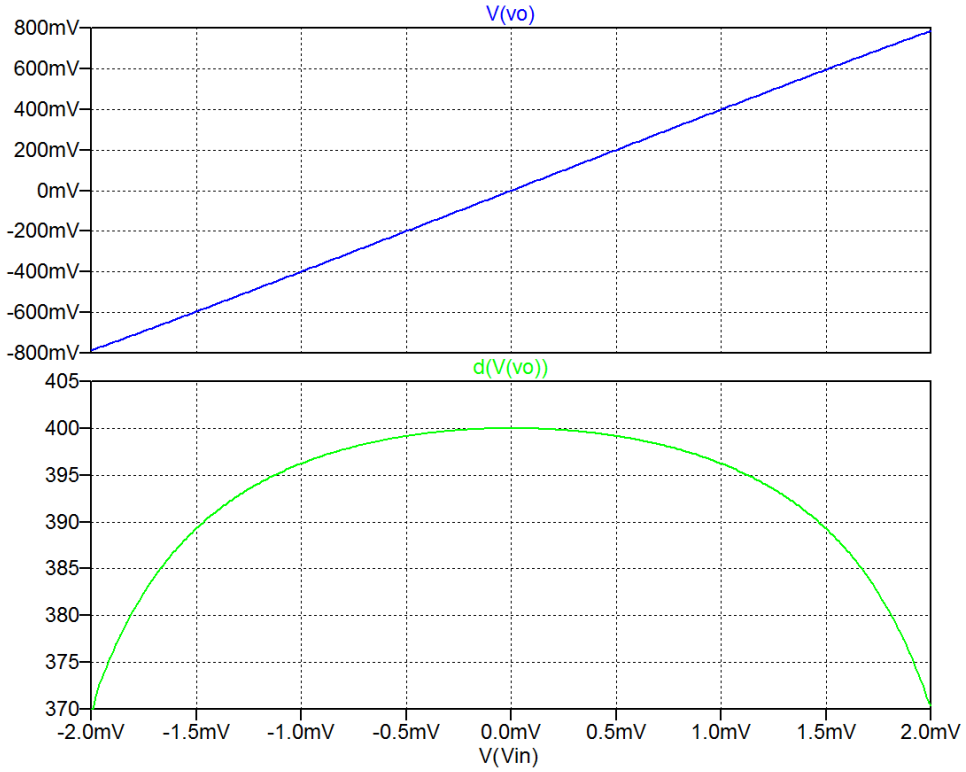


Figure 6.8: Output plots from a '.dc' simulation of the two-stage amplifier shown in Fig. 6.6.

shows the output voltage and the small-signal gain for an input voltage range from -2 mV to $+2$ mV. Clearly, the linearity of this amplifier is better than that of a single-stage amplifier without feedback. A transient simulation of the amplifier from Fig. 6.7 shows a distortion of 0.48% for an input amplitude

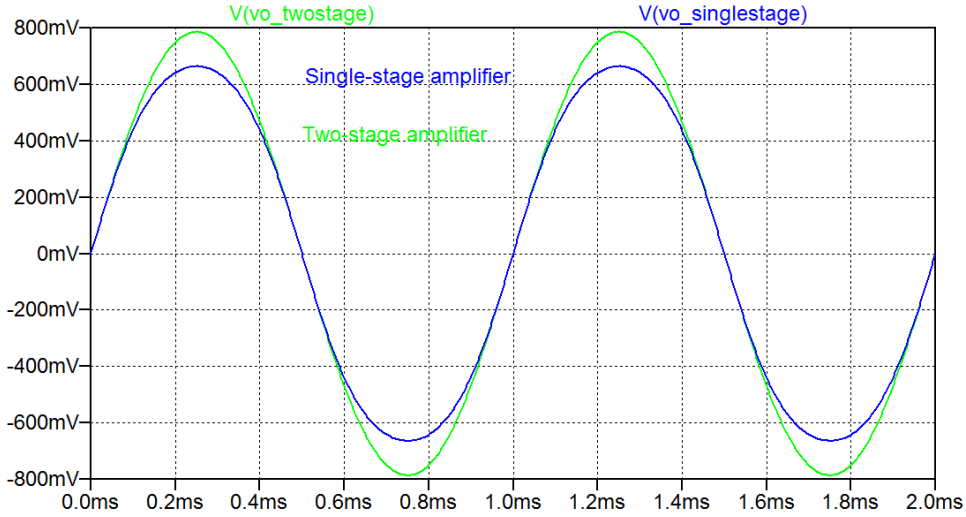


Figure 6.9: Output voltage from the single-stage amplifier and the two-stage amplifier with a sinusoidal input voltage with an amplitude of 2 mV.

of 2 mV, i.e., much lower than for the single-stage amplifier without feedback and about the same as found for the amplifier from Fig. 6.5, indicating that the distortion in the two-stage amplifier comes from the output stage.

The difference between the single-stage amplifier and the two-stage amplifier may be further illustrated by a plot of the output voltage for the two amplifiers with a sinusoidal input with an amplitude of 2 mV. This is shown in Fig. 6.9. Clearly, the single-stage amplifier cannot deliver an output amplitude of 800 mV and the distortion of the output is visible. The two-stage amplifier delivers an output amplitude very close to 800 mV, and the distortion is not visible.

Bandwidth. Assuming that the A-circuit is an amplifier where A is a transfer function with a low-frequency gain A_0 and a single pole at the frequency ω_p, we have

$$A(s) = \frac{A_0}{1+s/\omega_p} \tag{6.8}$$

Using Eq. (6.1), we find

$$A_{CL}(s) = \frac{A(s)}{1+\beta A(s)} = \frac{A_0/(1+s/\omega_p)}{1+\beta A_0/(1+s/\omega_p)} = \left(\frac{A_0}{1+\beta A_0}\right)\left(\frac{1}{1+s/(\omega_p(1+\beta A_0))}\right) \tag{6.9}$$

From Eq. (6.9), we find the low-frequency closed-loop gain

$$A_{0CL} = \frac{A_0}{1+\beta A_0} \tag{6.10}$$

and the pole frequency ω_{pCL} of the amplifier with feedback

$$\omega_{pCL} = \omega_p(1+\beta A_0) \tag{6.11}$$

Thus, the bandwidth of the amplifier is increased by a factor equal to the amount of feedback. The low-frequency gain is reduced by a factor equal to the amount of feedback, so the gain-bandwidth product is the same for the amplifier with feedback as for the amplifier without feedback. By applying feedback, we trade off gain in order to obtain bandwidth. This trade-off is shown in the Bode plot in Fig. 6.10.

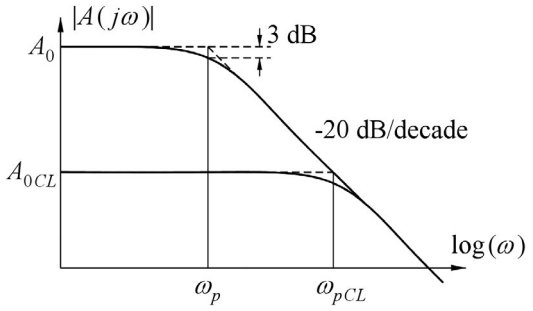


Figure 6.10: Bode plot of open-loop gain and closed-loop gain in a feedback amplifier with a single pole.

Output resistance. Let us consider a feedback system where the A-circuit is a voltage amplifier with a differential gain A, an input resistance r_{in} and an output resistance r_{out}. The feedback system can be modeled as shown in Fig. 6.11.

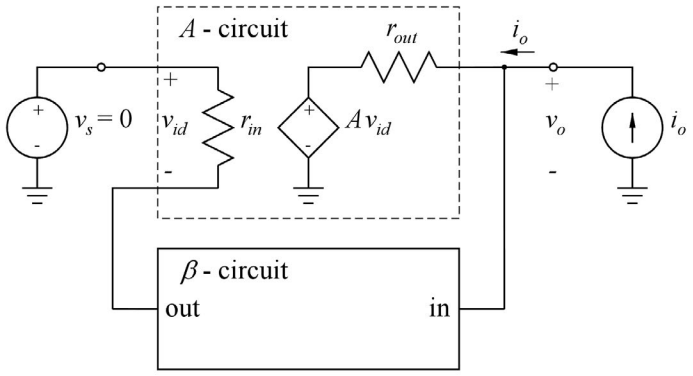


Figure 6.11: Circuit for finding the output resistance of a feedback amplifier with finite input resistance and output resistance of the A-circuit.

In order to find the output resistance of the amplifier with feedback, we reset the input voltage v_s, apply a current i_o to the output and calculate the resulting output voltage v_o. The output resistance is found as $r_{out\,CL} = v_o/i_o$. Assuming infinite input resistance and zero output resistance of the β-circuit, we find

$$v_o = A\,v_{id} + i_o\,r_{out} \tag{6.12}$$

and inserting $v_{id} = v_s - \beta\,v_o = -\beta\,v_o$, we find

$$v_o = -A\,\beta\,v_o + i_o\,r_{out} \;\Rightarrow\; r_{out\,CL} = \frac{v_o}{i_o} = \frac{r_{out}}{1+\beta\,A} \tag{6.13}$$

From Eq. (6.13), we see that the output resistance of the amplifier with feedback equals the output resistance of the basic amplifier reduced by a factor $(1+\beta\,A)$, i.e., the amount of feedback. For a voltage amplifier, the output resistance is ideally 0, so the decrease in output resistance is an improvement by a factor equal to the amount of feedback.

Input resistance. Also the input resistance is improved by feedback. Again, let us consider a system where the A-circuit is a voltage amplifier with a differential gain A. We assume the output resistance to be r_{out} and the input resistance to be r_{in}. Figure 6.12 shows a feedback system using this amplifier.

We can find the input resistance $r_{in\,CL}$ of the amplifier with feedback by applying an input voltage v_s and calculating $r_{in\,CL} = v_s/i_{in}$ where i_{in} is the input current. Assuming infinite input resistance and zero output resistance of the β-circuit, we find using Ohm's law

$$\begin{aligned} i_{in} &= \frac{v_{id}}{r_{in}} = \frac{v_s - \beta\,v_o}{r_{in}} = \frac{v_s - \beta\,A_{CL}\,v_s}{r_{in}} = \frac{v_s}{r_{in}}\left(1 - \frac{\beta\,A}{1+\beta\,A}\right) = \frac{v_s}{r_{in}}\left(\frac{1}{1+\beta\,A}\right) \\ \Rightarrow\; r_{in\,CL} &= r_{in}(1+\beta\,A) \end{aligned} \tag{6.14}$$

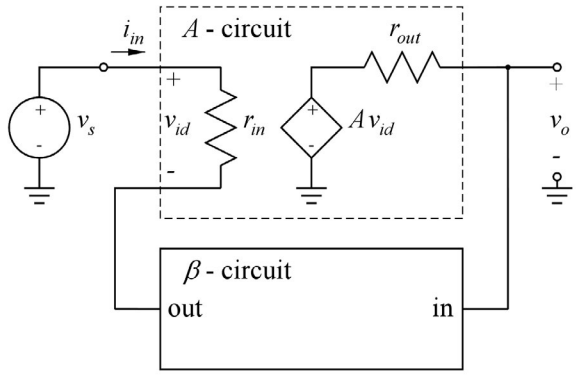


Figure 6.12: Circuit for finding the input resistance of a feedback amplifier with finite input resistance and output resistance of the A-circuit.

From Eq. (6.14), we see that the input resistance of the amplifier with feedback equals the input resistance of the basic amplifier multiplied by the amount of feedback. For a voltage amplifier, the input resistance is ideally infinite, so the increase in input resistance is an improvement by a factor equal to the amount of feedback.

Resistive β-circuit. In the analysis above, it was assumed that the β-circuit was an ideal voltage attenuator, i.e., a voltage-controlled voltage source with a gain smaller than 1. In practice, the β-circuit is often implemented using passive components (resistors), and it represents a load to the output of the amplifier. Also, the output of the β-circuit driving into the inverting amplifier input is not an ideal voltage source. It has a finite output resistance. With the β-circuit implemented as a resistive voltage divider, we have the circuit shown in Fig. 6.13.
The voltage attenuator has the attenuation ratio $\beta = R_1/(R_1+R_2)$, so we may modify the circuit as shown in Fig. 6.14 where the feedback circuit has been modeled by a Thévenin equivalent circuit.

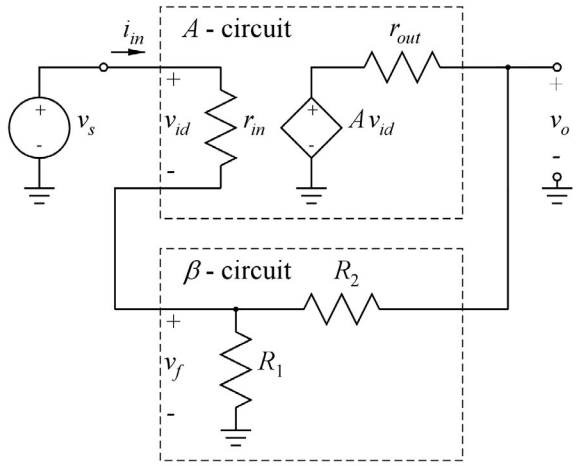


Figure 6.13: Feedback amplifier with finite input resistance and output resistance of the A-circuit and a resistive β-circuit.

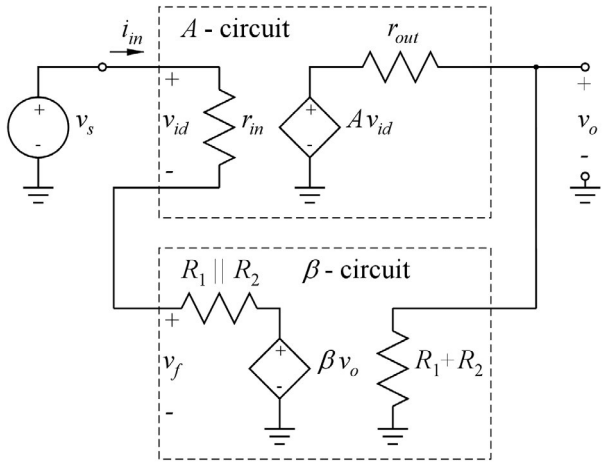


Figure 6.14: Feedback amplifier with Thévenin modeling of the β-circuit.

The resistors in the β-circuit may be moved into the A-circuit and by doing so, we see that $(R_1 \parallel R_2)$ is added to r_{in} and $(R_1 + R_2)$ is connected as a load to the output. This shows that we have voltage divisions at both the input and the output of the A-circuit, so the gain of the A-circuit is modified to $A[r_{in}/(r_{in} + R_1 \parallel R_2)][(R_1 + R_2)/(R_1 + R_2 + r_{out})]$.
Also, the input resistance of the A-circuit is $(r_{in} + R_1 \parallel R_2)$ and the output resistance of the A-circuit is $r_{out} \parallel (R_1 + R_2)$. However, the detailed calculations using this approach are beyond the scope of this book and the reader is referred to Sedra & Smith (2016).

6.3 Feedback topologies

The feedback amplifiers treated in Section 6.2 were all based on a voltage amplifier as the A-circuit and a voltage attenuator as the β-circuit. This configuration uses voltage sensing at the output and voltage mixing at the input. Referring to the electrical connections, we have a series connection at the input and a shunt connection at the output. It is called a series-shunt feedback configuration.

However, in the general structure shown in Fig. 6.1, the x-signals are not necessarily of the same type as the y-signal. Using current for the input signals and voltage for the output signal, we arrive at a feedback structure as shown in Fig. 6.15 where the A-circuit is a transresistance amplifier characterized by a transresistance A, an input resistance r_{in} and an output resistance r_{out}. For this configuration, the β-circuit is a voltage-controlled current source, and we have $i_f = \beta v_o$, so β is a transconductance. This configuration uses voltage sensing at the output and current mixing at the input. Referring to the electrical connections, we have shunt connections at both the input and the output. It is called a shunt-shunt feedback configuration.

Assuming that the β-circuit does not load the A-circuit, we find the closed-loop gain using Eq. (6.1):

$$A_{CL} = \frac{v_o}{i_s} = \frac{A}{1 + A\beta} \tag{6.15}$$

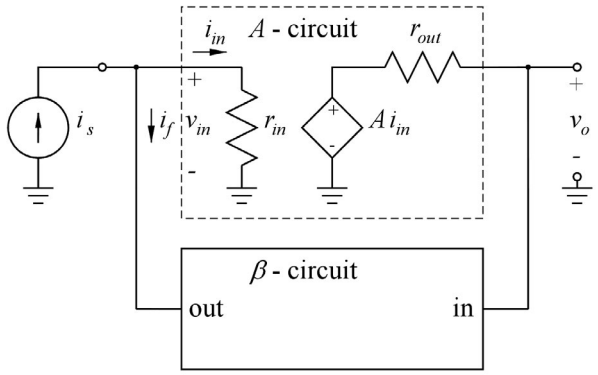


Figure 6.15: Feedback amplifier with voltage sensing and current mixing, shunt-shunt feedback.

The transresistance of the basic amplifier is reduced by a factor equal to the amount of feedback and for a large value of the loop gain βA, it is approximately equal to $1/\beta$.

We can find the output resistance of the amplifier with feedback by resetting the input signal i_s, applying a current i_o to the output and calculating the resulting output voltage v_o. The output resistance is found as $r_{out\,CL} = v_o/i_o$. We find

$$v_o = A\,i_{in} + i_o\,r_{out} = A(-i_f) + i_o\,r_{out} = A(-\beta\,v_o) + i_o\,r_{out} \;\Rightarrow\; r_{out\,CL} = \frac{v_o}{i_o} = \frac{r_{out}}{1+\beta A} \tag{6.16}$$

The output resistance of the amplifier with feedback equals the output resistance of the basic amplifier reduced by a factor $(1+\beta A)$, i.e., the amount of feedback. For a transresistance amplifier, the output resistance is ideally 0, so the decrease in output resistance is an improvement by a factor equal to the amount of feedback.

The input resistance of the amplifier with feedback is found by calculating the input voltage when a current i_s is applied to the input. We find

$$v_{in} = i_{in}\,r_{in} = (i_s - i_f)\,r_{in} = (i_s - \beta\,v_o)\,r_{in} = \left(i_s - \frac{\beta A}{1+\beta A} i_s\right) r_{in} \;\Rightarrow\; r_{in\,CL} = \frac{v_{in}}{i_s} = \frac{r_{in}}{1+\beta A} \tag{6.17}$$

The input resistance of the amplifier with feedback equals the input resistance of the basic amplifier reduced by a factor $(1+\beta A)$, i.e., the amount of feedback. For a transresistance amplifier, the input resistance is ideally 0, so the decrease in input resistance is an improvement by a factor equal to the amount of feedback.

The transresistance amplifier with a resistive feedback network. Ideally, the β-circuit in Fig. 6.15 is a voltage-controlled current source with infinite input resistance and infinite output resistance. In practice, it is often implemented just using a resistor R_f as shown in Fig. 6.16. In this case, $\beta = i_f/v_o = -1/R_f$, so we need an inverting A-circuit, i.e., $A = -R_T$ as shown in Fig. 6.16 in order to obtain a negative feedback.

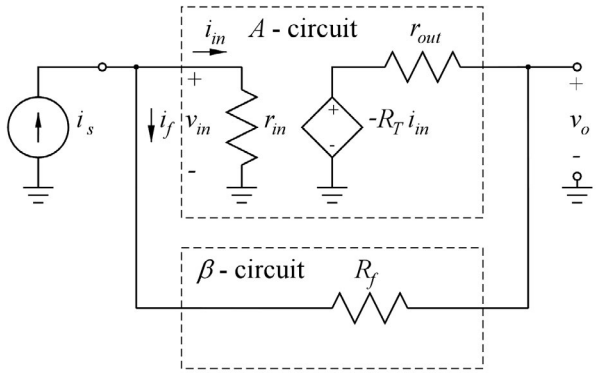


Figure 6.16: Feedback transresistance amplifier with a resistive feedback network.

As the feedback network is not an ideal voltage-controlled current source, it has some influence on the calculations above, similar to what we found for the resistive feedback network for the voltage amplifier. For the transresistance amplifier, we may modify the feedback circuit using a Norton transformation as shown in Fig. 6.17.

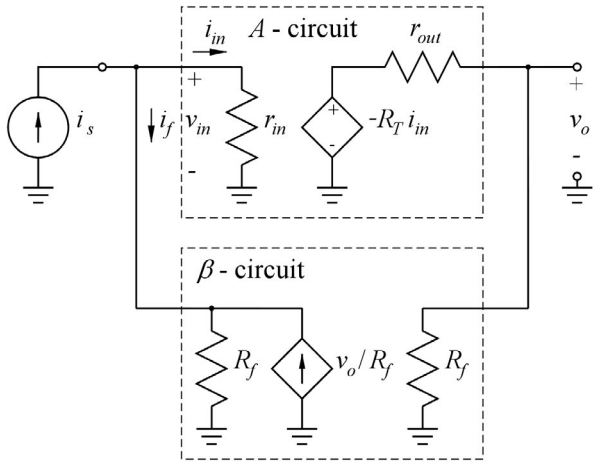


Figure 6.17: Feedback transresistance amplifier with Norton modeling of the β-circuit.

From this, we see that R_f appears in parallel with r_{in}, and R_f is connected as a load to the output. This causes a current division at the input and a voltage division at the output. We may take this into account by modifying the gain of the A-circuit to $-R_T[R_f/(r_{in}+R_f)][R_f/(R_f+r_{out})]$ and the input and output resistances of the A-circuit to $(r_{in} \parallel R_f)$ and $(r_{out} \parallel R_f)$, respectively. However, again the detailed calculations using this approach are beyond the scope of this book and the reader is referred to Sedra & Smith (2016).

Feedback amplifiers with current sensing. The feedback amplifiers examined so far have utilized voltage sensing at the output, meaning that the output voltage is fed into the β-circuit. Another possibility is to use current sensing where the output current is fed into the β-circuit. For this, the β-circuit is placed in series with the output and the load so that the output current flows through both the load and the β-network.

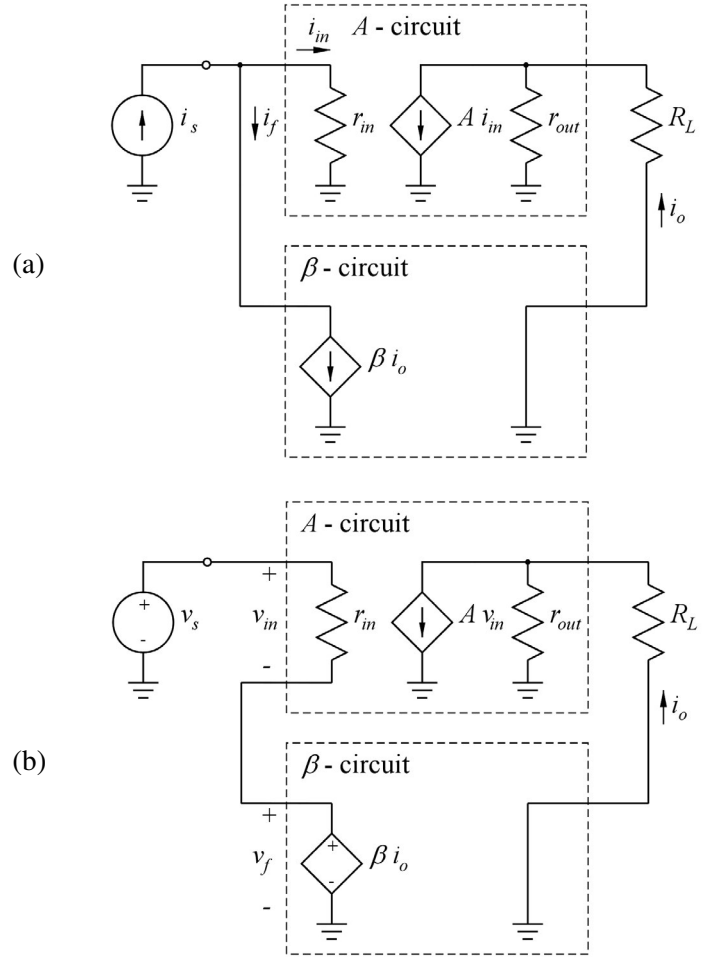


Figure 6.18: Feedback amplifiers with current sensing and current mixing (a) or current sensing and voltage mixing (b).

Figure 6.18 shows systems using current sensing. In Fig. 6.18(a), the input signal is a current i_s, so the feedback signal i_f is also a current, implying that current mixing is used at the input and that the β-circuit is a current divider, ideally a current-controlled current source as shown in Fig. 6.18(a). Electrically, this configuration uses a shunt connection at the input and a series connection at the output, so it is called a shunt-series feedback system.

In Fig. 6.18(b), the input signal is a voltage v_s, so the feedback signal v_f is also a voltage, implying that voltage mixing is used at the input and that the β-circuit is a transresistance, ideally a current-controlled voltage source as shown in Fig. 6.18(b). Electrically, this configuration uses a series connection at the input and a series connection at the output, so it is called a series-series feedback system.

For the circuits shown in Fig. 6.18, ideally we have $R_L = 0$, and it can be shown that the current sensing at the output causes an increase of the output resistance by a factor $(1 + \beta A)$. As a current-mode output ideally has an infinite output resistance, the feedback improves the output resistance by a factor equal to the amount of feedback.

Just as for the circuits using voltage sensing, the β-circuit is often realized using passive components, e.g., resistors, and the loading effects on the A-circuit caused by the β-circuit and a non-zero R_L can be taken into account in a way similar to what has been shown in Figs. 6.14 and 6.17 for the circuits using voltage sensing. The details of this are beyond the scope of this book and the reader is referred to Sedra & Smith (2016).

6.4 The inverting amplifier

A feedback amplifier frequently encountered in analog electronics is the inverting amplifier shown in Fig. 6.19. Using an ideal opamp, we find a gain of $A_v = v_o/v_s = -R_2/R_1$. The input resistance is R_1 and the output resistance is 0. Clearly, the system uses feedback since the resistor R_2 connects the output signal back to the input of the opamp.

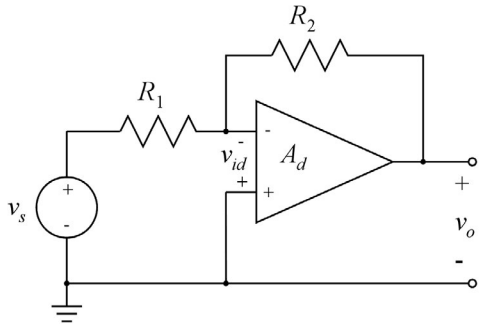


Figure 6.19: An inverting amplifier using an opamp.

The fact that both the input signal v_s and the output signal v_o are voltages may cause confusion. For a voltage input in a feedback system, we use voltage mixing corresponding to a series connection of the feedback signal and the input signal. However, this is not what we see in Fig. 6.19. Rather, we observe a shunt connection at the input of the opamp. This suggests that current mixing is used at the input to the opamp, so in terms of feedback analysis, the feedback amplifier is a transresistance amplifier with current input and voltage output.

In order to change the input signal into a current signal, we can apply a Norton transformation to v_s and R_1. The Norton current is $i_s = v_s/R_1$. Also, let us assume that the opamp has a finite differential voltage gain A_d. While still assuming an infinite input resistance and a zero output resistance for the opamp, we arrive at the equivalent model shown in Fig. 6.20. Here the resistor R_1 has been moved into the A-circuit.

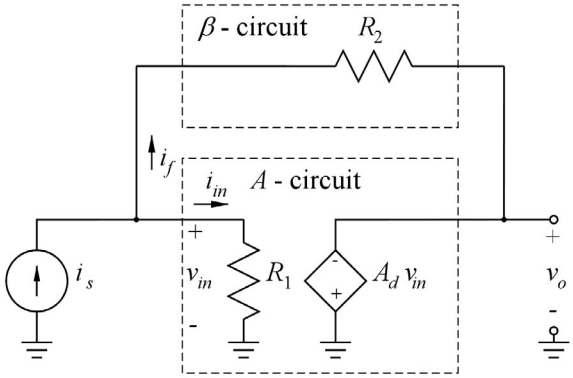


Figure 6.20: Equivalent model for the inverting amplifier.

Next, we may use a Norton transformation of the β-circuit as shown in Fig. 6.17. Moving the resistors from the Norton equivalent into the A-network, we find the circuit shown in Fig. 6.21. From this, we find $\beta = -1/R_2$, $A = v_o/i_{in} = -A_d\,(R_2 \parallel R_1)$ and $r_{in} = R_2 \parallel R_1$. We can then find the closed-loop gain $A_{CL} = v_o/i_s$ using Eq. (6.1):

$$A_{CL} = \frac{v_o}{i_s} = \frac{A}{1+A\beta} = \frac{-A_d\,(R_2 \parallel R_1)}{1+A_d\,(R_2 \parallel R_1)/R_2} \tag{6.18}$$

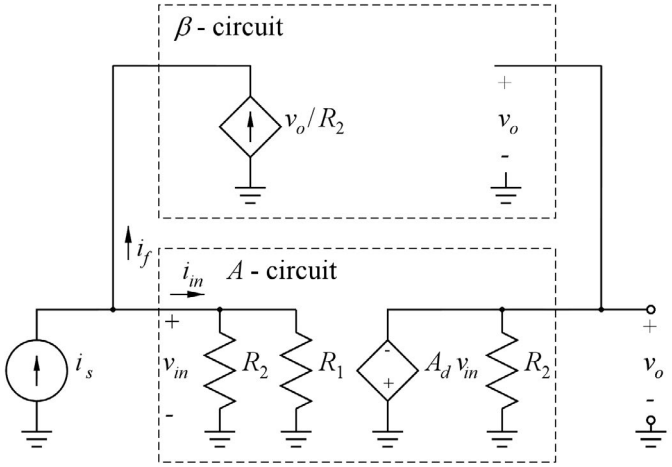


Figure 6.21: Equivalent model for the inverting amplifier with β-circuit transformed using a Norton equivalent.

Using $i_s = v_s/R_1$, we find

$$A_v = \frac{v_o}{v_s} = \frac{A_{CL}}{R_1} = \left(-\frac{R_2}{R_1}\right)\left(\frac{A_d\,(R_1/(R_1+R_2))}{1+A_d\,(R_1/(R_1+R_2))}\right) \tag{6.19}$$

Introducing $G_0 = -R_2/R_1$ as the gain obtained with an ideal opamp with infinite gain, Eq. (6.19) may be re-written

$$A_v = \frac{v_o}{v_s} = G_0\,\frac{1}{1+(1+|G_0|)/A_d} \tag{6.20}$$

Notice the similarity between this expression and the result found for the noninverting amplifier in Eq. (6.3).

We may also find the resistance seen by the current source i_s using Eq. (6.17). We find

$$r_{inCL} = \frac{r_{in}}{1+\beta A} = \frac{R_2 \parallel R_1}{1+A_d\,(R_2 \parallel R_1)/R_2} = R_2 \left(\frac{1}{A_d + |G_0| + 1} \right) \tag{6.21}$$

This resistance is equal to $(R_1 \parallel r_{id})$ where r_{id} is the resistance looking into the inverting input node of the opamp in Fig. 6.19. Using Eq. (6.21), we find

$$r_{inCL} = R_1 \parallel r_{id} \quad \Rightarrow \quad \frac{1}{r_{inCL}} = \frac{1}{R_1} + \frac{1}{r_{id}}$$

$$\Rightarrow \ \frac{1}{r_{id}} = \frac{1}{r_{inCL}} - \frac{1}{R_1} = \frac{A_d + |G_0| + 1}{R_2} \quad - \quad \frac{|G_0|}{R_2} = \frac{A_d + 1}{R_2} \ \Rightarrow \ r_{id} = \frac{R_2}{A_d + 1} \tag{6.22}$$

Incidentally, the feedback method as used above is not always a shortcut to the solution. You may verify Eq. (6.22) simply by using a loop equation to find r_{id} as v_{in}/i_s for the circuit shown in Fig. 6.16, see Problem 6.6.

The input resistance of the feedback amplifier is increased by r_{id} compared to the ideal value of R_1. Since this increase is normally small, the feedback ensures a low input resistance into the inverting opamp input even with an infinite input resistance of the opamp.

In the analysis above, we have assumed an infinite input resistance and a zero output resistance of the opamp. A finite input resistance of the opamp will appear in the A-circuit in parallel with R_1 and R_2 and modify r_{inCL} accordingly. A non-zero output resistance will appear in series with the controlled voltage source $A_d\,v_{in}$, causing a voltage division between R_2 and the output resistance which reduces the loop gain. Also, it implies that the output resistance of the amplifier with feedback is no longer zero but must be found by dividing the output resistance of the A-circuit by the amount of feedback. This is an exercise left for the reader.

A simulation example. We may illustrate the properties of the inverting amplifier by a simple LTspice simulation example. Assume that the inverting amplifier from Fig. 6.19 is built from an opamp with a gain A_d, an infinite input resistance and an output resistance r_{out}. In order to obtain a gain of -10 V/V, we select $R_2 = 10$ kΩ and $R_1 = 1$ kΩ. An LTspice schematic for this circuit is shown in Fig. 6.22. The gain of the opamp is defined as a parameter '`Ad`' which is stepped (logarithmically) from a value of 100 to a value of 10^5 and the output resistance of the opamp is defined as a parameter '`rout`' which is stepped (logaritmically) from a value of 100 Ω to a value of 10 kΩ.

For simulating gain, input resistance and output resistance of the amplifier with feedback, we run a '.tf' simulation. From this, we can plot gain, input resistance and output resistance versus the gain of the opamp, and we can plot curves for the different values of r_{out} specified by the '.step' directive for '`rout`'.

Figure 6.23 shows the plots where the green curves are for $r_{out} = 100$ Ω, the blue curves are for $r_{out} = 1$ kΩ and the red curves are for $r_{out} = 10$ kΩ. We see that the closed-loop gain is very close to the target value of -10 V/V if only $A_d > 10^4 \sim 80$ dB. Also input resistance and output resistance are very close to the ideal values when $A_d > 10^4$ but for smaller values of A_d, the gain falls below the target value and the input resistance and output resistance differ from the ideal values.

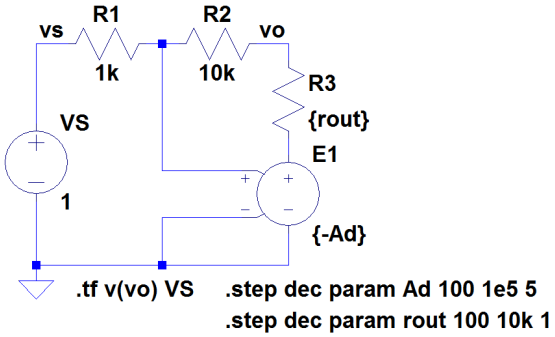


Figure 6.22: LTspice schematic for the inverting amplifier.

As an example, we may compare the simulated value of gain and input resistance to the values calculated from Eqs. (6.20) and (6.22) for $A_d = 100$. Using the simulations for $r_{out} = 100\ \Omega$, we find a gain of -9.0 V/V from the simulation and a gain of -9.0 V/V from Eq. (6.20). The input resistance found from the simulation is $r_{in} = 1.1\ \text{k}\Omega$, i.e., the increase in input resistance is 100 Ω compared to the ideal value. From Eq. (6.22), we may calculate the increase in input resistance to be 99 Ω. The difference is caused by the non-zero value of r_{out}. Taking this into account in Eq. (6.22), R_2 should be replaced by $R_2 + r_{out} = 10.1\ \text{k}\Omega$, resulting in an increase in input resistance of $10.1\ \text{k}\Omega/101 = 100\ \Omega$, exactly matching the simulated value.

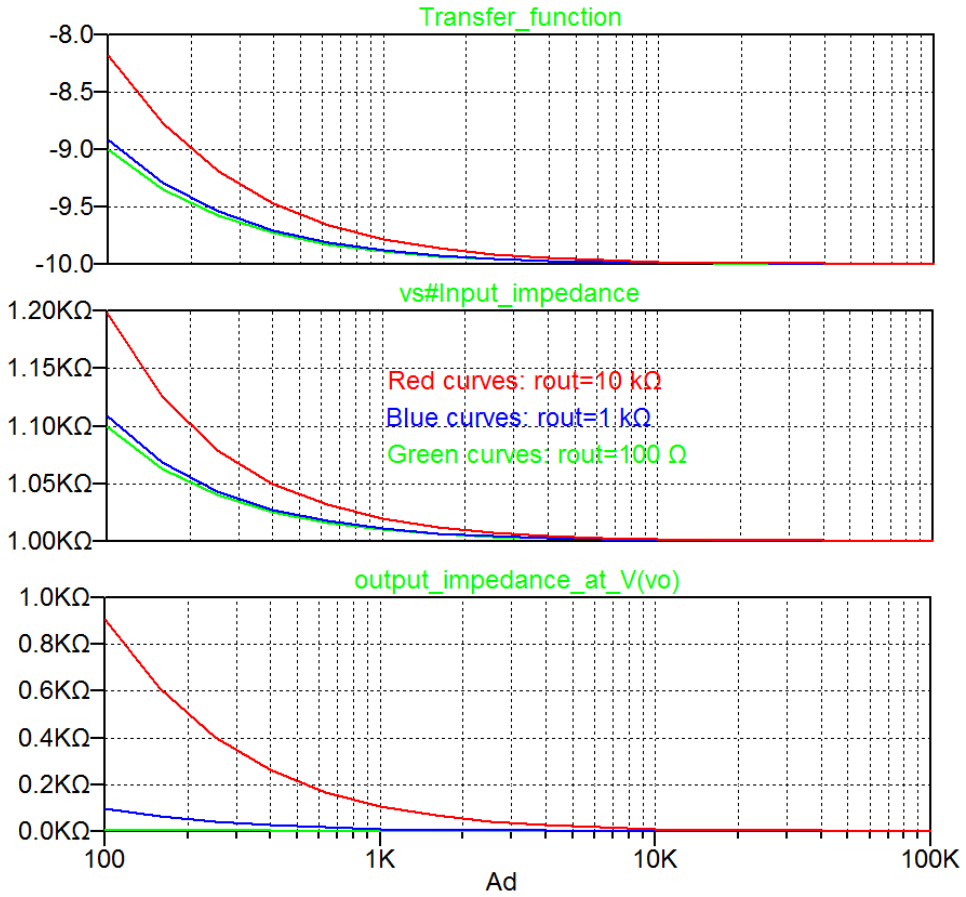


Figure 6.23: Gain, input resistance and output resistance for the inverting amplifier shown in Fig. 6.22.

6.5 Stability

We saw in Chapter 5.3 that feedback may cause instability if a phase shift of the loop gain causes the feedback to be positive rather than negative. A general requirement for stability of a linear amplifier system is that the poles of the transfer function are in the left half-plane of the complex frequency plane, the s-plane (Lathi 2009). For a system with feedback, the poles are found as the roots of the equation $(1+\beta(s)A(s))=0$. This has been reformulated by Harry Nyquist (1889-1976), an American electronics engineer, into what is known as the Nyquist stability criterion (Nyquist 1932).

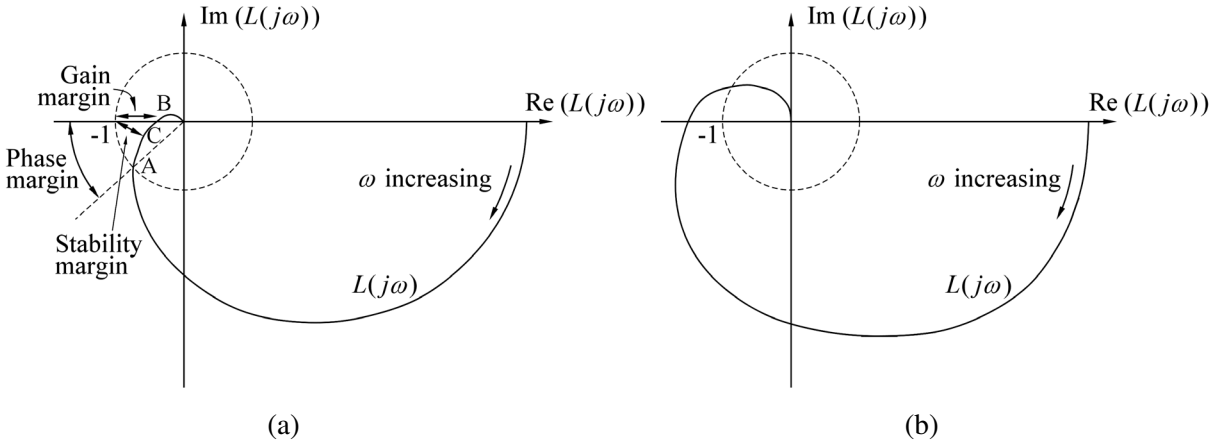


Figure 6.24: Nyquist plot of the loop gain for a stable amplifier (a) and an unstable amplifier (b).

A Nyquist plot of the loop gain $L(j\omega)=\beta(j\omega)A(j\omega)$ shows $L(j\omega)$ in the complex plane as a function of the angular frequency ω when ω is swept from 0 to ∞. The stability criterion states that if the plot of $L(j\omega)$ encircles the point $(-1,0)$, the amplifier is unstable. Figure 6.24 shows examples of loop gain plots for a stable amplifier (a) and an unstable amplifier (b). From the Nyquist plot, we can find the point where the loop gain plot intersects the unit circle, point A in Fig. 6.24(a). The phase angle left until a phase shift of 180° is reached is defined as the phase margin PM.

Point B in Fig. 6.24(a) is the intersection between the real axis and the loop gain plot. For a stable system, this point is reached at a higher frequency than point A and it falls within the unit circle. In point B, the loop gain is less than 0 dB and the absolute value of the loop gain measured in dB in point B is called the gain margin. The gain margin is only defined for systems where the phase shift of $L(j\omega)$ becomes larger than 180°.

Point C in Fig. 6.24(a) is the point on the loop gain plot that has the shortest distance to the point $(-1,0)$ and this distance is called the stability margin s_m (Åström & Murray 2021).

Normally in electronics engineering, Bode plots rather than Nyquist plots are used to show the frequency response. Figure 6.25 shows a Bode plot for a stable amplifier, corresponding to the Nyquist plot shown in Fig. 6.24(a). From the Bode plot, we find the phase margin as

$$\text{PM} = 180^\circ - (-\angle L(j\omega_t)) = 180^\circ + \angle L(j\omega_t) \tag{6.23}$$

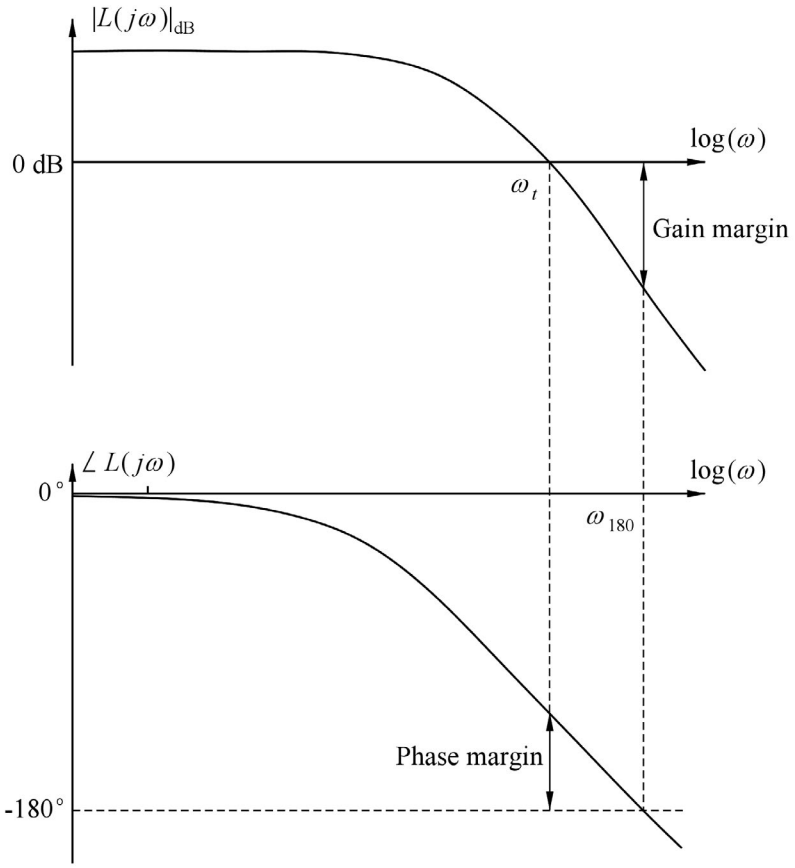


Figure 6.25: Bode plot of the loop gain for a stable amplifier.

where ω_t is the unity-gain frequency of the loop gain. From the Bode plot, we can also find the gain margin. It is equal to $-|L(\omega_{180})|_{\text{dB}}$ where ω_{180} is the frequency for which $\angle L(\omega) = -180°$.

We notice that the stability margin s_m cannot be found directly from the Bode plot of the loop gain but LTspice can be used to find the stability margin either from a Nyquist plot or by using a '.measure' directive to find s_m from a simulation of the loop gain.

In practice, both the stability margin and the phase margin are important parameters for characterizing the stability of a feedback amplifier. We will see in the following that both frequency response and transient response are closely related to the phase margin.

A first-order system. A system where the loop gain has only a single pole is called a first-order system. This is obtained from an amplifier with a single pole and a feedback circuit with a real, positive value of β. We examined such a system in Section 6.2 and we found that the amplifier with feedback is also a first-order system with a single pole at the frequency $\omega_{pCL} = \omega_p(1+\beta A_0)$ where A_0 and ω_p are the low-frequency gain and the pole frequency, respectively, of the basic amplifier.

Figure 6.26 shows a Bode plot of the open-loop gain A, the loop gain L and the closed-loop gain A_{CL}

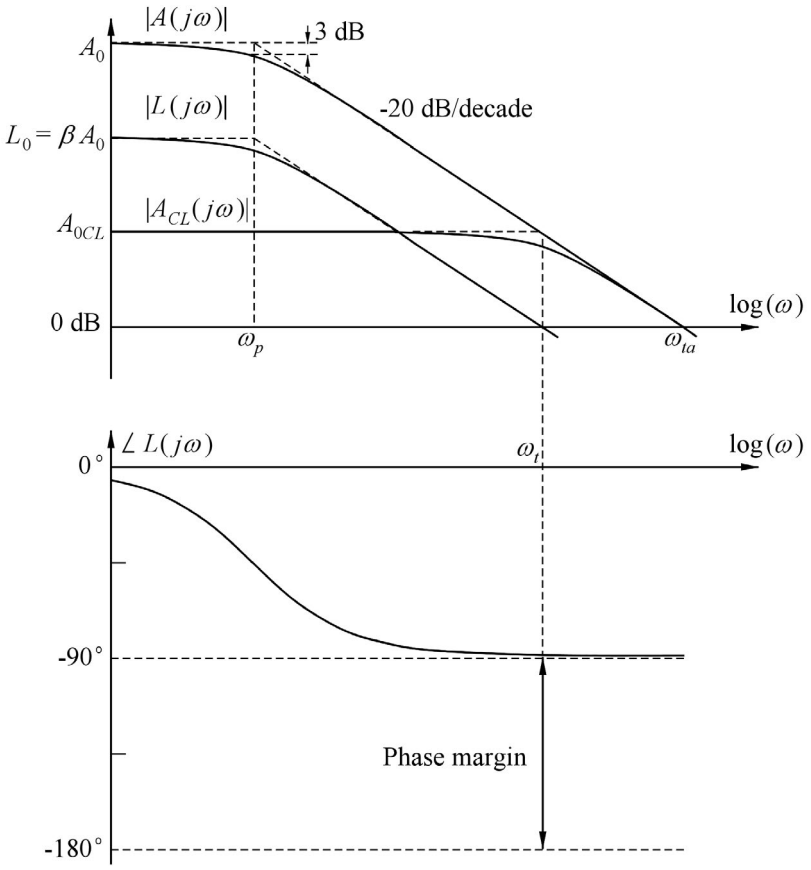


Figure 6.26: Bode plot for a first-order feedback system.

of the system. The low-frequency loop gain is $L_0 = \beta A_0$ and assuming $\beta A_0 \gg 1$, the unity-gain frequency of the loop gain is $\omega_t = \omega_p L_0 = \omega_p \beta A_0$. We notice that the unity-gain frequency of the basic amplifier equals the gain-bandwidth product $\omega_{ta} = A_0 \omega_p$, so $\omega_t = \beta \omega_{ta}$. Also, from Eq. (6.8), we find $L = L_0/(1 + s/\omega_p)$, giving a phase margin of

$$\text{PM} = 180° - \arctan(\omega_t/\omega_p) = 180° - \arctan(L_0) = 90° + \arctan(1/L_0) \tag{6.24}$$

so the phase margin is at least 90° and a first-order system is always stable. As the phase shift of the loop gain does not exceed $-90°$, it is also apparent that the shortest distance from the Nyquist plot of the loop gain to the point $(-1, 0)$ is equal to 1, so for a first-order system, the stability margin is $s_m = 1$.

For a basic amplifier with a very high low-frequency gain, often the exact values of gain and bandwidth are not very well defined, depending on transistor parameters, in particular small-signal output resistances, which exhibit large variations due to variations in the manufacturing process. However, the gain-bandwidth product GBW normally varies less with process variations, depending on capacitances and transconductances only, rather than output resistances. For a single-pole amplifier, we have $\text{GBW} = \omega_{ta}/(2\pi)$. With a closed-loop gain which is $\sim 1/\beta$, we thus find a closed-loop bandwidth which is the gain-bandwidth product of the basic amplifier divided by the low-frequency closed-loop gain.

The transient step response of a first-order system corresponds to the transient response of an *RC*-network with a time constant τ (see the analysis in Chapter 2.4) where

$$\tau = \frac{1}{\omega_{pCL}} = \frac{1}{\omega_p(1+\beta A_0)} \simeq \frac{1}{\omega_t} \tag{6.25}$$

With a transient step response given by $v_O = V_o(1-\exp(-t/\tau))$, we find that the time t_δ required to reach $(100-\delta)\%$ of the final value V_o is $t_\delta = -\tau\ln(\delta/100)$. With $\delta = 1\%$, we find $t_\delta = 4.6\cdot\tau$, so the time required to reach 99% of the final value is 4.6 times τ or 4.6 divided by the -3 dB bandwidth ω_{pCL} or 0.73 divided by the bandwidth f_{pCL} in Hz. Within a time equal to the reciprocal of the bandwidth in Hz, the output reaches about 99.8% of the final value.

Using the standard definition of rise time as the time required to go from 10% to 90% of the final value (Elmore 1948), we find the rise time

$$t_r = 2.197\cdot\tau = 2.197/\omega_{pCL} = 0.35/f_{pCL} \tag{6.26}$$

Since the bandwidth of the amplifier with feedback is larger than the bandwidth of the basic amplifier by a factor equal to the amount of feedback, the settling time for the amplifier with feedback is shorter than the settling time for the basic amplifier. It is improved by a factor equal to the amount of feedback.

A second-order system. A second-order system has two poles in the loop gain. We consider an example where β is real and positive and A is given by the transfer function

$$A(s) = \frac{A_0}{(1+s/\omega_{p1})(1+s/\omega_{p2})} \tag{6.27}$$

From this, we find

$$L(s) = \frac{L_0}{(1+s/\omega_{p1})(1+s/\omega_{p2})} \tag{6.28}$$

where $L_0 = \beta A_0$. The phase margin is

$$\text{PM} = 180^\circ - \arctan(\omega_t/\omega_{p1}) - \arctan(\omega_t/\omega_{p2}) = \arctan(\omega_{p1}/\omega_t) + \arctan(\omega_{p2}/\omega_t) \tag{6.29}$$

where ω_t is the unity-gain frequency of the loop gain.

A Bode plot of the open-loop gain A and the loop gain L is shown in Fig. 6.27 and an overview of the different parameters used for describing the frequency response of the feedback amplifier system is given in Table 6.1.

The plot also shows the closed-loop gain. We notice that at very high frequencies where $A\beta \ll 1$, the closed-loop gain approaches the open-loop gain of the amplifier as expected from Eq. (6.1). From this plot, we immediately see that the phase margin is always positive, i.e., the amplifier is stable. However, depending on the values of L_0, ω_{p1} and ω_{p2}, we may find a small phase margin and although the feedback

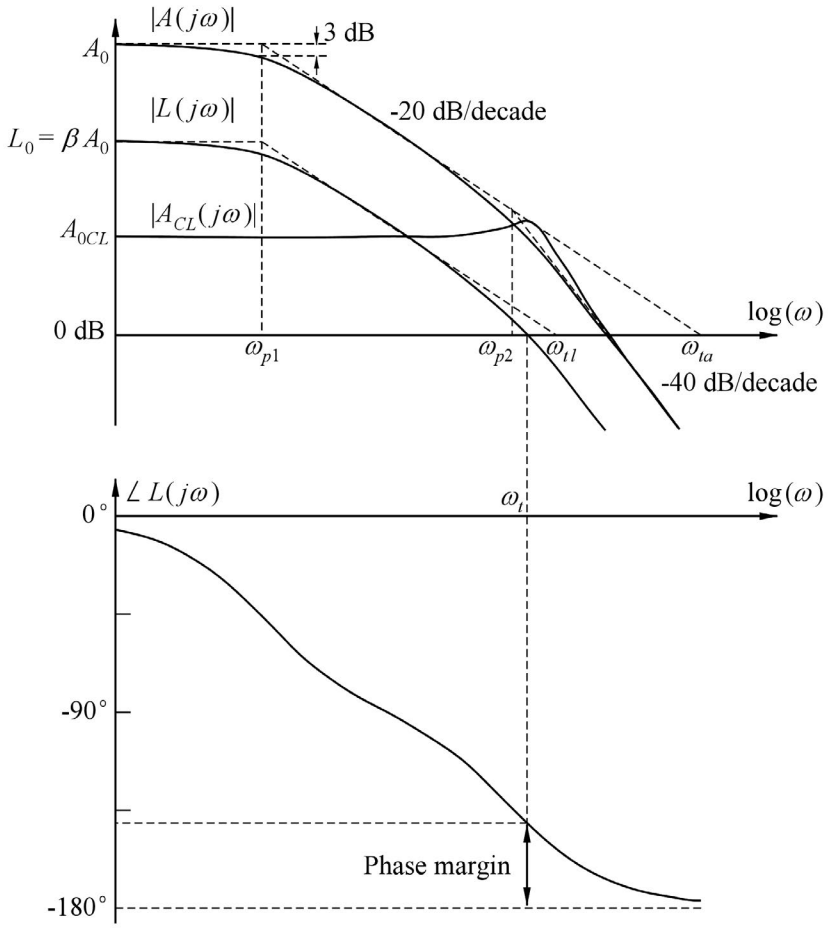


Figure 6.27: Bode plot for a second-order feedback system.

amplifier is stable, it may exhibit peaking in the frequency response as shown in Fig. 6.27 and overshoot in the step response, see for instance Fig. 5.16.

Overshoot occurs when the two poles of a second-order system are complex conjugate poles rather than two real poles, and peaking in the frequency response occurs with complex conjugate poles where the imaginary part is larger than the real part (Chan Carusone, Johns & Martin 2012). Thus, in order to investigate the frequency response and transient response of the second-order system in more detail, we find the poles of the system with feedback by solving the equation

$$1 + L(s) = 1 + \beta A(s) = 0 \tag{6.30}$$

Inserting $A(s)$ from Eq. (6.27), we find

$$\begin{aligned} 1 + \frac{\beta A_0}{(1 + s/\omega_{p1})(1 + s/\omega_{p2})} &= 0 \\ \Rightarrow \quad (1 + s/\omega_{p1})(1 + s/\omega_{p2}) + \beta A_0 &= 0 \\ \Rightarrow \quad s^2 + s(\omega_{p1} + \omega_{p2}) + (1 + \beta A_0)\omega_{p1}\,\omega_{p2} &= 0 \end{aligned} \tag{6.31}$$

Parameter	First-order system	Second-order system	Higher-order system
Amplifier low-frequency gain	A_0	A_0	A_0
Feedback factor	β	β	β
Loop gain low-frequency gain	L_0 $L_0 = \beta A_0$	L_0 $L_0 = \beta A_0$	L_0 $L_0 = \beta A_0$
Dominant pole of loop gain and amplifier	ω_p	ω_{p1}	ω_{p1}
Gain-bandwidth product of amplifier	ω_{ta} $\omega_{ta} = A_0\,\omega_p$	ω_{ta} $\omega_{ta} = A_0\,\omega_{p1}$	ω_{ta} $\omega_{ta} = A_0\,\omega_{p1}$
Gain-bandwidth product of loop gain	ω_{tl} $\omega_{tl} = \beta\,\omega_{ta}$	ω_{tl} $\omega_{tl} = \beta\,\omega_{ta}$	ω_{tl} $\omega_{tl} = \beta\,\omega_{ta}$
Non-dominant pole(s)	–	ω_{p2}	$\omega_{pi,\ i\geq 2}$, ω_{eq}
Unity-gain frequency of loop gain	ω_t $\omega_t = \omega_{tl}$	ω_t $\omega_t < \omega_{tl}$	ω_t $\omega_t < \omega_{tl}$

Table 6.1: Parameter definitions for the feedback amplifier system.

This is a quadratic equation with the solutions

$$s = -\frac{1}{2}(\omega_{p1}+\omega_{p2}) \pm \frac{1}{2}\sqrt{(\omega_{p1}+\omega_{p2})^2 - 4(1+\beta A_0)\omega_{p1}\,\omega_{p2}} \tag{6.32}$$

From Eq. (6.32), we see that the poles are complex if

$$(\omega_{p1}+\omega_{p2})^2 < 4(1+\beta A_0)\omega_{p1}\,\omega_{p2} \;\Rightarrow\; 1+\beta A_0 > \frac{(\omega_{p1}+\omega_{p2})^2}{4\omega_{p1}\,\omega_{p2}} \tag{6.33}$$

As for the first-order system, the exact values of low-frequency gain and bandwidth are often not very precisely defined whereas the product of low-frequency gain and bandwidth, the gain-bandwidth product GBW, is more well-defined. Using $\omega_{ta} = A_0\,\omega_{p1}$ and assuming $\omega_{p2} \gg \omega_{p1}$ and $\beta A_0 \gg 1$, we may reformulate Eq. (6.33) into

$$\omega_{p1}(1+\beta A_0) > \omega_{p2}/4 \;\Rightarrow\; \beta\,\omega_{ta} > \omega_{p2}/4 \;\Rightarrow\; \omega_{p2} < 4\,\omega_{tl} \tag{6.34}$$

where we have introduced ω_{tl} as the gain-bandwidth product of the loop gain. Notice that $\omega_{tl} = \beta A_0\,\omega_{p1}$ is larger than ω_t, the unity-gain frequency of the loop gain, see Fig. 6.27. Thus, Eq. (6.34) shows that overshoot in the transient response occurs for $\omega_{p2} < 4\,\omega_{tl}$. A system with overshoot is called an underdamped system. If $\omega_{p2} = 4\,\omega_{tl}$, the system is said to be critically damped.

With $\omega_{p2} \gg \omega_{p1}$ and $\omega_{tl} = \beta A_0\,\omega_{p1}$, we can also from Eq. (6.32) find that the real part of the solution is equal to the imaginary part of the solution for $\omega_{p2} = 2\,\omega_{tl}$. Thus, peaking in the frequency response occurs for $\omega_{p2} < 2\,\omega_{tl}$.

A system with no overshoot in the transient step response is called an overdamped system, and for a strongly overdamped system ($\omega_{p2} \gg \omega_{tl}$), the phase margin is rather close to 90° and the frequency response and transient response resemble the response of a first-order system with a closed-loop bandwidth equal to ω_{tl}, i.e., a closed-loop bandwidth equal to the gain-bandwidth product of the basic amplifier divided by the low-frequency closed-loop gain.

When ω_{p2} is reduced, both the frequency response and the transient step response are modified. We can analyze the responses as functions of the ratio ω_{p2}/ω_{tl}. For this analysis, we consider a system with the loop gain

$$L(s) = \frac{\omega_{tl}}{s(1+s/\omega_{p2})} \tag{6.35}$$

This describes a second-order system with $\omega_{p1} = 0$, having infinite gain for $\omega \to 0$ and a gain-bandwidth product of ω_{tl}. The closed-loop low-frequency gain is $1/\beta$. Even with finite values of ω_{p1} and A_0, Eq. (6.35) is a good approximation to the loop gain for frequencies around ω_t or larger if only the low-frequency loop gain $L_0 = \beta A_0$ is larger than about 30 dB. Actually, finite values of L_0 and ω_{p1} add the term $\arctan(\omega_{p1}/\omega_t)$ to the phase margin, see Eq. (6.29), and with ω_t typically being in the range $\omega_{tl}/2$ to ω_{tl}, the addition to the phase margin is in the range $\arctan(\omega_{p1}/\omega_{tl}) = \arctan(1/L_0)$ to $\arctan(2/L_0)$.

Using the filter blocks defined by Bruun (2020, Tutorial 5.3), we can model the system in LTspice as shown in Fig. 6.28. For the first filter block 'LP0', we define the gain-bandwidth product '`ftl`' = 1 in order to normalize the response of the system with respect to $f_{tl} = \omega_{tl}/(2\pi)$. For the second filter block 'LP1', we define the pole frequency '`fp2`' as a parameter which can be stepped through the range of $\omega_{p2}/\omega_{tl} = f_{p2}/f_{tl}$ to be investigated.

The output signal V_o is connected directly back into the inverting input of the amplifier so that the loop gain of the system is

$$L(jf) = \left(\frac{f_{tl}}{jf}\right)\left(\frac{1}{1+j(f/f_{p2})}\right) \tag{6.36}$$

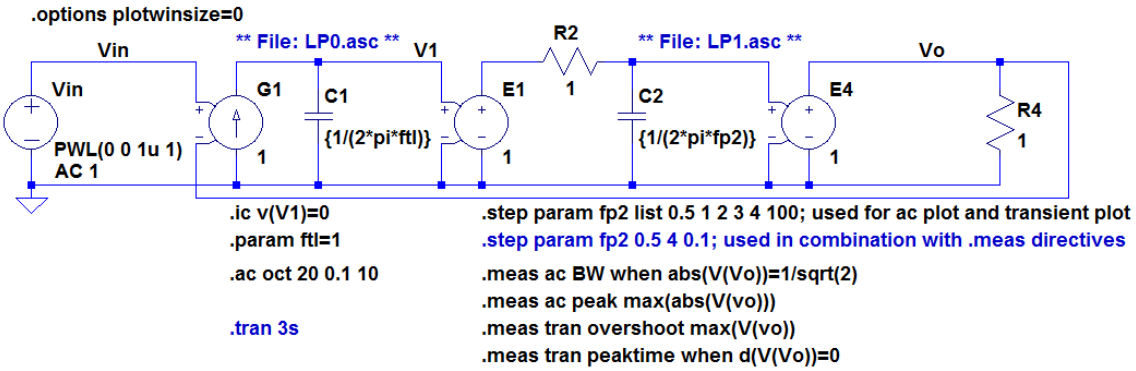


Figure 6.28: LTspice model for a second-order feedback system.

The frequency response has been simulated for a frequency range of $f_{tl}/10$ to $10\,f_{tl}$, i.e., in the '.ac' simulation, the start frequency is 0.1 Hz and the stop frequency is 10 Hz. The pole frequency f_{p2} has been stepped through selected values corresponding to specific values of ω_{p2}/ω_{tl}. The range of values for ω_{p2}/ω_{tl} is from 0.5 to 4, and also a value of 100 is included as this pushes ω_{p2} so high up in frequency that the system behaves like a first-order system with the closed-loop bandwidth equal to f_{tl}.

Figure 6.29 shows the simulated frequency response. We notice a peak in the response for $\omega_{p2}/\omega_{tl} < 2$. We also see that the -3 dB bandwidth is larger than f_{tl} for all values of ω_{p2}/ω_{tl}, except for $\omega_{p2}/\omega_{tl} = 0.5$ where the -3 dB bandwidth equals f_{tl}, the bandwidth for the first-order system. The maximum bandwidth occurs for $\omega_{p2}/\omega_{tl} = 2$ where the bandwidth is $1.41 \cdot f_{tl}$. Thus, selecting $f_{p2} = 2\,f_{tl}$ results in a frequency response without peaking and with the maximum achievable bandwidth of $1.41 \cdot f_{tl}$. This is called the maximally flat frequency response or the Butterworth response (Sedra & Smith 2016), named after the British physicist Stephen Butterworth (1885-1958). We also note from the figure that for frequencies above the -3 dB bandwidth, the gain falls off with 40 dB per decade of frequency for the two-pole system whereas the single-pole system (purple curve) falls off with 20 dB/decade.

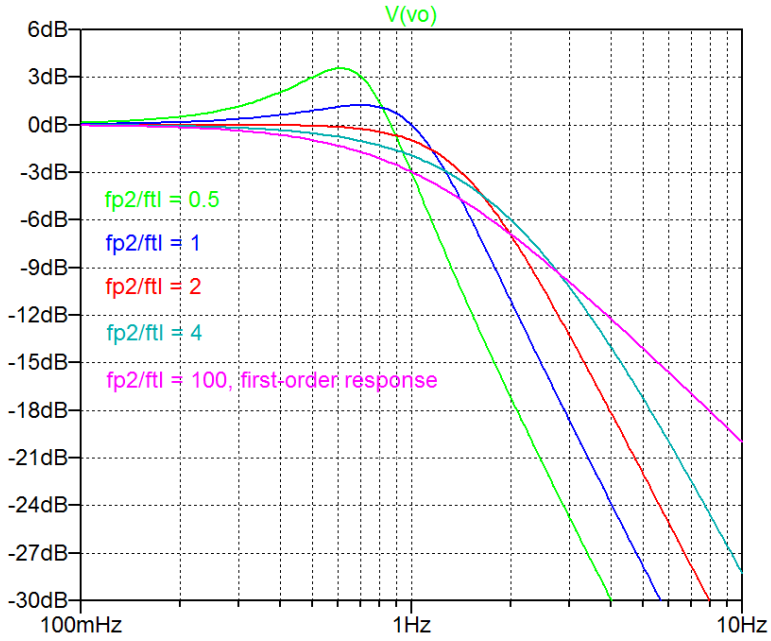
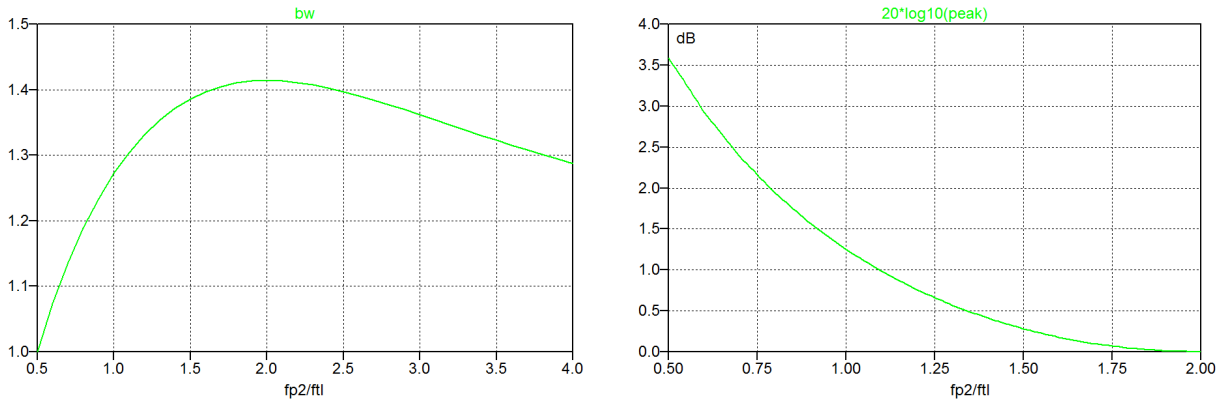


Figure 6.29: Simulated closed-loop frequency response for a second-order feedback system.

Using the '.meas ac' directives shown in Fig. 6.28, we can find the -3 dB bandwidth and the value of the peak as a function of ω_{p2}/ω_{tl}. From the error log file, we can plot the results calculated by the '.meas' directives (Bruun 2020). Figure 6.30 shows both the peaking in dB and the -3 dB bandwidth (normalized to f_{tl}) for ω_{p2}/ω_{tl} varying from 0.5 to 4. In order to show smooth curves, the '.step' directive has been modified to have small steps for this simulation.



(a) (b)

Figure 6.30: Simulated bandwidth (a) and amplitude peaking (b) for a second-order feedback system.

From Fig. 6.28, also a transient simulation with different values of ω_{p2}/ω_{tl} may be run. Figure 6.31 shows the simulated transient response. With $f_{tl} = 1$ Hz, the x-axis is normalized to $1/f_{tl}$. We observe an overshoot in the output for $\omega_{p2}/\omega_{tl} < 4$. For $\omega_{p2}/\omega_{tl} = 0.5$, the overshoot is about 30% and for $\omega_{p2}/\omega_{tl} = 2$ (the maximally flat frequency response), the overshoot is about 4.3%. We notice that for the responses with a significant overshoot, the settling time can be quite long, several times $1/f_{tl}$.

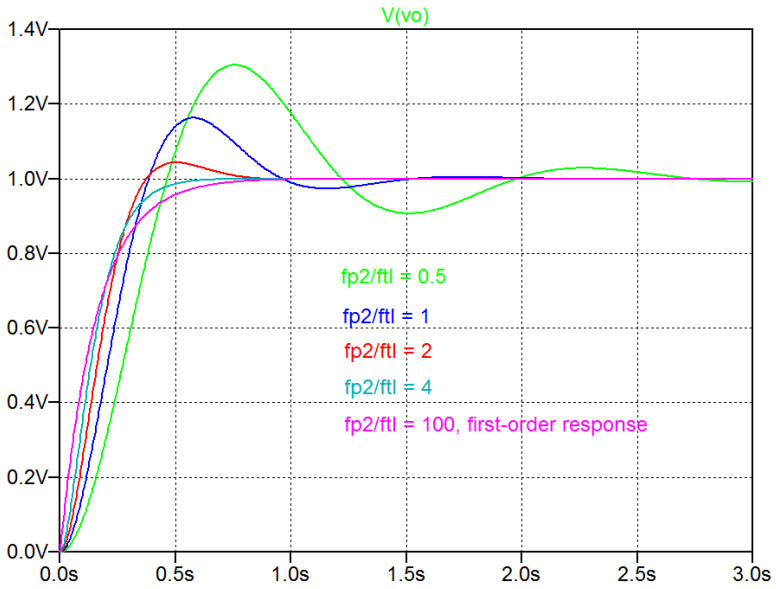
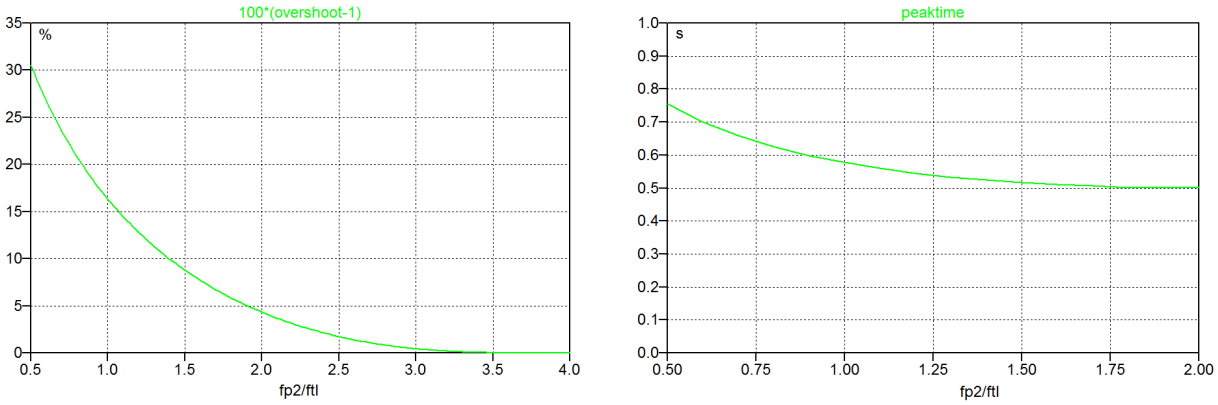


Figure 6.31: Simulated transient step response for a second-order feedback system.

For comparison, the first-order response is also shown (the purple trace), and for this, we recognize the transient response and the rise time described by Eqs. (6.25) and (6.26). For the overdamped response ($\omega_{p2}/\omega_{tl} > 4$) and the maximally flat frequency response ($\omega_{p2}/\omega_{tl} = 2$), we find rise times and settling times comparable to the first-order response, although with overshoot for the maximally flat frequency response.

A few simple parameters to characterize the underdamped response are the overshoot (in %) and the time at which the overshoot occurs. These parameters are simulated using the '.meas tran' directives shown in Fig. 6.28, again with small steps for f_{p2} in order to achieve smooth curves. The plots resulting from the error log file are shown in Fig. 6.32. We see that in general, an underdamped second-order system is slower than a first-order system.



(a) (b)

Figure 6.32: Simulated overshoot (a) and time for overshoot (b) for a second-order feedback system.

While requirements concerning peaking and overshoot may be applied directly in the design of a second-order feedback system, often the specification given for a system is the phase margin. The phase margin is also related to the position of the second pole frequency ω_{p2} relative to the gain-bandwidth product ω_{tl} of the loop gain. The smaller ω_{p2}/ω_{tl}, the smaller the phase margin.

Thus, from a relation between the phase margin PM and ω_{p2}/ω_{tl}, we can directly determine the required positioning of ω_{p2}, and using the results presented in Figs. 6.29 - 6.32, this can be related to the properties of the frequency response and the transient response. We can find a relation between the phase margin and ω_{p2}/ω_{tl} by simulating the loop gain of the system. Figure 6.33 shows the LTspice schematic for simulating the loop gain and Fig. 6.34 shows the resulting simulation of the frequency response.

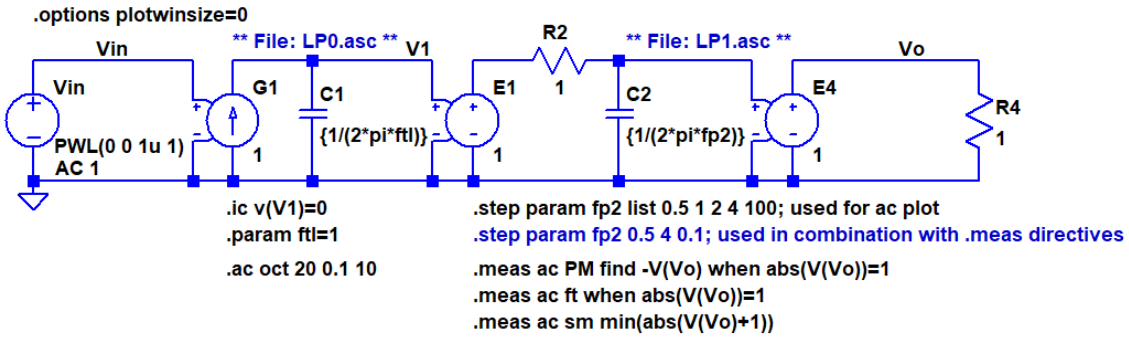


Figure 6.33: LTspice schematic for simulating the loop gain of a second-order feedback system.

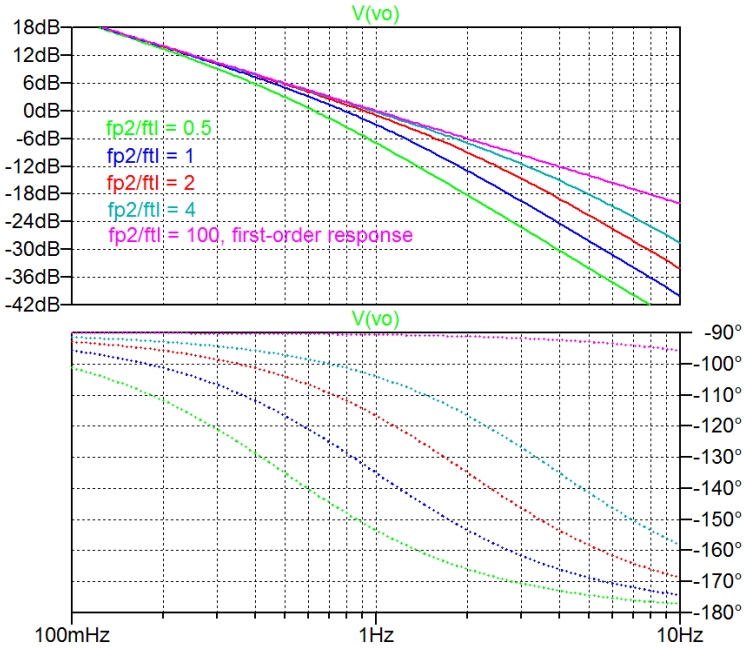


Figure 6.34: Simulated loop gain frequency response for a second-order feedback system.

The phase margin is calculated using the '.meas' directive shown in Fig. 6.33, and Fig. 6.35(a) shows the relation between the phase margin and ω_{p2}/ω_{tl}.

A few specific values are worth noting: A phase margin of 65.5° corresponds to $\omega_{p2}/\omega_{tl} = 2$, the maximally flat frequency response. This is a common goal in the design of a second-order feedback amplifier. Another design goal may be a design with no overshoot in the transient step response. This requires $\omega_{p2}/\omega_{tl} \geq 4$, corresponding to a phase margin of at least 76.3°.

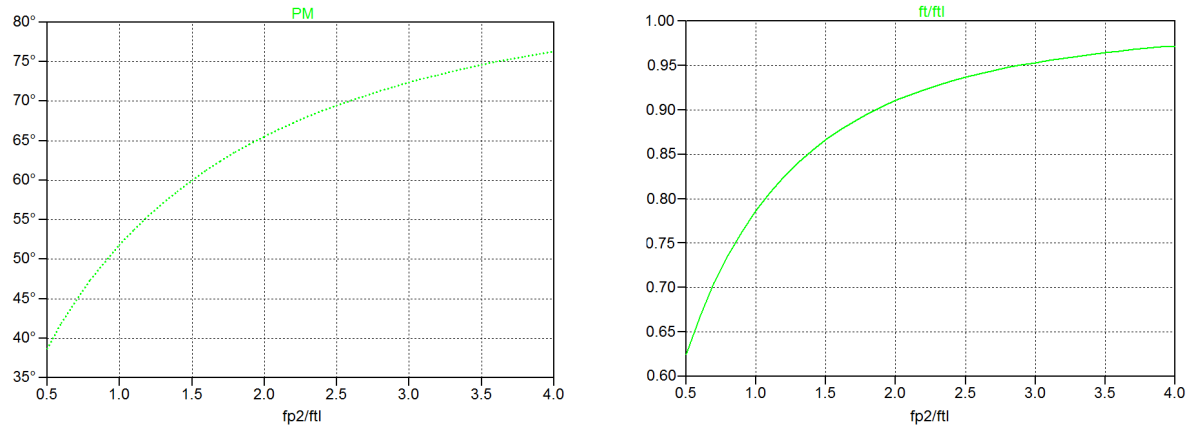
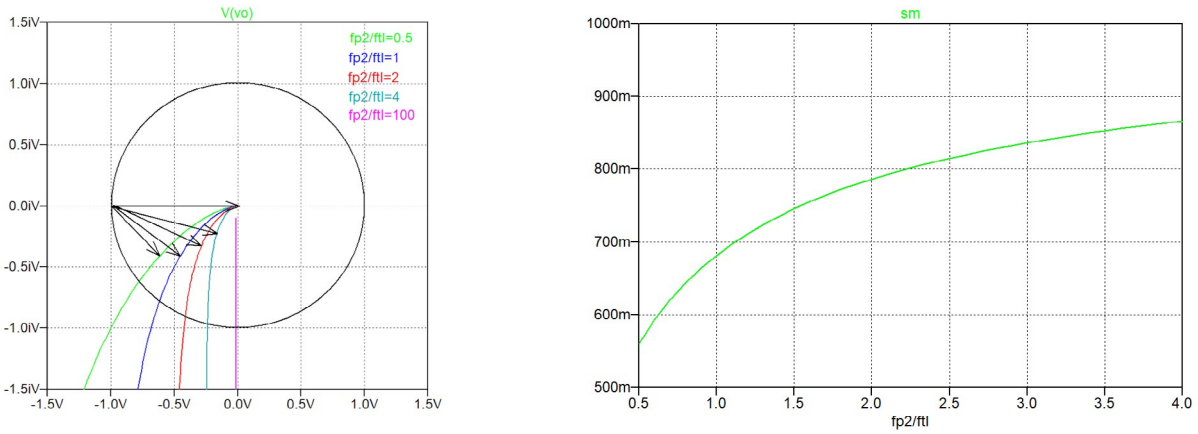


Figure 6.35: Simulated relation between pole location and phase margin (a) and between pole location and loop gain unity-gain frequency (b) for a second-order feedback system.

Extending the simulation in Fig. 6.33 to higher values of ω_{p2}/ω_{tl}, we find that PM = 80° requires $\omega_{p2}/\omega_{tl} = 5.6$ and PM = 85° requires $\omega_{p2}/\omega_{tl} = 11.4$.

From the simulation of the loop gain, we may also find the loop gain unity frequency $f_t = \omega_t/(2\pi)$. Using a '.meas' directive, f_t/f_{tl} can be plotted versus f_{p2}/f_{tl}. The resulting plot is shown in Fig. 6.35(b). We notice that $f_t < f_{tl}$, and for $f_{p2}/f_{tl} \to \infty$ corresponding to a first-order system, $f_t \to f_{tl}$.

The simulation of the loop gain may also be used to find the stability margin s_m. Instead of showing '`V(vo)`' as a Bode plot, Fig. 6.34, we select the Nyquist representation by pointing to the y-axis with a right-click in the plot and changing the representation to Nyquist. This results in the plot shown in Fig. 6.36(a) where both the x-axis and the y-axis have been changed to the range from -1.5 V to $+1.5$ V. We can find the stability margin for the simulated values of f_{p2}/f_{tl} as the lengths of the arrows shown in the plot. The length is the distance from the point $(-1,0)$ to the point $(\mathrm{Re}(L(j\omega)),\mathrm{Im}(L(j\omega)))$ which is closest to the point $(-1,0)$, i.e. $\min(|L(j\omega)+1|)$ when ω is swept from 0 to ∞. Thus, the stability margin can be found using the '.meas ac sm min(abs(V(Vo)+1))' directive shown in Fig. 6.33, and s_m can be plotted versus f_{p2}/f_{tl} as shown in Fig. 6.36(b). Just as the phase margin increases with increasing f_{p2}/f_{tl}, also the stability margin increases with increasing f_{p2}/f_{tl}, and for $f_{p2}/f_{tl} \to \infty$, corresponding to a first-order system, the stability margin approaches 1.



(a) (b)

Figure 6.36: Simulated Nyquist plot of the loop gain of the second-order system (a) and the relation between pole location and stability margin (b).

Combining the simulations of transient overshoot and amplitude peaking versus f_{p2}/f_{tl} and phase margin versus f_{p2}/f_{tl}, we may also plot the overshoot and the amplitude peaking versus the phase margin as shown in Fig. 6.37. This plot may be used to estimate the phase margin from a simulation or a measurement of the closed-loop transient response or the closed-loop frequency response of a feedback system.

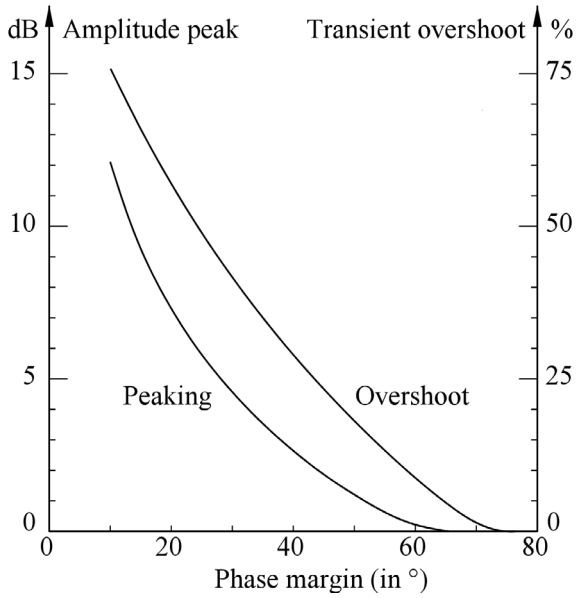


Figure 6.37: Simulated relation between transient overshoot, amplitude peaking and phase margin for a second-order feedback system.

In several textbooks, second-order systems are characterized by a quality factor Q and a resonance frequency ω_0 (Allen & Holberg 2012; Chan Carusone, Johns & Martin 2012; Sedra & Smith 2016).

Assuming $L_0 \gg 1$ and $\omega_{p2} \gg \omega_{p1}$, there is a simple relation between the resonance frequency ω_0 and the quality factor Q and ω_{tl} and ω_{p2} (Chan Carusone, Johns & Martin 2012):

$$\omega_0 \simeq \sqrt{\omega_{tl}\,\omega_{p2}} \tag{6.37}$$

$$Q \simeq \sqrt{(\omega_{tl}/\omega_{p2})} \tag{6.38}$$

For a system specified by a requirement for a specific value of Q, the required positioning of ω_{p2} relative to ω_{tl} is directly found using Eq. (6.38), and the value of ω_{tl} can be found from Fig. 6.30 based on the bandwidth requirements.

Third-order systems. For a third-order system, the loop gain has three poles, implying

$$L(s) = \frac{L_0}{(1+s/\omega_{p1})(1+s/\omega_{p2})(1+s/\omega_{p3})} \tag{6.39}$$

where L_0 is the loop gain at low frequencies. The phase margin is

$$\mathrm{PM} = 180^\circ - \arctan(\omega_t/\omega_{p1}) - \arctan(\omega_t/\omega_{p2}) - \arctan(\omega_t/\omega_{p3}) \tag{6.40}$$

where ω_t is the unity-gain frequency of the loop gain.

Let us assume that the first pole is a dominant pole, i.e. $\omega_{p1} \ll \omega_{p2}$ and $\omega_{p1} \ll \omega_{p3}$. For a system with a large low-frequency loop gain, we may approximate Eq. (6.40) by the expression

$$L(s) = \frac{\omega_{tl}}{s(1+s/\omega_{p2})(1+s/\omega_{p3})} \tag{6.41}$$

for frequencies around ω_t or larger, compare to Eq. (6.35). In Eq. (6.41), $\omega_{tl} = L_0\omega_{p1}$ is the gain-bandwidth product of the loop gain. We notice that the dominant pole causes a phase shift of 90° while the two non-dominant poles provide phase shifts of less than 90°. However, the sum of the phase shifts caused by the non-dominant poles may exceed 90°, causing the system to be unstable. The stability limit may be found by considering the situation $L(s) = -1$, causing the denominator in Eq. (6.1) to be 0 or the phase margin to be zero. This condition is also known as the Barkhausen criterion (Sedra & Smith 2016). Inserting $s = j\omega$, we find

$$\begin{aligned} L(j\omega) &= -1 \Rightarrow 1/L(j\omega) = -1 \\ &\Rightarrow \left(\frac{j\omega}{\omega_{tl}}\right)\left(1+\frac{j\omega}{\omega_{p2}}\right)\left(1+\frac{j\omega}{\omega_{p3}}\right) = -1 \\ &\Rightarrow j\omega\left(1+j\omega\left(\frac{1}{\omega_{p2}}+\frac{1}{\omega_{p3}}\right)-\frac{\omega^2}{\omega_{p2}\omega_{p3}}\right) = -\omega_{tl} \\ &\Rightarrow \left(-\omega^2\left(\frac{1}{\omega_{p2}}+\frac{1}{\omega_{p3}}\right)\right)+j\omega\left(1-\frac{\omega^2}{\omega_{p2}\omega_{p3}}\right) = -\omega_{tl} \end{aligned} \tag{6.42}$$

Equating the imaginary part of Eq. (6.42) to 0, we find for $\omega > 0$ that $\omega = \sqrt{\omega_{p2}\omega_{p3}}$. Inserting this value in Eq. (6.42), we find

$$-\omega_{p2}\omega_{p3}\left(\frac{1}{\omega_{p2}}+\frac{1}{\omega_{p3}}\right) = -\omega_{tl} \Rightarrow \omega_{p2}+\omega_{p3} = \omega_{tl} \tag{6.43}$$

Thus, for a third-order system to be stable, the sum of the pole frequencies for the non-dominant poles must be larger than the gain-bandwidth product of the loop gain.

We may examine the third-order system using simulations similar to those presented for the second-order system. Figure 6.38 shows the LTspice model for simulating the closed-loop gain of the system. As for the circuit shown in Fig. 6.28, the response has been normalized by defining the unity-gain frequency in the first filter block to be '`ftl`' = 1. In order to illustrate the stability requirement defined by Eq. (6.43), we run a transient simulation with a normalized value of 0.4 for the second pole frequency and the third pole frequency stepped through the values 0.5, 0.6 and 2.0. For $f_{p3} = 0.5$ and $f_{p3} = 0.6$, we expect the system to be unstable, implying that a transient step at the input will trigger an oscillation.

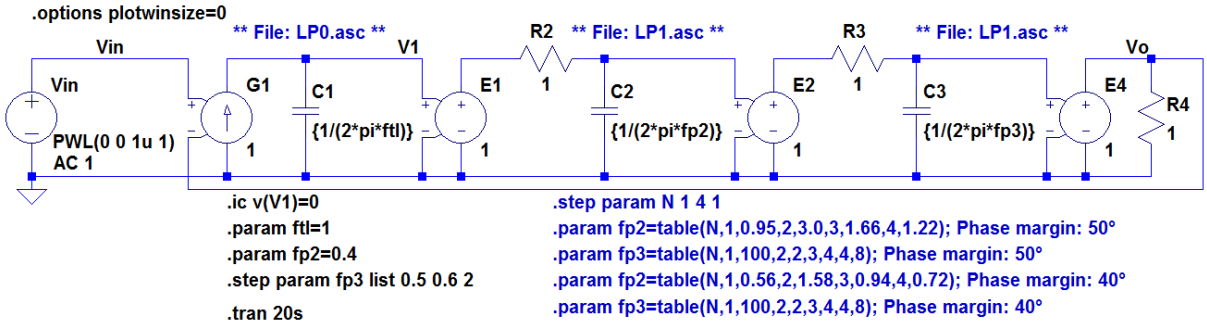


Figure 6.38: LTspice model for a third-order feedback system.

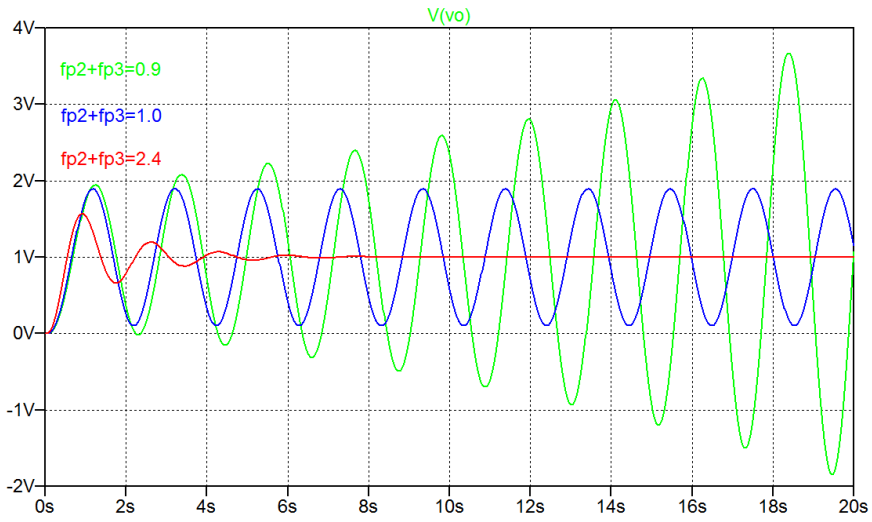


Figure 6.39: Transient response from a third-order feedback system.

For $f_{p3} = 0.5$, we expect an increasing amplitude of the oscillation. For $f_{p3} = 0.6$, we have $f_{p3} + f_{p2} = f_{tl}$, and we expect an oscillation with a constant amplitude and a normalized oscillation frequency of $f = \sqrt{0.4 \times 0.6} = 0.49$, corresponding to a normalized period of $T = 1/f = 2.04$. For $f_{p3} = 2.0$, we have $f_{p2} + f_{p3} > f_{tl}$, so we expect a transient response with a damped oscillation. The transient response shown in Fig. 6.39 confirms these expectation.

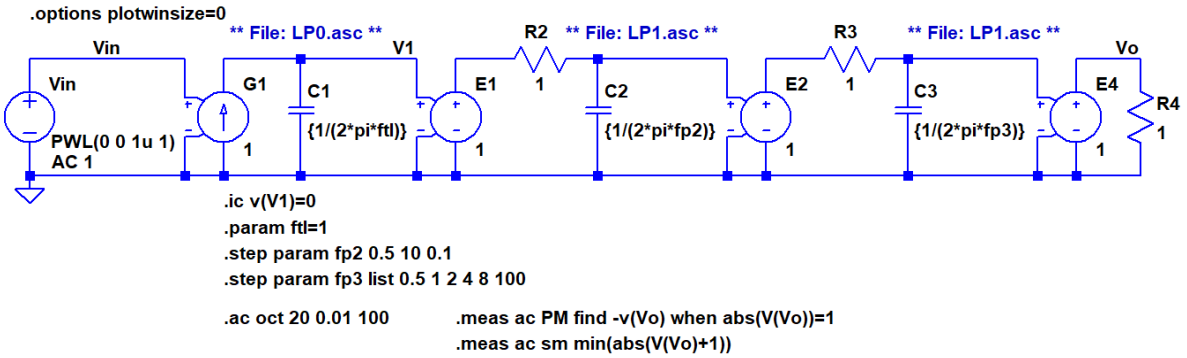


Figure 6.40: LTspice schematic for simulating the loop gain of a third-order feedback system.

For simulating the loop gain in order to evaluate the phase margin and the stability margin, we open the loop as shown in the schematic in Fig. 6.40. Using a '.meas' directive to calculate the phase margin in the same way as we did for the second-order system, we may plot the phase margin as a function of ω_{p2}/ω_{tl} for different values of ω_{p3}/ω_{tl}. Figure 6.41 shows the resulting plot for ω_{p3}/ω_{tl} in the range from 0.5 to 8. Also shown is the phase margin for $\omega_{p3}/\omega_{tl} = 100$, representing a system which is a good approximation to a second-order system, compare to Fig. 6.35.

We notice that for $f_{p2}/f_{tl} = f_{p3}/f_{tl} = 0.5$, i.e., $f_{p2} + f_{p3} = f_{tl}$, the phase margin is 0° as expected, indicating an unstable system. Generally, the presence of a third pole reduces the phase margin compared

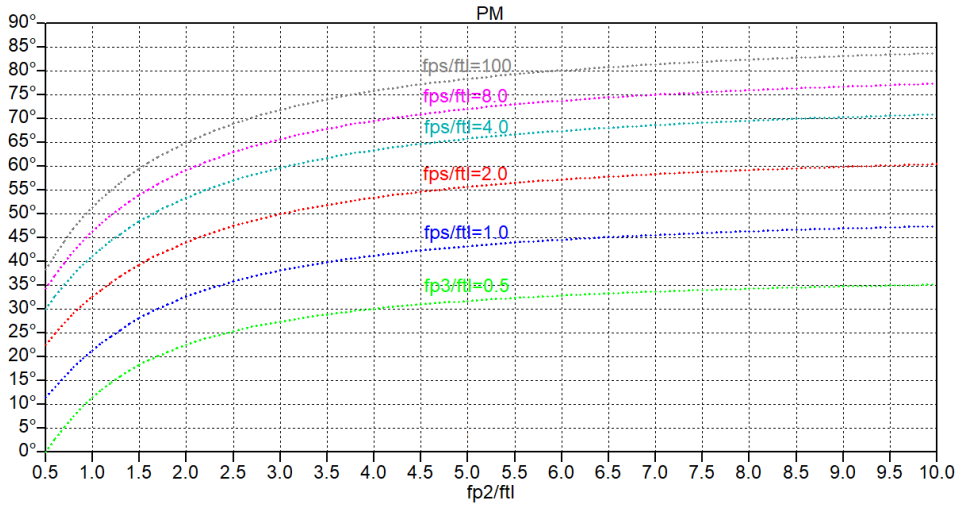


Figure 6.41: Simulated relation between pole location and phase margin for a third-order feedback system.

to a second-order system, and the closer the third pole is to the dominant pole and the second pole, the larger is the reduction of the phase margin.

Just as for the second-order systems, we may also use a '.meas' directive to find the stability margin s_m. Fig. 6.42 shows the resulting plot for ω_{p3}/ω_{tl} in the range from 0.5 to 8. Also shown is the stability margin for $\omega_{p3}/\omega_{tl} = 100$, representing a system which is a good approximation to a second-order system. We notice that for $f_{p2}/f_{tl} = f_{p3}/f_{tl} = 0.5$, i.e., $f_{p2} + f_{p3} = f_{tl}$, the stability margin is 0 as expected, indicating an unstable system. In general, the stability margin should be larger than 0.5 to ensure a stable and robust system (Åström & Murray 2021).

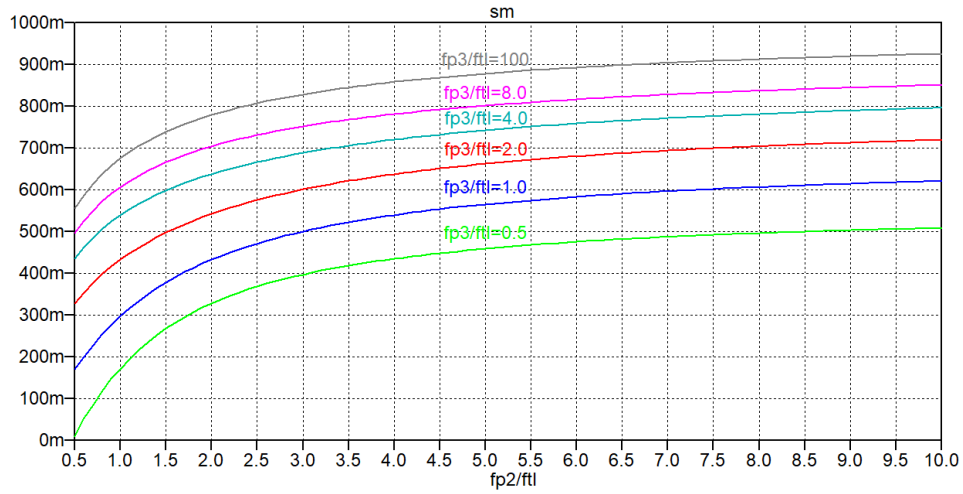
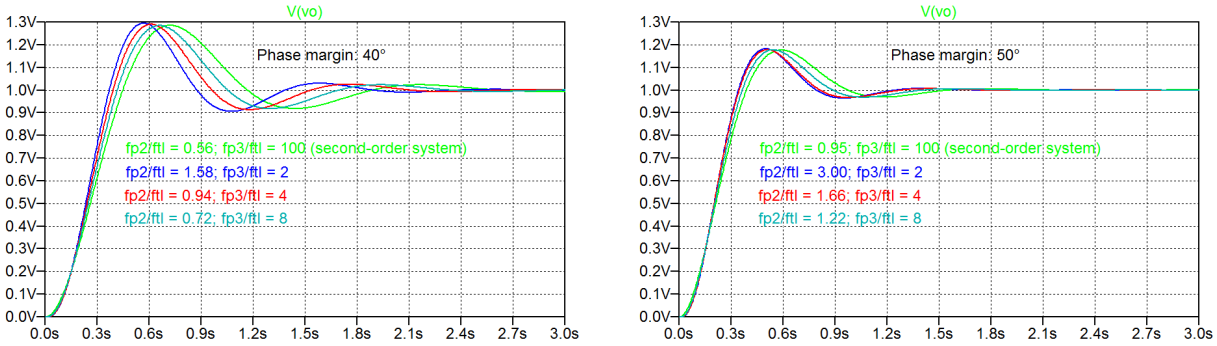


Figure 6.42: Simulated relation between pole location and stability margin for a third-order feedback system.

We may run a transient simulation for the closed-loop feedback system with various combinations of the positioning of the second and third pole resulting in a specific phase margin. For this, we use the '.step' directive and '.param' directives shown as comments in Fig. 6.38, defining tables for the parameters f_{p2} and f_{p3} (Bruun 2020). In Fig. 6.43(a), we show the transient response for some systems with a phase margin of 40°. We notice an overshoot of about 29%, independent of the combination of f_{p2} and f_{p3}

used to obtain the phase margin. This value of overshoot is also what is found from Fig. 6.37, although this figure was derived from a second-order feedback system. Similarly, Fig. 6.43(b) shows the transient response for systems with a phase margin of 50°. Again, we find approximately the same overshoot, about 18%, for the simulated combinations of f_{p2} and f_{p3}, and this value also fits well with Fig. 6.37. In both cases, the settling time depends somewhat on the combination of f_{p2} and f_{p3}.

Alternatively, we may verify that the peaking of the amplitude response may be used to estimate the phase margin, independent of the combination of f_{p2} and f_{p3} used to obtain the phase margin.



(a) (b)

Figure 6.43: Simulated transient response for systems with a phase margin of 40° (a) and a phase margin of 50° (b).

Higher-order systems. In practice, most feedback amplifiers have several poles in the loop gain $L(s)$, and $L(s)$ may also include zeros. In particular, zeros in the right half-plane of the complex frequency plane pose a problem because they cause an additional negative phase shift while also reducing the attenuation in gain. Clearly, a feedback amplifier using a basic amplifier with more than two poles and/or zeros may have a phase shift in excess of 180°, potentially causing instability in a feedback system.

An analysis like the analyses presented for second-order and third-order systems is not easy for higher-order systems and systems including zeros. A practical design approach is to place the higher-order poles and zeros at frequencies well above ω_{tl} so that they do not significantly affect the loop-gain response.

If the frequencies of higher-order poles and zeros are known, they may be taken into account by replacing the pole frequency ω_{p2} in the second-order system by an equivalent second-order pole frequency ω_{eq} given by (Chan Carusone, Johns & Martin 2012)

$$\frac{1}{\omega_{\text{eq}}} \simeq \sum_{i=2}^{n} \frac{1}{\omega_{pi}} - \sum_{j=1}^{m} \frac{1}{\omega_{zj}} \tag{6.44}$$

where ω_{pi} and ω_{zj} are the frequencies of higher-order poles and zeros, respectively. Notice that with right-half-plane zeros, ω_{zj} is negative in Eq. (6.44).

Alternatively, the phase shift caused by a higher-order pole or a zero may be estimated directly from $\phi_i = -\arctan(\omega_t/\omega_{pi})$ or $\phi_j = \arctan(\omega_t/\omega_{zj})$ where ω_t is found from ω_{tl} and Fig. 6.35(b). The phase shift thus estimated may be used to increase the requirements concerning phase margin and the design of gain-bandwidth product ω_{tl} and second pole frequency ω_{p2} may then use the relations for the second-order design, using the increased phase margin as a design target.

Finding the loop gain. In the previous section, we found that stability is related to the loop gain $L(s)$ of the feedback system, not to $A(s)$ and $\beta(s)$ separately. It may not be easy to identify β and A separately, and the loop gain βA may be easier to find. This can be done by breaking the feedback loop at a convenient point and analyzing the response of the signal path through the loop with the input signal x_s reset.

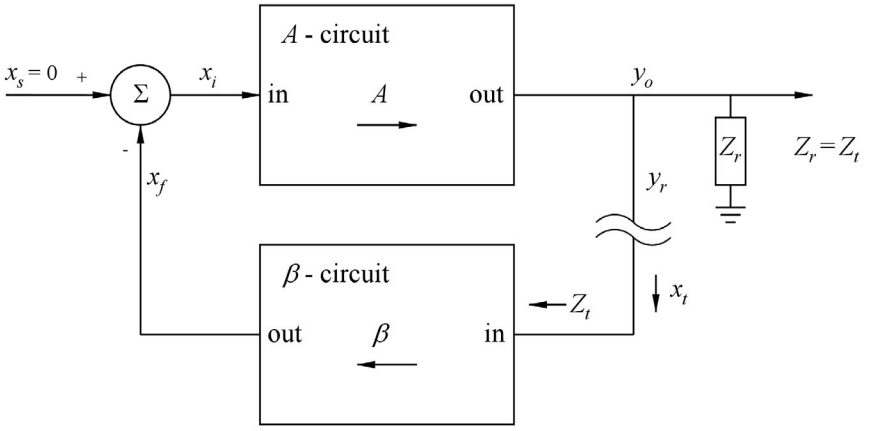


Figure 6.44: Feedback system with feedback loop broken for finding the loop gain.

Figure 6.44 shows a feedback system where the feedback loop is broken between the output of the A-circuit and the input of the β-circuit. A test signal x_t is applied as input to the β-circuit while the input signal x_s is reset to 0. With $x_s = 0$, the signal y_r returned to the breakpoint is

$$y_r = x_t \beta (-A) \tag{6.45}$$

and the loop gain is found as

$$L = -y_r/x_t \tag{6.46}$$

When breaking the loop, it does not have to be at a specific point such as the connection to the β-circuit as shown in Fig. 6.44. It can be at any point in the signal path of the loop but it is necessary to connect a load Z_r to the return point equal to the impedance Z_t looking into the input point of the loop. Often a convenient breakpoint can be identified where it is easy to find this impedance.

The test signal x_t may be either a voltage or a current and the return signal y_r will be of the same kind. If the impedance at the breakpoint cannot be found, a method is to find the open-circuit voltage loop gain T_{oc} with x_t as a voltage source and $Z_r = \infty$ and the short-circuit current loop gain T_{sc} with x_t as a current source and $Z_r = 0$. It can be shown (Rosenstark 1986) that the loop gain is

$$L = \beta A = \frac{1}{1/T_{\text{oc}} + 1/T_{\text{sc}}} \tag{6.47}$$

Another challenge in finding the loop gain by simulation is that when breaking the loop also the bias conditions for the transistors in the circuit are affected. It is necessary to run simulations to ensure proper bias conditions or to provide a dc feedback path ensuring proper bias conditions (Bruun 2020; Chan Carusone, Johns & Martin 2012). Alternatively, the loop gain may be simulated without breaking the loop using a method developed by Middlebrook (1975) (Post 2011; Tuinenga 1995).

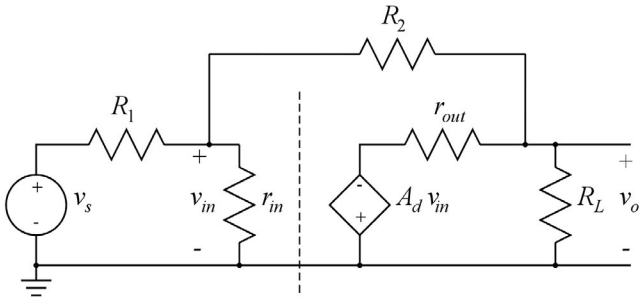


Figure 6.45: Equivalent circuit for an inverting amplifier.

As an example of finding the loop gain by breaking the loop, we reconsider the inverting amplifier from Section 6.4. We noticed that the identification of the A-circuit and the β-circuit was not quite straightforward. For this example, we assume an opamp modeled by an input resistance r_{in}, an output resistance r_{out} and a differential gain $A_d(s)$. The equivalent circuit corresponding to the inverting amplifier from Fig. 6.19 is shown in Fig. 6.45. A load resistor R_L is connected to the output of the amplifier.

A convenient place to break the loop is along the dotted line in the figure. A test voltage is then applied directly to the controlled voltage source of the opamp and the returned voltage is the voltage across r_{in}.

Figure 6.46 shows the circuit for finding the loop gain. Notice that v_s is reset, i.e., replaced by a short circuit. From this circuit, we can immediately find the returned voltage v_r and the loop gain using repeated voltage divisions:

$$\begin{aligned} v_r &= (-A_d\, v_t)\left(\frac{R_L \parallel (R_2 + (R_1 \parallel r_{in}))}{r_{out} + (R_L \parallel (R_2 + (R_1 \parallel r_{in})))}\right)\left(\frac{R_1 \parallel r_{in}}{R_2 + (R_1 \parallel r_{in})}\right) \\ \Rightarrow\ L &= -\frac{v_r}{v_t} = A_d\left(\frac{R_L \parallel (R_2 + (R_1 \parallel r_{in}))}{r_{out} + (R_L \parallel (R_2 + (R_1 \parallel r_{in})))}\right)\left(\frac{R_1 \parallel r_{in}}{R_2 + (R_1 \parallel r_{in})}\right) \end{aligned} \tag{6.48}$$

Incidentally, if you do the same exercise for a noninverting amplifier (resetting the input signal and breaking the loop in the opamp), you will find that the loop gain expression for the noninverting opamp

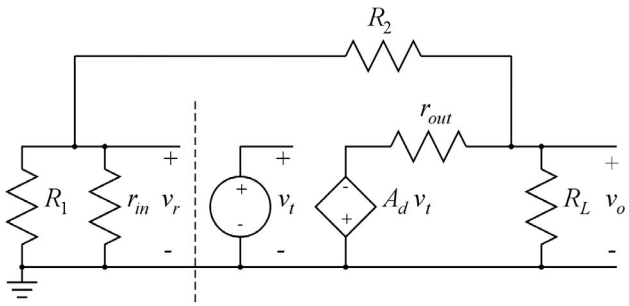


Figure 6.46: Equivalent circuit for finding the loop gain of an inverting amplifier.

is exactly the same as the expression for the inverting opamp, so the stability properties of the two circuits are the same.

In the examples above, β has been assumed constant. However, β may well be frequency-dependent so that the closed-loop gain given by (approximately) $1/\beta$ is also frequency-dependent. This is useful when designing a feedback system to have a desired frequency response. An example of this is the inverting integrator obtained by replacing the feedback resistor R_2 in the inverting amplifier shown in Fig. 6.19 by a capacitor C_2. In this case, $\beta = -sC_2$ and the closed-loop response is that of an inverting integrator (Sedra & Smith 2016). However, for stability, it is the loop gain $L(s) = \beta(s)A(s)$ which is important. It does not matter if the frequency dependency of the loop gain is caused by β or by A.

For some applications, the loop gain is deliberately designed to result in positive feedback in order to produce oscillations. However, the design of oscillators using positive feedback is beyond the scope of this book and the reader is referred to Sedra & Smith (2016).

6.6 Frequency compensation

We learned in Section 6.5 that designing a feedback amplifier for a required bandwidth specification or transient response involves designing the gain-bandwidth product ω_{tl} of the loop gain and designing the position ω_{p2}, ω_{p3} or ω_{eq} of higher-order poles and zeros relative to the gain-bandwidth product of the loop gain. In general, it is desirable to place the higher-order poles and zeros at frequencies considerably larger than the gain-bandwidth product of the loop gain. If higher-order poles and zeros are placed too close to ω_{tl}, a poor stability characterized by a small phase margin will result.

The process of designing the location of poles and zeros in order to fulfill stability requirements is called frequency compensation. This normally involves reducing the gain-bandwidth product ω_{tl} but as we will see in the following, it can also include an increase of ω_{eq}. Since ω_{tl} is defined by $\omega_{tl} = \omega_{p1}A_0\beta$, a reduction in ω_{tl} can be obtained by reducing any of the three factors β, A_0, or ω_{p1}.

Normally, β cannot be reduced since it is determined from the gain requirements of the closed-loop amplifier. For a general-purpose opamp, often a value of $\beta = 1$ is assumed, corresponding to a voltage follower configuration. If the amplifier is stable for this configuration, it will also be stable for configurations providing a higher gain, i.e., using $\beta < 1$.

Reducing the gain A_0 corresponds to shifting the amplitude plot of $L(j\omega)$ downwards in a Bode plot of $L(j\omega)$. This is normally not desirable because it reduces the amount of feedback $(1 + \beta A_0)$ at low frequencies, thereby reducing the advantages of feedback.

Dominant-pole compensation. Reducing the frequency of the dominant pole ω_{p1} corresponds to shifting the amplitude plot of $L(j\omega)$ to the left in a Bode plot of $L(j\omega)$. This is shown graphically in Fig. 6.47. The original frequencies are ω_{p1} and ω_{tl}, giving the phase margin PM. The frequencies after compensation are ω'_{p1} and ω'_{tl}, giving the phase margin PM$'$. Notice that the non-dominant pole remains unchanged at the frequency ω_{p2}.

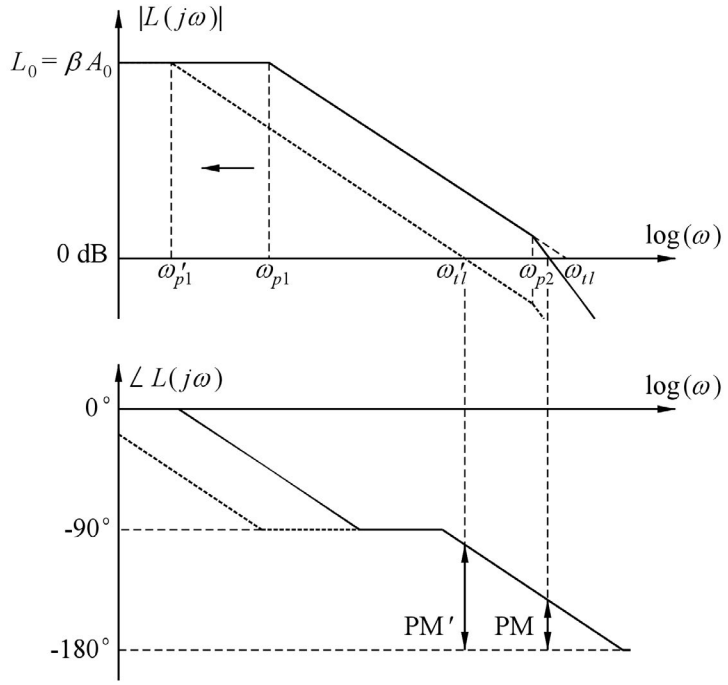


Figure 6.47: Frequency compensation by moving the dominant pole. The dotted lines show the response after compensation.

Often this frequency compensation can be achieved by placing or increasing a capacitor in the node giving rise to the dominant pole. Consider for instance an amplifier consisting of a differential input stage followed by a single-ended output stage as shown in Fig. 6.48. Here the dominant pole is assumed

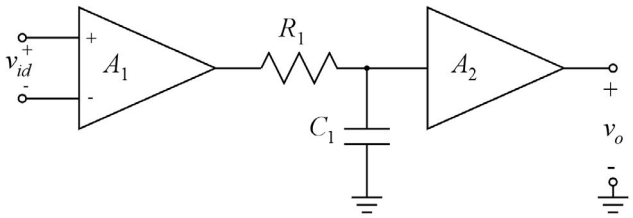


Figure 6.48: Dominant-pole compensation in a two-stage opamp.

to be caused by the output resistance R_1 of the differential stage and the capacitance C_1 to ground at the input node of the following stage. This creates a pole at the frequency $\omega_{p1} = 1/(R_1 C_1)$. By increasing C_1, the pole frequency is reduced without affecting the gain and the position of other poles and zeros in the circuit, so the gain-bandwidth product of the loop gain is reduced as shown in Fig. 6.47.

Miller compensation. Another way of reducing the frequency of the dominant pole is by introducing an extra capacitor C_c between input and output of the second stage as shown in Fig. 6.49. As explained in Chapter 4.5, the Miller effect causes this capacitor to appear as a capacitance $C_M = (1 - A_2)C_c$ in parallel with the capacitor C_1 to ground. With A_2 negative and $|A_2| \gg 1$, C_M causes the pole from the input node to the second stage to move to a lower frequency. Because of the factor $(1 - A_2)$, the size of C_c

required to move the pole to a specific frequency is much smaller than the size of a capacitor inserted in parallel with C_1 for moving the pole to the same frequency. The capacitor C_c is called the compensation capacitor.

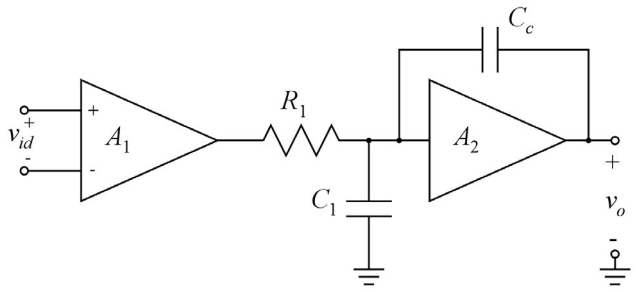


Figure 6.49: Miller compensation in a two-stage opamp.

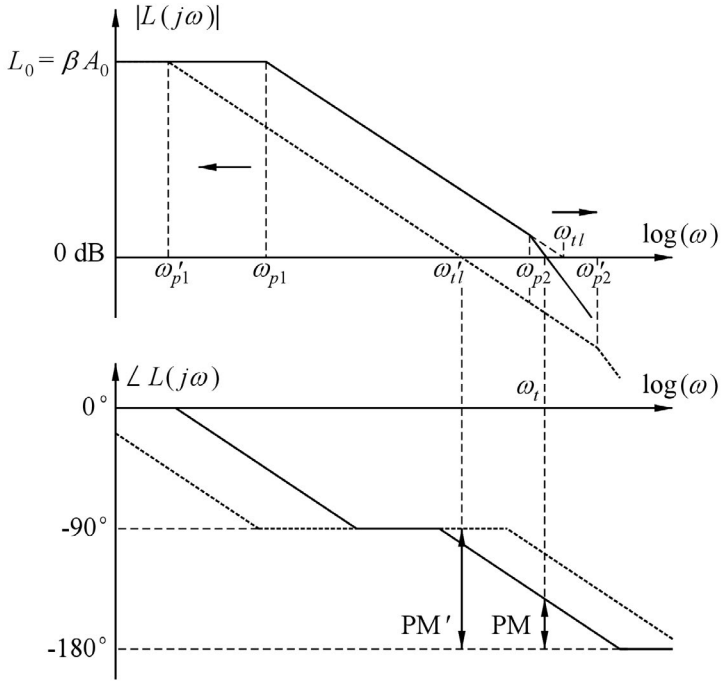


Figure 6.50: Frequency compensation from a Miller capacitor. The dotted lines show the response after compensation.

We also found previously in Chapter 4.5 that the pole coming from the output node moves to a higher frequency when inserting C_c, so ω_{p2}/ω_{tl} is increased both by an increase of ω_{p2} and by a reduction of ω_{p1}. This is shown in the Bode plot in Fig. 6.50.

We may examine the Miller compensation in more detail by assuming that the differential stage can be modeled as a voltage source $V_1 = A_1 V_{id}$ with an output resistance R_1 and that the second stage is a transconductance amplifier with a transconductance g_{m2} and an output resistance R_2. Figure 6.51 shows a circuit model including in addition to the compensation capacitor C_c also the capacitors C_1 and C_2 representing the capacitors to ground from the input and the output of the second stage.

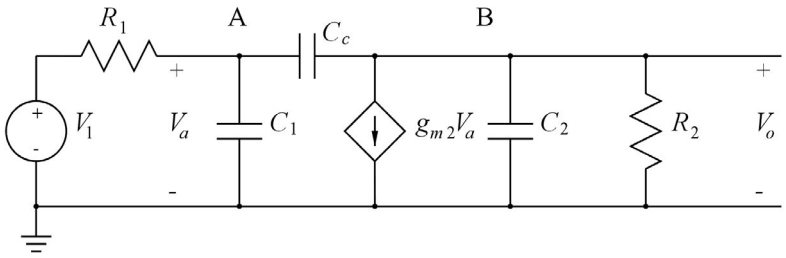


Figure 6.51: Circuit model for the Miller-compensated two-stage opamp.

In order to find poles and zeros for this configuration, we use node equations for the nodes A and B shown in Fig. 6.51.

$$\text{Node A:} \quad (V_a - V_1)/R_1 + V_a\, sC_1 + (V_a - V_o)\, sC_c = 0 \tag{6.49}$$

$$\text{Node B:} \quad (V_o - V_a)\, sC_c + g_{m2} V_a + V_o/R_2 + V_o\, sC_2 = 0 \tag{6.50}$$

Rearranging, we find

$$\text{Node A:} \quad V_a(1/R_1 + s(C_1 + C_c)) - V_o\, sC_c = V_1/R_1 \tag{6.51}$$

$$\text{Node B:} \quad V_a(g_{m2} - sC_c) + V_o(1/R_2 + s(C_2 + C_c)) = 0 \tag{6.52}$$

From Eqs. (6.51) and (6.52), we find the transfer function

$$\frac{V_o}{V_1} = \frac{-(g_{m2} - sC_c)/R_1}{(1/R_1 + s(C_1 + C_c))(1/R_2 + s(C_2 + C_c)) + sC_c(g_{m2} - sC_c)} \tag{6.53}$$

From Eq. (6.53), we identify a right-half-plane zero at the frequency

$$\omega_z = \frac{g_{m2}}{C_c} \tag{6.54}$$

For finding the poles, we must solve the equation

$$\begin{aligned} & (1/R_1 + s(C_1 + C_c))(1/R_2 + s(C_2 + C_c)) + sC_c(g_{m2} - sC_c) = 0 \\ \Rightarrow \quad & (C_1C_2 + C_1C_c + C_2C_c)s^2 + ((C_2 + C_c)/R_1 + (C_1 + C_c)/R_2 + g_{m2}C_c)s + 1/(R_1R_2) = 0 \end{aligned} \tag{6.55}$$

This is a quadratic equation, so it has two solutions, $s = -\omega_{p1}$ and $s = -\omega_{p2}$ where we assume that ω_{p1} is the frequency of the dominant pole and ω_{p2} is the frequency of the non-dominant pole.

Assuming $\omega_{p1} \ll \omega_{p2}$, we can write Eq. (6.55) as

$$\begin{aligned} N_0(s + \omega_{p1})(s + \omega_{p2}) &= 0 \\ \Rightarrow \quad N_0(s^2 + (\omega_{p1} + \omega_{p2})s + \omega_{p1}\,\omega_{p2}) &= 0 \\ \Rightarrow \quad N_0(s^2 + \omega_{p2}s + \omega_{p1}\,\omega_{p2}) &\simeq 0 \end{aligned} \tag{6.56}$$

When comparing Eqs. (6.55) and (6.56), we find

$$N_0\,\omega_{p2} \simeq (C_2+C_c)/R_1+(C_1+C_c)/R_2+g_{m2}C_c \;\Rightarrow \omega_{p2} \simeq \frac{(C_2+C_c)/R_1+(C_1+C_c)/R_2+g_{m2}C_c}{C_1C_2+C_1C_c+C_2C_c} \tag{6.57}$$

and

$$N_0\,\omega_{p1}\,\omega_{p2} \simeq 1/(R_1R_2) \;\Rightarrow \omega_{p1} \simeq \frac{1}{R_2(C_2+C_c)+R_1(C_1+C_c)+R_1R_2\,g_{m2}C_c} \tag{6.58}$$

Normally, the resistances R_1 and R_2 are output resistances from MOS transistors and they are much larger than $1/g_{m2}$ where g_{m2} is the transconductance of a transistor. Using $g_{m2} \gg 1/R_1$ and $g_{m2} \gg 1/R_2$, Eq. (6.57) can be simplified to

$$\omega_{p2} \simeq \frac{g_{m2}C_c}{C_1C_2+C_1C_c+C_2C_c} = \frac{g_{m2}}{C_1+C_2+C_1C_2/C_c} \tag{6.59}$$

We may recognize ω_{p2} as the reciprocal of the time constant τ_{pB} found in Eq. (4.87) when $R_2\,g_{m2} \gg 1+C_1/C_c$.

Using the same approximation, we find

$$\omega_{p1} \simeq \frac{1}{R_1R_2\,g_{m2}C_c} \tag{6.60}$$

We may recognize ω_{p1} as the pole caused by the Miller capacitance $C_M = (1-A_2)C_c = (1+g_{m2}R_2)C_c \simeq g_{m2}R_2C_c$ and the output resistance R_1 from the differential stage. We also note that without C_c, the second pole would have been at the frequency $1/(R_2C_2)$. With $C_1+C_1C_2/C_c$ comparable to C_2, we find from Eq. (6.59) that the frequency of the second pole is on the order of g_{m2}/C_2, and with $g_{m2} \gg 1/R_2$, we find that the compensation capacitor also moves the non-dominant pole up in frequency.

A penalty associated with the Miller compensation is the introduction of the right-half-plane zero ω_z which may completely spoil the frequency compensation by causing a further phase shift in combination with a reduced roll-off of the gain. The impact of the zero may be analyzed by introducing a zero ω_z in the second-order system shown in Fig. 6.33. Like the pole frequency f_{p2}, the zero frequency is normalized with respect to $f_{tl} = \omega_{tl}/(2\pi)$. Figure 6.52 shows the LTspice schematic where the zero frequency f_z/f_{tl} can be stepped through the values 1, 2, 4, 8, 100. The value $f_z/f_{tl} = 100$ corresponds to

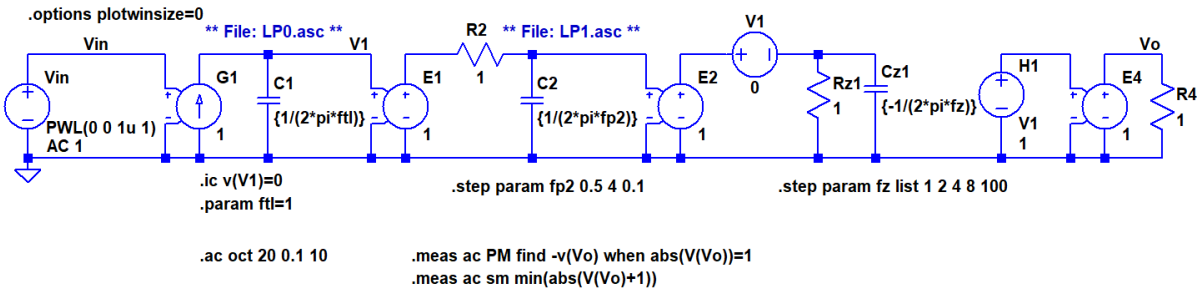
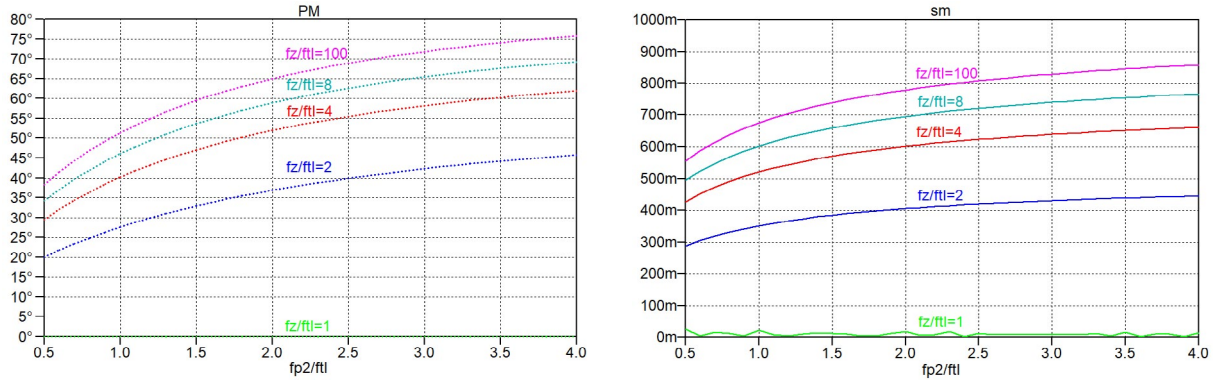


Figure 6.52: LTspice schematic for simulating the loop gain of a second-order feedback system with a right-half-plane zero.

a system where the zero frequency is so much higher than both f_{tl} and f_{p2} that it can be neglected when analyzing the phase margin.

The simulated phase margin and stability margin are shown in Fig. 6.53. The purple curves corresponding to $f_z/f_{tl} = 100$ are identical to the curves shown in Fig. 6.35(a) and Fig. 6.36(b) but we see that with smaller values of f_z/f_{tl}, the phase margin and the stability margin are reduced. For $f_z = f_{tl}$, both the phase margin and the stability margin are zero.



(a) (b)

Figure 6.53: Simulated relation between pole and zero location and phase margin (a) and between pole and zero location and stability margin (b) for a second-order feedback system with a right-half-plane zero.

Sometimes the circuit can be designed so that the zero is at a frequency much higher than f_{tl}. With f_z larger than about 8 to 10 times f_{tl}, Fig. 6.53(a) shows that the reduction of the phase margin is less than about 6°. If it is not possible to achieve $f_z \gg f_{tl}$, a resistor R_c can be inserted in series with C_c. It can be shown (Gray, Hurst, Lewis & Meyer 2009) that this moves the zero to a higher frequency given by

$$\omega_z = \frac{1}{C_c(1/g_{m2} - R_c)} \tag{6.61}$$

It is even possible to eliminate the zero by selecting $R_c = 1/g_{m2}$.

A simulation example. In order to illustrate the dominant-pole compensation and the Miller compensation by simulation, we examine the circuit shown in the LTspice schematic in Fig. 6.54. We assume that this circuit models the loop gain $L(s)$. Without compensation, the amplifier is a two-stage amplifier

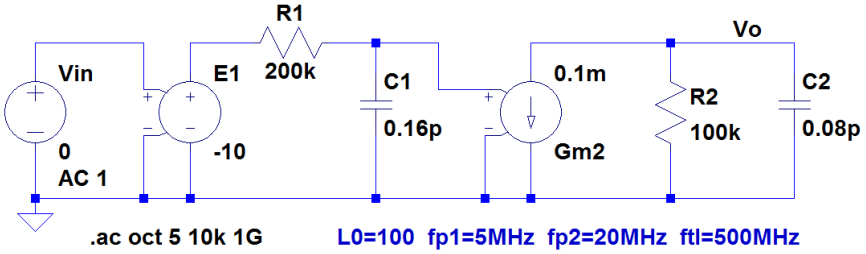


Figure 6.54: LTspice model of a two-stage amplifier without frequency compensation.

with a gain of -10 V/V in each stage. The first stage is modeled as a voltage-controlled voltage source with an output resistance $R_1 = 200\ \text{k}\Omega$. The second stage is modeled as a transconductance amplifier like in Fig. 6.51. The transconductance is $g_{m2} = 0.1$ mA/V and the output resistance is $R_2 = 100\ \text{k}\Omega$, so $A_2 = -g_{m2}R_2 = -10$ V/V.

The circuit also includes the capacitors C_1 and C_2, and with the values shown, we can calculate the frequency of the dominant pole $f_{p1} = 1/(2\pi R_1 C_1) = 5$ MHz, giving a gain-bandwidth product of $f_{tl} = A_1 A_2 f_{p1} = 500$ MHz. The second pole is caused by R_2 and C_2, and with the values shown, we find $f_{p2} = 1/(2\pi R_2 C_2) = 20$ MHz.

Thus, we have a non-dominant pole at a frequency which is much smaller than the gain-bandwidth product of the loop gain which clearly leads to an unacceptable phase margin. Figure 6.55 shows a plot from a '.ac' simulation of the loop gain. We find a phase margin of less than 15°, clearly inadequate for normal purposes.

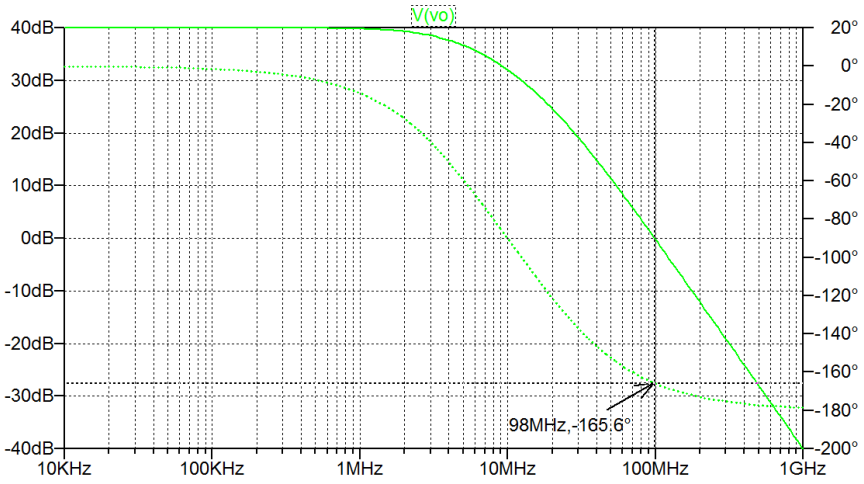


Figure 6.55: Plot of loop gain for the uncompensated amplifier from Fig. 6.54.

Next, we introduce dominant-pole compensation. Assume that the design target is a phase margin of 65.5°, resulting in a maximally flat frequency response. From Fig. 6.35(a), we find that this requires $f_{p2}/f_{tl} = 2$, corresponding to $f_{tl} = 10$ MHz. With a low-frequency gain of 100 V/V, this gives $f_{p1} = 0.1$ MHz, and for $R_1 = 200$ kΩ, we find $C_1 = 1/(2\pi R_1 f_{p1}) = 8$ pF.

Repeating the simulation with $C_1 = 8$ pF results in the Bode plot shown in Fig. 6.56 where we see that that phase margin is now about 66°, so the design target has been reached. As expected, the increased phase margin is obtained at the expense of a reduced gain-bandwidth product of the loop gain.

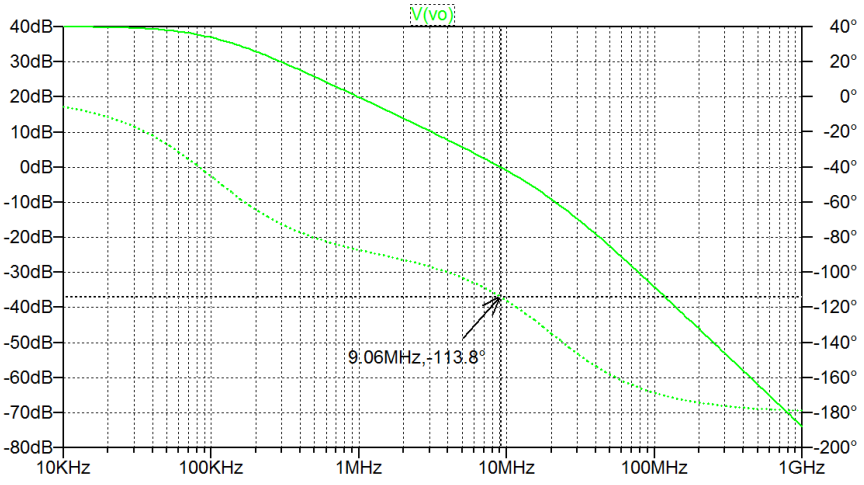


Figure 6.56: Plot of loop gain for the amplifier from Fig. 6.54 with the value of C_1 increased to 8 pF.

Alternatively, we may insert a compensation capacitor to create a Miller compensation. The dominant pole is found from

$$f_{p1} = \frac{1}{2\pi R_1(C_1 + C_M)} = \frac{1}{2\pi R_1(C_1 + (1 + |A_2|)C_c)} \tag{6.62}$$

Inserting component values and a target value of $f_{p1} = 0.1$ MHz, we find $C_c = 0.71$ pF. Repeating the '.ac' simulation with $C_1 = 0.16$ pF and $C_c = 0.71$ pF inserted between the input and output of the transconductance g_{m2} results in the Bode plot shown in Fig. 6.57. We see that the bandwidth is close to 0.1 MHz as expected but we find a phase margin of only about 57°.

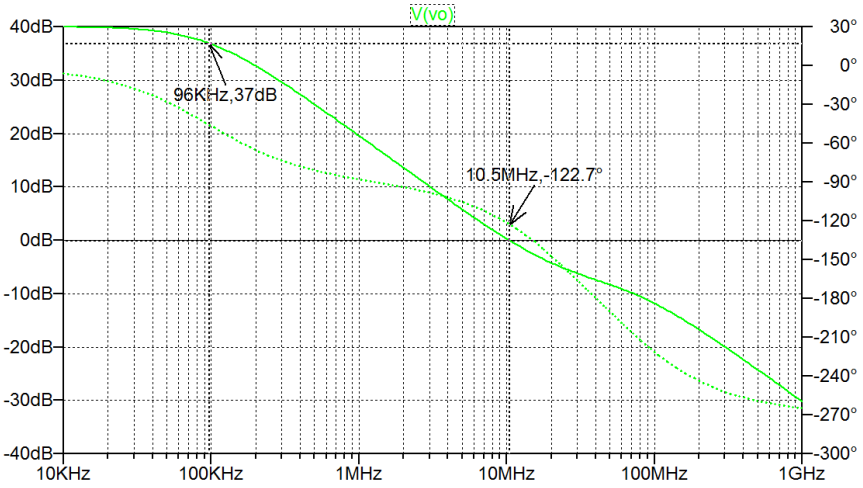


Figure 6.57: Plot of loop gain for the amplifier from Fig. 6.54 with a compensation capacitor of 0.71 pF inserted between input and output of g_{m2}.

Examining the Bode plot, we find that in a frequency range around 40 MHz, the slope of the gain plot is not -20 dB/dec, and this indicates the presence of a zero. It is not easy to estimate the frequency of the zero from the Bode plot shown in Fig. 6.57 but instead of plotting '`V(vo)`', we may plot '`V(vo)*frequency`'. This adds 20 dB/dec to the slope of the line segments in a straight-line approximation to the Bode plot so that the line segment after the first pole at about 0.1 MHz becomes a horizontal line and the line segment between the zero and the second pole has a slope of $+20$ dB/dec. The line segment after the second pole then becomes a horizontal line. This implies that the frequency of the zero and the frequency of the second pole can be found as breakpoints between long horizontal lines and a line with a slope of 20 dB/dec.

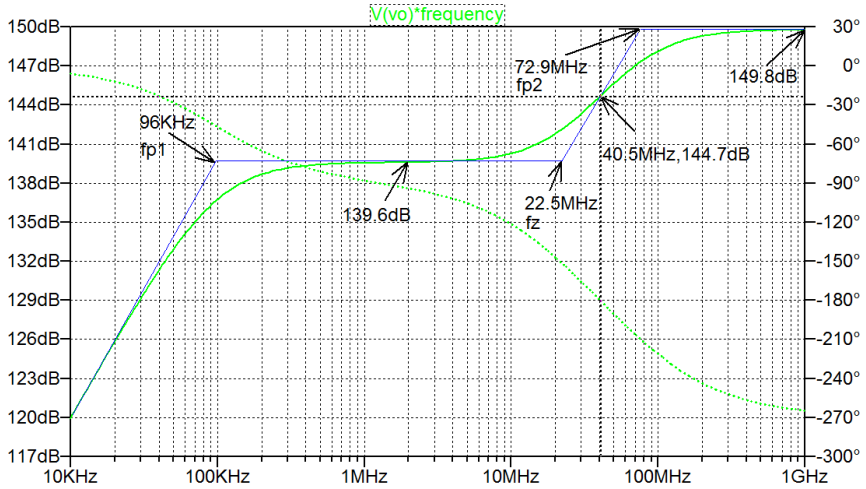


Figure 6.58: Plot of the loop gain multiplied by the frequency for the amplifier from Fig. 6.54 with a compensation capacitor of 0.71 pF inserted between input and output of g_{m2}.

From the plot shown in Fig. 6.58, we note that the zero and the second pole are not spaced far apart so the 3 dB frequencies do not give an accurate estimate of the zero frequency and the pole frequency. Instead, we locate the midpoint of the line segment between the zero and the pole as the point where the gain (in dB) is the average of the gains at the horizontal lines, see Fig. 6.58. From the midpoint at 40.5 MHz there is a factor of $(149.8\text{ dB} - 144.7\text{ dB}) = 5.1$ dB or 1.799 to the second pole, so the second pole is located at 72.9 MHz. Likewise, there is a factor of -5.1 dB or 0.556 to the zero, so the zero is located at 22.5 MHz. Also the piecewise linear approximation to the Bode plot is shown in Fig. 6.58. The line segment between the zero and the second pole has been drawn as a line through the midpoint of the segment and with a slope of 20 dB/dec.

From Eq. (6.54), we can also calculate the frequency of the zero to be $f_z = 22.4$ MHz, i.e., a value close to the value found from Fig. 6.58. This is only a factor of 2.2 larger than f_{tl}. From Eq. (6.59), we can further estimate the frequency of the second pole f_{p2} to about 62 MHz. Due to the approximations used when deriving Eq. (6.59), the calculated value is slightly smaller than the value found from Fig. 6.58 but still, the frequency of the second pole is more than adequate for achieving a phase margin of 65.5°.

However, in order to fulfill the specification for the phase margin, we need to modify the circuit so that the zero does not spoil the phase margin. We could either try to move the zero to a higher frequency, or we could insert a resistor in series with C_c. Moving the zero to a higher frequency is possible when designing the two stages of the amplifier. This is an exercise left for Chapter 7.

Here, we consider a resistor in series with C_c. From Eq. (6.61), we see that the zero is eliminated for $R_c = 1/g_{m2} = 10\,\text{k}\Omega$. Figure 6.59 shows the LTspice schematic modified to include both C_c and R_c and Fig. 6.60 shows the plot from the '.ac' simulation.

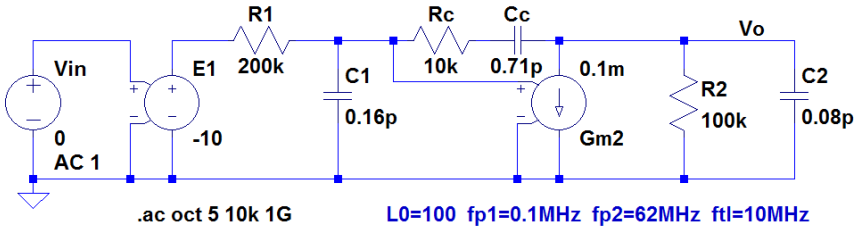


Figure 6.59: LTspice model of a two-stage amplifier with Miller compensation of the frequency compensation.

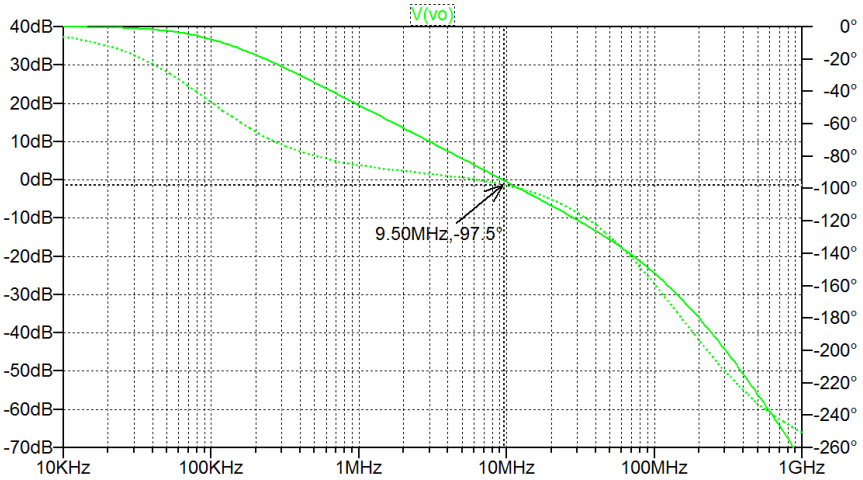


Figure 6.60: Plot of loop gain for the Miller-compensated amplifier from Fig. 6.59.

From the Bode plot, we find a phase margin of more than 82°. In a design iteration, the phase margin may be reduced by reducing C_c, resulting in a larger bandwidth and gain-bandwidth product of the amplifier. From Eq. (6.59), we see that the position of the second pole is only slightly affected by C_c, so with $f_{p2} \simeq 62$ MHz, the gain-bandwidth product may be increased from 10 MHz to about 30 MHz. Using Eq. (6.62), we find that this is achieved by reducing C_c to 0.22 pF. Figure 6.61 shows a Bode plot of the loop gain where C_c in Fig. 6.59 has been reduced to 0.22 pF. From the plot, we find a phase margin of about 66° and a bandwidth of 300 kHz, corresponding to a gain-bandwidth product $f_{tl} = 30$ MHz, i.e., a significant increase compared to the circuit using dominant-pole compensation.

With $f_{tl} \simeq 30$ MHz and a maximally flat frequency response, Fig. 6.30(a) shows that we can expect a closed-loop bandwidth of about 42 MHz. This may be verified by a simulation where the output voltage '`Vo`' is connected directly back to the inverting input of 'E1' in Fig. 6.59. This is an exercise left for the reader.

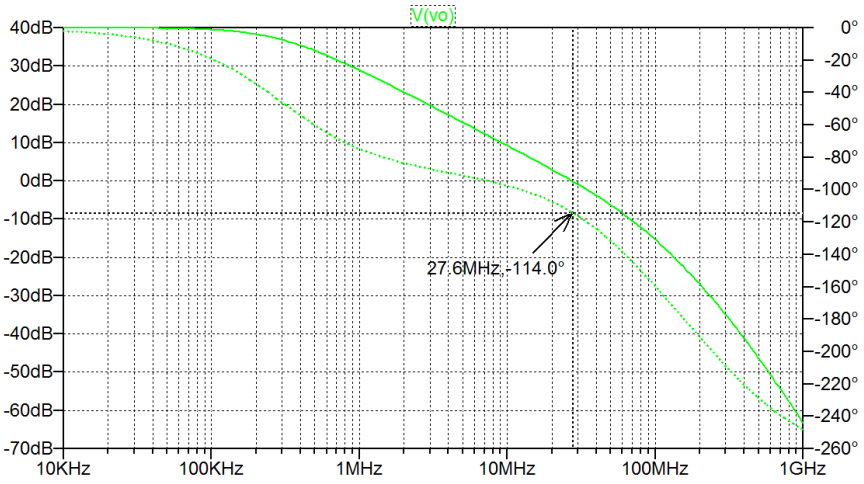


Figure 6.61: Plot of loop gain for the Miller-compensated amplifier from Fig. 6.59 with C_c reduced to 0.22 pF.

References

Allen, PE., & Holberg, DR., 2012, *CMOS Analog Integrated Circuit Design*, Third Edition, Oxford University Press, New York, USA.

Bruun, E. 2020, *CMOS Integrated Circuit Simulation with LTspice*, Third Edition, bookboon. Available from: http://bookboon.com/en/cmos-integrated-circuit-simulation-with-ltspice-ebook

Bruun, E. & Jørgensen, I., 2022, *CMOS Analog IC Design: Learning by Problem Solving*, bookboon. Available from: http://bookboon.com/en/cmos-analog-ic-design-ebook

Chan Carusone, T., Johns, D. & Martin, K. 2012, *Analog Integrated Circuit Design*, Second Edition, John Wiley & Sons, Inc., Hoboken, USA.

Elmore, WC. 1948, 'The Transient Response of Damped Linear Networks with Particular Regard to Wideband Amplifiers', *Journal of Applied Physics*, vol. 19, no. 1, pp. 55-63, January 1948.

Gray, PR., Hurst, PJ., Lewis, SH. & Meyer, RG. 2009, *Analysis and Design of Analog Integrated Circuits*, Fifth Edition, John Wiley & Sons, New York, USA.

Lathi, BP. 2009, *Signal Processing & Linear Systems*, International Second Edition, Oxford University Press, New York, USA.

Middlebrook, RD. 1975, 'Measurement of loop gain in feedback systems', *Int. J. Electronics*, vol. 38, no. 4, pp. 485-512, April 1975.

Nyquist, H. 1932, 'Regeneration Theory', *Bell System Technical Journal*, vol. 11, no. 1, pp. 126-147, January 1932.

Post, JE. 2011, 'On Determining Loop Gain through Circuit Simulation', *QEX*, pp. 3-8, March-April 2011.

Rosenstark, S. 1986, *Feedback Amplifier Principles*, Macmillan Publishing Company, New York, USA.

Sedra, AS. & Smith, KC. 2016, *Microelectronic Circuits*, International Seventh Edition, Oxford University Press, New York, USA.

Tuinenga, PW. 1995, *SPICE: A Guide to Circuit Simulation and Analysis Using PSpice*, Third Edition, Prentice-Hall, Englewood Cliffs, USA.

Åström, KJ. & Murray, RM. 2021, *Feedback Systems: An Introduction for Scientists and Engineers*, Second Edition, Princeton University Press, 2021.

Multiple-choice test

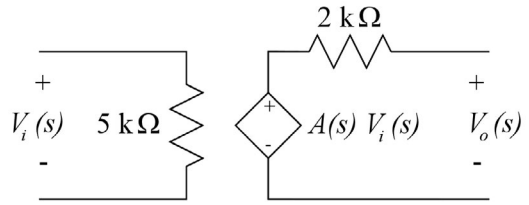
1. Complete the following statements by selecting the appropriate continuation from the table below.

A:	In an amplifier with feedback, the feedback network is called ...
B:	For an amplifier built from an ideal opamp and a feedback network, the closed-loop gain is determined by ...
C:	The product of the feedback factor and the open-loop gain is called ...
D:	For an amplifier with series-shunt feedback, the output resistance is ...
E:	For an amplifier with series-shunt feedback, the bandwidth is ...
F:	The stability of an amplifier with feedback is determined by ...
G:	The phase margin in a first-order feedback system is larger than ...
H:	The sign of the phase margin in a stable feedback system is ...
I:	A second-order feedback system with a phase margin between 65.5° and 76.3° shows ...
J:	A method for modifying the frequency response of the loop gain involves the insertion of a Miller capacitor. The Miller capacitor is inserted ...

Continuation:

1:	the A-network.
2:	the open-loop gain.
3:	from input to ground of an inverting gain stage.
4:	the β-network.
5:	the opamp.
6:	reduced compared to the amplifier without feedback.
7:	the loop gain.
8:	a maximally flat frequency response.
9:	negative.
10:	90°.
11:	180°.
12:	overshoot in the transient response.
13:	increased compared to the amplifier without feedback.
14:	from input to output of a noninverting gain stage.
15:	6 dB.
16:	the amount of feedback.
17:	from output to ground.
18:	positive.
19:	from input to output of an inverting gain stage.
20:	peaking in the frequency response.

2. The amplifier shown below has a transfer function $A(s) = 40/(1 + s/10^4\ \text{s}^{-1})$.



When inserted in an amplifier with series-shunt feedback and $\beta = 0.1$, the resulting closed-loop gain at low frequencies is

A: 4 V/V
B: 8 V/V
C: 40 V/V

3. When the amplifier shown above is inserted in an amplifier with series-shunt feedback and $\beta = 0.1$, the resulting -3 dB frequency is

A: 8 kHz
B: 10 kHz
C: 50 kHz

4. When the amplifier shown above is inserted in an amplifier with series-shunt feedback and $\beta = 0.1$, the resulting input resistance is

A: 1 kΩ
B: 5 kΩ
C: 25 kΩ

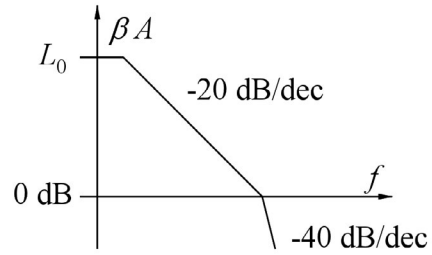
5. When the amplifier shown above is inserted in an amplifier with series-shunt feedback and $\beta = 0.1$, the resulting output resistance is

A: 200 Ω
B: 400 Ω
C: 500 Ω

6. Using the amplifier shown above in a series-shunt feedback configuration, the feedback factor required to give a resulting closed-loop gain of 10 V/V at low frequency is

A: 0.025
B: 0.075
C: 0.100

7. A feedback amplifier is assumed to have a loop gain as shown below where $L_0 = \beta A_0 \gg 1$.



The phase margin, estimated from the piecewise linear Bode plot approximation, is

A: 45°
B: 52°
C: 65°

8. For the feedback amplifier with the loop gain shown above, the phase margin, estimated from the actual Bode plot (without using the piecewise linear approximation), is

   A: 45°
   B: 52°
   C: 65°

9. For the amplifier with the loop gain shown above, the ratio between the closed-loop bandwidth and the frequency of the non-dominant pole is

   A: 1
   B: 1.27
   C: 1.41

10. The amplifier with the loop gain shown above may be compensated using a dominant-pole compensation. In order to achieve a phase margin of 65.5°, the dominant-pole compensation should reduce the frequency of the dominant pole by a factor of

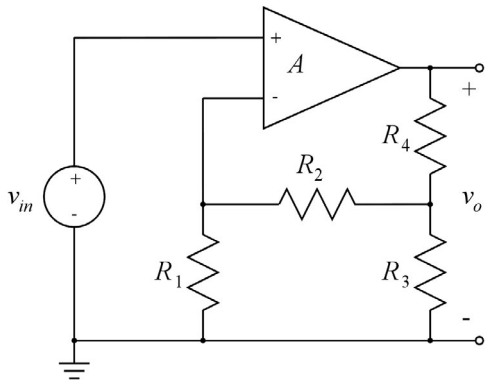
    A: 2
    B: 3
    C: 4

11. When the amplifier with the loop gain shown above is compensated to have a phase margin of 65.5°, the closed-loop bandwidth is reduced by a factor of approximately

    A: 0.25
    B: 0.33
    C: 0.56

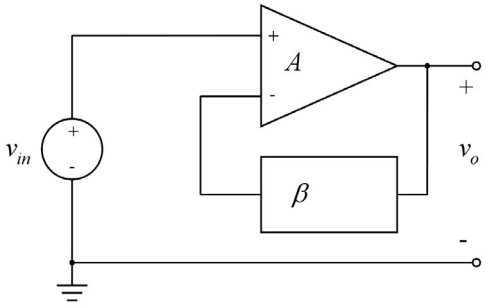
Problems

Problem 6.1



For the feedback amplifier shown above, find an expression for the feedback factor β and the closed-loop gain v_o/v_{in}, assuming that the amplifier A is an ideal opamp.
Calculate the closed-loop gain for $R_1 = 1$ kΩ, $R_2 = 10$ kΩ, $R_3 = 2$ kΩ and $R_4 = 8$ kΩ.
Calculate the closed-loop gain, assuming that the opamp has a finite gain of $A = 100$ V/V.

Problem 6.2



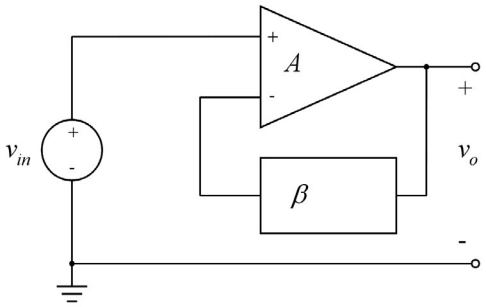
For the feedback amplifier shown above, assume that the amplifier is a voltage amplifier with a gain A, an infinite input resistance and an output resistance of 0. With A infinite, the closed-loop gain is $G_0 = v_o/v_{in} = 1/\beta$.
Show that the relative reduction in closed-loop gain caused by a finite value of A is $G_0/(A + G_0)$.

For the circuit from Problem 6.1, find the minimum value of A required to obtain a closed-loop gain which deviates at most 2% from the value obtained with an ideal opamp.

Problem 6.3

An amplifier with a specified gain of $A = 1000$ V/V is used to implement a feedback amplifier with a closed-loop gain of 100 V/V. Assuming a 30% tolerance in the specification of A, what is the tolerance of the closed-loop gain? What is the tolerance of the closed-loop gain if it is reduced to 10 V/V ?
What is the total gain and the tolerance of the total gain for an amplifier using two cascade-connected stages where each stage provides a closed-loop gain of 10 V/V ?

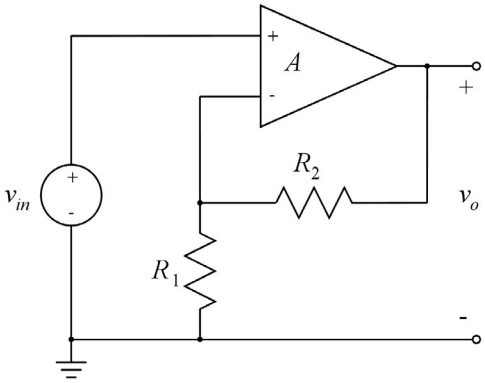
Problem 6.4



For the feedback amplifier shown above, assume that the amplifier is a voltage amplifier with a gain $A = 200$ V/V, an infinite input resistance and an output resistance of 10 kΩ. Also assume that the β-network does not load the amplifier.

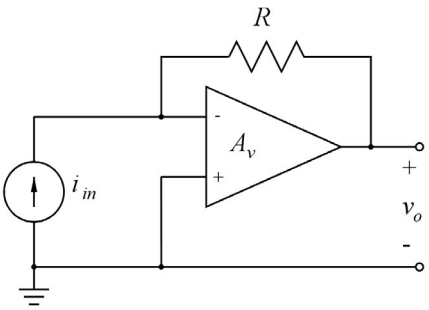
What is the maximum closed-loop gain that can be achieved if the closed-loop output resistance must not be larger than 1 kΩ ?
Which value of β is required to give the maximum closed-loop gain?

Problem 6.5



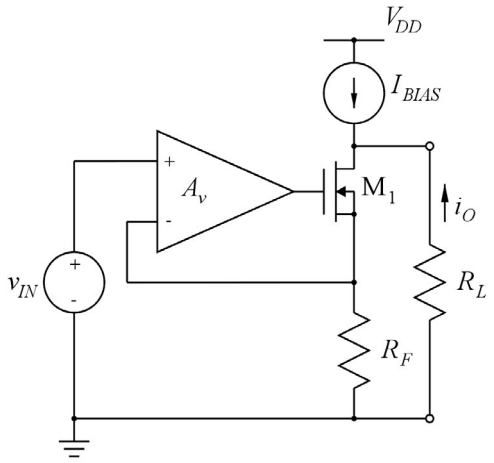
For the feedback amplifier shown above, assume that the amplifier is a voltage amplifier with a gain $A =$ 200 V/V, an infinite input resistance and an output resistance of 10 kΩ. Also assume that $R_1 = 0.45$ kΩ and $R_2 = 9.55$ kΩ.
Find the closed-loop gain v_o/v_{in} and the closed-loop output resistance. You may use LTspice for this.

Problem 6.6



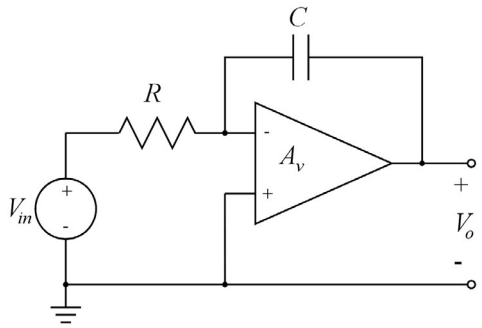
For the feedback amplifier shown above, assume that the amplifier is a voltage amplifier with a gain A_v, an infinite input resistance and an output resistance of 0. Use node equations and loop equations to find an expression for the closed-loop gain $A_{CL} = v_o/i_{in}$ and the closed-loop input resistance r_{inCL}. Calculate the values of closed-loop gain and input resistance for $A_v = 20$ V/V and $R = 10$ kΩ.
Repeat the problem, assuming a non-zero output resistance of $r_{out} = 100$ Ω for the amplifier.

Problem 6.7



The figure above shows a transconductance amplifier with series-series feedback in order to obtain a well-defined small-signal transconductance i_o/v_{in}. The amplifier is a voltage amplifier with $A_v = 60$ V/V, infinite input resistance and an output resistance of 0. Transistor M_1 is biased to have $g_{m1} = 1$ mA/V. The small-signal output resistance of M_1 can be neglected. The resistor R_F has a value of 5 kΩ.
Find the small-signal loop gain $A\beta$, the feedback factor β and the closed-loop transconductance i_o/v_{in}.

Problem 6.8

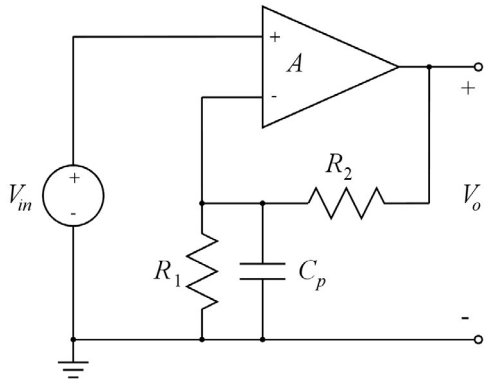


The figure above shows an inverting integrator built from an opamp, a resistor $R = 10$ kΩ and a capacitor $C = 16$ pF. The opamp has infinite input resistance, an output resistance of 0, and the transfer function $A_v(s)$ has a low-frequency gain of 1000 V/V and a single pole causing a gain-bandwidth product of 100 MHz.
Find an expression for the loop gain $L(s)$ and find the phase margin of the system.

Problem 6.9

Repeat Problem 6.8, assuming that the opamp has an additional pole at $f_{p2} = 100$ MHz. You may use LTspice to find the phase margin.

Problem 6.10



The figure above shows a feedback amplifier using an opamp with infinite input resistance, an output resistance of 0, and a transfer function with a low-frequency gain of 200 V/V and a single pole at the frequency $f_{p1} = 270$ kHz. The feedback resistors are $R_1 = 50$ kΩ and $R_2 = 400$ kΩ. A parasitic capacitance $C_p = 0.6$ pF is also included as shown.

Sketch a Bode plot of the loop gain and estimate the phase margin. Verify your result using LTspice.

In order to increase the phase margin, a dominant-pole frequency compensation is applied. Find the new pole frequency f'_{p1}, so that a phase margin of 65.5° is obtained. Verify your result using LTspice.

Problem 6.11

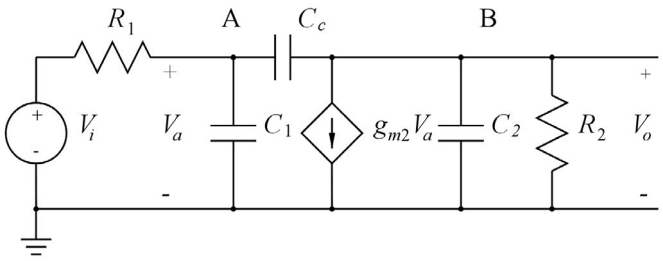
Assume now that the amplifier shown in Problem 6.10 has an additional pole at the frequency $f_{p2} =$ 12 MHz.

Find the phase margin when the dominant pole of the amplifier is located at $f_{p1} = 270$ kHz.

Find the phase margin when the dominant pole of the amplifier is located at the frequency f'_{p1} found in Problem 6.10 to give a phase margin of 65.5° for the single-pole amplifier.

Find the dominant-pole frequency required for a phase margin of 65.5° when using the two-pole amplifier with a dominant-pole compensation which leaves the frequency of the non-dominant pole unchanged at 12 MHz. You may use LTspice for solving this problem.

Problem 6.12



$R_1 = 200$ kΩ, $R_2 = 100$ kΩ, $C_1 = 0.16$ pF, $C_2 = 0.08$ pF, $C_c = 0.71$ pF and $g_{m2} = 0.1$ mA/V.

For the amplifier shown above, assume the device values listed with the figure. These values are the same as those used for the simulations in Figs. 6.57 and 6.58.
Use node equations for node A and node B to find exact values for the pole frequencies and the zero frequency of the amplifier and compare the calculated values to the simulated values and the values calculated when using the expressions given by Eqs. (6.54), (6.59), (6.60) and (6.62).

Chapter 7 – The Two-Stage Opamp

In this chapter, we present an example of the design of a two-stage opamp for application in a feedback amplifier. The design illustrates many of the concepts presented in the preceding chapters, and in the present chapter, we focus on synthesis rather than analysis. Additionally, we show how to proceed from a Shichman-Hodges transistor model to a BSIM transistor model. This is required to get a final circuit design which can be expected to fulfill the design specifications.

After having studied the chapter, you should be able to

- design a two-stage opamp to fulfill specifications concerning stability.
- identify slew-rate limitations in the two-stage opamp.
- design a two-stage opamp to fulfill specifications concerning bandwidth, slew rate, etc.
- determine transistor small-signal parameters by simulation using a BSIM transistor model.
- verify phase margin specification by a simulation of the loop gain.
- verify slew rate specification by a transient simulation.
- use LTspice simulations for performing design iterations.

When designing a two-stage opamp, many specifications may be considered. Listed below are some of the parameters for which specification requirements may be defined.

- Gain.
- Slew rate.
- Bandwidth.
- Stability in feedback systems.
- Power dissipation.
- Supply voltage.
- Output voltage swing.
- Input voltage swing.
- Noise.
- Common-mode rejection ratio.
- Power-supply rejection ratio.

Often only some of the parameters are specified while others may be allowed to vary over a reasonable range, thus providing flexibility in the design. Also, specifications may lead to conflicting requirements concerning the design, in which case a compromise will be necessary. For instance, it can be difficult to combine a large bandwidth with a low power consumption.

7.1 Specifications for a design example

For the design example treated in this chapter, we consider a two-stage opamp to be used in a noninverting feedback configuration as shown in Fig. 7.1. The feedback network is capacitive so that it does not provide a resistive load to the amplifier. This implies that the two-stage amplifier does not need an additional source-follower output stage.

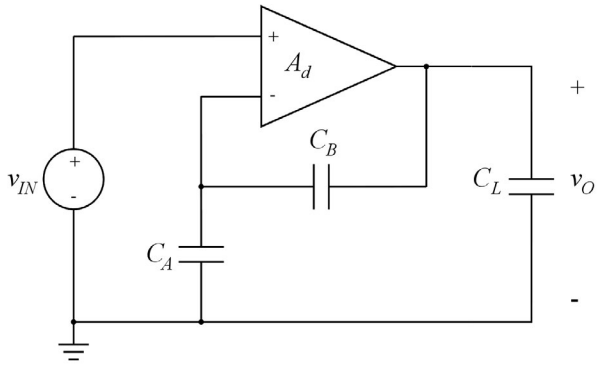


Figure 7.1: Noninverting amplifier circuit with capacitive feedback network.

The specifications which we will consider are the following:

- Midband closed-loop gain: 14 dB $\pm$ 0.3 dB.
- Bandwidth: $\geq$ 20 MHz.
- Load capacitance: 1.5 pF.
- Capacitive load from feedback network: $\leq$ 0.2 pF.
- Capacitor C_A must be much larger than the parasitic capacitance to ground from the inverting opamp input.
- Slew rate: $\geq$ 30 V/µs.
- Phase margin: $\geq 65°$.
- Positive supply voltage: 0.9 V.
- Negative supply voltage: -0.9 V.
- Output voltage range: -0.75 V to $+0.75$ V.
- Input voltage range: -0.15 V to $+0.15$ V.
- Technology: Generic 0.18 µm CMOS n-well process.

We notice that the requirement concerning input voltage range favors the use of a PMOS differential input pair when using an n-well process for the opamp because this makes it possible to avoid the bulk effect in the input transistors. The bulk effect increases the threshold voltage, and the voltage available for the gate-source voltage of the differential pair plus the drain-source voltage of the tail current source is only $0.9\text{ V} - 0.15\text{ V} = 0.75$ V, so there may not be enough headroom for a threshold voltage increased by the bulk effect.

A PMOS input pair is often the preferred choice anyway because bandwidth considerations favor the use of an NMOS common-source transistor for the second stage in order to achieve a large transconductance and a small gate-source capacitance. Also noise considerations may influence the choice between NMOS and PMOS transistors for the input stage (Chan Carusone, Johns & Martin 2012).

We can also immediately see that the requirements concerning midband gain (14 db $\sim$ 5 V/V) implies $1+C_A/C_B = 5 \Rightarrow C_A = 4C_B$, assuming that the opamp gain is large so that the closed-loop gain is (approximately) $1/\beta = (C_A^{-1}+C_B^{-1})/(C_A^{-1}) = 1+C_A/C_B$ and $\beta = 1/5$.

In order to ensure C_A much larger than the parasitic capacitances, we select C_A and C_B as large as possible without exceeding the limitation concerning the capacitive load from the feedback network. The capacitive load from C_A and C_B is the series connection of C_A and C_B which is then selected to the maximum value of 0.2 pF, giving $C_A^{-1}+C_B^{-1} = (0.2\text{ pF})^{-1}$, and with $C_A = 4C_B$, we find $C_A = 1$ pF and $C_B = 0.25$ pF.

For the frequency compensation, we select a Miller compensation with a capacitor C_c connected from gate to drain of the second-stage common-source transistor since this gives both a pole splitting and a small value of compensation capacitor as explained in Chapter 6.6.

With the opamp configuration and the capacitors in place, we now have the schematic of the feedback amplifier shown in Fig. 7.2.

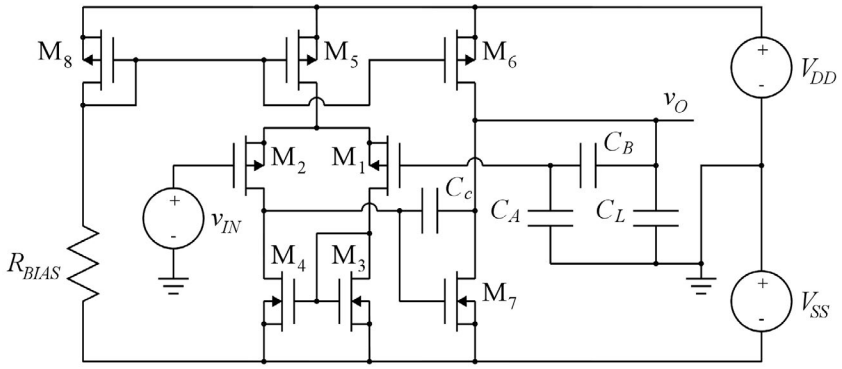


Figure 7.2: Transistor schematic of the noninverting opamp with feedback network and load capacitance.

The requirements determine the design as follows:

- The bandwidth specification, BW, puts constraints on the transconductance of the input transistors and the compensation capacitor which determine the gain-bandwidth product of the loop gain, in turn determining the bandwidth as shown in Fig. 6.30(a) in combination with the first non-dominant pole.
- The slew-rate specification, SR, puts constraints on the bias current for M_5 and M_6. The slew rate is defined as the maximum value of $|dv_O/dt|$ which can be achieved at the output of the opamp (Hambley 2014), and this is limited by the current available to change the voltage on a capacitor, see Eq. (2.11).
- The specification for the phase margin, PM, puts constraints on the location of non-dominant poles and zeros in the opamp.
- The specification for output voltage range puts constraints on the saturation voltage of the output transistors M_6 and M_7.
- The specification for input voltage range puts constraints on the saturation voltage of M_5 and the gate-source voltage of M_2.

Clearly, the constraints imposed by the specifications are not independent. In the design approach presented in the following section, we take the specifications concerning stability and phase margin as a starting point. This will determine the small-signal transconductances for the transistors in the signal path.

We can subsequently find transistor geometries and bias conditions, taking the other specifications into account and also using the BSIM transistor models from Fig. 3.44.

7.2 Bandwidth and stability requirements

For investigating bandwidth and stability, we must examine the loop gain of the feedback amplifier. The opamp is a two-stage opamp with the configuration shown in Figs. 6.49 and 6.51. For finding the loop gain, we break the feedback loop at the input to the differential stage and apply a test input voltage. Figure 7.3 shows a schematic for finding the loop gain. From Eq. (6.46), we find the loop gain as *minus* the returned voltage divided by the test voltage. The minus sign may be avoided by applying the test voltage to the positive input of the opamp while resetting the negative input as shown in Fig. 7.3.

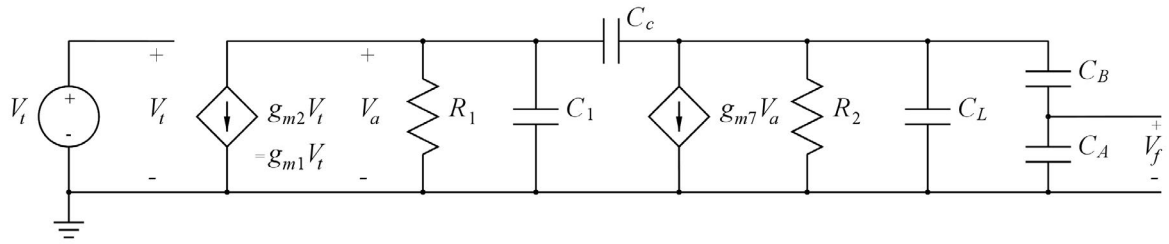


Figure 7.3: Open-loop circuit for finding the loop gain $L(s) = V_f(s)/V_t$.

In Fig. 7.3, the transistor capacitances at the output of the opamp are assumed to be much smaller than the load capacitance C_L and the capacitors C_A and C_B in the β-network. Capacitor C_1 is the parasitic transistor capacitance from the input of the second stage, often dominated by the gate-source capacitance of M_7. Resistor R_1 is the output resistance of the differential stage, i.e., $R_1 = (r_{ds2} \parallel r_{ds4}) = (g_{ds2} + g_{ds4})^{-1}$ and R_2 is the output resistance of the second stage, i.e., $R_2 = (r_{ds6} \parallel r_{ds7}) = (g_{ds6} + g_{ds7})^{-1}$.

Assuming that the transfer function of the opamp has a low-frequency gain A_0, a dominant pole at the frequency ω_{p1}, a non-dominant pole at the frequency ω_{p2} and a right-half-plane zero at the frequency ω_z, the loop gain is given by

$$L(s) = \left(\frac{C_B}{C_A + C_B}\right)\left(\frac{A_0(1 - s/\omega_z)}{(1 + s/\omega_{p1})(1 + s/\omega_{p2})}\right) \tag{7.1}$$

In this expression, all higher-order poles and zeros are neglected. With $\omega_{p2} \gg \omega_{p1}$, the loop gain can be approximated by

$$L(s) \simeq \left(\frac{C_B}{C_A + C_B}\right)\left(\frac{A_0\omega_{p1}(1 - s/\omega_z)}{s(1 + s/\omega_{p2})}\right) \tag{7.2}$$

for frequencies $\omega \gg \omega_{p1}$.

Using Eqs. (5.7) and (5.9), we find, assuming $1 \ll g_{m7}/(g_{ds6}+g_{ds7})$ and $C_1 \ll C_c g_{m7}/(g_{ds6}+g_{ds7})$

$$A_0 \simeq \left(\frac{g_{m1}}{g_{ds2}+g_{ds4}}\right)\left(\frac{g_{m7}}{g_{ds6}+g_{ds7}}\right) \tag{7.3}$$

$$\omega_{p1} \simeq \frac{(g_{ds2}+g_{ds4})(g_{ds6}+g_{ds7})}{g_{m7}C_c} \tag{7.4}$$

resulting in the gain-bandwidth product of the loop gain

$$\omega_{tl} = \beta A_0\, \omega_{p1} \simeq \beta\left(\frac{g_{m1}}{C_c}\right) = \left(\frac{C_B}{C_A+C_B}\right)\left(\frac{g_{m1}}{C_c}\right) \tag{7.5}$$

The non-dominant pole ω_{p2} can be estimated from Eq. (6.59). In this equation, g_{m2} should be replaced by g_{m7}, and C_2 is the total load capacitance at the output. Neglecting the transistor capacitances, we have $C_2 = C_L + C_A C_B/(C_A+C_B) = 1.7$ pF, so we find

$$\omega_{p2} \simeq \frac{g_{m7}}{C_1+C_2+C_1C_2/C_c} \tag{7.6}$$

The zero is found from Eq. (6.54) with g_{m2} replaced by g_{m7}, i.e.,

$$\omega_z = \frac{g_{m7}}{C_c} \tag{7.7}$$

In Chapter 6, we saw an example showing how the zero could be eliminated by inserting a resistor in series with C_c. For the two-stage opamp, the resistor is normally implemented as a transistor working in the triode region (Allen & Holberg 2012; Chan Carusone, Johns & Martin 2012). Another approach is to try to design the opamp so that the zero is at a frequency which is much higher than the unity-gain frequency ω_t of the loop gain and also modify the requirement concerning the phase margin to take the phase shift from the zero into account. With the loop gain given by Eq. (7.2), the phase margin is

$$\text{PM} = 180^\circ - 90^\circ - \arctan(\omega_t/\omega_z) - \arctan(\omega_t/\omega_{p2}) = 90^\circ - \arctan(\omega_t/\omega_z) - \arctan(\omega_t/\omega_{p2}) \tag{7.8}$$

The unity-gain frequency ω_t of the loop gain is somewhat lower than the gain-bandwidth product ω_{tl} of the loop gain, see Fig. 6.35(b). However, with a large phase margin requiring a large ratio between ω_{p2} and ω_{tl}, the difference between ω_t and ω_{tl} is not large, and anyway using ω_{tl} instead of ω_t in Eq. (7.8) gives a conservative estimation of the phase margin. So, with $\omega_z = 10\,\omega_{tl}$, the phase shift from the zero will be less than $\arctan(0.1) = 5.7^\circ$. This can also be seen from Fig. 6.53(a) which shows that with f_z larger than about 8 to 10 times f_{tl}, the phase margin reduction caused by the zero is less than about 6°.

From Eqs. (7.5) and (7.7), we find that

$$\frac{\omega_z}{\omega_{tl}} = \left(\frac{g_{m7}}{C_c}\right)\left(\frac{C_c}{\beta\, g_{m1}}\right) = \left(\frac{1}{\beta}\right)\left(\frac{g_{m7}}{g_{m1}}\right) = 10 \;\Rightarrow g_{m7} = 2\,g_{m1} \tag{7.9}$$

Allowing 6° phase shift from the zero and an additional phase shift of about 5° from higher order poles and zeros in the opamp, we may increase the phase margin requirement to 76°. From Fig. 6.35(a), we see that a phase margin of 76° is achieved with $\omega_{p2} = 4\,\omega_{tl}$. Using Eqs. (7.5) and (7.6), we find

$$\frac{\omega_{p2}}{\omega_{tl}} = \left(\frac{g_{m7}}{C_1+C_2+C_1C_2/C_c}\right)\left(\frac{C_c}{\beta\, g_{m1}}\right) \tag{7.10}$$

From Eq. (7.10), we can find C_c as a function of C_1, C_2, g_{m1}, g_{m7}, ω_{p2}/ω_{tl} and β. This results in

$$\frac{C_c}{C_2} = \frac{1+\frac{C_1}{C_2}+\sqrt{\left(1+\frac{C_1}{C_2}\right)^2+4\left(\frac{g_{m7}}{\beta\, g_{m1}}\right)\left(\frac{\omega_{tl}}{\omega_{p2}}\right)\left(\frac{C_1}{C_2}\right)}}{2\left(\frac{g_{m7}}{\beta\, g_{m1}}\right)\left(\frac{\omega_{tl}}{\omega_{p2}}\right)} \tag{7.11}$$

In Eq. (7.11), both C_c and C_1 have been normalized with respect to the load capacitance C_2 and Fig. 7.4 shows C_c/C_2 as a function of C_1/C_2 with $k = (g_{m7}/(\beta\, g_{m1}))(\omega_{tl}/\omega_{p2})$. With $\omega_{tl} \simeq \beta\, g_{m1}/C_C$ and $\omega_z = g_{m7}/C_C$, the parameter k is equal to ω_z/ω_{p2}, and using $\omega_z = 10\,\omega_{tl}$ and $\omega_{p2} = 4\,\omega_{tl}$, we find $k = 2.5$.

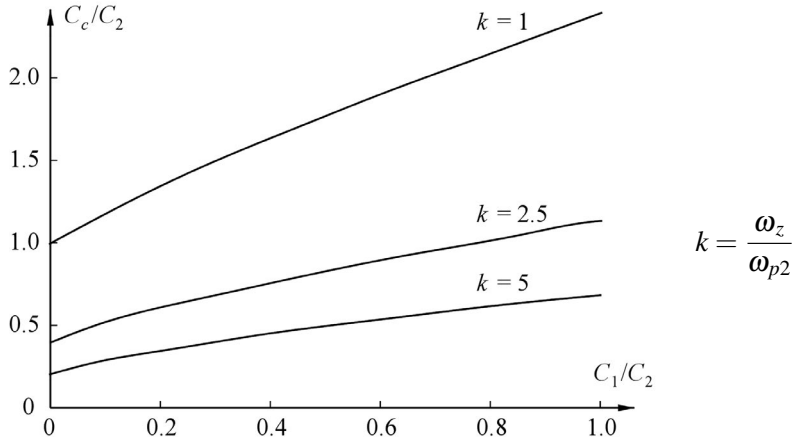


Figure 7.4: Compensation capacitor C_c versus parasitic transistor capacitance C_1.

Assuming that the transistor capacitances are much smaller than C_L, we may insert $C_1 = 0$ in Eq. (7.10), leading to

$$\frac{\omega_{p2}}{\omega_{tl}} = \left(\frac{g_{m7}}{C_2}\right)\left(\frac{C_c}{\beta\, g_{m1}}\right) = 4 \;\Rightarrow\; C_c = 0.4\, C_2 = 0.68 \text{ pF} \tag{7.12}$$

The same result is found from Fig. 7.4 with $k = 2.5$ and $C_1 = 0$.

Next, we may use the bandwidth specification to find the transconductances. The bandwidth depends on ω_{tl}. A conservative estimate is $\text{BW} = \beta\, g_{m1}/(2\,\pi C_c)$ from Eq. (7.5). From Fig. 6.30(a), we see that we may apply a factor between 1 and 1.41 but here we will use the conservative estimate and leave some margin for parasitic effect. This yields

$$2\pi \times 20 \text{ MHz} = \left(\frac{C_B}{C_A + C_B}\right)\left(\frac{g_{m1}}{C_c}\right) \;\Rightarrow\; g_{m1} = 0.43 \text{ mA/V} \tag{7.13}$$

and from Eq. (7.9), $g_{m7} = 2\, g_{m1} = 0.86$ mA/V.

We now have all the capacitor values and the transconductance values for the schematic in Fig. 7.3. We do not have the resistor values corresponding to the small-signal transistor drain-source resistances but since they do not enter into the design equations, the exact values are not important. Let us assume that the transistors can be designed to have g_m larger than about 50 times g_{ds}. For the resistors, we may then use values of about 25 times $1/g_m$, taking into account that each of the resistors is a parallel connection of two output resistors. With the values found for g_{m1} and g_{m7}, we may estimate $R_1 \simeq 70$ kΩ and $R_2 \simeq 35$ kΩ. With these values, we can now verify the bandwidth and phase-margin specifications using LTspice, even before starting the dimensioning of the transistors.

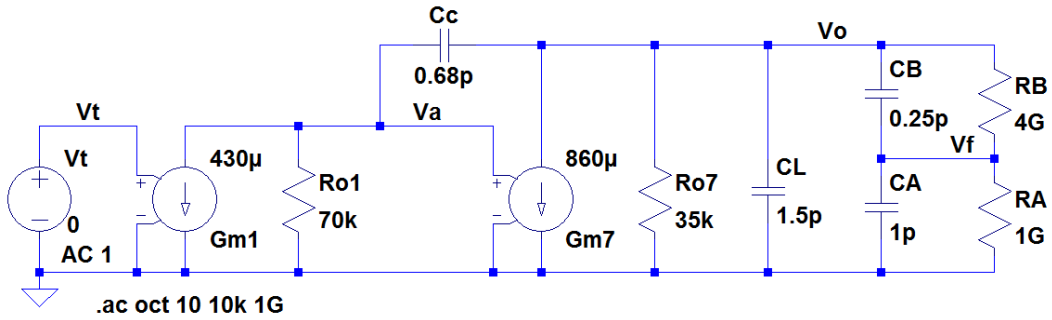


Figure 7.5: LTspice schematic for simulation of loop gain $L(s) = V_f(s)/V_t$.

Figure 7.5 shows an LTspice schematic corresponding to the circuit shown in Fig. 7.3 with the values of capacitors, transconductances and resistors inserted. In order to establish a dc path to the feedback node V_f, the large resistors R_A and R_B are connected in parallel with C_A and C_B. They have been selected such that $R_A/R_B = C_A^{-1}/C_B^{-2}$, and they are large enough that they do not influence the transfer function of the circuit in the frequency range of interest, but they ensure that LTspice can find the correct dc bias point for the '.ac' simulation.

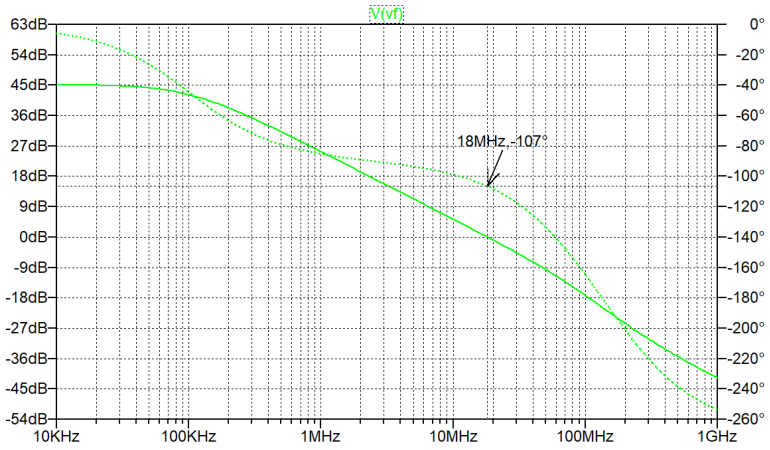


Figure 7.6: Plot of loop gain from LTspice simulation.

Figure 7.6 shows the simulated loop gain. We notice that the loop-gain unity-gain frequency is about 18 MHz, as expected slightly lower than the gain-bandwidth product $f_{tl} = \omega_{tl}/(2\pi) = 20$ MHz calculated using Eq. (7.5). We also see that the phase margin is about 73°, slightly higher than the phase margin estimated from Eq. (7.8) with ω_t replaced by ω_{tl}. At the frequency $f_{tl} = 20$ MHz, we find a phase shift of $-107°$ from Fig. 7.6, matching the calculated values reasonably well.

In the schematic in Fig. 7.5, we may also close the feedback loop by connecting V_f back to the inverting input of g_{m1} and run a '.ac' simulation to find the -3 dB frequency of V_o. Figure 7.7 shows the result of this simulation. From this, we can verify the low-frequency gain of 14 dB and the -3 dB frequency of 26 MHz, leaving some margin to the specification. Thus, we conclude that the values found for g_{m1}, g_{m7} and C_c constitute a good starting point for the transistor design.

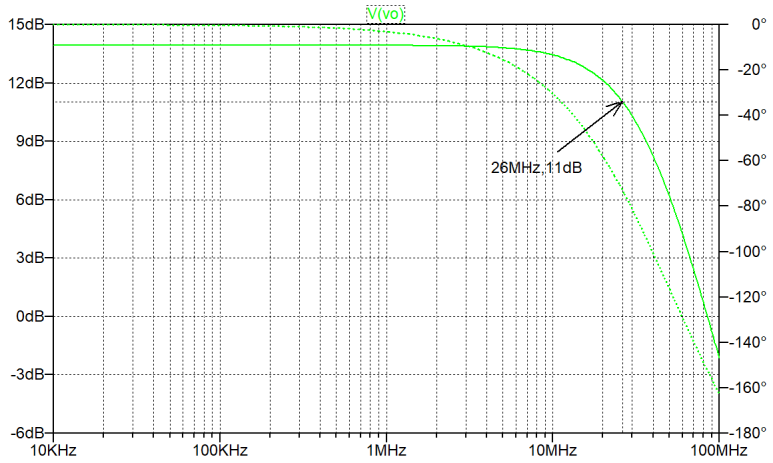


Figure 7.7: Plot of closed-loop gain from LTspice simulation.

7.3 Bias point and transistor dimensions

In order to design the transistors in the circuit shown in Fig. 7.2, we must find bias conditions for each transistor ensuring that the requirements concerning slew rate and input/output voltage range can be met while achieving the transconductances already found. Also the values of W and L must be designed for each transistor.

We may start by considering the slew rate. The slew rate is limited by the current available to charge and discharge the capacitors in the opamp when the differential input voltage is so large that it forces one of the input transistors into the cut-off region, implying that all of the tail current I_{D5} flows in the other input transistor, see Fig. 4.39. Thus, when $v_{G1} - v_{G2} > \sqrt{2}V_{\text{eff1}}$, the input stage is able to source the current I_{D5} into the capacitor C_c and when $v_{G2} - v_{G1} > \sqrt{2}V_{\text{eff1}}$, the input stage can sink the current I_{D5}, assuming an ideal current mirror M_3 - M_4.

A simplified model for analyzing the slew rate is shown in Fig. 7.8. The input current i_{IN} to the circuit is the output current from the differential input stage, i.e., $-I_{D5} \leq i_{IN} \leq +I_{D5}$ and the capacitors are those defined in the bandwidth analysis, i.e., $C_2 = C_L + C_A C_B/(C_A + C_B)$. For simplicity, M_6 has been replaced by the dc current source I_{D6}.

Assuming that all voltages change linearly with time, we have a constant dv/dt for all capacitors, leading to constant capacitor currents, so also the drain current i_{D7} must be constant. Neglecting the channel-length modulation, this implies that the voltage v_{GS7} is constant, hence the current $i_{C1} = C_1 dv_{GS7}/dt$ is 0 and $i_{Cc} = C_c d(v_{GS7} - v_O)/dt = -C_c dv_O/dt$. Thus, a node equation at the gate of M_7 shows $i_{IN} = -C_c dv_O/dt$.

With i_{IN} positive, we have a falling voltage at the output and using Kirchhoff's current law at the output node, we find

$$i_{D7} = i_{IN} + I_{D6} - i_{C2} \tag{7.14}$$

where $i_{C2} = C_2 dv_O/dt$ is negative. Thus, i_{D7} is positive and neglecting the channel-length modulation, the gate-source voltage v_{GS7} may be found from the simplified Shichman-Hodges model, Eq. (3.22).

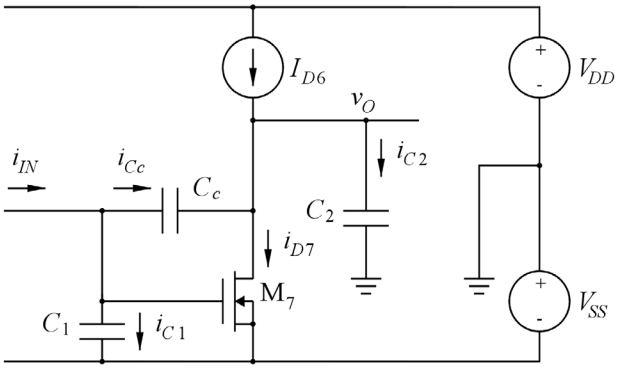


Figure 7.8: Circuit model for slew rate analysis.

With the maximum value of i_{IN} equal to I_{D5}, we find $|dv_O/dt|_{\max} = I_{D5}/C_c$, i.e., the slew rate for a falling output is

$$\mathrm{SR}^- = \left|\frac{dv_O}{dt}\right|_{\max} = I_{D5}/C_c \tag{7.15}$$

With i_{IN} negative, we have a rising voltage at the output and $i_{C2} = C_2 dv_O/dt$ is positive. Kirchhoff's current law, Eq. (7.14), still applies but it must be noted that i_{D7} cannot assume a negative value. Thus, with both i_{IN} and $-i_{C2}$ negative, Eq. (7.14) results in a positive value of i_{D7} only if $I_{D6} > |i_{IN}| + i_{C2}$. In this case, the slew rate is limited by the input current and we find

$$\mathrm{SR}^+ = \left|\frac{dv_O}{dt}\right|_{\max} = I_{D5}/C_c \tag{7.16}$$

The current in M_7 is

$$i_{D7} = I_{D6} - (C_c + C_2)\frac{dv_O}{dt} = I_{D6} - (C_c + C_2)\mathrm{SR}^+ = I_{D6} - (1 + C_2/C_c)I_{D5} \tag{7.17}$$

and using the simplified Shichman-Hodges model given by Eq. (3.22), the corresponding value of v_{GS7} may be found. It is smaller than the value found for a falling output voltage because i_{D7} is smaller during a rising output than during a falling output.

If $I_{D6} < I_{D5} + i_{C2}$, the current i_{D7} is 0 and the current i_{IN} is smaller than I_{D5}, so Eq. (7.14) reduces to

$$I_{D6} = -i_{IN} + i_{C2} = C_c\frac{dv_O}{dt} + C_2\frac{dv_O}{dt} = (C_c + C_2)\frac{dv_O}{dt} \tag{7.18}$$

implying that

$$\mathrm{SR}^+ = \left|\frac{dv_O}{dt}\right|_{\max} = \frac{I_{D6}}{C_c + C_2} \tag{7.19}$$

Combining Eqs. (7.16) and (7.19), we find

$$\mathrm{SR}^+ = \min\left(\frac{I_{D5}}{C_c}, \frac{I_{D6}}{C_c + C_2}\right) \tag{7.20}$$

Summarizing, Eqs. (7.15) and (7.20) impose the following constraints on the bias currents.

$$\begin{aligned} \mathrm{SR} &= \frac{I_{D5}}{C_c} \\ &\Rightarrow \quad I_{D5} = \mathrm{SR}\cdot C_c = 20.4\ \mu\mathrm{A} \end{aligned} \tag{7.21}$$

and

$$\begin{aligned} \text{SR} &= \frac{I_{D6}}{C_c+C_2} = \frac{I_{D6}}{C_c+C_L+C_AC_B/(C_A+C_B)} \\ &\Rightarrow \quad I_{D6} = \text{SR}\cdot(C_c+C_L+C_AC_B/(C_A+C_B)) = 71.4\ \mu\text{A} \end{aligned} \tag{7.22}$$

These values of bias current are minimum values, and it may be a good idea to select somewhat larger bias-current values in order to leave room for parasitic capacitances.

The bias current for M_1 and M_2 is obviously $I_{D5}/2$, and the bias current for M_7 is $I_{D7} = I_{D6}$, so for M_1, M_2 and M_7 we have found both the transconductances and the minimum bias currents. For M_1 and M_2, we must also ensure that V_{GS} is small enough to allow a maximum common-mode input voltage of 0.15 V. For the current-source transistor M_6, we also have a requirement for the output voltage range. With $V_{DD} - v_{O\,\max} = 0.15$ V, we may select the saturation voltage $|V_{GS} - V_{tp}|$ of M_6 to be no larger than 0.15 V, and M_5 and M_6 have the same gate-source voltage. This implies that the maximum voltage for $|V_{GS1}| = |V_{GS2}|$ is $V_{DD} - |V_{DS\text{sat}5}| - v_{IN\,\max} = 0.6$ V. With $V_{tp} = -0.42$ V from Table 3.1, this leaves a maximum of 0.18 V for $|V_{DS\text{sat}1}|$. Using Eq. (3.65), we find that $I_{D1} = 10.2\ \mu$A and $|V_{GS1} - V_{tp}| = |V_{DS\text{sat}1}| = 0.18$ V gives a transconductance of 113 μA/V.

From Eq. (3.71), the ratio W/L may be found from the Shichman-Hodges transistor model. With $g_m = 113\ \mu$A/V and $|V_{GS} - V_t| = 0.18$ V, we find $W/L \simeq 14$. However, a transconductance of 113 μA/V is not enough. It is a factor of 4 too small, so we must increase I_{D1} and/or decrease $|V_{DS\text{sat}1}|$. Using Eqs. (3.70) and (3.71), we may find a combination of drain current, saturation voltage and transistor geometry resulting in the required value of g_{m1}. However, we noticed in Chapter 3 that it is not easy to obtain a good match over a wide range of variation in g_m between the Shichman-Hodges model and the BSIM3 transistor models specified for the opamp. So, rather than using the Shichman-Hodges parameters, we may use LTspice with the BSIM model to find the transconductance g_m versus channel width W for selected values of bias current. First, we select a value for the channel length, and for this design, we choose (somewhat arbitrarily) $L_1 = 1$ μm, i.e., about five times the minimum dimension for the process.

In order to simulate g_m versus W, we use the circuit shown in Fig. 7.9(a). This is a diode-connected transistor where the small-signal resistance from drain to source is $(r_{ds} \parallel (1/g_m)) \simeq 1/g_m$ as g_m is much larger than g_{ds} with the transistor in the active region. A '.tf' simulation with the drain current as input and the drain voltage as output results in g_m as '`1/id#input_resistance`'. With the channel width W defined as a parameter stepped from 50 μm to 150 μm, we can directly plot g_m versus W, Fig. 7.9(b).

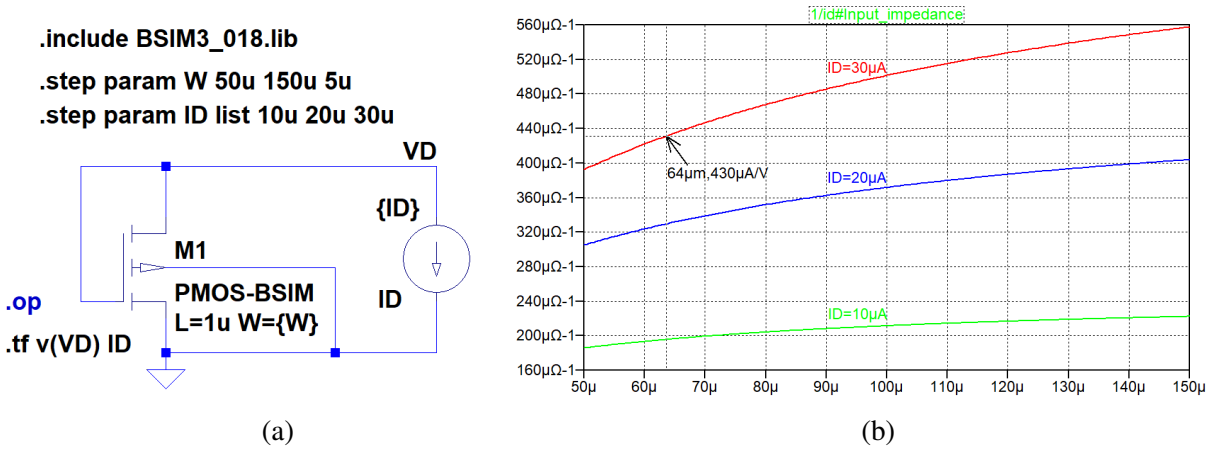


Figure 7.9: Simulation showing g_m versus W for a PMOS transistor with $L = 1$ µm using the BSIM3 transistor model.

In Fig. 7.9, also the drain current is defined as a parameter, and g_m is shown versus W for three different values of the bias current I_D. From the plot, we see that $g_m = 0.43$ mA/V cannot be achieved with a drain current of 10 µA with a reasonable transistor geometry. Instead, we select a drain current of 30 µA so that we can achieve a g_m of 0.43 mA/V with a transistor with $W/L = 64$. Next, we may run a '.op' simulation (shown as a comment in Fig. 7.9(a)) to find $|V_{GS}|$ for $W/L = 64$ and $I_D = 30$ µA. This results in $|V_{GS}| = 586$ mV, leaving more than 150 mV for the drain-source voltage of M_5 when the input voltage reaches its maximum value. This is enough to maintain M_5 in the active region, even with the input voltage at its maximum value.

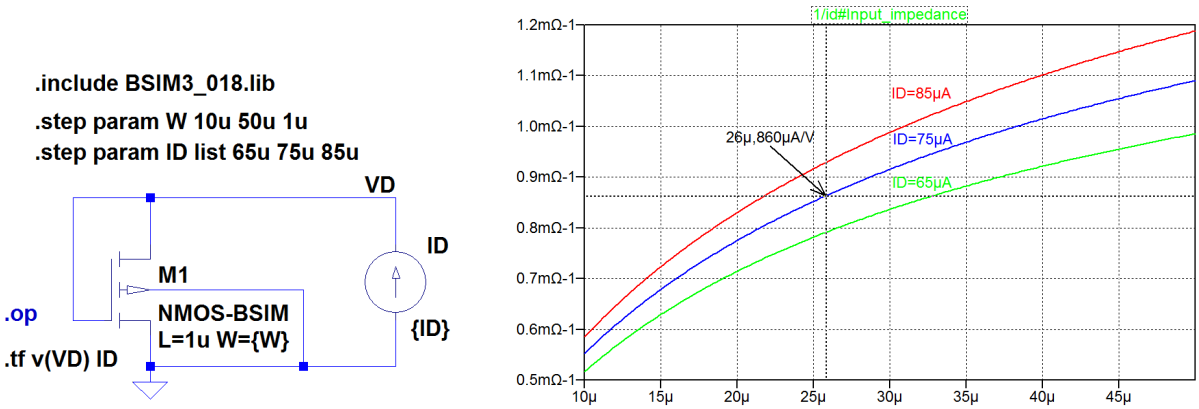


Figure 7.10: Simulation showing g_m versus W for an NMOS transistor with $L = 1$ µm using the BSIM3 transistor model.

Next, we may design M_7 to give the desired transconductance using a similar approach. Figure 7.10 shows a diode-connected NMOS transistor for simulating g_m versus channel width and drain current. From the simulation plot, we select a drain current of 75 µA, slightly higher than the minimum value found from Eq. (7.22). With $W/L = 26$, this gives a transconductance of 0.86 mA/V.

Again selecting a channel length of 1 µm, we may run a '.op' simulation with $W = 26$ µm and $I_D = 75$ µA to find the saturation voltage for M_7. This shows a saturation voltage of 0.137 V (listed in the error log file) which is small enough that the minimum output voltage can be reached with M_7 in the active region.

The active-load transistors M_3 and M_4 may now be designed using the scaling rule defined by Eq. (5.5). With $L_3 = L_4 = 1$ µm, this results in $W_3 = W_4 = 10.4$ µm.

The only transistors still to be designed are the bias transistors M_5, M_6 and M_8. For these, the requirement is a very low saturation voltage, less than 0.15 V, so that the upper limit of the output voltage can be reached with M_6 in the active region.

Using the Shichman-Hodges model, we may calculate W/L as

$$\frac{W_6}{L_6} = \frac{2I_{D6}}{\mu_p C_{ox}|V_{DS\,\mathrm{sat6}}|^2} \simeq 150\ \mu\mathrm{m} \tag{7.23}$$

In Eq. (7.23), the channel-length modulation has been neglected. We may also simulate the current

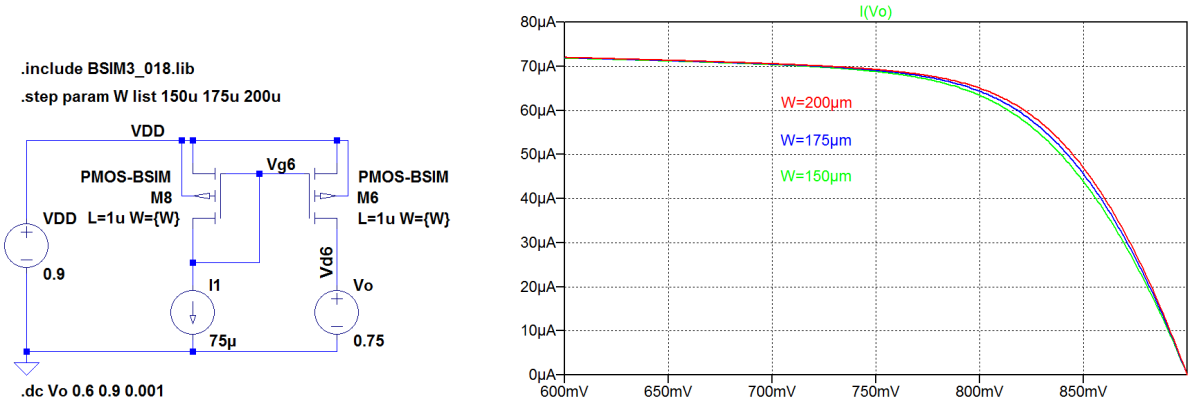


Figure 7.11: Simulation of PMOS current mirror showing output current versus output voltage.

mirror M_8 - M_6 in order to check the upper limit. Figure 7.11 shows a schematic and the plot from a '.dc' simulation where the output voltage is swept towards the positive supply voltage. For comparison, three different values of channel width have been simulated. We see that at an output voltage of 0.75 V, there is only a minor difference between the characteristics corresponding to three different channel widths, so we select $W_6 = 150$ µm and $L_6 = 1$ µm. From a '.op' simulation of the circuit, we may verify that the saturation voltage of M_6 is 0.111 V with $W = 150$ µm, i.e., smaller than 0.15 V. From the '.op' simulation, we can also find the gate voltage $V_{G6} = 0.310$ V.

Transistors M_5 and M_8 may then be designed using the scaling rule defined by Eq. (5.6). With $L_5 = 1$ µm and $I_{D5} = 60$ µA, we find $W_5 = 120$ µm. For M_8, we can select a smaller current. Using $I_{D8} = 10$ µA, and $L_8 = 1$ µm, we find $W_8 = 20$ µm.

Finally, the bias resistor R_{BIAS} can be found as $R_{BIAS} = (V_{G6} - (-V_{SS}))/I_{D8} = 121$ kΩ.

With all transistor dimensions in place, the complete opamp is now ready for simulation. The transistor channel widths are summarized in Table 7.1 which also shows the calculated bias currents for each transistor and values of g_m for the transistors where g_m is included in the design equations.

Transistor number	1	2	3	4	5	6	7	8
Channel width	64 µm	64 µm	10.4 µm	10.4 µm	120 µm	150 µm	26 µm	20 µm
Bias current	30 µA	30 µA	30 µA	30 µA	60 µA	75 µA	75 µA	10 µA
Transconductance	0.43 mA/V	0.43 mA/V					0.86 mA/V	

Table 7.1: Calculated transistor parameters for initial simulations. All transistors have a channel length of 1 µm.

7.4 Design verification and iteration

Figure 7.12 shows the LTspice schematic for the complete amplifier with feedback using the transistor geometries listed in Table 7.1. The drain and source areas are calculated as $(W \times 0.5\ \mu\text{m})$ and perimeters are calculated as $(W + 1\ \mu\text{m})$. Also shown are directives for '.op' simulation, '.dc' simulation, '.ac' simulation and '.tran' simulation. For '`Vin`', a 'pulse' specification for the transient simulation is shown. For the '.op' simulation and the '.ac' simulation, this must be replaced by a dc value of 0.

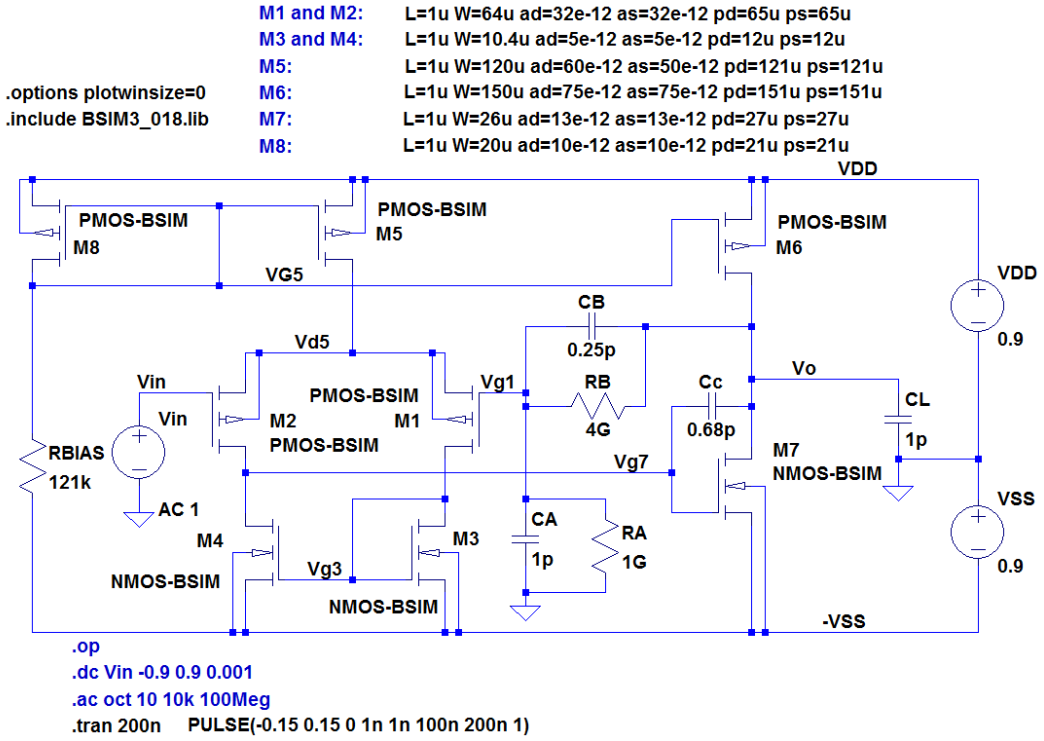


Figure 7.12: LTspice schematic for simulation of closed-loop performance.

Again, large resistors are inserted in parallel with C_A and C_B to ensure a dc feedback path. They are required not only in the simulation model but also in a real implementation of the opamp, and it is a non-trivial task to design such large resistors (Tajalli, Leblebici & Brauer 2008; Bikumandla, et al. 2004). However, this is beyond the scope of this book.

From the '.op' simulation, we find the values listed in Table 7.2. We also find that with an input voltage of 0, the output voltage is 0.3 mV and all node voltages assume reasonable values.

We see that the simulated values match the calculated values reasonably well. The minor differences are probably due to the channel-length modulation which has been neglected in the calculations.

Transistor number	1	2	3	4	5	6	7	8
Channel width	64 µm	64 µm	10.4 µm	10.4 µm	120 µm	150 µm	26 µm	20 µm
Bias current	29 µA	29 µA	29 µA	29 µA	58 µA	78 µA	78 µA	10 µA
Transconductance	0.42 mA/V	0.42 mA/V					0.89 mA/V	

Table 7.2: Simulated bias currents and transconductances from initial simulation.

Next, we run a '.dc' simulation. The simulation has been run with a sweep of the input voltage over the entire supply voltage range in order to see the saturation limits. From the output plot, it is apparent that the output voltage saturates at 0.9 V for high values of input voltage and at -0.88 V for low values of input voltage. Figure 7.13 shows the plot of the output voltage for $|v_{IN}| \leq 300$ mV. From the plot, we see that the specifications concerning input voltage range and output voltage range are fulfilled.

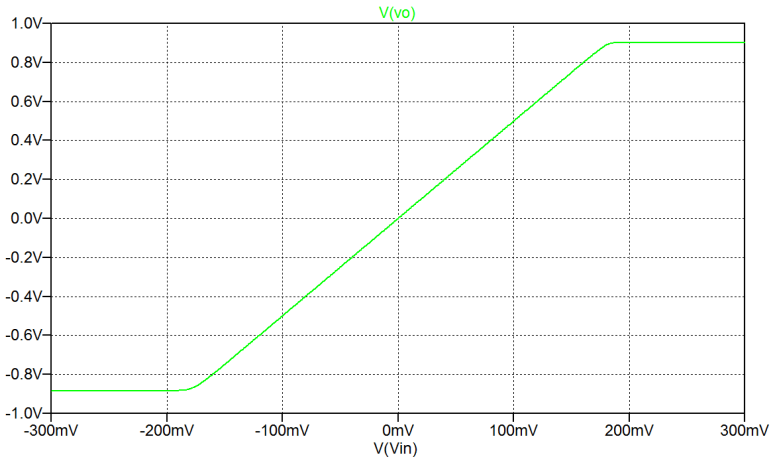


Figure 7.13: Plot of closed-loop dc transfer characteristics from LTspice simulation.

For checking the bandwidth, we run a '.ac' simulation. The result of this is shown in Fig. 7.14. We see that the bandwidth requirement is fulfilled and there is a margin by a factor of about 1.2, corresponding very well to the factor indicated in Fig. 6.30(a). We also note that the frequency response shows a slight peaking, indicating that the phase margin requirement may not be fulfilled. Also, the gain at low frequencies is slightly higher than 14 dB but still within the specification. This may be caused by the input capacitance of the inverting opamp input. This capacitance appears in parallel with C_A in the feedback network, thus increasing the ratio $(C_A + C_B)/C_B$.

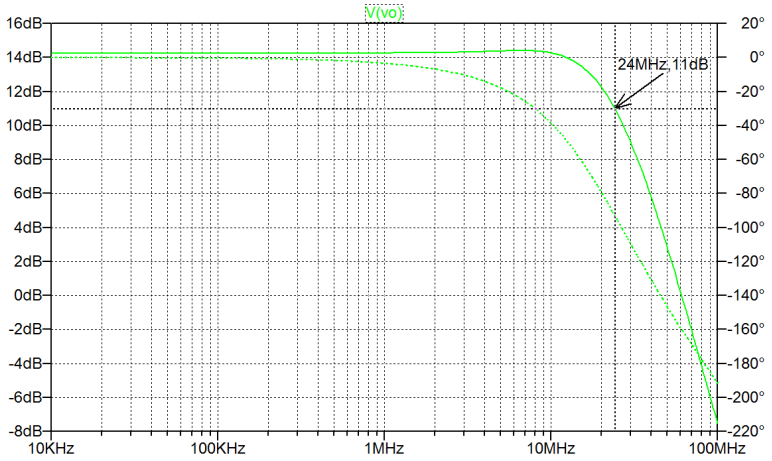


Figure 7.14: Plot of closed-loop ac characteristics from LTspice simulation.

For checking the slew rate, we run a transient simulation with the input defined as a pulse signal with very short rise time and fall time. The result of this is shown in Fig. 7.15. We see that the slew rate for the falling edge of the output voltage is much higher than 30 V/µs but for the rising edge, the slew rate is about 27.5 V/µs, slightly below the requirement of 30 V/µs. This is most likely caused by parasitic capacitances in the opamp which have been neglected in the calculations in Section 7.3.

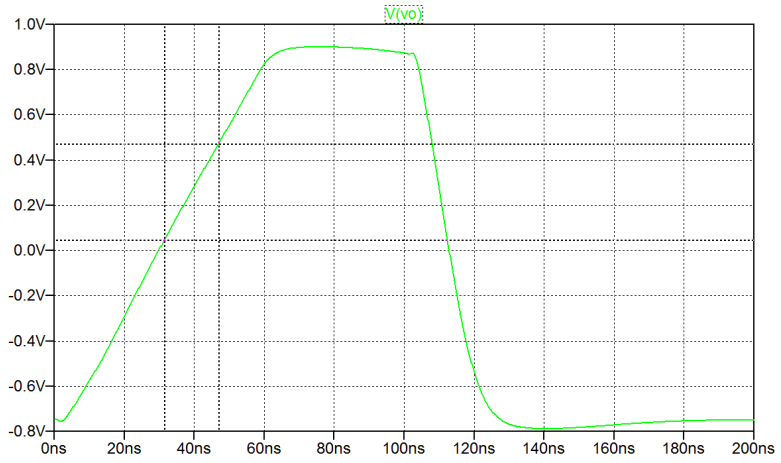


Figure 7.15: Plot of closed-loop transient response from LTspice simulation.

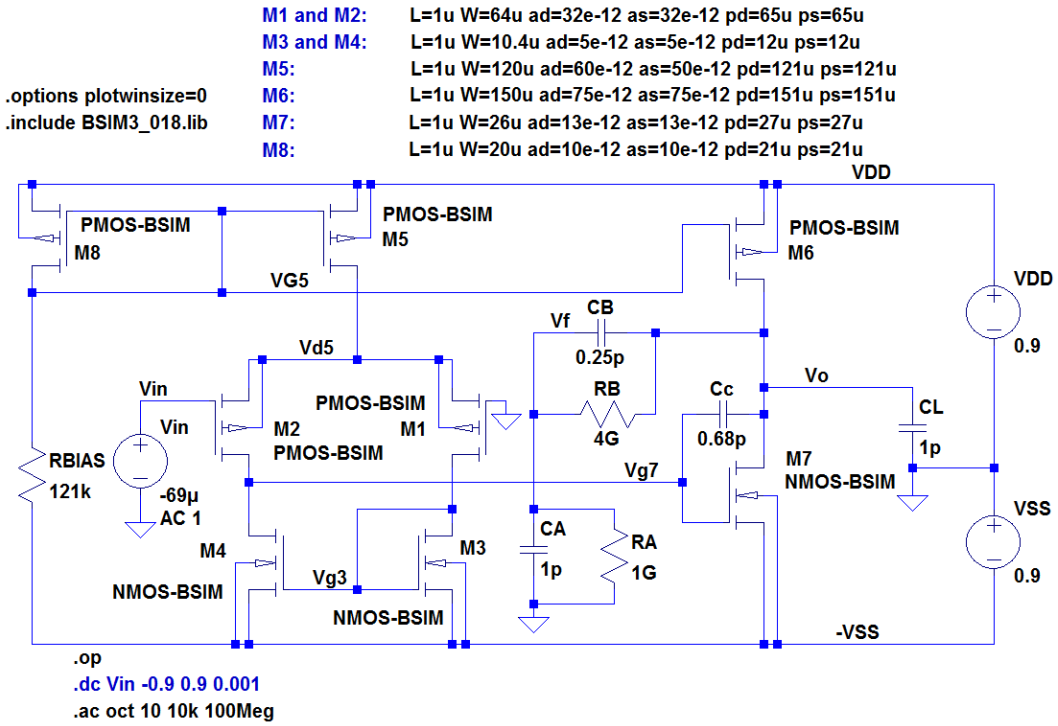


Figure 7.16: LTspice schematic for simulation of open-loop performance.

For checking the phase margin, we run a simulation of the open-loop gain. Figure 7.16 shows the LTspice schematic for this. The first simulation is a '.dc' simulation in order to find the dc value of '`Vin`' resulting in an output voltage of 0. This is found to be -69 µV, and with this value inserted, a '.op' simulation can be used to verify the bias point.

Next, a '.ac' simulation is run. The result of this is shown in Fig. 7.17. From the plot of the feedback voltage V_f, we see that the phase margin is about 65°, just fulfilling the specified requirement. We also see that the unity gain frequency of the loop gain is 16.7 MHz which is somewhat below the closed-loop bandwidth, again in accordance with the results found in Chapter 6 for a second-order feedback system.

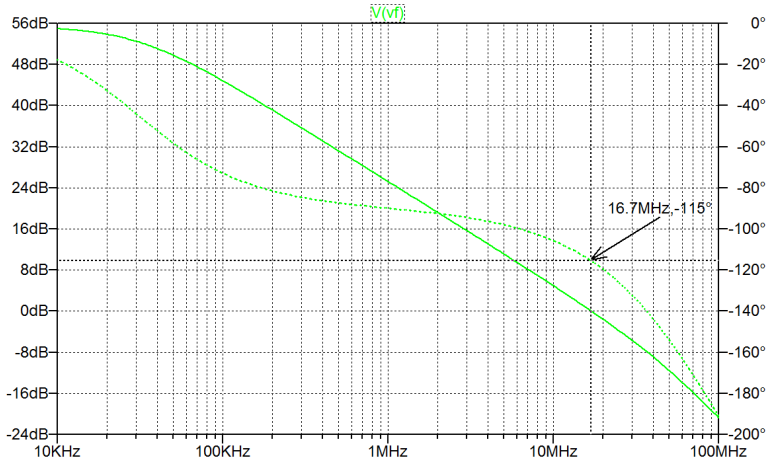


Figure 7.17: Plot of open-loop ac characteristics from LTspice simulation.

Summarizing, the only specification which has not been met is the rising-edge slew rate. The design equation for this is Eq. (7.22). The reason for the smaller slew rate appears to be an additional parasitic capacitance at the output of the opamp. We cannot change the capacitances, so the obvious way to increase the slew rate is to increase the current in M_6. This will also increase the transconductance of M_7 as we can see from Fig. 7.10. According to Eqs. (7.6) and (7.7), this will increase both ω_{p2} and ω_z, thereby increasing the phase margin. This is fine as long as it does not decrease the closed-loop bandwidth too much (see Fig. 6.30(a)), and the simulations of the closed-loop bandwidth showed some margin, so most likely, an increase of the bias current in M_6 and M_7 will bring all specifications within the required limits. The currents in M_6 and M_7 are increased by a scaling of the channel widths of the transistors. Using the same scale factor, W_6 and W_7 are increased to $W_6 = 173$ µm and $W_7 = 30$ µm. New simulations of both closed-loop performance and open-loop performance show a phase margin of 67°, a bandwidth of 23 MHz and a slew rate of 31 V/µs, fulfilling the required specification.

Thus, with this modification of W_6 and W_7, all specification requirements are fulfilled and this concludes the design of the two-stage opamp. The resulting specifications are summarized in the following table.

It should be emphasized that for a real implementation of the circuit, further simulations and design iterations are necessary to ensure correct operation also when process variations, supply voltage variations and temperature variations are taken into account (Bruun 2020).

	Gain	Bandwidth	Slew rate	Phase margin	Output range	Input range
Requirement	14 ± 0.3 dB	$\geq$20 MHz	$\geq$30 V/µs	$\geq$65°	$\pm$0.75 V	$\pm$0.15 V
Simulation	14.2 dB	23 MHz	31 V/µs	67°	$\simeq \pm$0.8 V	$\simeq \pm$0.16 V

Table 7.3: Specification requirements and simulated performance.

References

Allen, PE., & Holberg, DR. 2012, *CMOS Analog Integrated Circuit Design*, Third Edition, Oxford University Press, New York, USA.

Bikumandla, M., Ramírez-Angulo, J., Urquidi, C, Carvajal, RG. & Lopez-Martin, AJ. 2004, 'Biasing CMOS amplifiers using MOS transistors in subthreshold region', *IEICE Electronics Express*, vol. 1, no. 12, pp. 339-345, 25 September 2004.

Bruun, E. 2020, *CMOS Integrated Circuit Simulation with LTspice*, Third Edition, bookboon.
Available from: http://bookboon.com/en/cmos-integrated-circuit-simulation-with-ltspice-ebook

Chan Carusone, T., Johns, D. & Martin, K. 2012, *Analog Integrated Circuit Design*, Second Edition, John Wiley & Sons, Inc., Hoboken, USA.

Hambley, A. R. 2014, *Electrical Engineering, Principles and Applications*, Sixth Edition, Pearson Education Ltd., Harlow, UK.

Tajalli, A., Leblebici, Y. & Brauer, EJ. 2008, 'Implementing ultra-high-value floating tunable CMOS resistors', *Electronics Letters*, vol. 44, no. 5, pp. 349-350, 28 February 2008.

Multiple-choice test

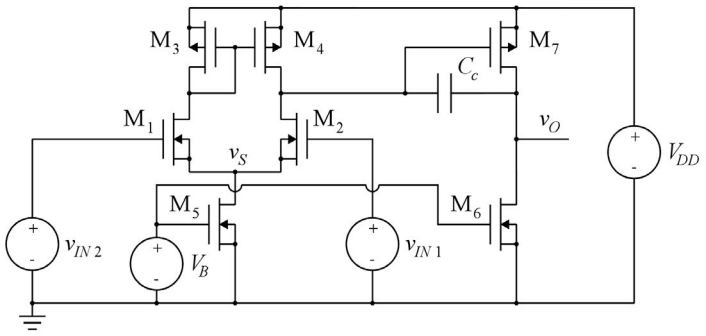
1. Complete the following statements by selecting the appropriate continuation from the table below.

A:	The gain-bandwidth product of a Miller-compensated two-stage opamp is given by ...
B:	A Miller-compensated two-stage opamp has a right-half-plane zero at a frequency given by ...
C:	The frequency of the right-half-plane zero in a Miller-compensated two-stage opamp may be pushed to a higher frequency relative to the gain-bandwidth product by ...
D:	When the compensation capacitor in a Miller-compensated two-stage opamp is increased, the phase margin is ...
E:	When the feedback factor for an opamp with feedback is reduced, the phase margin is ...
F:	When the channel-width-to-length ratio of the bias transistor providing the tail current to the input stage is decreased, the input voltage range is ...
G:	When the load capacitor in a Miller-compensated two-stage opamp is increased, the phase margin is ...
H:	When the compensation capacitor in a Miller-compensated two-stage opamp is increased, the slew rate is ...
I:	The channel-width-to-length ratio for the transistors forming the active load to the differential input pair is designed from a scaling of ...

Continuation:

1:	increasing the compensation capacitor.
2:	increased.
3:	the ratio between the transconductance of the input transistors and the compensation capacitor.
4:	decreasing the transconductance of the second-stage gain transistor.
5:	decreased.
6:	the second-stage gain transistor and the bias currents in the input stage and the second stage.
7:	the input transistors and the bias current in the input stage.
8:	negative.
9:	the ratio between the transconductance of the second-stage gain transistor and the compensation capacitor.
10:	the transistors defining the bias currents.
11:	the product of the output resistance of the input stage and the input capacitance of the second gain stage.
12:	unchanged.
13:	the ratio between the transconductance of the input transistors and the load capacitance.
14:	the product of the output resistance of the input stage and the compensation capacitor.
15:	increasing the transconductance of the second-stage gain transistor.
16:	decreasing the compensation capacitor.
17:	increasing the transconductance of the input transistors.

2. The amplifier shown below has a supply voltage of $V_{DD} = 1.8$ V, and the absolute value of the overdrive voltage is 0.2 V for all transistors. All NMOS transistors have a threshold voltage of 0.40 V without bulk effect and all PMOS transistors have a threshold voltage of -0.42 V without bulk effect. The bias current in M_5 and M_6 is $I_{D5} = I_{D6} = 0.1$ mA.



The minimum value of the common-mode input voltage with all transistors in the active region is

A: 0.2 V

B: 0.6 V

C: 0.8 V

3. The bias voltage V_B for the amplifier shown above has a value of

A: 0.2 V

B: 0.6 V

C: 0.8 V

4. The maximum value of the common-mode input voltage for the amplifier shown above with all transistors in the active region is

   A: 1.58 V
   B: 1.60 V
   C: 1.81 V

5. Assume now that M_1 and M_2 in the amplifier above has bulk connected to ground rather than to source, causing their threshold voltage to increase to 0.63 V. The maximum value of the common-mode input voltage for the amplifier shown above with all transistors in the active region is now

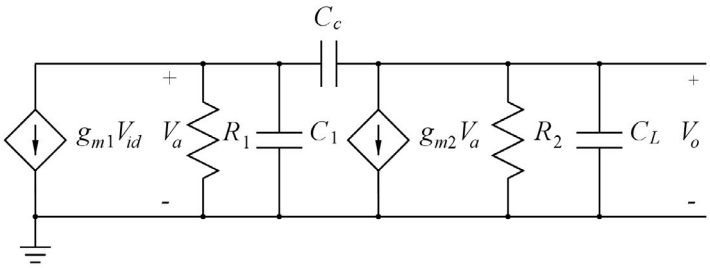
   A: 1.58 V
   B: 1.60 V
   C: 1.81 V

6. In order to achieve a gain-bandwidth product of 20 MHz for the amplifier shown above, the capacitor C_c should have a value of approximately

   A: 4 pF
   B: 8 pF
   C: 25 pF

Problems

Problem 7.1



The amplifier shown above is to be used in a feedback amplifier with the feedback factor $\beta = 1$. The transfer function V_o/V_{id} where V_{id} is a differential input voltage has a zero at the frequency ω_z and a non-dominant pole at ω_{p2}, and the gain-bandwidth product is ω_{ta}. You may assume $g_{m1}R_1 \gg 1$ and $g_{m2}R_2 \gg 1$.

Find a relation between g_{m1} and g_{m2} so that $\omega_z \simeq 10\,\omega_{tl}$ where ω_{tl} is the gain-bandwidth product of the loop gain. Also find an expression for C_c as a function of g_{m1}, g_{m2}, C_1 and C_L so that $\omega_{p2} \simeq 2\,\omega_{tl}$.

Hint: Use Eq. (7.11).

With $g_{m1} = 0.5$ mA/V, $C_L = 5$ pF and $C_1 = 1.7$ pF, calculate g_{m2} and C_c so that $\omega_z \simeq 10\,\omega_{tl}$ and $\omega_{p2} \simeq 2\,\omega_{tl}$ and use LTspice to find phase margin and bandwidth for the feedback amplifier.

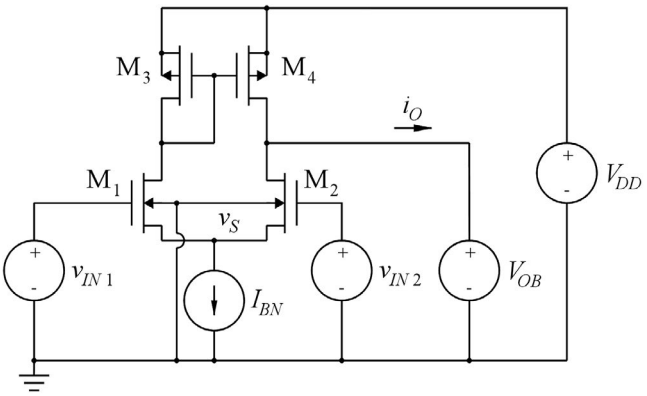
Problem 7.2

Redesign the circuit from Problem 7.1 to have a phase margin of 70° by increasing the value of C_c. Find the new value of C_c and find the bandwidth of the modified circuit.

Problem 7.3

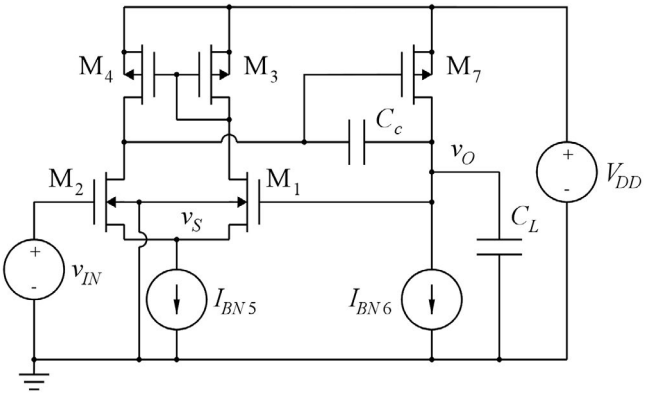
Redesign the circuit from Problem 7.1 to have a phase margin of 70° by increasing the value of g_{m2}. Find the new value of g_{m2} and find the bandwidth of the modified circuit.

Problem 7.4



For the differential gain stage shown above, assume $V_{DD} = 1.8$ V, $V_{OB} = 1.0$ V and $I_{BN} = 0.1$ mA. Also, all transistors have a channel length of $L = 1$ μm and they all have the same channel width W.
Design the transistors M_1 and M_2 to provide a transconductance $g_m = i_o/v_{id} = 0.5$ mA/V where $v_{ID} = v_{IN1} - v_{IN2}$.
Assume a common-mode input voltage of 1 V and use the BSIM3 transistor models shown in Fig. 3.44.

Problem 7.5



Design a two-stage opamp with feedback as shown above. The load capacitor C_L has a value of 5 pF which is considered to be much larger than all transistor capacitances, except C_{gs7}. The bias value of the input voltage is 1 V and the supply voltage is $V_{DD} = 1.8$ V. Use the input stage designed in Problem 7.4, and design I_{BN6} and M_7 to provide a g_{m7} of 5 mA/V. The transistors are modeled by the BSIM3 transistor models shown in Fig. 3.44.
Estimate C_{gs7} using $C_{ox} = 8.5$ fF/(μm)^2.
Design C_c so that the non-dominant pole is at a frequency which is $\simeq 2\,\omega_{tl}$ where ω_{tl} is the gain-bandwidth product of the loop gain.

Problem 7.6

Use LTspice to simulate the circuit designed in Problem 7.5. Use drain areas and source areas of $(W \times 0.5\ \mu\text{m})$ and drain perimeters and source perimeters of $(W + 1\ \mu\text{m})$. Find the phase margin and the bandwidth for the amplifier.

Problem 7.7

Suggest at least one modification to the circuit designed in Problem 7.5 in order to increase the phase margin to 65°.
Verify your suggestion using LTspice and find the resulting bandwidth.

Problem 7.8

For the circuit designed in Problem 7.5, use LTspice to find the input voltage range for which all transistors operate in the active region. Assume that the voltage across the current sources I_{BN5} and I_{BN6} must be at least 0.2 V.

Problem 7.9

For the circuit designed in Problem 7.5, use LTspice to find the slew rate of the output voltage when applying an input voltage step with the maximum possible value within the input voltage range found in Problem 7.8.

Chapter 8 – Bias Circuits, Bandgap References and Voltage Regulators

We have seen in the previous chapters that most of the circuit configurations used for gain stages need dc bias currents and/or dc bias voltages to establish the desired bias conditions for the transistors providing the signal processing. In several of the circuits, we have used a single transistor with a dc voltage at the gate to generate a dc bias current, or we have used a simple current mirror with a diode-connected transistor and a resistor connected to one of the supply rails for the input side of the current mirror.

In this chapter, we present an introduction to the design of circuits for generating bias currents, and we introduce circuits for generating dc voltages using the so-called bandgap principle. Finally, we present a few voltage regulator configurations which can be used to generate stable dc supply voltages and bias voltages from an unregulated voltage.

After having studied the chapter, you should be able to

- analyze and design a dc bias current source with a simple current mirror and a resistor.
- analyze and design a dc bias current source using a cascode current mirror.
- analyze and design a self-biased dc bias current circuit with reduced supply-voltage dependency.
- explain and analyze the basic operation of a bandgap voltage reference circuit.
- explain and analyze the basic operation of a linear voltage regulator.

Several of the circuits presented in this chapter are shown only at a level sufficient to describe the principle of operation. For further details, the reader may consult more advanced textbooks (Allen & Holberg 2012; Chan Carusone, Johns & Martin 2012; Gray, Hurst, Lewis & Meyer 2009; Sansen 2006) or scientific journal papers.

8.1 Current mirrors

The simplest version of a bias current circuit is a current mirror where the input side is a diode-connected transistor with a resistor to set the current in the mirror. This bias current circuit is shown in an NMOS version in Fig. 8.1. By using more than one output transistor, several bias current sources are obtained using just a single diode-connected input transistor. With different scaling factors for the channel widths of the output transistors, different bias current values can be obtained.

The precision of the currents depends on the matching of the transistors. Ideally, all transistor parameters have identical values except the channel width which is used to scale the output currents. In practice, the transistor parameters show some variation across the chip and a good matching is obtained only with

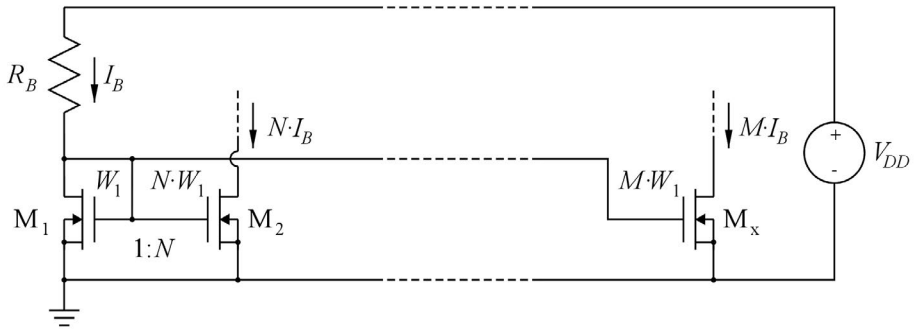


Figure 8.1: Bias current circuit with NMOS transistors. Notice the notation $1:N$ for the scaling ratio of the current mirror M_1 - M_2.

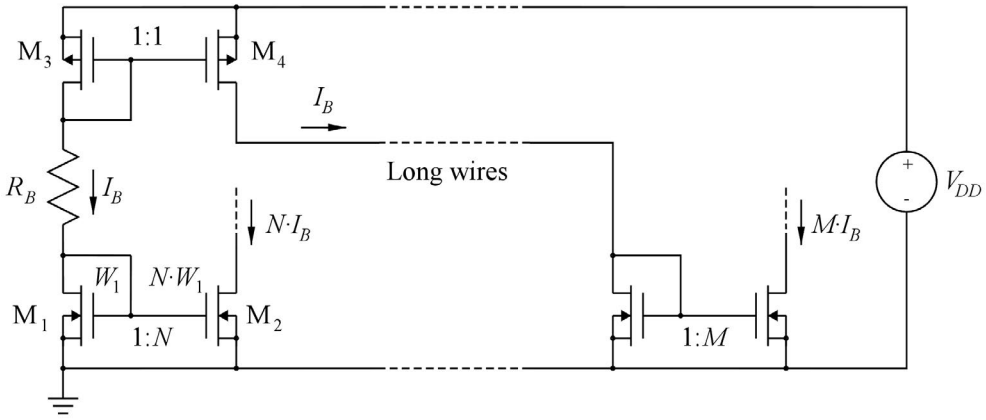
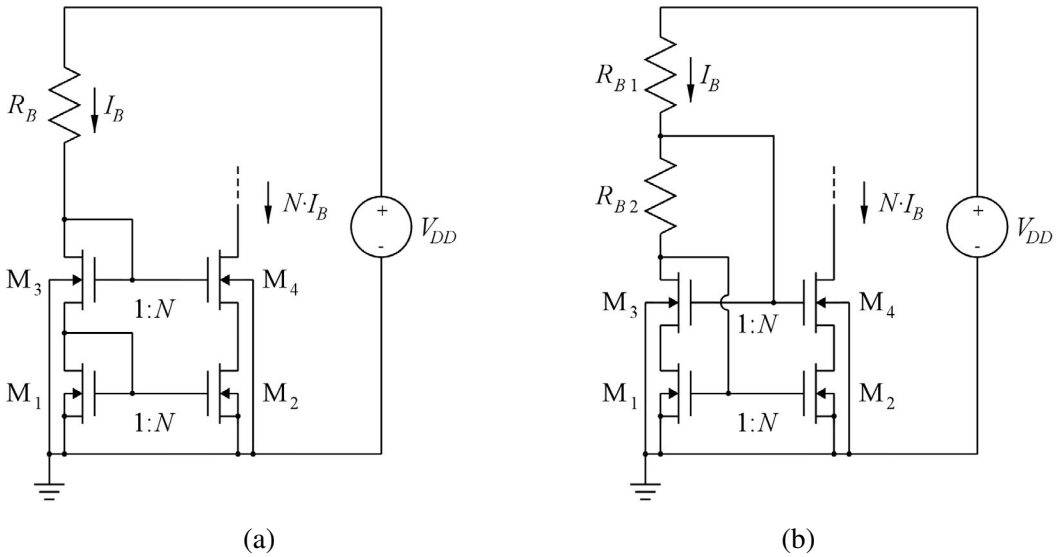


Figure 8.2: Bias current circuit with additional current mirrors to distribute the bias current over a long distance.

transistors located close to each other on the chip. For transistors located far apart, the matching is worse. If bias current sources located far apart are needed on a chip, it may be a good idea to use an extra current mirror and distribute a current to the remote location where a local current mirror can be designed using transistors close to each other as shown in Fig. 8.2. This approach also eliminates voltage drops caused by the resistance in the long wires to the remote location, see Problems 8.1 and 8.2.

Another issue with the bias current source is the output characteristics of the MOS transistor. The finite value of the small-signal output resistance limits the load regulation, causing the bias current to vary with a varying voltage at the output of the current source, see for instance Fig. 3.19. This strongly affects the performance of circuits such as a cascode gain stage, see Chapter 4.3. Also, the output voltage of the current source must be larger than the saturation voltage of the current-source transistor. Otherwise the transistor enters the triode region where it is no longer a good approximation to a constant current source. An example of the design considerations related to this issue was given in Chapter 7 in the design of transistors M_5 and M_6 in the two-stage opamp shown in Fig. 7.2.

A way to increase the output resistance of a bias current source is to use transistors with a long channel because this reduces the channel-length modulation. However, for the same current, then also the channel width must be increased, and although this improves the matching, it also implies an increase in the area



(a) (b)

Figure 8.3: Bias current circuits with cascoding in order to obtain a higher output resistance.

of the drain diffusion, hence of the parasitic capacitance at the output of the current mirror. Another way is to use a cascode transistor as shown in Fig. 4.26. This increases the output resistance by a factor which is approximately $(g_m + g_{mb})r_{ds}$ of the cascode transistor. However, it also implies that a higher voltage is needed across the bias current source in order to keep both the current-mirror transistor and the cascode transistor in the active region.

A simple way to bias the cascode transistor is shown in Fig. 8.3(a) but it has the disadvantage that the output voltage must be larger than

$$V_{D4\,\min} = V_{GS2} + V_{GS4} - V_{t4} = V_{t2} + V_{DS\,\mathrm{sat}2} + V_{DS\,\mathrm{sat}4} \tag{8.1}$$

which is fairly large due to the term V_{t2}, the threshold voltage of M_2. An alternative biasing is shown in Fig. 8.3(b) (Allen & Holberg 2012).

Here we find

$$V_{G1} = V_{G2} = V_{D3} = V_{t2} + V_{DS\,\mathrm{sat}2} \tag{8.2}$$

$$V_{G3} = V_{G4} = V_{G2} + R_{B2} I_B = V_{t2} + V_{DS\,\mathrm{sat}2} + R_{B2} I_B \tag{8.3}$$

$$V_{D2} = V_{G3} - V_{GS4} = V_{t2} + V_{DS\,\mathrm{sat}2} + R_{B2} I_B - V_{t4} - V_{DS\,\mathrm{sat}4} \tag{8.4}$$

In order to have M_1 and M_2 in the active region, we need

$$\begin{aligned} V_{D2} \geq V_{G2} - V_{t2} &\Rightarrow V_{t2} + V_{DS\,\mathrm{sat}2} + R_{B2} I_B - V_{t4} - V_{DS\,\mathrm{sat}4} \geq V_{t2} + V_{DS\,\mathrm{sat}2} - V_{t2} \\ &\Rightarrow R_{B2} I_B \geq V_{DS\,\mathrm{sat}4} + V_{t4} - V_{t2} \end{aligned} \tag{8.5}$$

In order to have M_3 in the active region, we need

$$V_{D3} \geq V_{G3} - V_{t3} \Rightarrow V_{t2} + V_{DS\,\mathrm{sat}2} \geq V_{t2} + V_{DS\,\mathrm{sat}2} + R_{B2} I_B - V_{t3} \Rightarrow V_{t3} = V_{t4} \geq R_{B2} I_B \tag{8.6}$$

Thus, with $V_{DS\,\mathrm{sat}4} + V_{t4} - V_{t2} \leq R_{B2} I_B \leq V_{t4}$, both M_1, M_2 and M_3 are in the active region. With

$$R_{B2} I_B = V_{t4} - V_{t2} + V_{DS\,\mathrm{sat}4} \tag{8.7}$$

we find $V_{DS2} = V_{DSsat2}$, giving a minimum output voltage of

$$V_{D4\,min} = V_{DSsat2} + V_{DSsat4} \tag{8.8}$$

Notice that the bulk effect causes the threshold voltages $V_{t3} = V_{t4}$ to be somewhat larger than $V_{t1} = V_{t2}$.

An alternative biasing method for a low-voltage cascode current mirror is shown in Problem 8.5. This method has the advantage of using no resistors, only transistors.

Another issue with the current-mirror bias current source is the variation of the current with variations in the supply voltage. For the circuit in Fig. 8.1, the bias current is

$$I_B = \frac{V_{DD} - V_{GS1}}{R_B} \tag{8.9}$$

so variations in V_{DD} directly affect I_B. With V_{GS1} being a considerable fraction of V_{DD}, the relative variation in I_B is larger than the relative variation in V_{DD}. This problem is even more pronounced for the circuit in Fig. 8.2 where $I_B = (V_{DD} - V_{GS1} - |V_{GS3}|)/R_B$.

A simulation example. We may illustrate some of the properties explained above by LTspice simulations. As an example, we show a circuit including both the simple NMOS current mirror from Fig. 8.1 and the circuit from Fig. 8.2 with both an NMOS current mirror and a PMOS current mirror. The LTspice schematic is shown in Fig. 8.4.

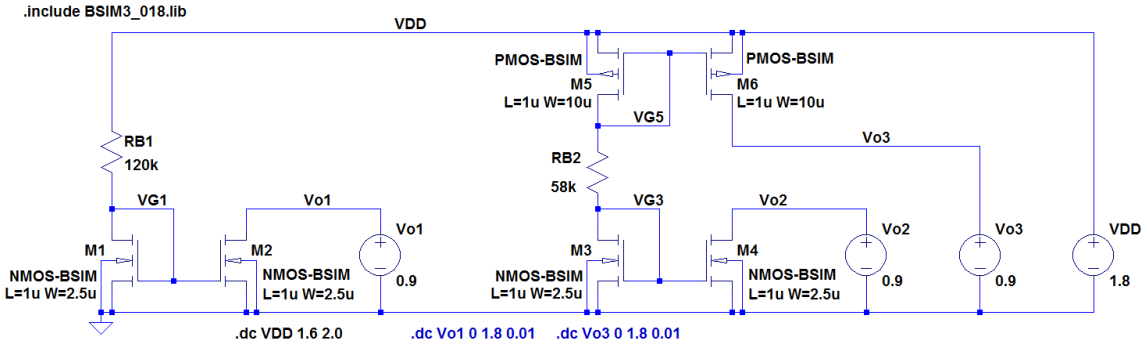


Figure 8.4: LTspice schematic for simple bias current circuits.

Using the Shichman-Hodges model with the transistor parameters from Table 3.1, the transistors have been sized to have an effective gate voltage of about 200 mV and a drain current of about 10 µA. For this, we find $R_{B1} = 120$ kΩ and $R_{B2} = 58$ kΩ.

For the simulations, the transistors are modeled by the BSIM3 models from Fig. 3.44.

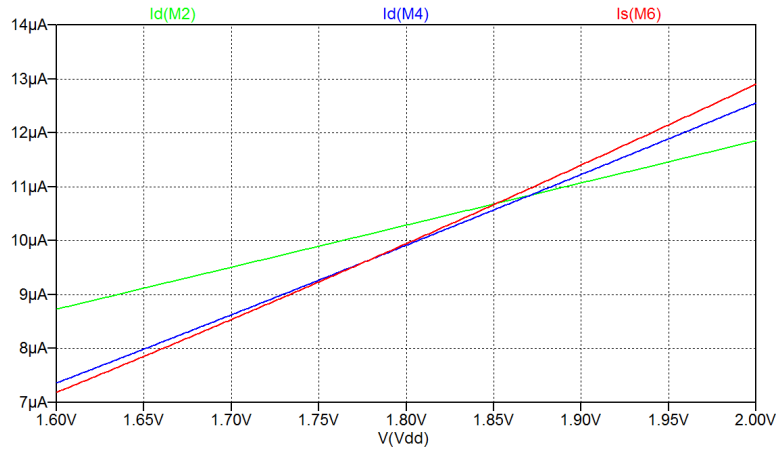
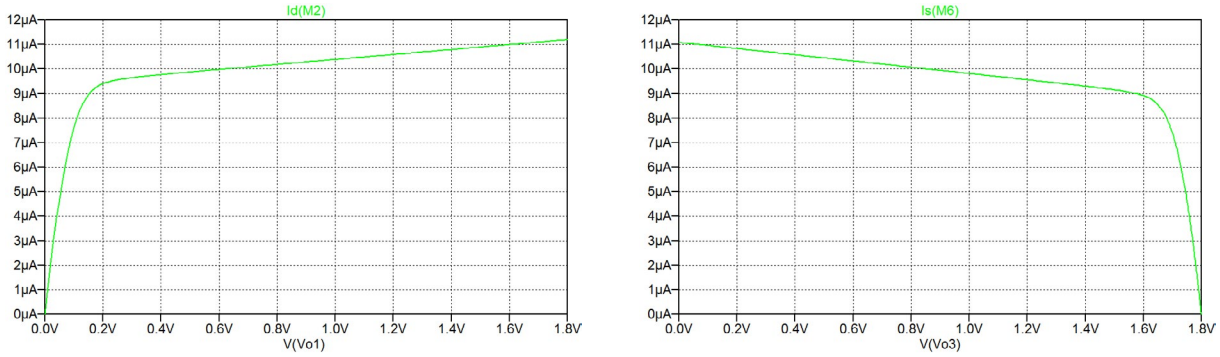


Figure 8.5: Plot of bias currents from the circuit from Fig. 8.4 versus variations of the supply voltage.

First, we run a '.dc' simulation with a sweep of the supply voltage V_{DD}. With V_{DD} swept from 1.6 V to 2.0 V, i.e., slightly more than $\pm 10\%$, we find the variation of the output currents I_{D2}, I_{D4} and I_{D6} shown in Fig. 8.5.

We notice that the value of the currents is fairly close to 10 µA for $V_{DD} = 1.8$ V, even though the resistors were calculated simply using the Shichman-Hodges model. We also see that the relative variation in the bias currents exceeds the relative variation in V_{DD}, and as expected, the circuit with both an NMOS transistor and a PMOS transistor in the input branch shows the largest variation.

Another property which may be shown from the circuit in Fig. 8.4 is the output resistance and the output voltage range of the current sources. For examining the output properties of the NMOS transistors M_2 and M_4, we sweep the output voltage from 0 to V_{DD}. The result for M_2 is shown in Fig. 8.6(a).



(a) (b)

Figure 8.6: Plot of bias currents from the circuit from Fig. 8.4 versus variations of output voltage. NMOS transistor (a) and PMOS transistor (b).

For the PMOS transistor M_6, a similar sweep is shown in Fig. 8.6(b). We see that both the NMOS current source and the PMOS current source require a voltage of about 0.2 V across the current source in order to function properly. This is also the value of the effective gate voltage used for the design. We can also find a load regulation of 1.0 µA/V for the NMOS transistor and 1.3 µA/V for the PMOS transistor, corresponding to output resistances of 1.0 MΩ and 0.77 MΩ, respectively. As expected from the Shichman-Hodges parameters, the output resistance is smaller for the PMOS transistor than for the NMOS transistor because the channel-length modulation parameter λ is larger for PMOS transistors.

8.2 Bias current circuits with reduced supply voltage dependency

One of the problems with the bias circuits presented in Figs. 8.1 and 8.2 is the dependency on the supply voltage as shown in Fig. 8.5. Several approaches may be taken to generate a current with less dependency on the supply voltage. One method is to let the bias current depend on a gate-source voltage rather than on the supply voltage. Figure 8.7(a) shows a circuit where the current I_B is given by (neglecting the channel-length modulation)

$$I_B = \frac{V_{GS1}}{R_B} = \frac{V_{t1} + V_{\text{eff1}}}{R_B} = \frac{V_{t1} + \sqrt{\dfrac{2I_{D1}}{\mu_n C_{ox}(W_1/L_1)}}}{R_B} \tag{8.10}$$

If I_{D1} is independent of V_{DD}, Eq. (8.10) shows that also I_B is independent of V_{DD}. An obvious way of providing a supply-voltage independent I_{D1} is by using a mirror of I_B for I_{D1}, leading to the bias current circuit shown in Fig. 8.7(b). Here, additional transistors M_5 and M_6 are inserted to show how I_B is mirrored and scaled to provide bias currents for other subcircuits. Inserting $I_{D1} = I_B$ in Eq. (8.10) and solving for R_B, we find

$$R_B = \frac{V_{t1}}{I_B} + \sqrt{\frac{2}{\mu_n C_{ox}(W_1/L_1) I_B}} \tag{8.11}$$

We notice that the supply voltage does not enter into the equation. However, in the derivation of Eqs. (8.10) and (8.11), we neglected the channel-length modulation. With this taken into account, the currents depend on the drain-source voltages, and they depend on the supply voltage.

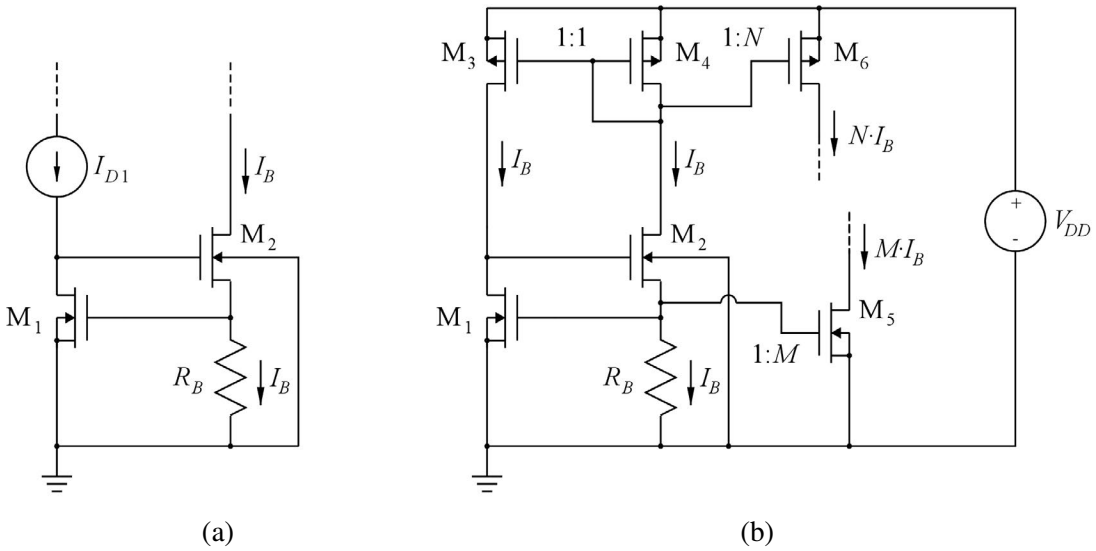


Figure 8.7: Bias current circuits where the bias current shows a small dependency on the supply voltage.

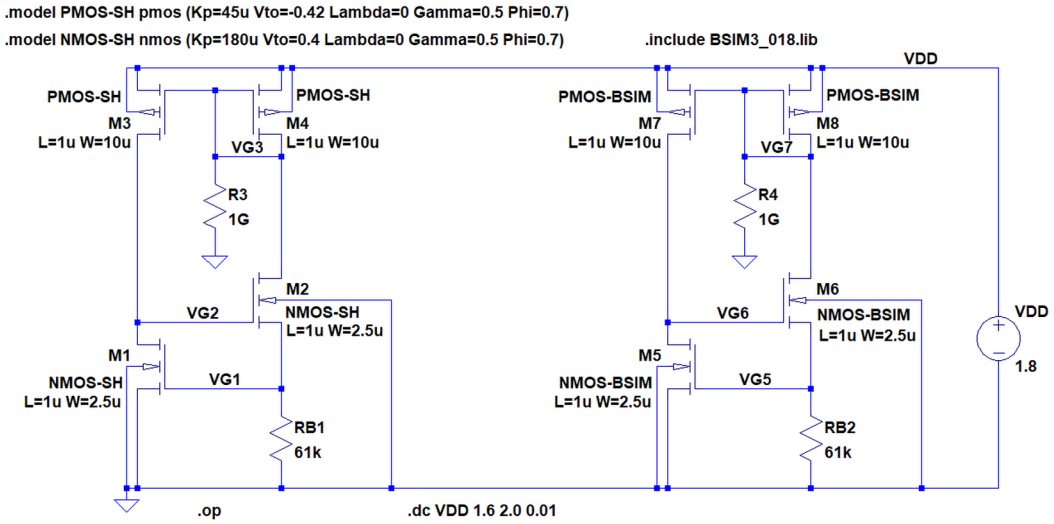


Figure 8.8: LTspice schematic for the bias circuit from Fig. 8.7(b). Left circuit: Shichman-Hodges transistor models without channel-length modulation. Right circuit: BSIM3 transistor models.

We may illustrate this problem by a simulation of the circuit. Figure 8.8 shows an LTspice schematic of the circuit with transistors using the Shichman-Hodges model without channel-length modulation in the left part of the circuit and transistors using the BSIM3 model from Fig. 3.44 in the right part of the circuit. The transistors have been sized to have an effective gate voltage of about 200 mV for a drain current of about 10 µA. The resistors R_{B1} and R_{B2} are found from Eq. (8.11) to give a bias current of about 10 µA. The two large resistors R_3 and R_4 are inserted to ensure a correct start-up of the circuits. Both of the circuits have an undesired operating point with zero current in all of the transistors as a stable point, so some sort of start-up circuit is necessary to bring the circuits into the desired operating point with $I_B \simeq 10$ µA. The large resistors shown here are suitable for simulations but for a practical circuit implementation, more sophisticated circuits are used (Gray, Hurst, Lewis & Meyer 2009; Sansen 2006).

Figure 8.9 shows the plot from a '.dc' simulation with a sweep of the supply voltage V_{DD}. With V_{DD} swept from 1.6 V to 2.0 V, i.e., slightly more than $\pm 10\%$, we find the output currents I_{D1} and I_{D5} shown in Fig. 8.9. Evidently, without the channel-length modulation, the circuit generates a bias current I_{D1} which is independent of the supply voltage. With the channel-length modulation included, the bias current I_{D5} shows some dependency on the supply voltage but we see that the relative variation of the bias current is smaller than the relative variation of the supply voltage, and this is an improvement compared to the results shown in Fig. 8.5.

The dependency on the supply voltage can be further reduced by replacing the current mirror M_3 - M_4 by a cascode current mirror as shown in Fig. 8.3. In Fig. 8.10(a), we show an LTspice schematic including the cascode transistors and a resistor to generate the gate bias voltage for the cascode transistors. The voltage drop across this resistor should be at least equal to the saturation voltage $|V_{DSsat4}|$. We have inserted 25 kΩ, giving a voltage drop of 250 mV with $I_B = 10$ µA.

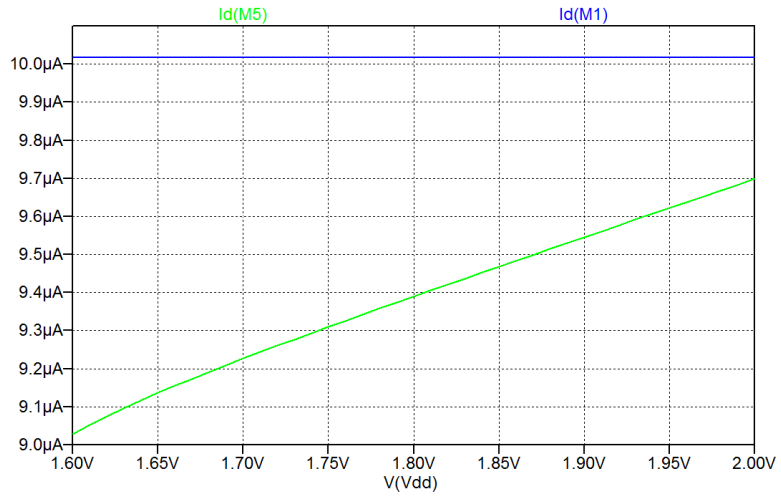
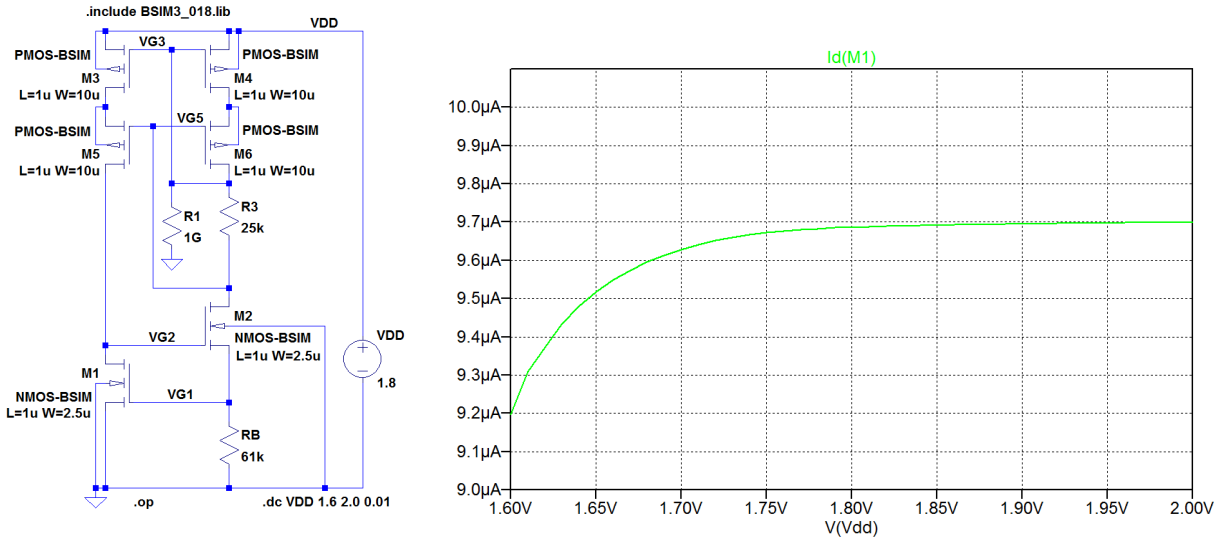


Figure 8.9: Plot of bias currents from the circuit from Fig. 8.8 versus supply voltage.

The plot from the '.dc' simulation with V_{DD} swept from 1.6 V to 2.0 V is shown in Fig. 8.10(b). Comparing to Fig. 8.9, we find that the bias current shows much less variation, especially for V_{DD} in the range 1.80 V to 2.0 V. For smaller values of V_{DD}, the current drops. This drop occurs because M_3 (and M_2) operate at the edge of saturation. For supply voltages below 1.8 V, the drain-source voltage of M_3 drops below 200 mV and it is no longer in strong saturation. For the circuit to operate with all transistors in the active region, the supply voltage must be larger than $V_{GS1} + V_{GS2} + |V_{DS5}| + |V_{DS3}|$ and with gate-source voltages of about 0.6 V (somewhat more for M_2 because of the bulk effect) and drain-source saturation voltages of about 0.2 V, operation at a supply voltage of 1.6 V is marginal.



(a) (b)

Figure 8.10: Bias current circuit with cascode transistors for reduced supply-voltage sensitivity (a) and plot of bias current versus supply voltage (b).

The self-biased Widlar current source. An alternative circuit capable of operating at a lower supply voltage is a circuit where the bias current is derived from the difference between two gate-source voltages rather than from a single gate-source voltage. Figure 8.11(a) shows a MOS version of a Widlar current source (Widlar 1965), originally developed for bipolar circuits by Robert Widlar, an American electronics engineer (1937-1991). The transistors are scaled by a factor $K = (W_2/L_2)/(W_1/L_1)$ so that $V_{GS1} > V_{GS2}$, causing a voltage drop across R_B. The method of inserting a resistor in series with the source is called source degeneration (Sedra & Smith 2016). Assuming $I_{D1} = I_{D2} = I_B$ and neglecting the channel-length modulation (i.e., $\lambda = 0$) and the bulk effect (i.e., $V_{t1} = V_{t2}$), we find

$$\begin{aligned} R_B = \frac{V_{GS1} - V_{GS2}}{I_B} &= \left(\frac{1}{I_B}\right)\left(V_{t1} + \sqrt{\frac{2I_B}{\mu_n C_{ox}(W_1/L_1)}} - V_{t2} - \sqrt{\frac{2I_B}{\mu_n C_{ox}(W_2/L_2)}}\right) \\ &= \left(1 - \frac{1}{\sqrt{K}}\right)\sqrt{\frac{2}{\mu_n C_{ox}(W_1/L_1)I_B}} \end{aligned} \tag{8.12}$$

Just as we did for the circuit in Fig. 8.7, we can derive I_{D1} as a mirror of I_B, leading to the bias current circuit shown in Fig. 8.11(b). Again, additional transistors M_5 and M_6 are inserted to show how I_B is mirrored and scaled to provide bias currents for other subcircuits.

With a scale factor $K = 4$ for M_1 and M_2, we find from Eq. (8.12), using Eq. (3.71)

$$R_B = \frac{1}{g_{m1}} \tag{8.13}$$

Thus, the transconductances of transistors biased by this circuit are scaled with respect to R_B. For the

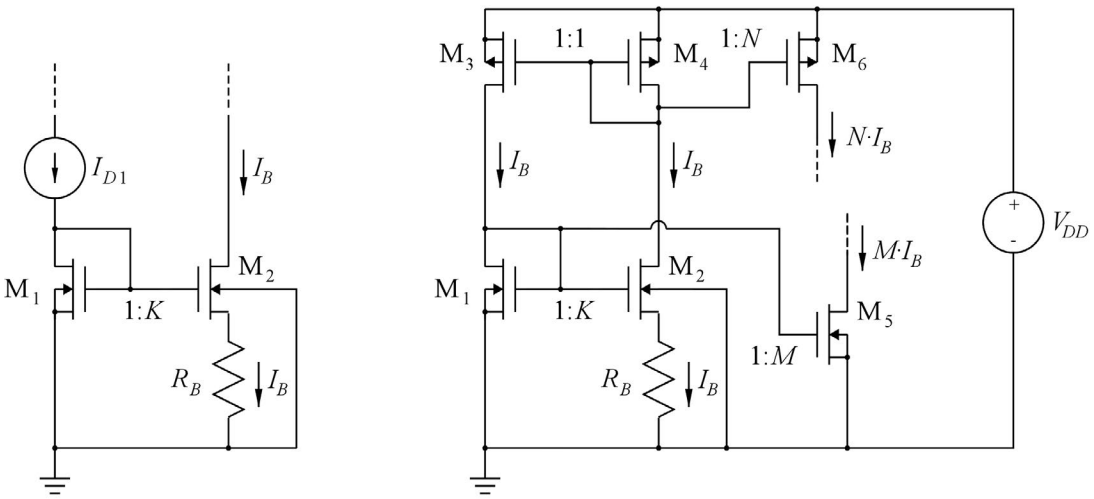


Figure 8.11: Widlar current source (a) and self-biased Widlar current source (b).

NMOS transistor M_5 in Fig. 8.11(b), we find using Eq. (3.71)

$$g_{m5} = g_{m1}\sqrt{\frac{(W_5/L_5)I_{D5}}{(W_1/L_1)I_{D1}}} = M\,g_{m1} \tag{8.14}$$

where M is the scale factor between M_5 and M_1. Likewise

$$g_{m6} = N\,g_{m3} \tag{8.15}$$

where N is the scale factor between M_6 and M_3. The transconductance g_{m3} is related to g_{m1} by

$$g_{m3} = g_{m1}\sqrt{\frac{\mu_p\,(W_3/L_3)}{\mu_n\,(W_1/L_1)}} \tag{8.16}$$

We may show some of the properties of the self-biased Widlar current source by simulation. For this, we choose a complementary version of the circuit so that the Widlar current source is implemented with PMOS transistors so that we can avoid the bulk effect for M_2 (assuming a standard n-well process).

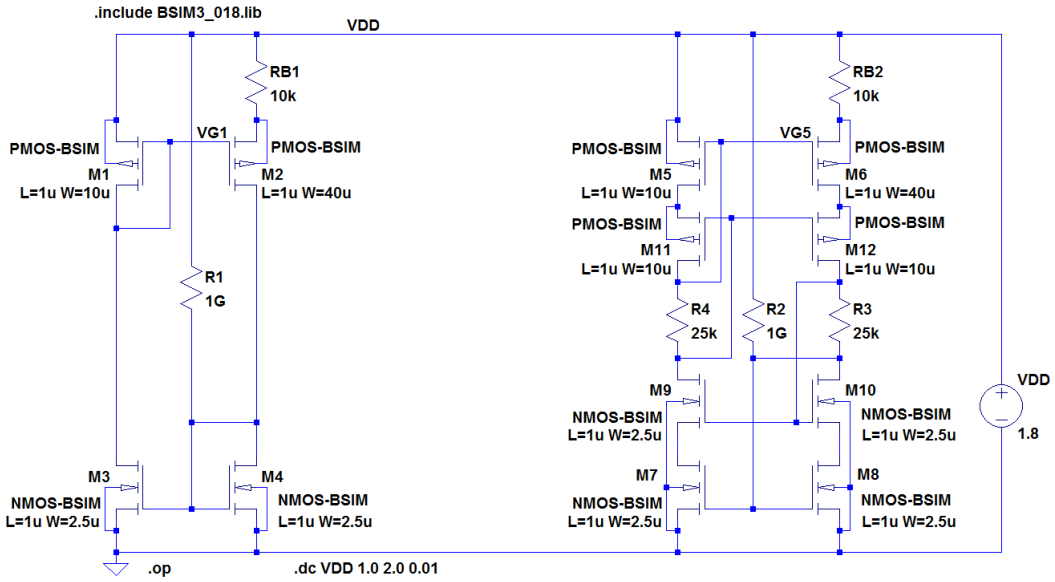


Figure 8.12: LTspice schematic for the self-biased Widlar current source. Left circuit: Simple model without cascode transistors. Right circuit: Circuit including cascode transistors.

Figure 8.12 shows the LTspice schematic of the circuit in a version with a simple current mirror to bias the Widlar current source and in a version with cascode transistors for M_2 and M_3. The transistor geometries shown are the same as used for the circuit in Fig. 8.8, except M_2 which is scaled by $K = 4$, and for a current of about 10 µA and an absolute value of the effective gate-source voltage of about 200 mV for M_1 and 100 mV for M_2, we find a voltage of about 100 mV across R_B. Thus, $R_B = 10\,\text{k}\Omega$ and we would expect $g_{m1} \simeq 0.1$ mA/V.

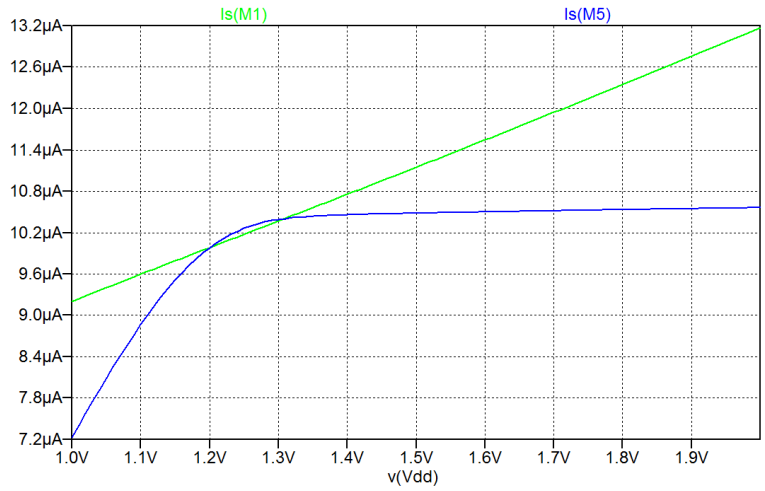


Figure 8.13: Plot of bias currents from the circuit from Fig. 8.12 versus supply voltage.

Figure 8.13 shows a plot of I_{D1} and I_{D5} for a variation of V_{DD} from 1 V to 2 V. We see that the current is close to 10 µA. For the simple version of the circuit (left circuit, M_1), the relative variation of I_{D1} is about 6% for a 10% relative variation of V_{DD}. For the improved version (right circuit, M_5), the relative variation of I_{D5} is only about 0.3% for a 10% relative variation of V_{DD}. Running a '.op' simulation, we may also from the error log file verify the transconductance $g_{m1} \simeq 0.11$ mA/V and $g_{m5} \simeq 0.10$ mA/V as expected from Eq. (8.13). Finally, we see that the circuit operates reasonably well down to a supply voltage of about 1.3 V, i.e., significantly lower than the circuit in Fig. 8.10.

The self-biased Widlar current source also requires a start-up circuit, and it can be further improved by replacing the cascode transistors with an amplifier-based control of the bias point of the transistors. However, this is beyond the scope of this book and the reader is referred to Chan Carusone, Johns & Martin (2012).

8.3 Bandgap voltage references

Voltage references are widely used in analog circuits, both for the control of regulated supply voltages and for establishing fixed dc bias voltages. Often circuits with a low dependency on temperature are desirable, and in this section we present some reference circuits based on the so-called bandgap principle for achieving a low temperature dependency. The bandgap principle was first presented in 1964 by David Hilbiber from Fairchild Semiconductor (Hilbiber 1964).

The starting point is a pn diode. As explained in Chapter 3, a silicon diode has a relation between diode current I_D and diode voltage V_D given by

$$I_D = I_S \exp\left(\frac{V_D}{nV_T} - 1\right) \simeq I_S \exp\left(\frac{V_D}{V_T}\right) \tag{8.17}$$

where the approximation is valid for an emission coefficient $n = 1$ and $V_D \gg V_T$. Recall that V_T is the thermal voltage, not to be confused with a transistor threshold voltage V_t. The thermal voltage is given by $V_T = kT/q$ where T is the absolute temperature, k is Boltzmann's constant and q is the magnitude of the electron charge. At room temperature, we find $V_T \simeq 26$ mV.

From Eq. (8.17), we find the diode voltage when the diode is in the forward direction, $V_D \gg V_T$

$$V_D = V_T \ln\left(\frac{I_D}{I_S}\right) \tag{8.18}$$

For a silicon diode, the voltage V_D typically has a value in the range 500 mV to 800 mV for currents below a few mA and it shows only a small variation with I_D as illustrated in Fig. 3.2. However, since the saturation current I_S is strongly temperature dependent, V_D has a temperature dependency with a negative temperature coefficient $\partial V_D/\partial T$ of about -2 mV/K, so it is not directly suited as a voltage reference.

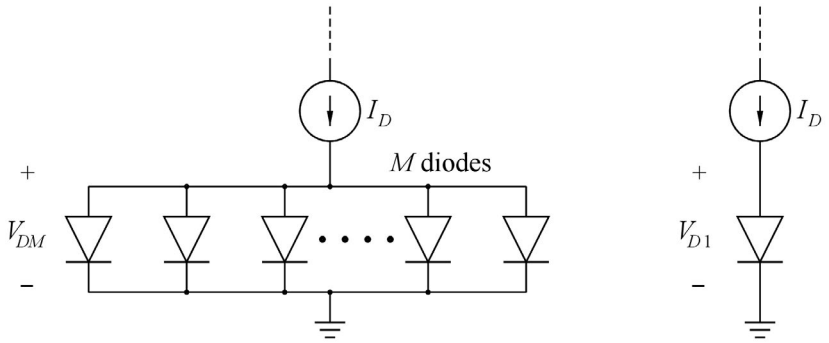


Figure 8.14: Diodes connected to obtain different diode voltages V_{D1} and V_{DM}.

Now, connecting M identical diodes in parallel and supplying them with the same current I_D as supplied to a single diode, see Fig. 8.14, we find the voltage V_{D1} across the single diode given by Eq. (8.18) and the voltage across the parallel connection given by

$$V_{DM} = V_T \ln\left(\frac{I_D/M}{I_S}\right) = V_T \ln\left(\frac{I_D}{M I_S}\right) \tag{8.19}$$

since each of the diodes carry a current I_D/M.

From Eqs. (8.18) and (8.19), we find

$$\Delta V_D = V_{D1} - V_{DM} = V_T \left(\ln\left(\frac{I_D}{I_S}\right) - \ln\left(\frac{I_D}{M I_S}\right)\right) = V_T \ln(M) = \frac{kT}{q}\ln(M) \tag{8.20}$$

Equation (8.20) shows that ΔV_D has a positive temperature coefficient of

$$\frac{\partial \Delta V_D}{\partial T} = \frac{k}{q}\ln(M) = 0.0861 \text{ mV/K} \times \ln(M) \tag{8.21}$$

As an example, let us assume $M = 10$. From Eq. (8.20), we find $\Delta V_D = 26 \text{ mV} \times \ln(10) = 60 \text{ mV}$ at room temperature (27°C), and from Eq. (8.21), we find $\partial \Delta V_D/\partial T = 0.1983$ mV/K.

Incidentally, it may be noted that ΔV_D could also be obtained from just two identical diodes where the ratio of the currents in the diodes is $I_{D1} = M I_{DM}$ so that the current density (current per unit area of the pn junction) in the diodes is scaled by a factor M.

By combining ΔV_D and V_D we may now establish a reference voltage

$$V_{\text{ref}} = G\Delta V_D + V_D \tag{8.22}$$

which can be made independent of temperature. From Eq. (8.22), we find

$$\frac{\partial V_{\text{ref}}}{\partial T} = G\frac{\partial \Delta V_D}{\partial T} + \frac{\partial V_D}{\partial T} \tag{8.23}$$

and with $\partial V_D/\partial T = -2$ mV/K, we find that $\partial V_{\text{ref}}/\partial T = 0$ is obtained for

$$G = -\frac{\partial V_D/\partial T}{\partial \Delta V_D/\partial T} = \frac{2\text{ mV/K}}{0.0861\text{ mV/K} \times \ln(M)} = \frac{23.2}{\ln(M)} \tag{8.24}$$

For $M = 10$, we find $G = 10.1$. Using Eq. (8.22) with G given by Eq. (8.24) and ΔV_D given by Eq. (8.20), and assuming $V_D = 650$ mV and $V_T = 26$ mV at room temperature, we then find

$$V_{\text{ref}} = G\Delta V_D + V_D = \left(\frac{23.2}{\ln(M)}\right) \times 26\text{ mV} \times \ln(M) + 650\text{ mV} = 1.25\text{ V} \tag{8.25}$$

We notice that V_{ref} is independent of M and G which are related by Eq. (8.24). It can be shown that for achieving $\partial V_{\text{ref}}/\partial T = 0$, the reference voltage assumes a value equal to the bandgap voltage V_{G0} of silicon (Hilbiber 1964), hence the name bandgap voltage reference.

The bandgap voltage reference requires access to diodes which can be biased in the forward direction. In a standard n-well process, the pn junction from a drain/source diffusion into a well may be used for this with the n-well shorted to substrate, see Fig. 8.15 showing both a cross-sectional view of the structure and the corresponding bipolar transistor symbol. Essentially, the structure corresponds to a pnp bipolar transistor with the source/drain diffusion as the emitter, the well as the base and the substrate as the collector.

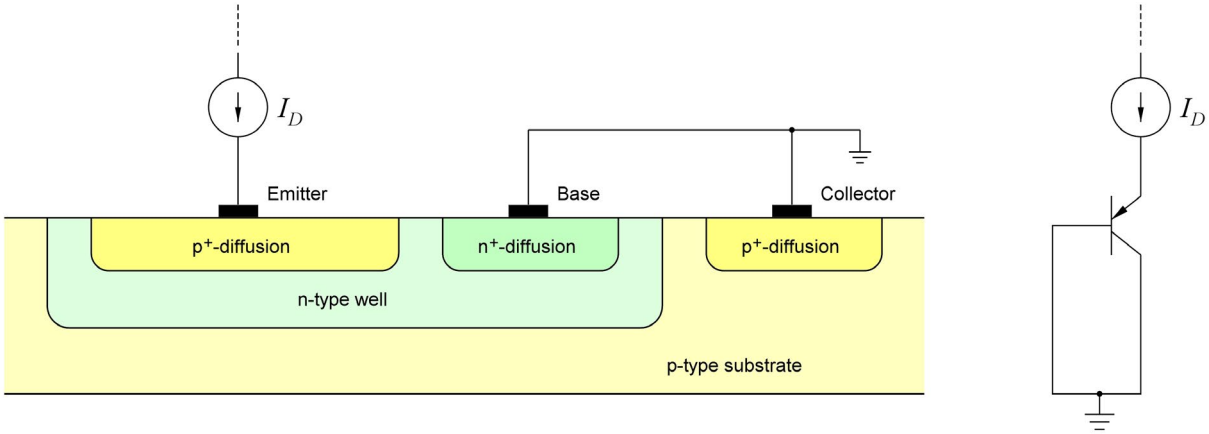


Figure 8.15: A diode-connected bipolar pnp transistor in an n-well CMOS process. Cross-sectional view (a) and schematic (b).

We may verify the bandgap principle by simulating the diode characteristics in LTspice using the BSIM3 model from Fig. 3.44.

Figure 8.16 shows the LTspice schematic corresponding to the circuits from Fig. 8.14. For the diodes, we use PMOS transistors with drain and source connected together to form the anode of the diode and gate and bulk connected together to form the cathode. The transistors have been specified with drain and source areas of 30 µm by 30 µm and the scaling of transistor M_2 is obtained by specifying the scale factor $M = 10$ for the transistor, see Fig. 3.12. For sweeping the temperature from −20°C to +80°C, we specify the directive '`.step temp -20 80 1`'. This steps the temperature with a step size of 1°C. We simply use a '.op' simulation to find the diode voltages V_{D1} and V_{DM}.

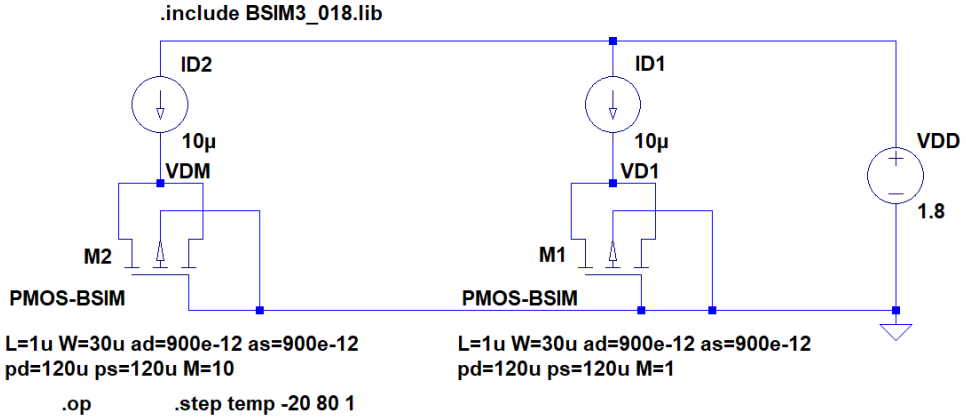
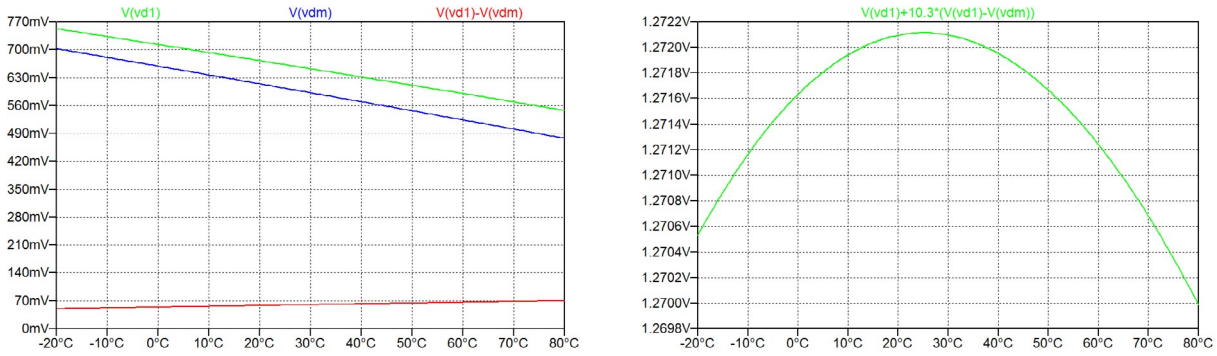


Figure 8.16: LTspice schematic corresponding to Fig. 8.14 for simulating diode characteristics.

The result of the simulation is shown in Fig. 8.17. The plot in Fig. 8.17(a) shows the diode voltages V_{D1} and V_{DM} and the difference $\Delta V_D = V_{D1} - V_{DM}$. We see that the diode voltages have a negative temperature coefficient whereas the difference has a positive temperature coefficient. From the plot of V_{D1}, we may find $V_{D1} = 659$ mV at 27°C and we find a temperature coefficient of -2.05 mV/K. Using these values



(a) (b)

Figure 8.17: Plot of diode voltages from Fig. 8.16 versus temperature (a) and calculated reference voltage versus temperature (b).

in Eqs. (8.24) and (8.25), we find $G = 10.3$ and $V_{\text{ref}} = 1.27$ V. The plot in Fig. 8.17(b) shows V_{ref} calculated as $V_{D1} + 10.3 \times (V_{D1} - V_{D2})$. We see that the reference voltages varies only about 2.1 mV over the temperature interval from -20°C to $+80$°C and the simulated value of the reference voltage matches the calculated value.

After having verified the principle in the bandgap voltage reference generation, we need a practical circuit implementation which can perform the addition of a V_D-voltage and a ΔV_D-voltage multiplied by a gain factor. In the literature, many such circuits have been described. Several textbooks (Allen & Holberg 2012; Chan Carusone, Johns & Martin 2012; Gray, Hurst, Lewis & Meyer 2009; Sansen 2006) provide examples of bandgap reference circuits and many more are described in journal and conference publications. An overview of design considerations with an emphasis on low-voltage circuits is given by Mok and Leung (2004). Here we present configurations which can operate with a supply voltage of 1.0 V to 2.0 V, and we assume the 0.18 µm technology used throughout this book.

The core of the circuit is shown in Fig. 8.18. We have the two diodes scaled by a factor M, resulting in the diode voltages V_{D1} and V_{DM}, and the difference ΔV_D appears across resistor R_1. The amplifier with a voltage gain A_v establishes a feedback loop together with the PMOS transistors M_1 and M_2, and with a high loop gain, this ensures equal voltages at the inputs to the amplifier, i.e.,

$$V_{D1} = V_{DM} + \Delta V_D = V_{DM} + R_1 I_D \tag{8.26}$$

The two transistors M_1 and M_2 are identical and have identical bias conditions, so their drain currents I_D are identical. From Eq. (8.26), we find

$$I_D = \frac{V_{D1} - V_{DM}}{R_1} = \frac{\Delta V_D}{R_1} \tag{8.27}$$

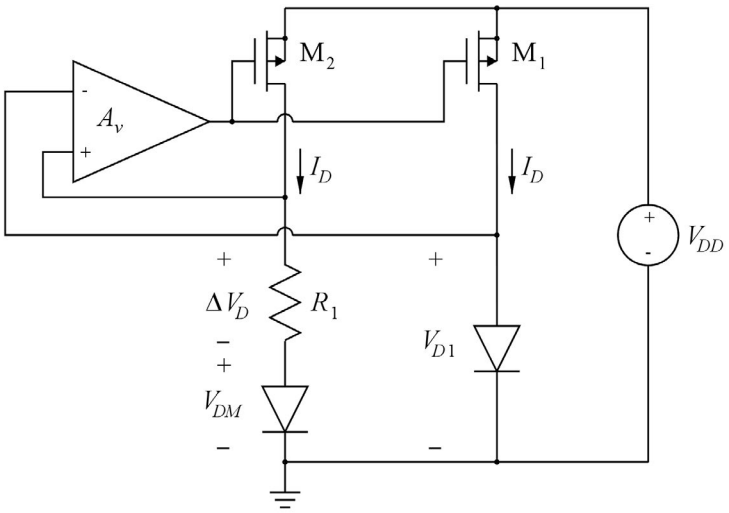


Figure 8.18: The core circuit for biasing the diodes for the bandgap reference.

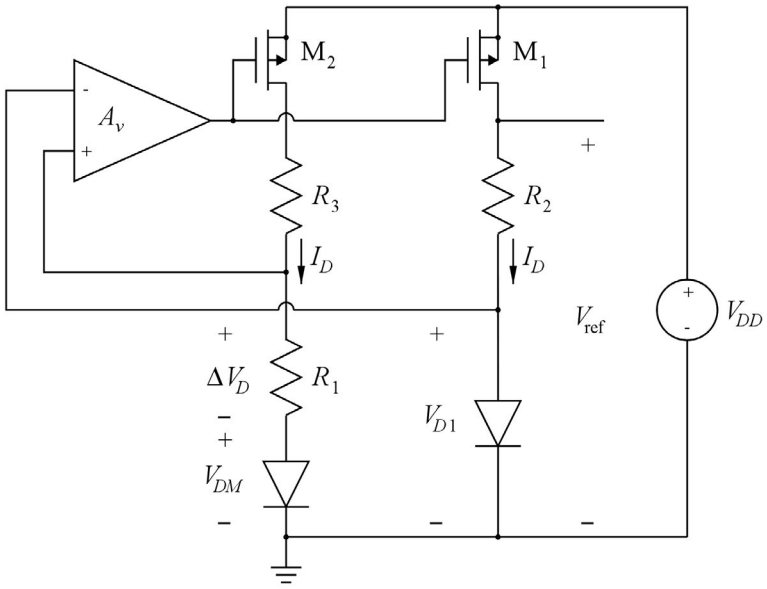


Figure 8.19: A bandgap reference circuit.

Now, inserting a resistor R_2 in series with diode D_1 as shown in Fig. 8.19 adds a voltage $R_2 I_D$ to the voltage V_{D1}, so the voltage across diode D_1 and resistor R_2 is

$$V_{\text{ref}} = V_{D1} + R_2 I_D = V_{D1} + \left(\frac{R_2}{R_1}\right)\Delta V_D \tag{8.28}$$

From Eq. (8.28), we see that V_{ref} is equal to the bandgap voltage V_{G0} if R_2/R_1 is selected to be equal to the factor G defined in Eq. (8.24). In order to ensure the same bias conditions for transistors M_1 and M_2, a resistor $R_3 = R_2$ is also inserted in series with M_2 as shown in Fig. 8.19.

The circuit shown in Fig. 8.19 can operate with supply voltages down to $V_{\text{ref}} + |V_{DS1}|$. With $|V_{DS1}| \geq 0.25$ V, this gives a minimum supply voltage of about 1.5 V.

For operation at even lower supply voltages, clearly we cannot add ΔV_D and V_D because this gives a voltage of about 1.25 V. However, from the circuit in Fig. 8.18, we have a current I_D which is proportional to ΔV_D, so if we can create a current which is proportional to V_D, we can add these currents to achieve a total current which is temperature-insensitive. A current proportional to V_D is obtained from the circuit in Fig. 8.18 simply by connecting a resistor R_4 in parallel with the diode D_1 as shown in Fig. 8.20.

The current in R_4 is V_{D1}/R_4, so the total current in M_1 is

$$I_{D1} = \frac{V_D}{R_4} + \frac{\Delta V_D}{R_1} = \left(\frac{1}{R_4}\right)\left(V_D + \left(\frac{R_4}{R_1}\right)\Delta V_D\right) \tag{8.29}$$

By selecting $R_4/R_1 = G$, we achieve a temperature-independent current I_{D1}. In order to ensure the same current in M_1 and M_2, we must also connect a resistor $R_5 = R_4$ from the drain of M_2 to ground. Finally,

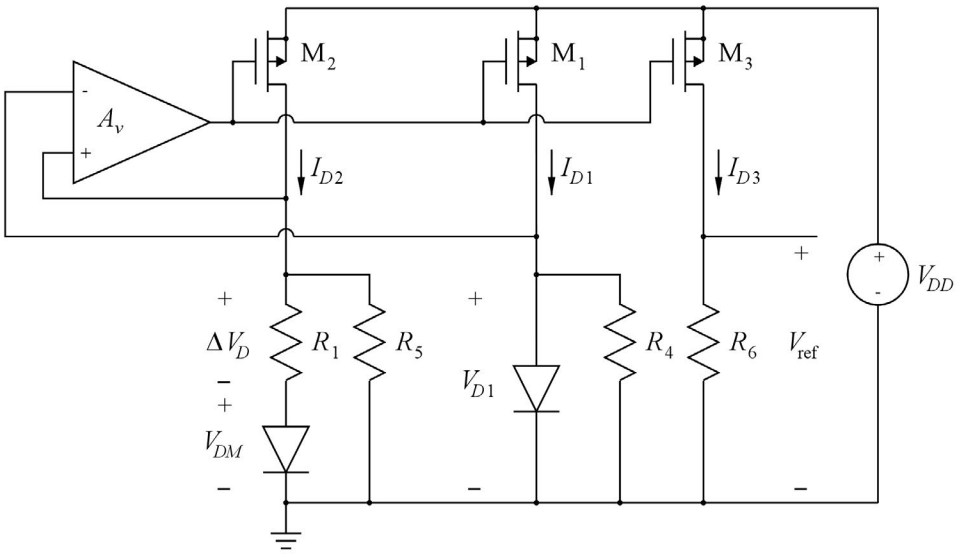


Figure 8.20: A low-voltage bandgap reference circuit.

we may mirror the current in M_1 and M_2 to a transistor M_3 with the same size as M_1 and M_2 as shown in Fig. 8.20 to generate a reference voltage as

$$V_{\text{ref}} = R_6 I_{D3} = \left(\frac{R_6}{R_4}\right)\left(V_D + \left(\frac{R_4}{R_1}\right)\Delta V_D\right) = \left(\frac{R_6}{R_4}\right) V_{G0} \tag{8.30}$$

This circuit can operate with a supply voltage of $V_{D1} + |V_{DS1}|$, i.e., down to about 1 V. However, it can be quite challenging to design an amplifier to operate at a very low supply voltage.

An advantage of the bandgap reference circuits is that the value of V_{ref} depends only on the bandgap voltage and resistor ratios, not on the resistor values. A resistor *ratio* is significantly less affected by process and temperature variations than a resistor *value*.

Simulation examples. We complete the investigation of bandgap reference circuits by LTspice simulations of the circuits from Figs. 8.19 and 8.20. For this, we use the diodes from Fig. 8.16 with a current of 10 µA at room temperature. Thus, $\Delta V_D \simeq 60$ mV, and we select $R_1 = 60\text{ mV}/10\ \mu\text{A} = 6\ \text{k}\Omega$.

For both of the circuits, we would expect the gain factor G to be about 10. However, the biasing of the diodes is not a constant current as applied in Fig. 8.16, so we cannot expect exactly the same value of G as found for this circuit. In order to find G for each of the circuits, we define G as a parameter.

Figure 8.21 shows the LTspice schematic corresponding to the circuit from Fig. 8.19. The transistors M_1 and M_2 have been designed for an effective gate voltage of about 200 mV for a drain current of 10 µA.

For determining G so that $\partial V_{\text{ref}}/\partial T \simeq 0$ at room temperature, we run a '.op' simulation where the temperature is stepped from -20°C to $+80$°C and G is stepped from 9.5 to 10.5 in steps of 0.2. The result of this simulation is shown in Fig. 8.22(a). We see that $G = 9.9$ results in $\partial V_{\text{ref}}/\partial T \simeq 0$ for a

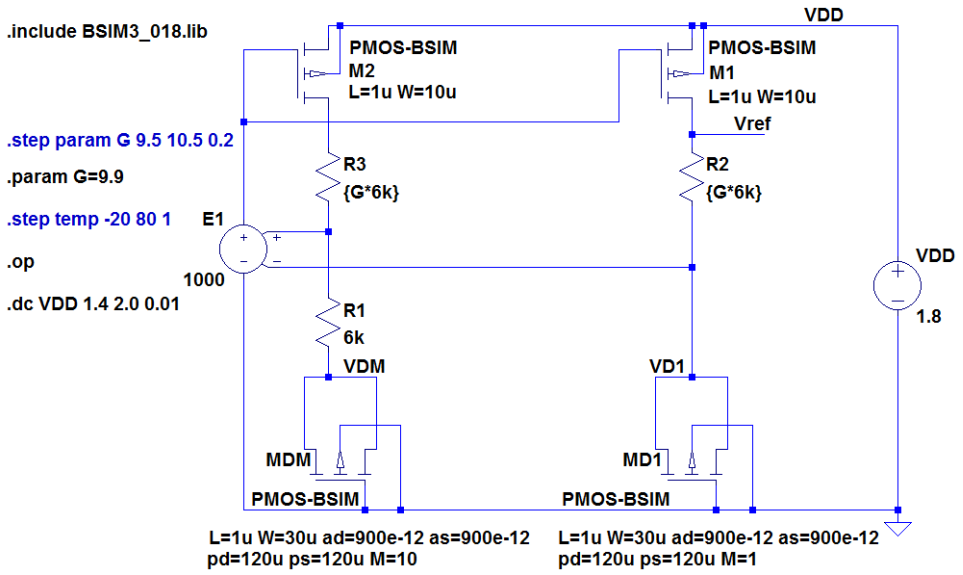


Figure 8.21: LTspice schematic corresponding to the bandgap reference from Fig. 8.19.

temperature of 27°C, the default temperature for LTspice simulations. We also find that the reference voltage is 1.260 V at 27°C, and the variation over a temperature range from -20°C to $+80$°C is less than 1.3 mV, i.e., about 0.1%.

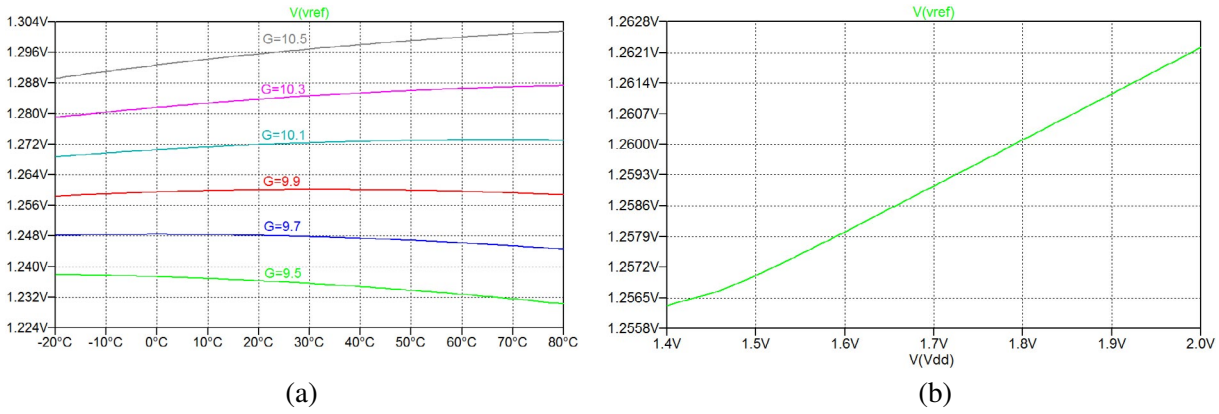


Figure 8.22: Reference voltage versus temperature (a) and reference voltage versus supply voltage (b) for the bandgap reference from Fig. 8.21.

We may also sweep the supply voltage. Figure 8.22(b) shows the reference voltage at 27°C for V_{DD} in the range from 1.4 V to 2.0 V. We see that the circuit operates down to about 1.5 V with a variation in V_{ref} of about 4 mV for 1.6 V $\leq V_{DD} \leq$ 2.0 V, i.e., a line regulation defined as $\partial V_{\text{ref}}/\partial V_{DD}$ of about 10 mV/V.

By changing the gain of the voltage-controlled voltage source 'E1', i.e., the gain A_v of the amplifier in Fig. 8.18, you will find that the line regulation is inversely proportional to A_v. This is not surprising since A_v is a factor in the gain of the loop controlling V_{ref}, and the feedback reduces the gain $\partial V_{\text{ref}}/\partial V_{DD}$ by a factor approximately equal to the loop gain, see details in Problem 8.9.

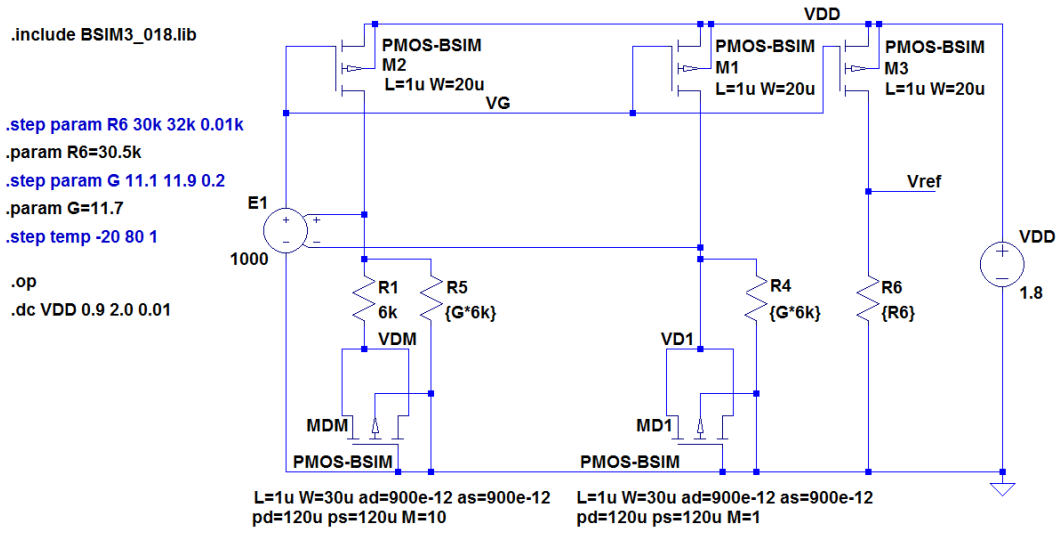


Figure 8.23: LTspice schematic corresponding to the low-voltage bandgap reference from Fig. 8.20.

Figure 8.23 shows the LTspice schematic for the circuit from Fig. 8.20. In this circuit, the bias current in M_1 - M_3 is about 20 µA because of the additional current in R_4 and R_5, so M_1 - M_3 have a channel width of 20 µm instead of the value of 10 µm used in Fig. 8.21.

Again the gain parameter G may be found by a '.step' directive to give $\partial V_{\text{ref}}/\partial T \simeq 0$ at room temperature. This results in $G = 11.7$. For $G = 11.7$, the resistor R_6 has then been found to be 30.5 kΩ for a reference voltage of 600 mV at room temperature.

Figure 8.24(a) shows V_{ref} versus temperature, and we find a variation of less than 0.6 mV, i.e., less than 0.1% over a temperature range from $-20°C$ to $+80°C$.

Figure 8.24(b) shows V_{ref} versus the supply voltage for V_{DD} in the range from 0.9 V to 2.0 V. We see that the circuit operates down to about 1 V with a variation in V_{ref} of about 2 mV for $1.0\text{ V} \leq V_{DD} \leq 1.5\text{ V}$, and the line regulation $\partial V_{\text{ref}}/\partial V_{DD}$ is found to be about 5 mV/V.

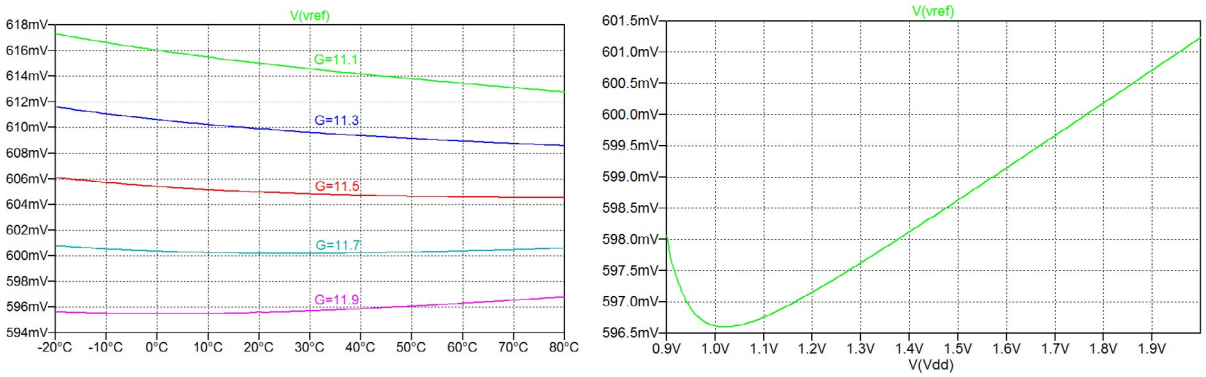


Figure 8.24: Reference voltage versus temperature (a) and reference voltage versus supply voltage (b) for the bandgap reference from Fig. 8.23.

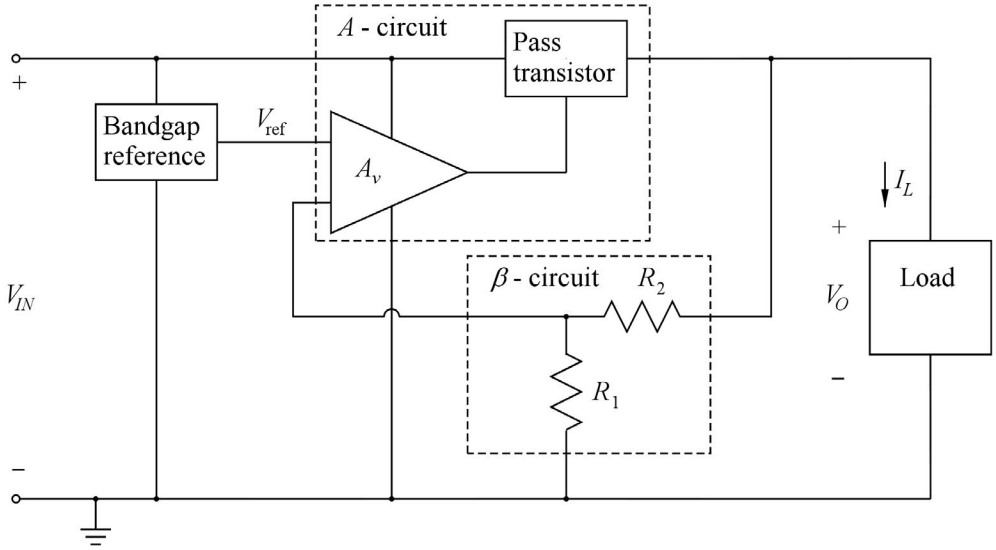


Figure 8.25: Block diagram of a linear regulator for positive voltages.

8.4 Voltage regulators

While bandgap reference circuits are very well suited for providing a stable reference voltage, they cannot be used directly as supply voltage sources. A supply voltage source typically has to deliver a voltage different from the bandgap reference voltage, and it has to be able to deliver a supply current. Conversely, the circuits shown in Figs. 8.19 and 8.20 are assumed to deliver zero current from the output terminal.

A popular approach to the design of high-efficiency power supplies capable of delivering a considerable dc power is the use of switching regulators (Erickson & Maksimovic 2001). However, this is a topic which is outside the scope of this book. Another approach is the design of linear regulators where an unregulated, unipolar voltage V_{IN} is used as the input to a linear regulator circuit comprising a voltage reference and a feedback loop to control the output voltage. Figure 8.25 shows a version of such a circuit for a positive input voltage V_{IN} and a positive output voltage. The circuit delivers a regulated output voltage $V_O < V_{IN}$ to a load assumed to draw a current I_L.

The current I_L is delivered from the input via a pass transistor controlled by the feedback loop including an amplifier with the gain A_v and a β-circuit consisting of the resistive voltage divider R_1 and R_2. With a high loop gain in the feedback loop, the output voltage is

$$V_O = V_{\text{ref}}/\beta = V_{\text{ref}}(1 + R_2/R_1) \tag{8.31}$$

Efficiency and dropout voltage. The efficiency η of the voltage regulator is defined as the ratio between the power $P_L = V_O I_L$ consumed by the load and the total power delivered to the load and the regulator. The power consumed by the regulator is often dominated by the power dissipated in the pass transistor, i.e., $(V_{IN} - V_O) I_L$, and in this case, the efficiency is $\eta \simeq V_O/V_{IN}$.

The minimum value of the voltage difference $V_{IN} - V_O$ required for the regulator to work properly is called the dropout voltage, and for achieving a high efficiency, a small dropout voltage V_{DO} is desirable.

Line regulation and load regulation. Two parameters used to describe the ability of the regulator to maintain a constant output voltage are the line regulation and the load regulation (Sedra & Smith 2016).

The line regulation describes the change in output voltage versus the change in input voltage (ripple) and as for the bandgap reference circuits, it is defined as $\partial V_O/\partial V_{IN}$. It is normally expressed in mV/V. Alternatively, the ability to suppress ripple from the input voltage may be described by the power supply rejection ratio PSRR $= 20\log(\partial V_{IN}/\partial V_O)$. It is expressed in dB, and it is frequency-dependent, i.e., depending of the frequency of the variations in the input voltage. Clearly, a small value of the line regulation is desirable, corresponding to a large value of PSRR.

The load regulation describes the change in output voltage versus the change in load current and is defined as $\partial V_O/\partial I_L$. It is normally expressed in mV/mA. It can also be interpreted as the small-signal output resistance $r_{out} = -\partial V_O/\partial I_L$of the voltage regulator and a small value is desirable.

Using an NMOS pass transistor. For the pass transistor, either an NMOS transistor or a PMOS transistor may be selected. Figure 8.26 shows a regulator with an NMOS transistor M_1. The source of M_1 is connected to the load and the drain is the input terminal. For simplicity, the simplified three-terminal symbol from Fig. 3.11(b) has been used. Whether the transistor can have bulk and source connected depends on the process used for implementing the regulator. Some processes provide special high-voltage transistors with different options, and for the pass transistor in a regulator circuit, it may be necessary to select a transistor type which can operate with a higher drain voltage than the standard NMOS transistors, depending on the maximum value of the input voltage V_{IN}.

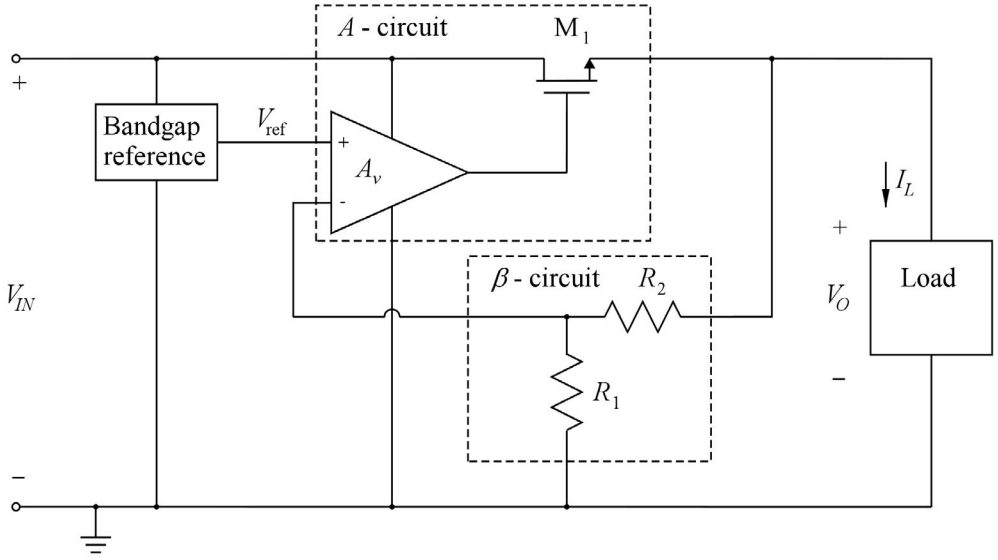


Figure 8.26: Linear voltage regulator with NMOS pass transistor.

Essentially, M_1 is a source follower buffering the output of the differential amplifier A_v. Often the differential amplifier is supplied by the unregulated input voltage V_{IN} and is designed to be able to deliver a maximum output voltage close to V_{IN}. Typically, the amplifier can deliver a maximum output voltage of $V_{IN} - |V_{DS\,\mathrm{satP}}|$ where $V_{DS\,\mathrm{satP}}$ is the saturation voltage for a PMOS transistor driving the output voltage of the amplifier high. This is the maximum gate voltage for M_1, and with a gate-source voltage of M_1 which is $V_{GS1} = V_{t1} + V_{DS\,\mathrm{satN}}$ where V_{t1} and $V_{DS\,\mathrm{satN}}$ are the threshold voltage and the saturation voltage, respectively, the maximum output voltage which can be obtained is

$$V_{O\max} = V_{IN} - V_{t1} - V_{DS\,\mathrm{satN}} - |V_{DS\,\mathrm{satP}}| \tag{8.32}$$

Thus, the dropout voltage for the regulator is

$$V_{DO} = V_{t1} + V_{DS\,\mathrm{satN}} + |V_{DS\,\mathrm{satP}}| \tag{8.33}$$

The regulator provides a good line regulation since the input voltage is connected to the drain of the pass transistor, implying that variations in V_{IN} are strongly attenuated at the output. Depending on the circuit design, the major contribution to ripple at the output may come from the bandgap reference circuit if this is supplied from an unregulated supply voltage. In the simulation examples in Section 8.3, we found the line regulation of the bandgap circuits to be on the order of 10 mV/V and the ripple on the bandgap reference voltage is amplified by the factor $(1 + R_2/R_1)$ according to Eq. (8.31).

The regulator also provides a good load regulation. Neglecting the bulk effect, the output resistance of the source-follower transistor M_1 is on the order of $1/g_{m1}$ according to Eq. (4.16), and this is further attenuated by the amount of feedback in the feedback loop. With a high gain A_v in the amplifier and a gain close to 1 in the source follower M_1, the amount of feedback is almost equal to the loop gain which is on the order of $L_0 = A_v R_1/(R_1 + R_2)$.

The feedback loop in the circuit comprises the resistors R_1 and R_2, the amplifier A_v, the pass transistor M_1 and the load. Clearly, stability of this feedback system is an important design issue but it is beyond the scope of this book and the reader is referred to Chan Carusone, Johns & Martin (2012).

Using a PMOS pass transistor. The dropout voltage of a regulator with an NMOS pass transistor always exceeds the NMOS transistor threshold voltage, and for a regulator intended to provide a supply voltage of 1.8 V or less, the power loss in the pass transistor is a substantial fraction of the total power consumption, leading to a low efficiency. Some processes may offer the option of a special NMOS transistor with a low threshold voltage but if this is not available, a PMOS pass transistor may be chosen instead in order to achieve a smaller dropout voltage, see Fig. 8.27. In this circuit, the difference between V_{IN} and V_O may be as small as $|V_{DS\,\mathrm{satP}}|$, i.e., considerably smaller than with an NMOS pass transistor. The configuration shown in Fig. 8.27 is often referred to as a low dropout (LDO) voltage regulator.

For the low dropout regulator, the source of the pass transistor is connected to the input and the drain is connected to the output. Hence, both line regulation and load regulation differ from those of a regulator

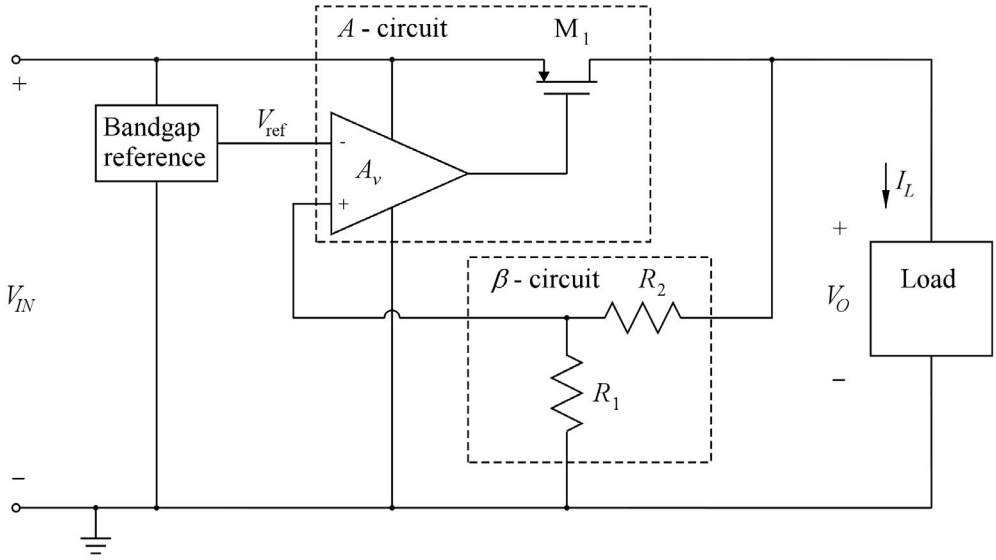


Figure 8.27: Low dropout voltage regulator with PMOS pass transistor.

with an NMOS pass transistor. Where the regulator with an NMOS pass transistor had a fairly good line regulation and load regulation, even with a rather small gain A_v in the amplifier, the line regulation and the load regulation of the low dropout regulator rely heavily on the gain in the feedback loop.

Concerning the line regulation, input voltage ripple now appears as an input signal to a common-gate configuration which has a high gain, on the order of $g_{m1}\,(r_{ds1} \parallel r_L)$, where r_L is the small-signal resistance of the load. However, it is attenuated by the feedback loop, so with a high loop gain, the output ripple is attenuated well below the input ripple. With a loop gain of about $L_0 = A_v\, g_{m1}\,(r_{ds1} \parallel r_L) R_1/(R_1+R_2)$, the resulting line regulation is

$$\frac{\partial V_O}{\partial V_{IN}} \simeq \frac{g_{m1}\,(r_{ds1} \parallel r_L)}{L_0} \simeq \left(\frac{1}{A_v}\right)\left(1+\frac{R_2}{R_1}\right) \tag{8.34}$$

Concerning the load regulation, the output resistance of the regulator is determined by the output resistance of M_1 and the loop gain. With an output resistance r_{ds1} of M_1, we find a load regulation on the order of

$$r_{out} = \left|\frac{\partial V_O}{\partial I_L}\right| \simeq \frac{r_{ds1}}{A_v\, g_{m1}\, r_{ds1}\, R_1/(R_1+R_2)} \simeq \left(\frac{1}{g_{m1}}\right)\left(\frac{1}{A_v}\right)\left(1+\frac{R_2}{R_1}\right) \tag{8.35}$$

The fact that the line regulation and the load regulation of the low dropout regulator rely on the loop gain implies that the frequency response of the loop gain is important for the characteristics and it may be more difficult to achieve good frequency characteristics for the low dropout regulator than for the standard regulator (Chan Carusone, Johns & Martin 2012). Also for the low dropout regulator, the stability of the feedback loop is an important design issue but it is beyond the scope of this book and the reader is referred to Chan Carusone, Johns & Martin (2012).

Simulation examples. We complete the presentation of voltage regulators by LTspice simulations of the circuits from Figs. 8.26 and 8.27. The LTspice schematic of the regulator with an NMOS pass transistor is shown in Fig. 8.28. For the simulations, we use the BSIM3 model from Fig. 3.44 even though this is for a 0.18 µm CMOS process and may not be representative of a pass transistor suited for the high input voltage required. The circuit has been designed for a regulated output voltage of 1.5 V and is assumed to deliver a load current of 10 mA.

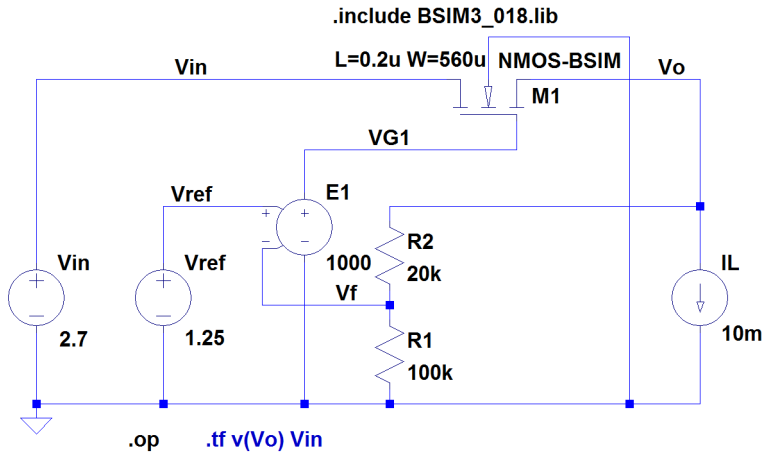


Figure 8.28: LTspice schematic corresponding to the voltage regulator from Fig. 8.26.

Using the simple Shichman-Hodges transistor model given by Eq. (3.10) with the transistor parameters from Table 3.1, we find $W/L = 2800$ for a current of 10 mA and an effective gate voltage of 200 mV. This is a fairly large aspect ratio, so in order to achieve a compact transistor structure, we select a small channel length, $L = 0.2$ µm, giving $W = 560$ µm.

With a source-bulk voltage equal to $V_O = 1.5$ V, we expect a threshold voltage for M_1 of about 0.72 V calculated from Eq. (3.17) with the transistor parameters from Table 3.1. With a saturation voltage of 0.2 V for both M_1 and the amplifier output, we find the dropout voltage from Eq. (8.33) to be 1.12 V, so the minimum input voltage is about 2.62 V. Allowing for some ripple at the input, we assume $V_{IN} = 2.7$ V.

For the amplifier, we select a gain of 1000, the order of magnitude achieved for the two-stage opamps presented in Chapters 5 and 7.

The reference voltage is assumed to have a value of $V_{\text{ref}} = 1.25$ V, so in order to achieve $V_O = 1.5$ V, we need $(1 + R_2/R_1) = 1.5/1.25 \Rightarrow R_2/R_1 = 0.2$. We select $R_1 = 100$ kΩ and $R_2 = 20$ kΩ, implying that they draw a negligible current compared to I_L.

The reference voltage is assumed to have a value of $V_{\text{ref}} = 1.25$ V, so in order to achieve $V_O = 1.5$ V, we need $(1 + R_2/R_1) = 1.5/1.25 \Rightarrow R_2/R_1 = 0.2$. We select $R_1 = 100$ kΩ and $R_2 = 20$ kΩ, implying that they draw a negligible current compared to I_L.

Since the line regulation is the small-signal gain from the input voltage to the output voltage and the absolute value of load regulation is the small-signal output resistance, we can find these parameters by running a '.tf' simulation with '`Vin`' as the source and '`v(Vo)`' as the output. From the output file, we find the line regulation as the transfer function and the absolute value of the load regulation as the output impedance. In order to find the dc output voltage, we also run a '.op' simulation. Figure 8.29 shows the simulation results.

Results from '.op' simulation.

```
        --- Operating Point ---

V(vin):         2.7             voltage
V(vg1):         2.2776          voltage
V(vo):          1.49727         voltage
```

Results from '.tf' simulation.

```
        --- Transfer Function ---

Transfer_function:              7.09391e-005    transfer
vin#Input_impedance:            1.68874e+009    impedance
output_impedance_at_V(vo):      0.0107543       impedance
```

Figure 8.29: Results from the output files from a '.op' simulation and a '.tf' simulation of the voltage regulator from Fig. 8.28.

We notice that the dc value of the output voltage is 1.497 V, i.e., very close to 1.5 V as expected. We also see that the gate voltage of M_1 is more than 0.4 V below V_{IN} which is enough headroom for the amplifier driving the gate. From the '.tf' simulation, we find the line regulation to be 0.07 mV/V, corresponding to PSRR = 83 dB, which is much better than the line regulation found for the bandgap reference circuit in Fig. 8.21. Thus, even with a smaller gain A_v in the opamp, we can achieve an adequate line regulation. The output resistance is found to be 11 mΩ, corresponding to a load regulation of −11 μV/mA.

The LTspice schematic for the low dropout regulator is shown in Fig. 8.30. The regulator has been designed to provide an output voltage of 1.5 V to a load of 10 mA. Compared to the regulator from Fig. 8.28, the pass transistor has been changed to a PMOS transistor which is four times as wide as the NMOS pass transistor in order to compensate for the difference between hole mobility and electron mobility. The dropout voltage is now equal to the saturation voltage of the pass transistor, and with this voltage selected to 200 mV, the minimum input voltage is 1.7 V. In order to allow for some input ripple, the input voltage has been selected to be 1.8 V.

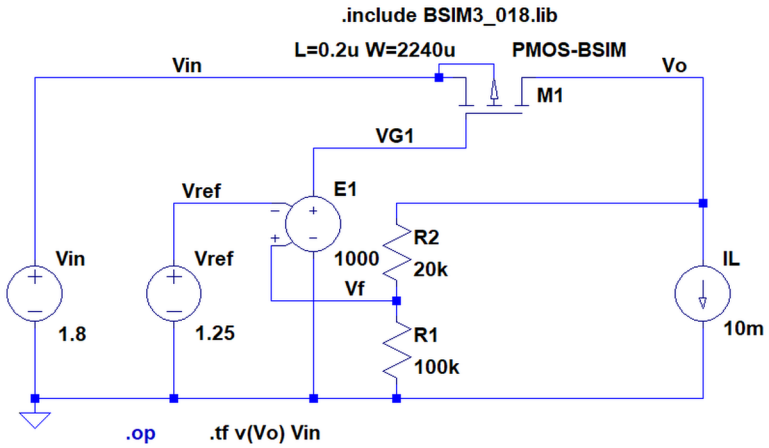


Figure 8.30: LTspice schematic corresponding to the low dropout voltage regulator from Fig. 8.27.

For this regulator, we run the same simulations as before. Figure 8.31 shows the output voltage, the line regulation and the load regulation. Again, the output voltage is very close to 1.5 V as expected, and we also note that the output voltage of the amplifier, i.e., the gate voltage of M_1 is well within the voltage range which can be delivered by the amplifier.

Results from '.op' simulation.

```
        --- Operating Point ---

V(vin):         1.8             voltage
V(vg1):         1.29246         voltage
V(vo):          1.50155         voltage
```

Results from '.tf' simulation.

```
        --- Transfer Function ---

Transfer_function:              0.00127657      transfer
vin#Input_impedance:            9.40018e+007    impedance
output_impedance_at_V(vo):      0.0094622       impedance
```

Figure 8.31: Results from the output files from a '.op' simulation and a '.tf' simulation of the low dropout voltage regulator from Fig. 8.30.

We find a line regulation of 1.28 mV/V, corresponding to PSSR = 58 dB, i.e., considerably worse than for the regulator with an NMOS pass transistor. Estimating the regulation from Eq. (8.34), we find $\partial V_O/\partial V_{IN} \simeq 1.2$ mV/V, agreeing well with the simulated value.

The output resistance is found to be 9.5 mΩ, corresponding to a load regulation of -9.5 μV/mA. This is about the same as for the regulator with an NMOS pass transistor. From the error log file from the '.op' simulation, we may find $g_{m1} = 127$ mA/V, and using Eq. (8.35), we can estimate the load regulation to be 9.4 μV/mA, again a result showing good agreement with the simulated value.

You may notice that the output voltage of the regulator with an NMOS pass transistor is slightly lower than 1.500 V whereas the output voltage for the regulator with a PMOS pass transistor is slightly higher than 1.500 V. The reason for this is that in the feedback circuit, a dc bias voltage is needed for the gate of the pass transistor. This requires a small dc input voltage to the gain stage '`E1`'. For the regulator with an NMOS pass transistor, this dc voltage must positive in order to provide a positive gate voltage for M_1. Thus, the value of V_{ref} should be slightly larger than 1.250 V to produce an output voltage of 1.500 V. For the regulator with a PMOS pass transistor, this dc voltage must negative in order to provide a positive gate voltage for M_1. Thus, the value of V_{ref} should be slightly smaller than 1.250 V to produce an output voltage of 1.500 V.

References

Allen, PE., & Holberg, DR. 2012, *CMOS Analog Integrated Circuit Design*, Third Edition, Oxford University Press, New York, USA.

Chan Carusone, T., Johns, D. & Martin, K. 2012, *Analog Integrated Circuit Design*, Second Edition, John Wiley & Sons, Inc., Hoboken, USA.

Erickson, RW. & Maksimovic, D. 2001: *Fundamentals of Power Electronics*, Second Edition, Kluwer Academic Publishers, New York, USA.

Gray, PR., Hurst, PJ., Lewis, SH. & Meyer, RG. 2009, *Analysis and Design of Analog Integrated Circuits*, Fifth Edition, John Wiley & Sons, New York, USA.

Hilbiber, DF. 1964, 'A New Semiconductor Voltage Standard', *Digest of Technical Papers, 1964 IEEE International Solid-State Circuits Conference*, pp. 32-33, Philadelphia, USA.

Mok, PKT. & Leung KN. 2004, 'Design Considerations of Recent Advanced Low-Voltage Low-Temperature-Coefficient CMOS Bandgap Voltage Reference', *Proc. IEEE 2004 Custom Integrated Circuits Conference*, pp. 635-642, Orlando, USA.

Sansen, WMC. 2006, *Analog Design Essentials*, Springer, Dordrecht, The Netherlands.

Sedra, AS. & Smith, KC. 2016, *Microelectronic Circuits*, International Seventh Edition, Oxford University Press, New York, USA.

Widlar, RJ. 1965, 'Some Circuit Design Techniques for Linear Integrated Circuits', *IEEE Trans. Circuit Theory*, vol. CT-12, no. 4, pp. 586-590, December 1965.

Multiple-choice test

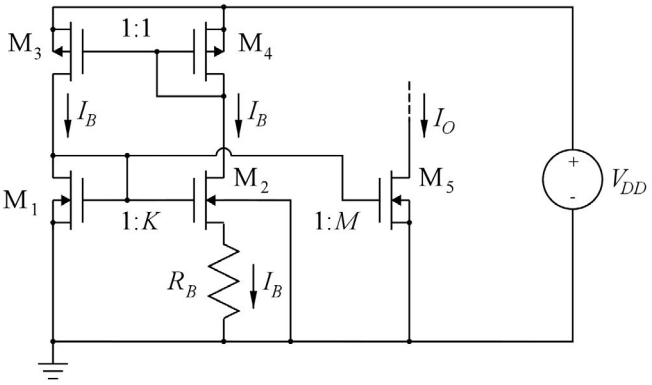
1. Complete the following statements by selecting the appropriate continuation from the table below.

A:	The minimum voltage needed across the output transistor of a simple current mirror bias source is ...
B:	The minimum voltage needed across the output transistors of a cascode current mirror bias source biased for low-voltage operation is ...
C:	The small-signal output resistance of a simple current mirror bias source is ...
D:	The small-signal output resistance of a cascode current mirror bias source is approximately ...
E:	In a MOS Widlar current source, the current is determined by a resistor and ...
F:	A bandgap voltage reference obtains a very small temperature coefficient of the reference voltage by combining ...
G:	The temperature coefficient of the voltage across a forward-biased diode is approximately ...
H:	When two identical diodes are forward-biased with different currents, the temperature coefficient of the voltage difference between the diodes is ...
I:	In a low dropout voltage regulator, the pass transistor terminal connected to the output of the regulator is ...

Continuation:

1:	the threshold voltage of a transistor.
2:	the sum of two saturation voltages.
3:	the transconductance of the transistor.
4:	the product of a transistor intrinsic gain and a small-signal output resistance.
5:	the gate-source voltage of a transistor.
6:	-2 mV/K.
7:	the voltage of a forward-biased diode and the voltage difference between two diodes which are forward biased with different current densities.
8:	the difference between the gate-source voltages of two transistors biased with different current densities.
9:	the gate.
10:	$+2$ mV/K.
11:	positive.
12:	the sum of three saturation voltages.
13:	the drain.
14:	the small-signal output resistance of the transistor.
15:	the source.
16:	the transistor saturation voltage.
17:	the bulk.

2. For the bias circuit shown below, we assume that the transistors are modeled by the Shichman-Hodges model and that the channel-length modulation can be neglected.



The scaling ratio K between transistors M_1 and M_2 must be

A: smaller than 1

B: equal to 1

C: larger than 1

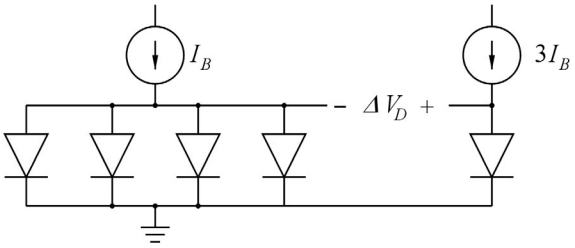
3. With $K = 4$ and $M = 3$ in the circuit above and an effective gate voltage of 0.4 V for M_1, the value of R_B required for $I_O = 0.3$ mA is

A: 1 kΩ

B: 2 kΩ

C: 4 kΩ

4. For the circuit shown below, we assume that the diodes are identical and that the relation between diode current I_D and diode voltage V_D is $I_D = I_S \exp(V_D/V_T)$ where I_S is the diode saturation current and V_T is the thermal voltage.



At room temperature (27°C), the voltage ΔV_D is approximately

A: 64 mV

B: 36 mV

C: 26 mV

5. For the circuit above, the temperature coefficient $\partial \Delta V_D / \partial T$ is approximately

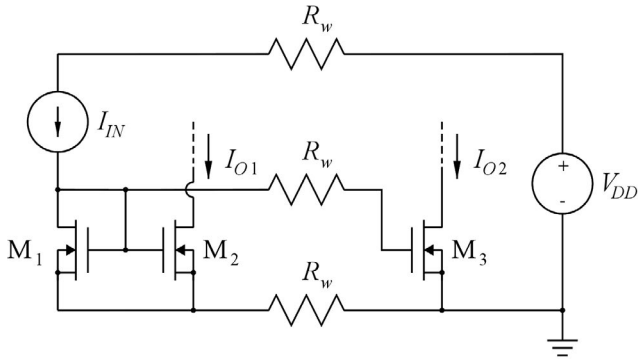
A: 0.086 mV/K

B: 0.119 mV/K

C: 0.214 mV/K

Problems

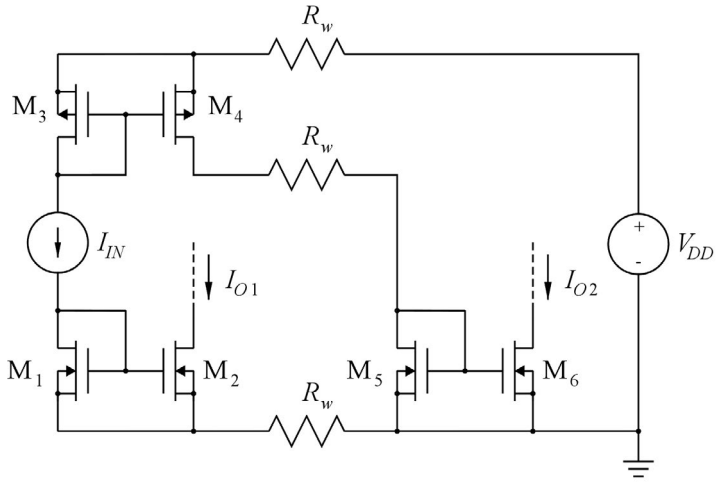
Problem 8.1



The circuit shown above is used to mirror the current $I_{IN} = 90$ µA to the output currents I_{O1} and I_{O2}. M_1, M_2 and M_3 are identical with $V_{tn} = 0.40$ V, $\mu_n C_{ox} = 180$ µA/V^2 and $L = 1$ µm. The channel-length modulation can be neglected. Design M_1 to have an effective gate-source voltage of 200 mV.

The resistors R_w represent the wiring resistance of aluminum connections between M_3 and M_1 - M_2 which are located far away from M_3. Calculate R_w, assuming a wiring length of 2 mm, width of 1 µm and thickness of 0.25 µm. The resistivity of aluminum is $\rho = 2.7 \times 10^{-8}$ Ωm. Calculate the output currents I_{O1} and I_{O2}, assuming that all transistors are in the active region.

Problem 8.2

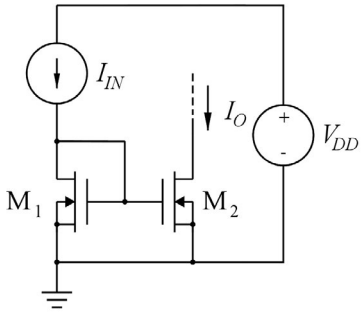


The circuit shown above is used to mirror the current $I_{IN} = 90$ µA to the output currents I_{O1} and I_{O2}. The four NMOS transistors are identical and have $V_{tn} = 0.40$ V, $\mu_n C_{ox} = 180$ µA/V^2 and $L = 1$ µm. The two PMOS transistors are also identical and have $V_{tp} = -0.42$ V, $\mu_p C_{ox} = 45$ µA/V^2 and $L = 1$ µm. The channel-length modulation can be neglected for all transistors. Design M_1 and M_3 to have an absolute value of the effective gate-source voltage of 200 mV.

The wiring resistance is $R_w = 216\ \Omega$. Calculate the output currents I_{O1} and I_{O2}, assuming that all transistors are in the active region.

Verify your results from Problems 8.1 and 8.2 using LTspice simulation.

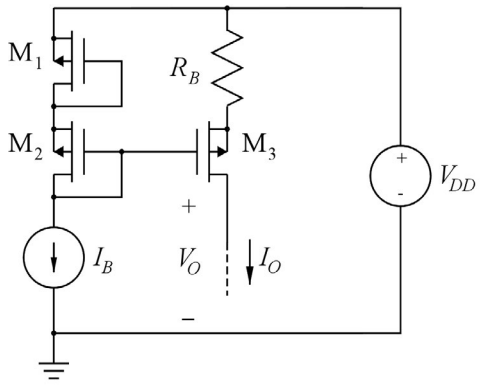
Problem 8.3



The current mirror above consists of two NMOS transistors with $W/L = 25$, $\mu_n C_{ox} = 180$ µA/V^2 and a nominal value of the threshold voltage V_{tn} of 400 mV. However, due to spread in the manufacturing process, V_{tn} may differ by 5 mV for the two transistors, i.e., $V_{tn1} = 400$ mV and $V_{tn2} = 400\ \text{mV} \pm 5$ mV. The channel-length modulation can be ignored and both transistors are in the active region.

Ideally, $I_O = I_{IN}$. However, the difference in V_{tn} causes a deviation from the ideal value of the output current. Calculate the maximum relative error in I_O for $I_{IN} = 100$ µA. Repeat the calculation for $I_{IN} = 400$ µA.

Problem 8.4

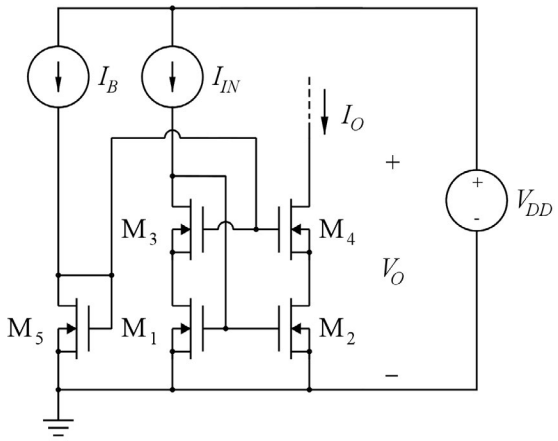


The circuit shown above is used to generate a constant current I_O. All transistors are assumed to be identical with $\mu_p C_{ox} = 45\ \mu\text{A/V}^2$, $V_{tp} = -0.42$ V and $W/L = 10$. The supply voltage is $V_{DD} = 1.8$ V. The current I_B is $I_B = 9\ \mu$A.

Calculate R_B so that $I_O = 100\ \mu$A, assuming that all transistors are in the active region.

What is the maximum value of V_O for which M_3 is in the active region?

Problem 8.5



The figure above shows a cascode current mirror biased for operation with a low output voltage and all transistors in the active region. All transistors have $\mu_n C_{ox} = 180\ \mu\text{A/V}^2$ and $V_{tn} = 0.4$ V. The channel-length modulation and the body effect can be neglected. Transistors M_1 to M_4 are assumed to be identical with $W/L = 25$. The input current I_{IN} is 90 μA and also M_5 is biased by a current $I_B = 90\ \mu$A.

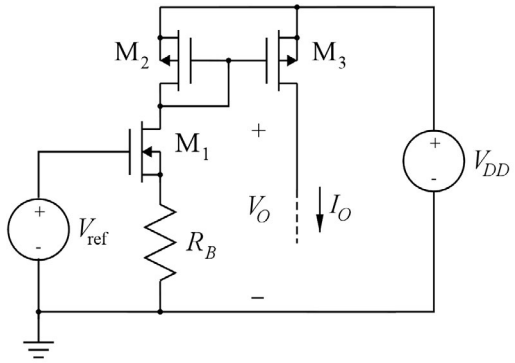
Calculate W_5/L_5 so that the gate voltage for M_3, M_4 and M_5 equals $V_{GS3} + V_{DS\text{sat}1}$.

Find the minimum output voltage for which all transistors are in the active region.

Now assume that the channel-length modulation parameter is $\lambda = 0.1\ \text{V}^{-1}$ for all transistors. Use LTspice to adjust W_5/L_5 so that M_1 and M_2 remain in the active region and find the output resistance of the current mirror when the output voltage is $V_O = 0.9$ V.

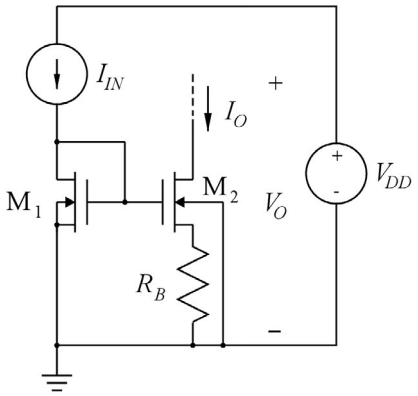
Next, assume that M_3 and M_4 have their bulk connected to ground so that the bulk effect cannot be neglected. Use LTspice with the transistor model from Table 3.1 and a channel length of 1 μm to find the maximum value of W_5/L_5 ensuring that M_1 and M_2 are in the active region. Also find the minimum output voltage for which all transistors are in the active region and find the output resistance when the output voltage is $V_O = 0.9$ V.

Problem 8.6



The figure above shows a bias circuit which generates a constant bias current I_O from a bandgap reference voltage $V_{\text{ref}} = 1.25$ V. The supply voltage is $V_{DD} = 1.8$ V and the transistors have the following parameters: $\mu_n C_{ox} = 180$ μA/V^2 and $V_{tn} = 0.4$ V, $\mu_p C_{ox} = 45$ μA/V^2 and $V_{tp} = -0.42$ V.
The channel-length modulation and the body effect can be neglected. All transistors have $W/L = 25$.
Calculate the value of R_B so that the output current is $I_O = 90$ μA, assuming that all transistors are in the active region. Find the maximum output voltage for which all transistors are in the active region.
Now, assume that M_1 has its bulk connected to ground so that the bulk effect cannot be neglected. Also assume that the channel-length modulation must be included for all transistors. Use LTspice with the transistor models from Table 3.1 and a channel length of 1 μm to find a new value for R_B resulting in $I_O = 90$ μA when $V_O = 0.9$ V. Also find the output resistance of the current source when $V_O = 0.9$ V and find the maximum output voltage for which all transistors are in the active region.

Problem 8.7

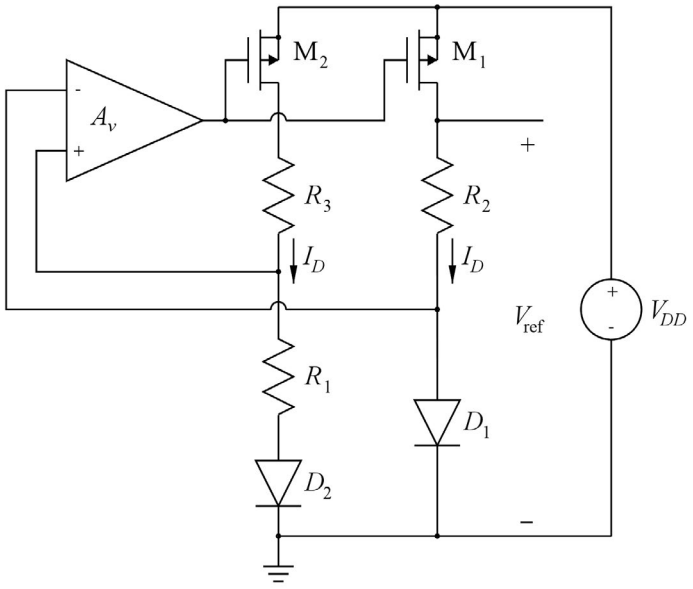


The figure above shows a Widlar current source used to generate a small dc output current I_O. Assume that M_1 and M_2 are two identical transistors with a channel length of 1 μm and a channel width of 5 μm and they have transistor parameters as specified in Table 3.1. The input current is $I_{IN} = 50$ μA and the supply voltage is $V_{DD} = 1.8$ V.
Use LTspice to find the value of R_B resulting in an output current $I_O = 5$ μA when $V_O = 0.9$ V.
From the LTspice simulation, find g_{m2}, g_{mb2} and r_{ds2} and calculate the small-signal output resistance, assuming an output voltage of $V_O = 0.9$ V. Also find the output resistance by an LTspice simulation.

Problem 8.8

Using the LTspice default diode model, find the diode voltage V_{D1} and the temperature coefficient $\partial V_{D1}/\partial T$ for a single diode with $I_D = 10$ μA at room temperature, i.e., a temperature of 27°C. Also find the diode voltage V_{DM} for 10 parallel-connected diodes with a total current of 10 μA, and find the temperature coefficient of $V_{D1} - V_{DM}$ at room temperature.
For $V_{\text{ref}} = V_{D1} + G(V_{D1} - V_{DM})$, find G so that $\partial V_{\text{ref}}/\partial T = 0$ at room temperature and find the resulting value of V_{ref}.

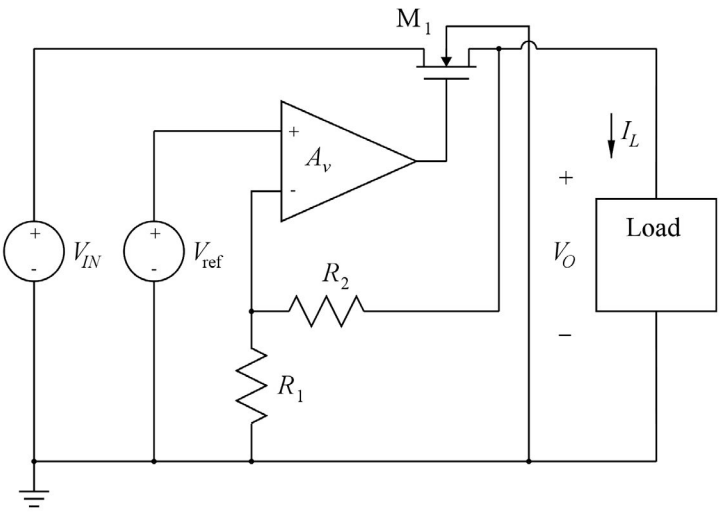
Problem 8.9



The figure above shows a bandgap reference circuit. An amplifier with the voltage gain A_v establishes a feedback loop controlling the current in the two identical PMOS transistors M_1 and M_2. The diode D_2 is composed of 10 parallel-connected diodes, each identical to diode D_1. The resistors are $R_1 = 6$ kΩ and $R_2 = R_3 = 65$ kΩ, and the amplifier has a gain $A_v = 1000$ V/V. The supply voltage is $V_{DD} = 1.8$ V. For the transistors, assume a drain current $I_D = 10$ μA and an effective gate voltage $|V_{GS} - V_{tp}| = 0.2$ V. The transistors are in the active region and the channel-length modulation can be neglected.
Find an expression for the loop gain by breaking the loop at the output of the amplifier and applying a test voltage V_{test} to the gate of M_1 and M_2, resulting in a returned voltage V_r at the output of the amplifier. The diodes have identical small-signal resistances $r_{d1} = r_{d2} = V_T/I_D$ where V_T is the thermal voltage.
Calculate the numerical value of the loop gain.

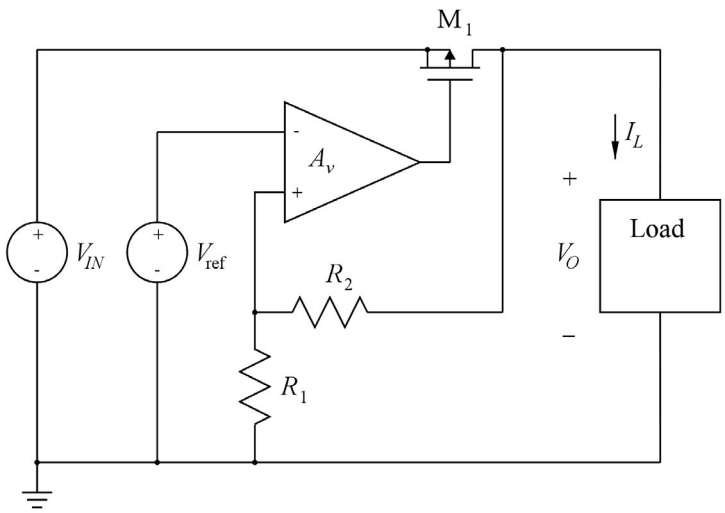
Find an expression for the small-signal gain $\partial V_{\text{ref}}/\partial V_{DD}$ and calculate the numerical value when the feedback loop is open and the dc value of V_{test} is applied to the gate of M_1 and M_2.
Find an expression for $\partial V_{\text{ref}}/\partial V_{DD}$ when the feedback loop is closed and calculate the numerical value.

Problem 8.10



The figure above shows a voltage regulator with an NMOS pass transistor M_1. Assume that the channel-length modulation for M_1 can be neglected and that it has a gate transconductance of $g_{m1} = 100$ mA/V and a bulk transconductance of $g_{mb1} = 20$ mA/V. The resistors are $R_1 = 100$ kΩ and $R_2 = 20$ kΩ, i.e., both R_1 and R_2 are much larger than $1/g_{m1}$. The load can be modeled as a dc current source with a value of 10 mA and the amplifier has a gain of $A_v = 1000$ V/V.
Find an expression for the loop gain in the feedback loop controlling V_O, and find an expression for the load regulation $\partial V_O/\partial I_L$ and calculate the numerical results.

Problem 8.11



The figure above shows a voltage regulator with a PMOS pass transistor M_1. The load is a dc current source with a value of 10 mA and the amplifier has a gain of $A_v = 1000$ V/V. The transistor has $\mu_p C_{ox}(W/L) = 500$ mA/V^2, $V_t = -0.42$ V and $\lambda = 0.7$ V^{-1}. The resistors are $R_1 = 100$ kΩ and $R_2 = 20$ kΩ. V_{ref} is a bandgap reference voltage with a value of 1.25 V and a line regulation of $\partial V_{\text{ref}}/\partial V_{IN} = 10$ mV/V. The input voltage V_{IN} has a nominal value of 1.8 V with a ripple of 0.2 V superimposed on the nominal value.

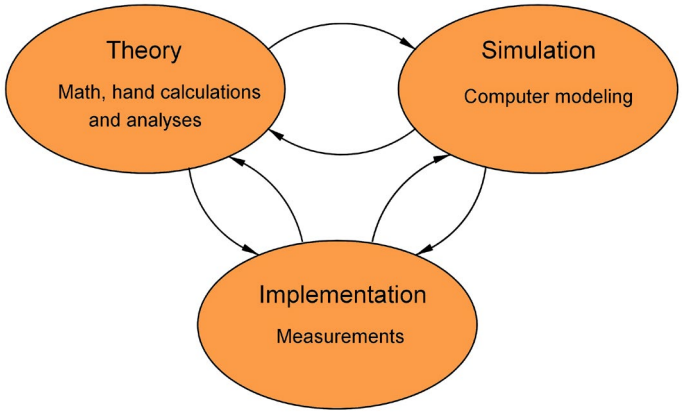
Find the nominal value of the regulated output voltage V_O and find the ripple on the output voltage. Also find the load regulation and the line regulation for the voltage regulator, including the bandgap reference circuit.

Chapter 9 – Essential Results and Equations

This final chapter of the book is a summary of the most important results and equations presented in the previous chapters. It is intended to serve as a reference guide which will - hopefully - provide a quick overview of principles and equations frequently used when designing analog CMOS circuits. References are given to the text in the preceding chapters where details can be found.

9.1 Design methodology

Design phases:

The design process includes both analytical methods using math and calculations by hand and circuit simulation by computer. Also implementation and evaluation by measurements is part of the design process, but this is not covered by the present book.	

Simulation methods in LTspice:

Dc operating point. '.op'	This command calculates an operating point and the small-signal transistor parameters in the operating point using non-linear circuit equations. The small-signal parameters are listed in the error log file, opened by 'Ctrl-L'.
Dc sweep. '.dc'	This command computes dc currents and voltages over a range of values for one, two or three independent current sources or voltage sources.
Transient analysis. '.tran'	This command computes currents and voltages as a function of time in a circuit with one or more sources specified as time-varying sources.
Dc transfer analysis. '.tf'	This command computes the small-signal input resistance, small-signal output resistance and transfer function from an (independent) input source to an output at a frequency of 0 Hz.
Ac analysis. '.ac'	This command computes the small-signal ac behavior of the circuit linearized about its dc operating point. This is used for finding the frequency response of a circuit, e.g., the Bode plot of a gain function.

Fundamental principles:

The two fundamental theorems in circuit analysis are:

Kirchhoff's current law (KCL), see Chapter 2.3.	The algebraic sum of currents flowing into a node equals zero.
Kirchhoff's voltage law (KVL), see Chapter 2.3.	The algebraic sum of voltages across circuit elements connected in a closed loop equals zero.

Additional methods:

Methods for simplifying linear circuits:

Thévenin equivalent circuit, see Chapter 2.3.	A linear two-terminal circuit consisting of impedances, voltage sources and current sources can be replaced by an equivalent circuit consisting of an independent voltage source in series with an impedance.
Norton equivalent circuit, see Chapter 2.3.	A linear two-terminal circuit consisting of impedances, voltage sources and current sources can be replaced by an equivalent circuit consisting of an independent current source in parallel with an impedance.
Superposition, see Chapter 2.3.	In a linear circuit with more than one independent source, the total response is the sum of the responses to each of the independent sources acting alone with all other independent sources being reset.

Signal notation, see Chapter 2.1.

Time-domain notation, see Fig. 2.4:

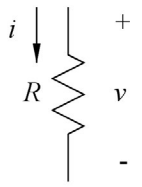
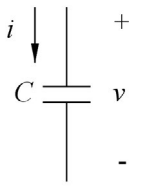
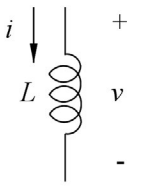
Instantaneous value.	Lowercase letter, uppercase subscript, e.g., v_A.
Dc value or bias value.	Uppercase letter, uppercase subscript, e.g., V_A.
Ac value or small-signal value.	Lowercase letter, lowercase subscript, e.g., v_a.
Amplitude.	Uppercase letter, lowercase subscript, e.g., V_a.

Frequency-domain notation:

Angular frequency ω.	Uppercase letter, lowercase subscript, e.g., $V_a(j\omega)$.

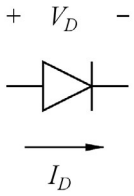
9.2 Device models, linear passive devices

See Chapter 2.3.

	Resistor	Capacitor	Inductor
Diagram symbol.			
Time domain.	$v(t) = Ri(t)$	$i(t) = C\dfrac{dv(t)}{dt}$	$v(t) = L\dfrac{di(t)}{dt}$
Frequency domain.	$V(s) = RI(s)$	$I(s) = sCV(s)$	$V(s) = sLI(s)$

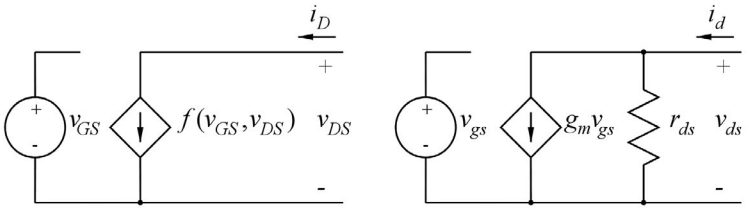
9.3 Device model, pn diode

See Chapter 3.1.

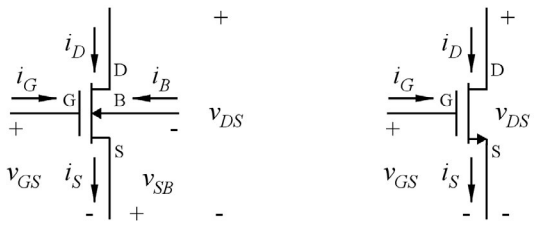
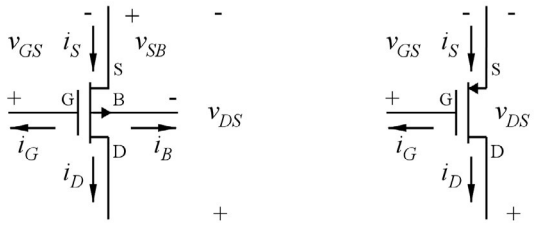
Diagram symbol.	
Shockley diode model, see Chapter 3.1.	$I_D = I_S \exp\left(\dfrac{V_D}{nV_T} - 1\right) \simeq I_S \exp\left(\dfrac{V_D}{nV_T}\right)$ for $V_D \gg V_T = \dfrac{kT}{q}$.
Small-signal model.	$r_d = \left(\dfrac{\partial i_D}{\partial v_D}\right)^{-1} \simeq \dfrac{nV_T}{I_D}$

9.4 Small-signal models

A small-signal model is a linearized model of a nonlinear device equation, see Chapter 3.5. The small-signal parameters are calculated as partial derivatives in a specific bias point (operating point). Example:

Small-signal parameters: $g_m = \left.\dfrac{\partial i_D}{\partial v_{GS}}\right\|_{\substack{\text{bias}\\\text{point}}}$ $g_{ds} = \left.\dfrac{\partial i_D}{\partial v_{DS}}\right\|_{\substack{\text{bias}\\\text{point}}} = 1/r_{ds}$	 Nonlinear large-signal model Linear small-signal model

9.5 Device models, MOS transistors

Diagram symbol, NMOS transistor, see Chapter 3.3.	
Diagram symbol, PMOS transistor, see Chapter 3.3.	
Transistor parameters.	Threshold voltage V_t, positive for NMOS transistors, negative for PMOS transistors. Channel width W. Channel length L. Channel-length modulation parameter λ, inversely proportional to L. Electron mobility μ_n (NMOS transistors). Hole mobility μ_p (PMOS transistors). Gate capacitance per unit area C_{ox}. Bulk threshold parameter (body effect constant) γ. Fermi potential of the body $\lvert\Phi_F\rvert$.
LTspice transistor parameters	Zero-bias threshold voltage (V_{to}): `Vto` Transconductance parameter (μC_{ox}): `Kp` Channel-length modulation parameter (λ): `Lambda` Bulk threshold parameter (γ): `Gamma` Surface inversion potential ($2\lvert\Phi_F\rvert$): `Phi`

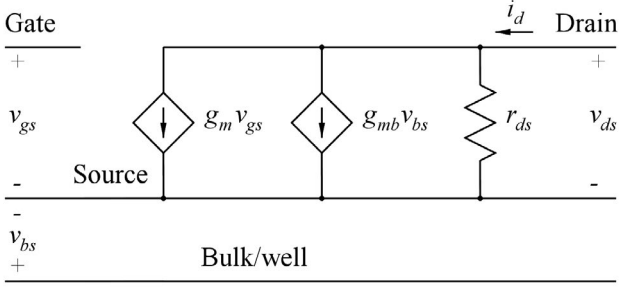
Nonlinear device models for an NMOS transistor:

Shichman-Hodges transistor model, see Chapter 3.3.	Cut-off region: $v_{GS} \leq V_t$: $i_D = 0$ Triode region: $0 \leq v_{DS} \leq v_{GS} - V_t$: $i_D = \mu_n C_{ox} \left(\frac{W}{L}\right) [(v_{GS} - V_t)v_{DS} - v_{DS}^2/2](1 + \lambda v_{DS})$ Active region (saturation region): $0 \leq v_{GS} - V_t \leq v_{DS}$: $i_D = \frac{\mu_n C_{ox}}{2} \left(\frac{W}{L}\right) (v_{GS} - V_t)^2(1 + \lambda v_{DS})$
Simplified Shichman-Hodges transistor model without channel-length modulation, see Chapter 3.3. Using $1 + \lambda v_{DS} \simeq 1$ is often a reasonable approximation for hand calculations.	Cut-off region, $v_{GS} \leq V_t$: $i_D = 0$ Triode region: $0 \leq v_{DS} \leq v_{GS} - V_t$: $i_D = \mu_n C_{ox} \left(\frac{W}{L}\right) [(v_{GS} - V_t)v_{DS} - v_{DS}^2/2]$ Active region (saturation region): $0 \leq v_{GS} - V_t \leq v_{DS}$: $i_D = \frac{\mu_n C_{ox}}{2} \left(\frac{W}{L}\right) (v_{GS} - V_t)^2$
Body effect, change in threshold voltage when source and bulk are not connected, see Chapter 3.3.	$V_t = V_{to} + \gamma(\sqrt{v_{SB} + \lvert 2\Phi_F \rvert} - \sqrt{\lvert 2\Phi_F \rvert})$ where V_{to} is the threshold voltage with $v_{SB} = 0$, γ is the bulk threshold parameter (or body effect constant) and $\lvert \Phi_F \rvert$ is the Fermi potential of the body.

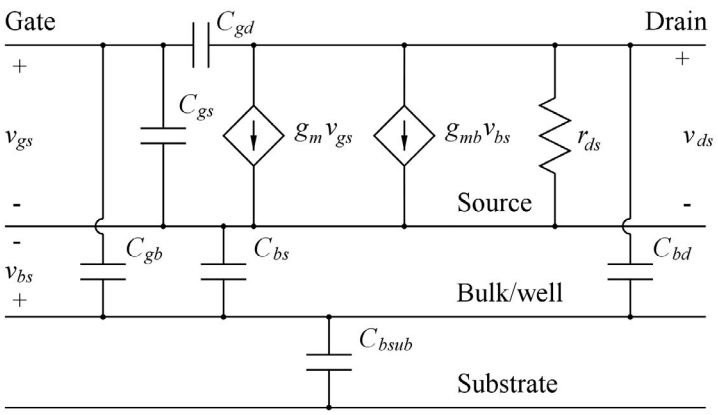
The voltage $(v_{GS} - V_t)$ is called the overdrive voltage, v_{ov}, or the effective gate voltage, v_{eff}, because it is the voltage effectively available for creating the channel beneath the gate electrode, see Chapter 3.3. It is also called the saturation voltage, $v_{DS\,\text{sat}}$, because it is the border between the triode region and the saturation region (active region) for the transistor characteristics, see Chapter 3.3.

The nonlinear device equations for a PMOS transistor are the same as for an NMOS transistors, provided absolute values for $v_{GS} - V_t$, v_{DS} and v_{SB} are used and μ_n is replaced by μ_p, see Chapter 3.3.
For a PMOS transistor, the body effect causes V_t to be more negative than V_{to}.
For a PMOS transistor in the active region, $v_{GS} - V_t$ is negative and v_{DS} is more negative than $v_{GS} - V_t$, i.e., $0 \geq v_{GS} - V_t \geq v_{DS}$ and $0 \leq |v_{GS} - V_t| \leq |v_{DS}|$. For a PMOS transistor in the triode region, $v_{GS} - V_t$ is negative and v_{DS} is negative, but less negative than $v_{GS} - V_t$, i.e., $0 \geq v_{DS} \geq v_{GS} - V_t$ and $0 \leq |v_{DS}| \leq |v_{GS} - V_t|$.

Low-frequency small-signal model in the active region:

Small-signal diagram, see Chapter 3.5. The small-signal model is the same for NMOS and PMOS transistors.	
Transconductance g_m, see Chapter 3.5. The equations are the same for NMOS and PMOS transistors, but for PMOS transistors you must use absolute values of $V_{GS}-V_t$ and V_{DS} and replace μ_n by μ_p.	$g_m = \dfrac{\partial i_D}{\partial v_{GS}} = \mu_n C_{ox}\left(\dfrac{W}{L}\right)(V_{GS}-V_t)(1+\lambda V_{DS})$ $= \dfrac{2I_D}{V_{GS}-V_t}$ $= \sqrt{2\mu_n C_{ox}\left(\dfrac{W}{L}\right) I_D (1+\lambda V_{DS})}$
Approximate expressions for g_m using $1+\lambda V_{DS} \simeq 1$ which is often a reasonable approximation for hand calculations, see Chapter 3.5. For PMOS transistors, use $\lvert V_{GS}-V_t\rvert$ and μ_p.	$g_m = \dfrac{2I_D}{V_{GS}-V_t}$ $\simeq \mu_n C_{ox}\left(\dfrac{W}{L}\right)(V_{GS}-V_t)$ $\simeq \sqrt{2\mu_n C_{ox}\left(\dfrac{W}{L}\right) I_D}$
Output conductance g_{ds}, output resistance $r_{ds} = 1/g_{ds}$, see Chapter 3.5. For PMOS transistors, use $\lvert V_{DS}\rvert$ and $\lvert V_{GS}-V_t\rvert$ and μ_p. Using $1+\lambda V_{DS} \simeq 1$ is often a reasonable approximation for hand calculations.	$g_{ds} = 1/r_{ds} = \dfrac{\partial i_D}{\partial v_{DS}} = \lambda \dfrac{\mu_n C_{ox}}{2}\left(\dfrac{W}{L}\right)(V_{GS}-V_t)^2$ $= \dfrac{\lambda I_D}{1+\lambda V_{DS}}$ $\simeq \lambda I_D$
Bulk transconductance g_{mb}, see Chapter 3.5. For PMOS transistors, use $\lvert V_{SB}\rvert$.	$g_{mb} = \dfrac{\partial i_D}{\partial v_{BS}} = g_m \dfrac{\gamma}{2\sqrt{V_{SB}+\lvert 2\Phi_F\rvert}}$
Low-frequency figure of merit, intrinsic voltage gain A_{vi}, see Chapter 3.5.	$A_{vi} = \dfrac{g_m}{g_{ds}} = \dfrac{2(1+\lambda V_{DS})}{\lambda (V_{GS}-V_t)} \simeq \dfrac{2}{\lambda (V_{GS}-V_t)}$

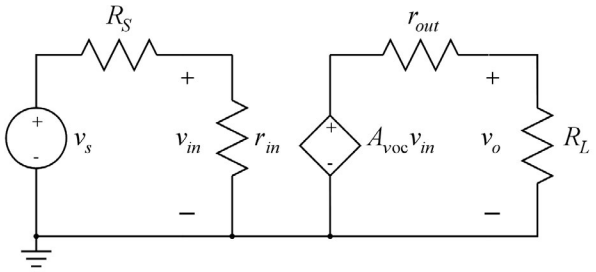
High-frequency small-signal model in the active region:

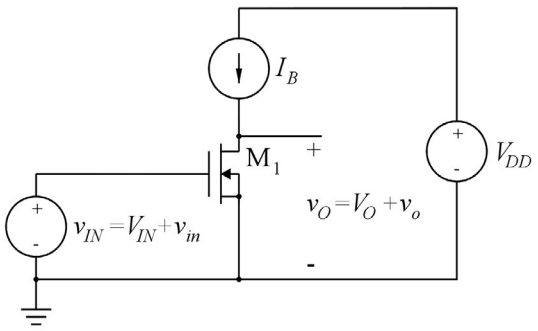
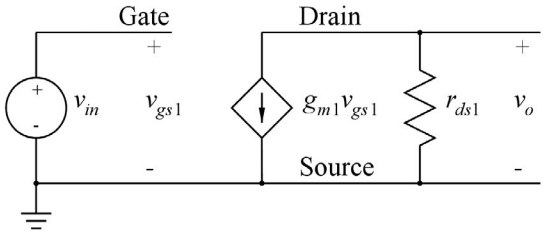
High-frequency small-signal diagram, see Chapter 3.5. The small-signal model is the same for NMOS and PMOS transistors.	
Gate-source capacitance C_{gs}, see Chapter 3.5.	Oxide capacitance. $C_{gs} \simeq (2/3)WLC_{ox}$. An overlap capacitance proportional to channel width may be added.
Gate-drain capacitance C_{gd}, see Chapter 3.5.	Overlap capacitance. Proportional to channel width. Normally $C_{gd} \ll C_{gs}$.
Gate-bulk capacitance C_{gb}, see Chapter 3.5.	Overlap capacitance. Proportional to channel length. Normally $C_{gb} \ll C_{gs}$.
Bulk-source capacitance C_{bs}, see Chapter 3.5.	Junction capacitance, approximately proportional to area of source diffusion.
Bulk-drain capacitance C_{bs}, see Chapter 3.5.	Junction capacitance, approximately proportional to area of drain diffusion.
Bulk/well-substrate capacitance C_{bsub}, see Chapter 3.5.	Junction capacitance, approximately proportional to area of well diffusion.
High-frequency figure of merit, unity-gain frequency f_T of small-signal current gain, see Chapter 3.5.	$f_T \simeq \dfrac{g_m}{2\pi C_{gs}} \simeq \dfrac{3\mu (V_{GS} - V_t)}{4\pi L^2}$

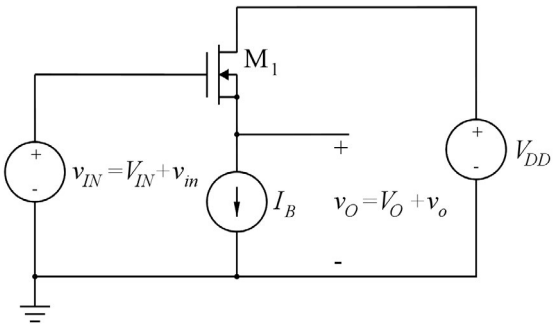
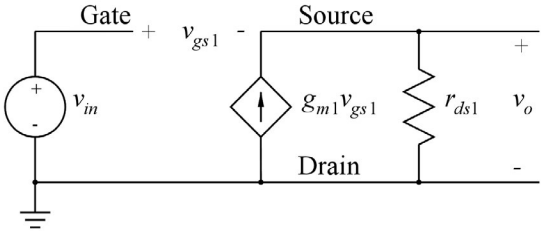
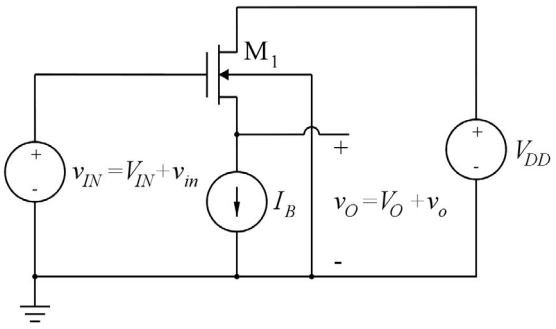
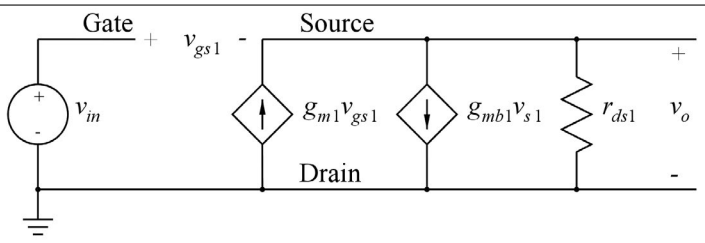
Typical Shichman-Hodges transistor parameters for a generic 0.18 µm CMOS process. Note that λ must be calculated according to channel-length from $\lambda = \lambda' / L$, see Chapter 3.3.

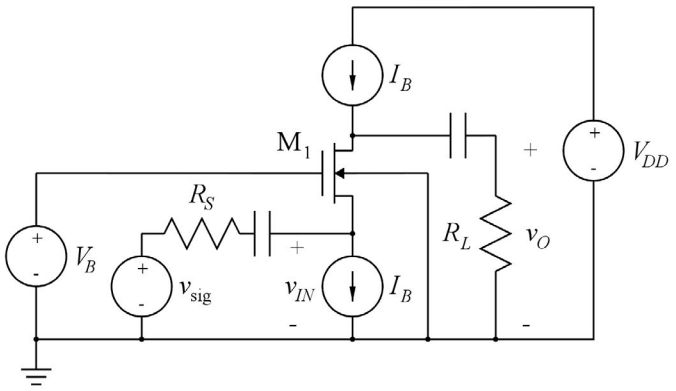
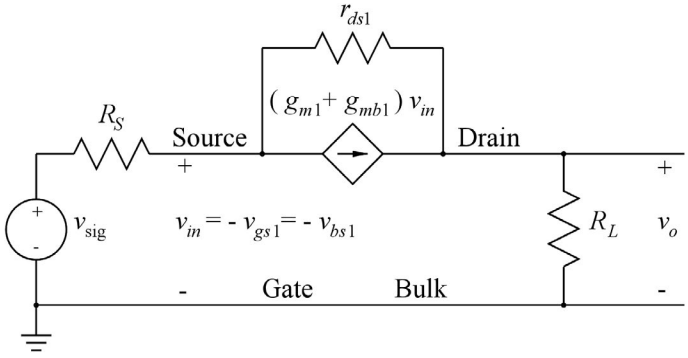
Parameter:	μC_{ox}	V_{to}	$\lambda' = \lambda L$	γ	$\lvert 2\Phi_F \rvert$
NMOS:	180 µA/V^2	0.40 V	0.10 µm/V	0.5 $\sqrt{\text{V}}$	0.7 V
PMOS:	45 µA/V^2	−0.42 V	0.14 µm/V	0.5 $\sqrt{\text{V}}$	0.7 V

9.6 Basic gain stages at low frequency

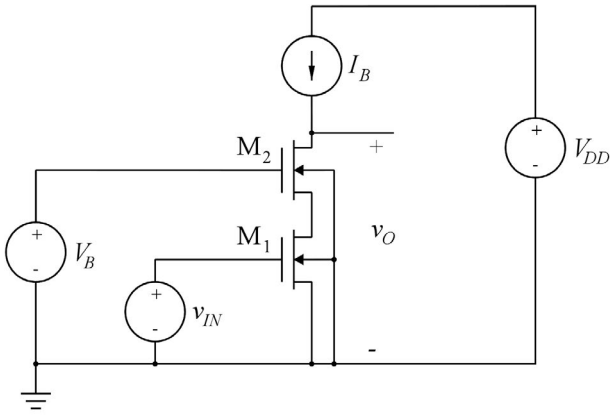
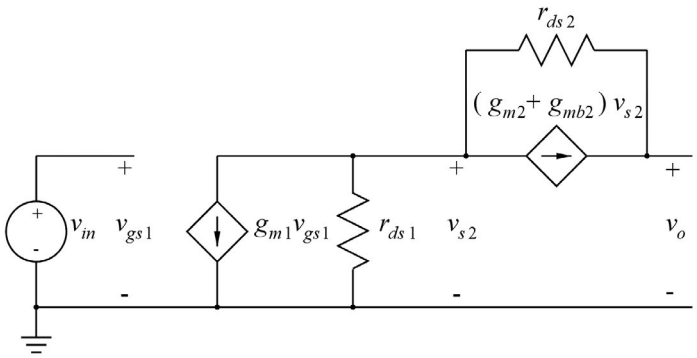
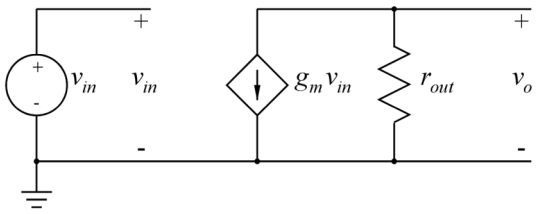
General amplifier model with a voltage-controlled voltage source, see Chapter 4.	
Small-signal parameters: Open-circuit voltage gain A_{voc}. Input resistance r_{in}. Output resistance r_{out}.	Small-signal voltage gain A_v with a load resistor R_L: $A_v = \dfrac{v_o}{v_{in}} = A_{voc}\dfrac{R_L}{R_L + r_{out}}$ Small-signal voltage gain A_{vs} with a load resistor R_L and a signal source resistor R_S: $A_{vs} = \dfrac{v_o}{v_s} = A_{voc}\left(\dfrac{R_L}{R_L + r_{out}}\right)\left(\dfrac{r_{in}}{R_S + r_{in}}\right)$

Common-source stage with ideal bias current source, see Chapter 4.1.	
Small-signal equivalent for the common-source stage.	
Input resistance r_{in}.	$r_{in} = \infty$
Output resistance r_{out}.	$r_{out} = r_{ds1}$
Open-circuit voltage gain A_{voc}.	$A_{voc} = -g_{m1}r_{ds1}$

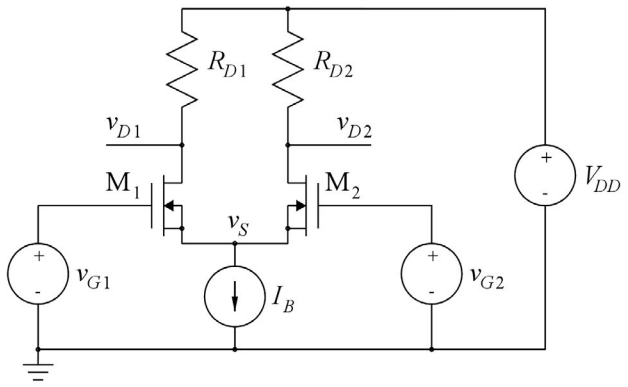
Common-drain stage (source follower) with ideal bias current source and no bulk effect, see Chapter 4.2.	
Small-signal equivalent for the common-drain stage.	
Input resistance r_{in}.	$r_{in}=\infty$
Output resistance r_{out}.	$r_{out}=\dfrac{1}{g_{m1}}\parallel r_{ds1}=\dfrac{r_{ds1}}{1+g_{m1}r_{ds1}}\simeq\dfrac{1}{g_{m1}}$
Open-circuit voltage gain A_{voc}.	$A_{voc}=\dfrac{g_{m1}r_{ds1}}{1+g_{m1}r_{ds1}}$
Common-drain stage (source follower) with ideal bias current source and bulk effect, see Chapter 4.2.	
Small-signal equivalent for the common-drain stage.	
Input resistance r_{in}.	$r_{in}=\infty$
Output resistance r_{out}.	$r_{out}=\dfrac{1}{g_{m1}+g_{mb1}}\parallel r_{ds1}=\dfrac{r_{ds1}}{1+(g_{m1}+g_{mb1})r_{ds1}}\simeq\dfrac{1}{g_{m1}+g_{mb1}}$
Open-circuit voltage gain A_{voc}.	$A_{voc}=\dfrac{g_{m1}r_{ds1}}{1+(g_{m1}+g_{mb1})r_{ds1}}\simeq\dfrac{g_{m1}}{g_{m1}+g_{mb1}}$

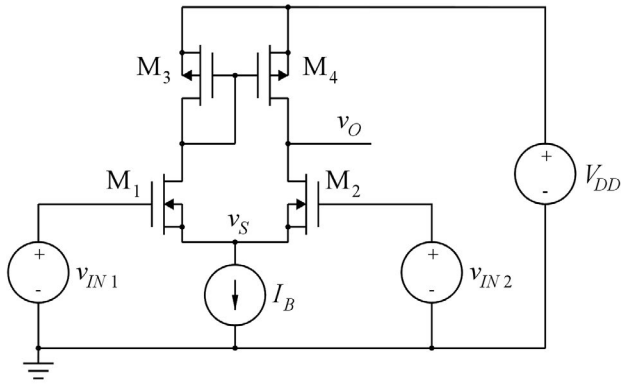
Common-gate stage with ideal bias current sources and bulk effect, see Chapter 4.3. The coupling capacitors in the diagram are considered ideal dc open circuits and ideal ac short circuits.	
Small-signal equivalent for the common-gate stage, including a source resistor R_S and a load resistor R_L.	
Input resistance r_{in}. Note that r_{in} depends on R_L.	$r_{in} = \dfrac{r_{ds1}+R_L}{1+(g_{m1}+g_{mb1})r_{ds1}} \simeq \dfrac{1}{g_{m1}+g_{mb1}} + \dfrac{R_L}{(g_{m1}+g_{mb1})r_{ds1}}$
Output resistance r_{out}. Note that r_{out} depends on R_S.	$r_{out} = R_S + (1+(g_{m1}+g_{mb1})R_S)\,r_{ds1}$ $\simeq r_{ds1} + (g_{m1}+g_{mb1})r_{ds1}\,R_S$
Open-circuit voltage gain A_{voc}.	$A_{voc} = 1+(g_{m1}+g_{mb1})r_{ds1} \simeq (g_{m1}+g_{mb1})r_{ds1}$

For a common-gate stage without bulk effect (i.e., bulk and source connected), use $g_{mb} = 0$ in the equations above.

Cascode stage with ideal bias current source and bulk effect, see Chapter 4.3.	
Small-signal equivalent for the cascode stage.	
Input resistance r_{in}.	$r_{in} = \infty$
Output resistance r_{out}.	$r_{out} = r_{ds1} + (1 + (g_{m2} + g_{mb2})r_{ds1})\, r_{ds2}$ $\simeq (g_{m2} + g_{mb2})r_{ds2}\, r_{ds1}$
Open-circuit voltage gain A_{voc}.	$A_{voc} = -(1 + (g_{m2} + g_{mb2})r_{ds2})g_{m1}r_{ds1}$ $\simeq -g_{m1}(g_{m2} + g_{mb2})r_{ds1}\, r_{ds2}$
Transconductance amplifier model for the cascode stage, see Chapter 4.3.	 $g_m \simeq g_{m1}$ $r_{out} \simeq (g_{m2} + g_{mb2})r_{ds2}\, r_{ds1}$

The cascode stage may alternatively be implemented with an NMOS/PMOS combination. This is called a folded cascode, see Chapter 4.3.

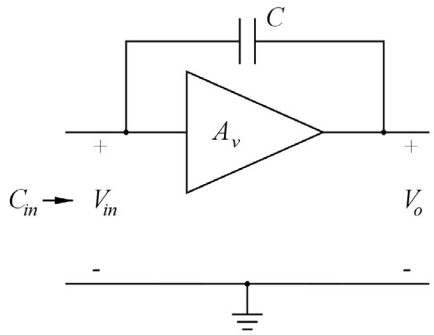
Differential pair with resistive load, see Chapter 4.3.	
Input voltages.	Differential input: $v_{ID} = v_{G1} - v_{G2}$ Common-mode input: $v_{ICM} = (v_{G1} + v_{G2})/2$ $v_{G1} = v_{ICM} + v_{ID}/2$ $v_{G2} = v_{ICM} - v_{ID}/2$
Output voltages.	Differential output voltage: $v_{OD} = v_{D1} - v_{D2}$
Small-signal transistor parameters.	$g_{m1} = g_{m2}$ $g_{ds1} = g_{ds2};\ r_{ds1} = r_{ds2}$
Small-signal differential gain.	$A_d = v_{od}/v_{id} = -(R_D \parallel r_{ds1})\, g_{m1};\ R_D = R_{D1} = R_{D2}$

Differential pair with active load and single-ended output, see Chapter 4.4.	
Small-signal transistor parameters.	$g_{m1} = g_{m2}$ $g_{ds1} = g_{ds2};\ r_{ds1} = r_{ds2}$ $g_{m3} = g_{m4}$ $g_{ds3} = g_{ds4};\ r_{ds3} = r_{ds4}$
Small-signal differential gain.	$A_d = v_o/v_{id} = v_o/(v_{in1} - v_{in2}) = (r_{ds2} \parallel r_{ds4})\, g_{m1}$
Small-signal output resistance.	$r_{out} = (r_{ds2} \parallel r_{ds4})$

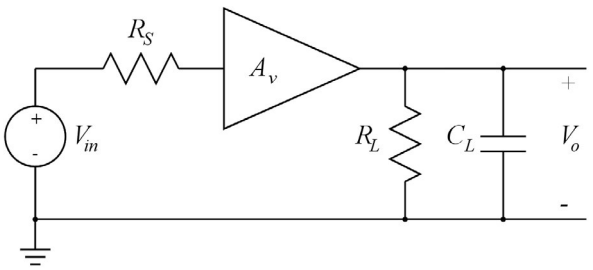
9.7 Frequency response of basic gain stages

Exact method: Find the transfer function in the frequency domain (s-domain) using node equations and normalize numerator and denominator to be polynomials in s. Find zeros by setting the numerator equal to zero and poles by setting the denominator equal to zero and solve for s.

Approximation methods for estimating pole frequencies:

Dominant time constant method, see Chapter 4.5.	For each node in the signal path, find the small-signal resistance to ground and the small-signal capacitance to ground and calculate the time constant as the product of the resistance and the capacitance. The node with the largest time constant is causing the dominant pole, i.e., the lowest pole frequency.
Miller equivalent, see Chapter 4.5: The input capacitance C_{in} of a gain stage with a voltage gain A_v and a capacitance C between input and output is given by $C_{in} = C_M \simeq (1 - A_v)C$. C_M is called the Miller capacitance.	

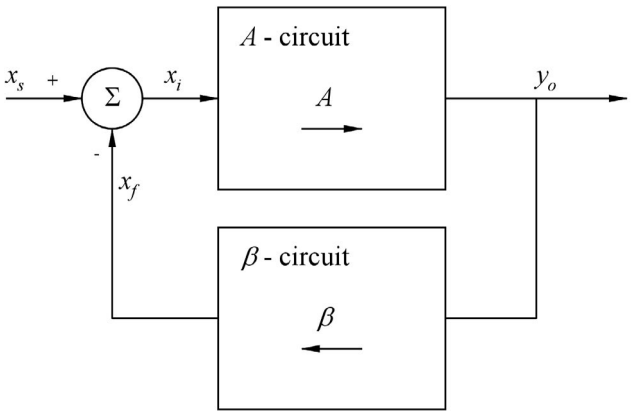
Qualitative explanations for the basic gain stages:

General model.	
Common-source stage.	The common-source stage has a gate-drain capacitance C_{gd} from input to output which is subject to the Miller effect and may cause a fairly large input capacitance. With R_S large, the dominant pole is often caused by R_S and the input capacitance. With R_S small, the dominant pole is often caused by $(R_L \parallel r_{out})$ and C_L. The common-source stage also has a zero in the transfer function. For a detailed calculation of the zero and the poles, see Chapter 6.6.

Common-drain stage.	The common-drain stage has a gate-source capacitance C_{gs} from input to output which is subject to the Miller effect and since A_v is positive and close to one, this normally causes the input capacitance to be small, whereas R_S may be large. At the output node, R_L is often small whereas C_L may be large, see Chapter 4.5.
Common-gate stage.	The common-gate stage has a small input resistance and a large output resistance, so often, the dominant pole comes from the output node, see Chapter 4.5.
Cascode stage.	The cascode stage has a very large output resistance and normally, the dominant pole comes from the output node. However, the input resistance is infinite, so with a large value of R_S, the input time constant may also be large. For a detailed discussion, see Chapter 4.5.
Differential pair.	The differential pair resembles the common-source stage and has a high output resistance when using an active load. Often the dominant pole comes from the output node with a capacitive load which is the input capacitance of a subsequent common-source stage, see Chapter 4.5.

9.8 Feedback

Basic structure of a system with negative feedback:

General model, see Chapter 6.1.	
System parameters: Open-loop gain: A Feedback factor: β	Loop gain: $L = A\beta$ Amount of feedback: $1 + L = 1 + A\beta$ Closed-loop gain: $A_{CL} = \dfrac{A}{1 + A\beta}$ For $A\beta \gg 1$, $A_{CL} \simeq \dfrac{1}{\beta}$

Benefits of feedback:

Using feedback, the gain is reduced by a factor $(1+A\beta)$ from A to A_{CL}. By sacrificing gain, we achieve improvements in several other system parameters. For the linear system parameters, the factor of improvement is $(1+A\beta)$.
In the following table, r_{in} is the input resistance of the amplifier A without feedback and r_{out} is the output resistance of the amplifier A without feedback.

Relative gain error, $\frac{dA_{CL}}{A_{CL}}$, see Chapter 6.2.	$\frac{dA_{CL}}{A_{CL}} = \left(\frac{1}{1+A\beta}\right)\frac{dA}{A}$
Closed-loop bandwidth ω_{pCL} of a system with a single pole ω_p in A, see Chapter 6.2.	$\omega_{pCL} = (1+A_0\beta)\omega_p$ where A_0 is the low-frequency value of A. With $A_0\beta \gg 1$, $\omega_{pCL} \simeq \beta A_0\omega_p = \beta\omega_{ta}$ where ω_{ta} is the gain-bandwidth product of the amplifier.
Input resistance r_{inCL} of a voltage-mode input, ideally infinite, see Chapter 6.2.	$r_{inCL} = (1+A\beta)r_{in}$
Input resistance r_{inCL} of a current-mode input, ideally zero, see Chapter 6.3.	$r_{inCL} = \frac{r_{in}}{1+A\beta}$
Output resistance r_{outCL} of a voltage-mode output, ideally zero, see Chapter 6.2.	$r_{outCL} = \frac{r_{out}}{1+A\beta}$
Output resistance r_{outCL} of a current-mode output, ideally infinite, see Chapter 6.3.	$r_{outCL} = (1+A\beta)r_{out}$

Stability of systems with feedback:

The stability of a system with feedback depends on the frequency response of the loop gain $L(s)$. With increasing frequency, $|L(s)|$ decreases and $\angle L(s)$ turns negative. With ω_t defined as the unity-gain frequency of the loop gain, the phase margin PM is the phase angle left until a phase shift of 180° in $L(s)$ is reached, i.e.,

$$\text{PM} = 180° - (-\angle L(j\omega_t)) = 180° + \angle L(j\omega_t)$$

Another stability measure is the stability margin s_m which is the shortest distance from the loop gain plot to the point $(-1,0)$ in a Nyquist plot of the loop gain, see Fig. 6.24(a). For a system with feedback to be stable, the phase margin must be positive, see Fig. 6.25.
In practice, a stability margin of more than 0.5 and a phase margin of at least 45° is normally required.

First-order system, see Chapter 6.5:

Loop gain, $L(s) = \beta A(s)$.	$L(s) = \dfrac{L_0}{1+s/\omega_p}$
Unity-gain frequency ω_t of loop gain.	$\omega_t = L_0\omega_p$, i.e., ω_t is equal to the gain-bandwidth product ω_{tl} of the loop gain.
Phase margin PM.	$\text{PM} = 180° - \arctan(\omega_t/\omega_p) = 180° - \arctan(L_0)$ $= 90° + \arctan(1/L_0); \quad \text{PM} > 90°$

Second-order system, see Chapter 6.5:
For a second-order system with a dominant pole ω_{p1} at a much lower frequency than the non-dominant pole ω_{p2}, the phase margin depends on the location of the non-dominant pole relative to the gain-bandwidth product $\omega_{tl} = L_0\omega_{p1}$ of the loop gain. The dependency is shown graphically in Fig. 6.35(a) and in tabular form below for selected values of the phase margin PM.

Loop gain, $L(s) = \beta A(s)$.	$L(s) = \dfrac{L_0}{(1+s/\omega_{p1})(1+s/\omega_{p2})} \simeq \dfrac{\omega_{tl}}{s(1+s/\omega_{p2})}; \quad \omega_{p2} \gg \omega_{p1}$
Gain-bandwidth product ω_{tl} of loop gain.	$\omega_{tl} = L_0\omega_{p1}$

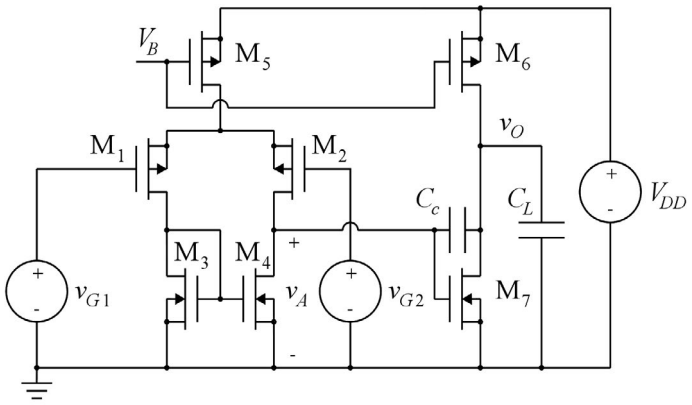
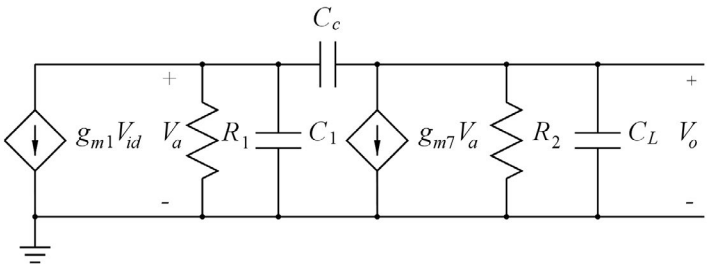
Phase margin PM.	45°	50°	55°	60°	65°	65.5°	70°	75°	76.3°
ω_{p2}/ω_{tl}.	0.71	0.91	1.17	1.50	1.94	2.00	2.58	3.60	4.00

For $\omega_{p2}/\omega_{tl} < 2$, corresponding to a phase margin of less than 65.5°, the closed-loop frequency response shows peaking. For $\omega_{p2}/\omega_{tl} < 4$, corresponding to a phase margin of less than 76.3°, the closed-loop transient response shows overshoot.
The closed-loop bandwidth is also depending on ω_{p2}/ω_{tl} and is between ω_{tl} and $1.41 \times \omega_{tl}$ for $0.5 \leq \omega_{p2}/\omega_{tl} \leq 4$, see Fig. 6.30.

Frequency compensation:

The process of designing the location of poles and zeros in order to fulfill stability requirements is called frequency compensation. Two common types of frequency compensation are dominant-pole compensation and Miller compensation.
In dominant-pole compensation, the frequency ω_{p1} of the dominant pole is reduced in order to reduce the gain-bandwidth product of the loop gain. This is typically achieved by increasing the capacitance to ground in the node causing the dominant pole, see Chapter 6.6.
In Miller compensation, a capacitor is placed from input to output of an inverting gain stage to create a dominant pole ω_{p1} from the input of the gain stage. An advantage of the Miller compensation is that it not only reduces ω_{p1} but also increases the frequency ω_{p2} of the non-dominant pole from the output of the gain stage. A disadvantage is that it introduces a right-half-plane zero causing a phase shift which reduces the phase margin, see Chapter 6.6.

9.9 The two-stage opamp

A two-stage opamp with a PMOS differential input pair. The inverting input is v_{G1} and the noninverting input is v_{G2}, so the differential input voltage is $v_{ID} = v_{G2} - v_{G1}$.	
Simplified small-signal diagram of the two-stage opamp. $R_1 = r_{ds2} \parallel r_{ds4}$; $R_2 = r_{ds6} \parallel r_{ds7}$. C_1 represents the sum of the parasitic capacitances to ground at the output of the first stage.	

Frequency compensation of the two-stage opamp for a feedback system with a feedback factor β, see Chapter 7.2.

Loop gain transfer function.	$L(s) \simeq \beta \left(\dfrac{A_0(1-s/\omega_z)}{(1+s/\omega_{p1})(1+s/\omega_{p2})} \right)$
Low-frequency loop gain.	$L_0 = \beta A_0 \simeq \beta \left(\dfrac{g_{m1}}{g_{ds2}+g_{ds4}} \right) \left(\dfrac{g_{m7}}{g_{ds6}+g_{ds7}} \right)$
Frequency of dominant pole.	$\omega_{p1} \simeq \dfrac{(g_{ds2}+g_{ds4})(g_{ds6}+g_{ds7})}{g_{m7}C_c}$
Gain-bandwidth product of loop gain.	$\omega_{tl} = \omega_{p1}L_0 \simeq \beta \dfrac{g_{m1}}{C_c}$
Frequency of non-dominant pole.	$\omega_{p2} \simeq \dfrac{g_{m7}}{C_1 + C_L + C_1C_L/C_c}$
Frequency of right-half-plane zero.	$\omega_z = \dfrac{g_{m7}}{C_c}$

Compensation approach: Make $\omega_z \gg \omega_{tl}$ so that the phase shift from from the zero is small, e.g., $\omega_z = 10\,\omega_{tl}$. This will cause a phase shift of approximately 6° from the zero at the unity-gain frequency of the loop gain.

Design the position of ω_{p2} relative to ω_{tl} for a phase margin which is at least 6° larger than the specification requirement for the phase margin using the table in Chapter 9.9 or Fig. 6.35(a).

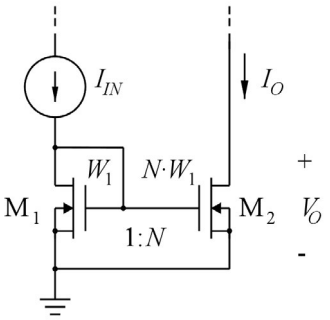
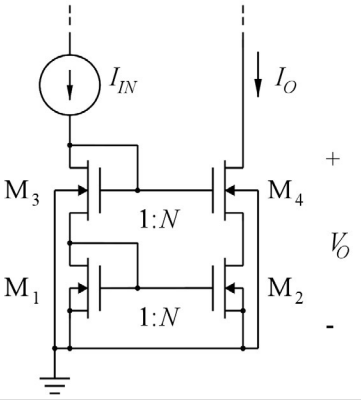
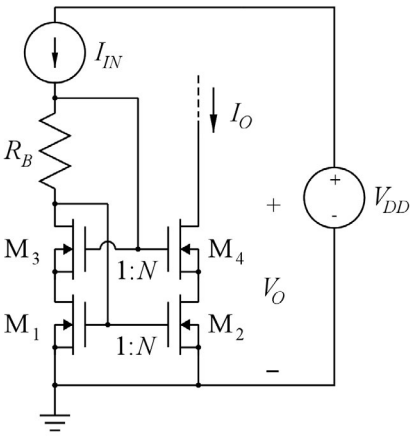
As an example, assume $C_1 \ll C_L$ and $C_1 \ll C_c$ and a design requirement for a phase margin of at least 70°. This implies that we design the non-dominant pole location corresponding to a phase margin of $\sim 76°$ in the table in Chapter 9.9, so we select $\omega_{p2}/\omega_{tl} = 4$.

Zero frequency.	$\dfrac{\omega_z}{\omega_{tl}} \simeq \left(\dfrac{1}{\beta}\right)\left(\dfrac{g_{m7}}{g_{m1}}\right) = 10 \Rightarrow g_{m7} = 10\,\beta\, g_{m1}$
Non-dominant pole frequency.	$\dfrac{\omega_{p2}}{\omega_{tl}} \simeq \left(\dfrac{g_{m7}}{C_L}\right)\left(\dfrac{C_c}{\beta\, g_{m1}}\right) = 4 \Rightarrow C_c = 0.4\, C_L$

From the frequency compensation, we have determined a relation between g_{m1} and g_{m7} and we have found the compensation capacitor C_c as a function of the load capacitance C_L.

For other design considerations, see Chapter 7.3.

9.10 Current mirrors and current sources

Simple current mirror with transistor scaling. $I_O = N\,I_{IN}$. $V_{O\min} = V_{DS\text{sat}2}$. $r_{out} = r_{ds2}$. See Chapter 3.4 and Chapter 8.1.		$W_2/L_2 = N(W_1/L_1)$ Use $L_2 = L_1$ for good matching.
Cascode current mirror with transistor scaling, shown with bulk effect for M_3 and M_4. $I_O = N\,I_{IN}$. $V_{O\min} = V_{t2} + V_{DS\text{sat}2} + V_{DS\text{sat}4}$. $r_{out} \simeq r_{ds4} + (g_{m4} + g_{mb4}) r_{ds4}\, r_{ds2}$. See Chapter 8.1.		$W_4/L_4 = N(W_3/L_3)$ $W_2/L_2 = N(W_1/L_1)$ Use $L_4 = L_3$ and $L_2 = L_1$ for good matching.
Low-voltage cascode current mirror with transistor scaling, shown without bulk effect for M_3 and M_4. $I_O = N\,I_{IN}$. $R_B\, I_{IN} = V_{t4} - V_{t2} + V_{DS\text{sat}4}$. $V_{O\min} = V_{DS\text{sat}2} + V_{DS\text{sat}4}$. $r_{out} \simeq r_{ds4} + g_{m4}\, r_{ds4}\, r_{ds2}$. See Chapter 8.1.		$W_4/L_4 = N(W_3/L_3)$ $W_2/L_2 = N(W_1/L_1)$ Use $L_4 = L_3$ and $L_2 = L_1$ for good matching.

Alternative low-voltage cascode current mirror with transistor scaling, shown without bulk effect for M_3 and M_4.

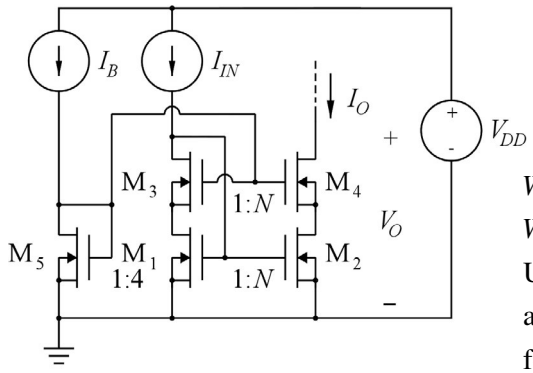
$I_O = N I_{IN}$.

$I_B = I_{IN}$; $W_5/L_5 = (W_1/L_1)/4$.

$V_{O\min} = V_{DS\text{sat}2} + V_{DS\text{sat}4}$.

$r_{out} \simeq r_{ds4} + g_{m4} r_{ds4} r_{ds2}$.

See Chapter 8.1 and Problem 8.5.



$W_4/L_4 = N(W_3/L_3)$

$W_2/L_2 = N(W_1/L_1)$

Use $L_4 = L_3$ and $L_5 = L_2 = L_1$ for good matching.

Source-degenerated current mirror, Widlar current source, shown without bulk effect for M_2.

$I_O = (V_{GS1} - V_{GS2})/R_B$.

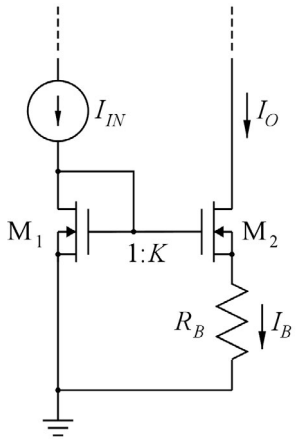
$V_{O\min} = V_{GS1} - V_{t2}$.

$r_{out} \simeq r_{ds2} + g_{m2} r_{ds2} R_B$.

$I_O = I_{IN}$ for $K = 4$ and

$$R_B = \sqrt{\frac{1}{2\mu_n C_{ox}(W_1/L_1) I_B}} = \frac{1}{g_{m1}}.$$

See Chapter 8.2.



$W_2/L_2 = K(W_1/L_1)$

Use $L_2 = L_1$ for good matching.

9.11 Bandgap reference principle

A bipolar diode forward voltage V_D has a negative temperature coefficient of about -2 mV/K.

The difference ΔV_D between the diode voltages for two diodes with current densities (diode current I_D relative to diode saturation current I_S) scaled by a factor M has a positive temperature coefficient of $\partial \Delta V_D/\partial T = (k/q)\ln(M) = 0.0861 \text{ mV/K} \times \ln(M)$.

By adding V_D and $G\Delta V_D$ with $G = 23.2/\ln(M)$, a temperature-independent voltage is obtained. The voltage is $V_{\text{ref}} = G\Delta V_D + V_D \simeq 1.25$ V and is equal to the bandgap voltage of silicon, see Chapter 8.3.

A bandgap reference circuit, see Chapter 8.3.

$$V_{D1} = V_T \ln(\frac{I_D}{I_S}); V_T = \frac{kT}{q}$$

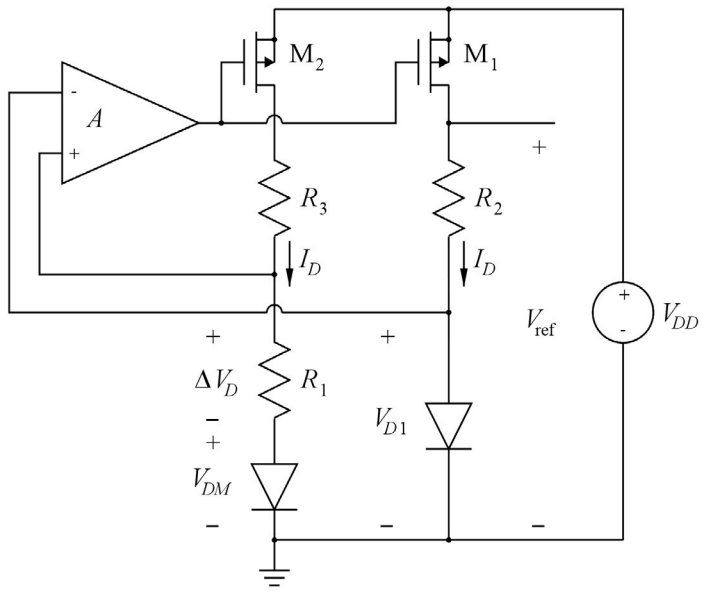
$$V_{DM} = V_T \ln(\frac{I_D}{MI_S})$$

$$\Delta V_D = V_T \ln(M)$$

$$R_2 = R_3$$

$$I_D = \frac{\Delta V_D}{R_1} = \frac{V_T}{R_1}\ln(M)$$

$$V_{\text{ref}} = V_{D1} + R_2 I_D$$
$$= V_{D1} + \left(\frac{R_2}{R_1}\right)\Delta V_D$$



9.12 Voltage regulators

A linear series voltage regulator with an NMOS pass transistor, see Chapter 8.4.

$$V_O = \left(1 + \frac{R_2}{R_1}\right) V_{\text{ref}}$$

Dropout voltage:

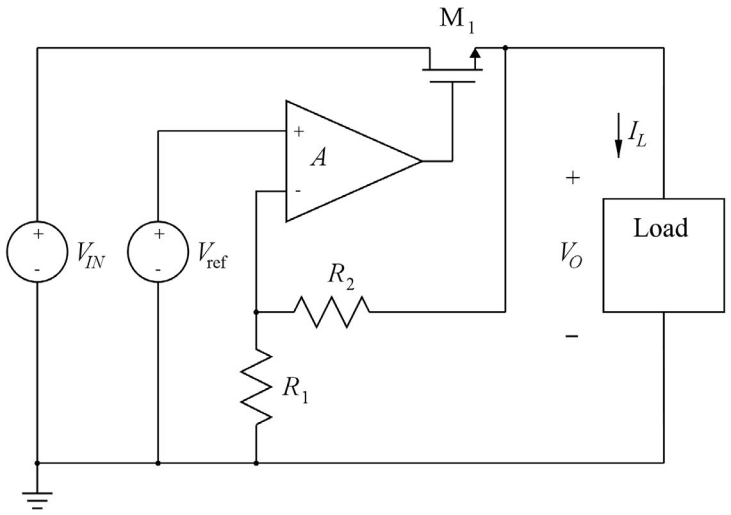
$$V_{DO} = V_{t1} + V_{DS\,\text{satN}} + V_{DS\,\text{satP}}$$

Line regulation:

$$\frac{\partial V_O}{\partial V_{IN}} \simeq \left(\frac{1}{A\, g_{m1} r_{ds1}}\right)\left(1 + \frac{R_2}{R_1}\right)$$

Load regulation:

$$\frac{\partial V_O}{\partial I_L} \simeq -\left(\frac{1}{A\, g_{m1}}\right)\left(1 + \frac{R_2}{R_1}\right)$$



A linear series voltage regulator with a PMOS pass transistor, i.e., a low dropout (LDO) regulator, see Chapter 8.4.

$$V_O = \left(1 + \frac{R_2}{R_1}\right) V_{\text{ref}}$$

Dropout voltage:

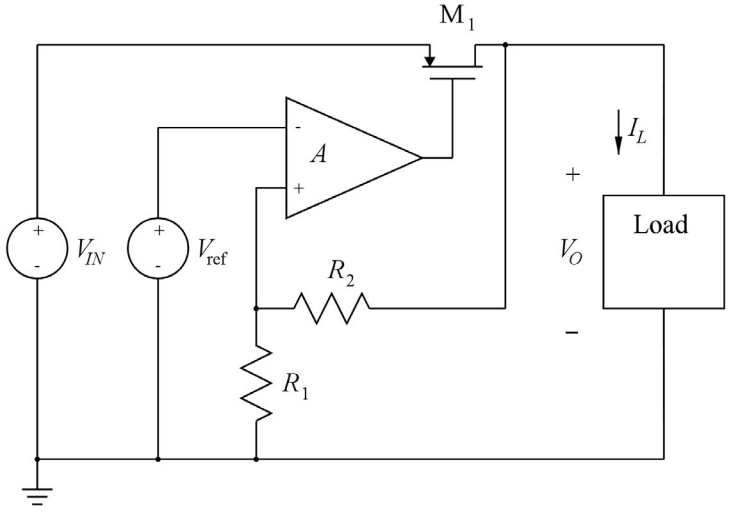
$$V_{DO} = V_{DS\,\text{satP}}$$

Line regulation:

$$\frac{\partial V_O}{\partial V_{IN}} \simeq \left(\frac{1}{A}\right)\left(1 + \frac{R_2}{R_1}\right)$$

Load regulation:

$$\frac{\partial V_O}{\partial I_L} \simeq -\left(\frac{1}{A\, g_{m1}}\right)\left(1 + \frac{R_2}{R_1}\right)$$



Appendix A

Answers to Multiple-Choice Tests

Chapter 1:

1: A-11; B-17; C-4; D-13; E-16; F-10; G-1; H-14

2: B

3: A

4: B

Chapter 2:

1: A-3; B-1; C-9; D-6; E-8; F-15; G-11; H-16; I-14; J-10

2: A

3: A

4: B

5: B

6: B

Chapter 3:

1: A-4; B-9; C-7; D-11; E-14; F-16; G-20; H-6; I-1; J-8

2: C

3: A

4: B

5: B

6: A

Chapter 4:

1: A-7; B-8; C-5; D-1; E-2; F-11; G-10; H-15; I-17; J-6

2: A-8; B-1; C-7; D-2; E-3; F-10; G-6; H-16; I-17; J-11

3: C

4: C

5: B

6: A

7: B

8: B

9: C

10: C

11: A

Chapter 5:

1: A-5; B-2; C-6; D-11; E-7; F-14; G-18; H-3; I-12; J-20

2: C

3: A

4: B

5: A

6: B

Chapter 6:

1: A-4; B-4; C-7; D-6; E-13; F-7; G-10; H-18; I-12; J-19

2: B

3: A

4: C

5: B

6: B

7: A

8: B

9: B

10: A

11: C

Chapter 7:

1: A-3; B-9; C-15; D-2; E-2; F-2; G-5; H-5; I-6

2: C

3: B

4: A

5: C

6: A

Chapter 8:

1: A-16; B-2; C-14; D-4; E-8; F-7; G-6; H-11; I-13

2: C

3: B

4: A

5: C

Appendix B

Answers to End-of-Chapter Problems

In several problems, you are encouraged to use reasonable approximations. The answers to the problems may depend somewhat on the approximations used for the calculations.

Chapter 2:

2.1: $P_{\text{avg}} = V_f^2/R$; Fundamental: 81%; 3rd harmonic: 9%; 5th harmonic: 3.2%.

2.2: $V_A = 500$ mV; $V_a = 200$ mV; $\omega = 18.85 \times 10^3$ rad/s; $\theta = 120°$.

2.3: $|V_o(j2\omega_0)| = V_a/\sqrt{5} = 0.447 V_a$; $\angle V_o(j2\omega_0) = -\arctan(2) = -63.4°$;
$v_O(t) = 0.447 V_a \cos(2\omega_0 t - 63.4°)$

2.4: $R_m = -R_2 \parallel [A_v(R_1 \parallel R_2)] = -36.4$ kΩ; $r_{in} = (R_1 \parallel R_2) \parallel (R_2/A_v) = 727\ \Omega$.

2.5: $v_C(t=0) = 0$; $v_C(t \to \infty) = 10$ V; $\tau = 0.3$ ms.

2.6: $H(s) = -\dfrac{G_m R_2 R_3 s C_2}{(1+sC_2(R_2+R_3))(1+sC_1R_1)}$; $f_{p\text{low}} = 19.9$ Hz; $f_{p\text{high}} = 24.9$ kHz.

2.7: Bias point: $I_D = 0.125$ mA; $V_O = 3.75$ V.
Small-signal gain: $dv_O/dv_{IN} = 5$ V/V $- 10$ V$^{-1} \cdot V_{IN}$; $dv_O/dv_{IN}|_{V_{IN}=1\text{ V}} = -5$ V/V.

2.8: $\dfrac{v_o}{v_{in}} = \dfrac{50\exp(-5\text{ V}^{-1}\cdot V_{IN})}{(1+\exp(-5\text{ V}^{-1}\cdot V_{IN}))^2}$.
$V_{IN} = 0$ V: $\dfrac{dv_O}{dv_{IN}} = 12.5$ V/V; $V_{IN} = \pm 0.3$ V: $\dfrac{dv_O}{dv_{IN}} = 7.46$ V/V.

2.9: $|H(jf)| = A_0 \dfrac{\sqrt{1+(f/f_z)^2}}{\sqrt{(1+(f/f_{p1})^2)(1+(f/f_{p2})^2)}}$;
$\angle H(jf) = -\arctan(f/f_z) - \arctan(f/f_{p1}) - \arctan(f/f_{p2})$;

$f = 0.4$ MHz:	$\lvert H(jf)\rvert = 44.7$ V/V	$\sim$ 33.1 dB;	$\angle H(jf) = -63.9°$
$f = 4$ MHz:	$\lvert H(jf)\rvert = 5.0$ V/V	$\sim$ 13.9 dB;	$\angle H(jf) = -94.0°$
$f = 40$ MHz:	$\lvert H(jf)\rvert = 0.36$ V/V	~ -8.9 dB;	$\angle H(jf) = -146,0°$
$f = 400$ MHz:	$\lvert H(jf)\rvert = 0.011$ V/V	~ -39.1 dB;	$\angle H(jf) = -237.7°$

Chapter 3:

3.1: $V_{D1} = -0.250$ V; $V_{S2} = -0.650$ V; $V_{D3} = 0.650$ V.

3.2: $V_{D1} = 0.864$ V; $V_{D2} = -0.200$ V; $V_{D3} = 0.230$ V.

3.3: $V_{D1} = -0.254$ V; $V_{S2} = -0.641$ V; $V_{D3} = -0.406$ V.

3.4: $R_1 = 110$ kΩ. For $\lambda = 0$: $W_1 = 17.0$ µm; For $\lambda L = 0.14$ µm/V: $W_1 = 16.5$ µm.

3.5: $W_1 = 2.05$ µm; $v_{IN} = -0.2$ V: $v_O = 0.810$ V; $v_{IN} = 0.2$ V: $v_O = -0.810$ V.

3.6: $39.7\ \text{k}\Omega \le R_{\text{eq}} \le 111.1\ \text{k}\Omega$; $v_{DS\max} = 80$ mV.

3.7: $W_1 = 8.0$ µm; $W_2 = 16.0$ µm; $r_o = 118$ kΩ.

3.8: $v_o/v_{in} = -g_m R_D = -2V_{SS}/|V_{IN} - V_{DD} - V_t|$; $V_{IN} = 0.18$ V; $R_D = 11.25$ kΩ; $W = 39.5$ µm.

3.9: $g_{m1} = 0.219$ mA/V; $r_{ds1} = 86$ kΩ; $g_{m2} = 0.210$ mA/V; $r_{ds2} = 111$ kΩ; $v_o/v_{in} = -20.8$ V/V.

3.10: $A_v(j\omega) = -\dfrac{(g_m - j\omega C_{gd})\, r_{ds}}{1 + j\omega\,(C_{gd} + C_{bd})\, r_{ds}}$; $A_0 = -75$ V/V $\sim$ 37.5 dB; $f_p = 32.2$ MHz.

3.11: Without overlap capacitances: $C_{gs} = 6.12$ fF; $C_{bd} = 4.33$ fF.
Including overlap capacitances: $C_{gd} = 0.54$ fF; $C_{gs} = 6.66$ fF.
$g_m = 248$ µA/V; $f_T = 5.5$ GHz.

3.12: Without overlap capacitances: $C_{gs} = 6.22$ fF; $C_{bd} = 4.33$ fF.
Including overlap capacitances: $C_{gd} = 0.52$ fF; $C_{gs} = 6.74$ fF.
$g_m = 248$ µA/V; $f_T = 5.3$ GHz.
According to Eq. (3.93), f_T is approximately inversely proportional to L^2, so f_T may be increased by a factor of 4 by reducing the channel length with a factor of 2. An LTspice simulation with $L = 0.3$ µm shows a value of 21.1 GHz for f_T.

Chapter 4:

4.1: $I_{D1} = I_{D2} = 10.8$ µA; $V_{IN} = 0.6$ V; 0.2 V $\leq v_O \leq 1.4$ V.

4.2: $A_v = v_o/v_{in} \simeq -62.5$ V/V. Simulated value: $A_v = -66.9$ V/V.

4.3: $V_{GS1} = V_{DS1} = 0.658$ V; $v_o/v_{in} = -(g_{m1} - 1/R_G)\,(R_L \parallel R_G) = -14$ V/V.

4.4: $W = 50$ µm; $V_S = 0.3$ V; $V_{IN} = 0.9$ V;
$v_d/v_{in} = -R_D\, g_{m1}/(1 + R_S\, g_{m1}) = -1.5$ V/V; $v_s/v_{in} = R_S\, g_{m1}/(1 + R_S\, g_{m1}) = 0.75$ V/V.

4.5: $R_D = 18$ kΩ; $W_1 = 68.9$ µm; $V_{IN} = 1.126$ V.

4.6: $V_{IN} = -0.668$ V; $A_{voc} = 0.985$ V/V; $r_{out} = 978$ Ω; $A_v = 0.824$ V/V;
-0.192 V $< v_O <$ 0.600 V.

4.7: $V_{IN} = 545.38$ mV; $A_{voc} = -13326$ V/V; $r_{out} = 48.8$ MΩ; $r_x = 368$ kΩ; $A_{vx} = -101$ V/V.

4.8: $V_{IN} = 537.66$ mV; $A_{voc} = -84.6$ V/V; $r_{out} = 327$ kΩ; $r_x = 5.33$ kΩ; $A_{vx} = -1.39$ V/V.

4.9: $I_{D1} = I_{D2} = I_{D3} = I_{D4} = 17.6$ µA; $I_{D5} = 35.3$ µA; $g_{m1} = g_{m2} = 178$ µA/V;
$g_{m3} = g_{m3} = 356$ µA/V; $g_{m5} = 252$ µA/V; $r_{ds1} = r_{ds2} = 406$ kΩ; $r_{ds3} = r_{ds4} = 568$ kΩ;
$r_{ds5} = 203$ kΩ; $A_d = 42.1$ V/V.
LTspice gives slightly larger values of bias currents and transconductances due to the channel-length modulation which has been neglected in the hand calculations. Also, simulated r_{ds} values are slightly different due to the factor $(1 + \lambda V_{DS})$ and the larger values of I_D. The differences cause the simulated value of the small-signal gain to be slightly larger than the calculated value.
Simulated input common-mode voltage range: 0.07 V to 0.91 V.

4.10: $V_o/V_{in} = -[(g_m - sC_2)\, r_{ds}]/[1 + s\, r_{ds}\,(C_2 + C_3)]$; $p = -[r_{ds}\,(C_2 + C_3)]^{-1}$; $z = g_m/C_2$.

4.11: $r_{out} = 250$ kΩ; $A_v = -50$ V/V; $f_{-3\,\mathrm{dB}} = 127.3$ kHz; GBW = 6.37 MHz.

4.12: $r_{out} = 15$ MΩ; $A_v = -3000$ V/V; $f_{-3\,\mathrm{dB}} = 2.12$ kHz; GBW = 6.37 MHz.

4.12: $r_{out} = 15$ MΩ; $A_v = -3000$ V/V; $f_{-3\,\text{dB}} = 2.12$ kHz; GBW = 6.37 MHz.

4.13: $g_m = 0.141$ mA/V; $r_{ds} = 500$ kΩ; $r_{out} = 500$ kΩ; $A_v = 5000$ V/V; $f_{-3\,\text{dB}} = 63.7$ kHz; GBW = 316 MHz.

4.14: $V_o/V_{in} = (g_m + sC_{gs})/(g_m + s(C_L + C_{gs}))$; $p = -g_m/(C_L + C_{gs})$; $z = -g_m/C_{gs}$.

4.15: $g_{m1} = g_{m2} = 0.628$ mA/V; $A_v = v_o/(v_{in1} - v_{in2}) = -6.28$ V/V; $I_{D1} = I_{D2} = I_{D3} = I_{D4} = 94$ µA; $I_{D5} = 188$ µA; $W_1 = W_2 = 11.6$ µm; $W_3 = W_4 = 46.4$ µm; $W_5 = 23.2$ µm.

4.16: $A_v = v_o/v_{in} = g_{m1}R_D/2$; $A_s = v_s/v_{in} = 1/2$; $C_{in} = C_{gd} + C_{gs}/2$.

4.17: Common-source stage: Simulated values: $r_{out} = 44.9$ kΩ, $C_{out} = 0.113$ pF.
Calculated values from small-signal parameters: $r_{out} = 44.9$ kΩ, $C_{out} = 0.113$ pF.

Low-gain cascode stage: Simulated values: $r_{out} = 106$ kΩ, $C_{out} = 0.177$ pF.
Calculated values from small-signal parameters: $r_{out} = 106$ kΩ, $C_{out} = 0.15$ pF.

High-gain cascode stage: Simulated values: $r_{out} = 5.05$ MΩ, $C_{out} = 0.413$ pF.
Calculated values from small-signal parameters: $r_{out} = 5.0$ MΩ, $C_{out} = 0.16$ pF.

Note that while the simulated and calculated resistor values match very well, some of the simulated capacitor values are significantly higher than the calculated values. This is due to the difference between the simple capacitor modeling used for the hand calculations and the advanced capacitor modeling used in LTspice.

Chapter 5:

5.1: $A_d = 1540$ V/V; $f_{-3\,\text{dB}} = 82$ kHz.

5.2: $W_1 = W_2 = 10$ µm; $W_3 = W_4 = 16$ µm; $W_5 = W_6 = 9.88$ µm; $W_7 = 32$ µm.

5.3: $A_1 = 30.7$ dB; $A_2 = 27.4$ dB. Gains are larger than 30 dB and 26 dB because the factor $(1 + \lambda V_{DS})$ has been neglected in the calculations in Problem 5.2.

5.4: Calculated −3 dB frequency: 75.8 kHz. Simulated −3 dB frequency: 63.6 kHz.
The difference is due to the factor $(1 + \lambda V_{DS})$ which has been neglected in the calculations.

5.5: $W_1 = W_2 = 20$ µm; $W_3 = W_4 = 5$ µm; $W_5 = 40$ µm; $W_6 = 200$ µm; $W_7 = 50$ µm; $V_B = 1.08$ V; $A_d = 625$ V/V ∼ 55.9 dB; $r_{out} = 8.33$ kΩ.

5.6: $g_{m1} = 0.326$ mA/V; $g_{m7} = 3.35$ mA/V; $A_d = 762$ V/V ∼ 57.6 dB; $r_{out} = 8.34$ kΩ.
Simulated values are different from calculated values because the factor $(1 + \lambda V_{DS})$ has been neglected in the calculations in Problem 5.5.

5.7: $C_c = 1.2$ pF. Simulated GBW: 40 MHz, corresponding to the calculated value.

5.8: $A_v = 1$ V/V. Frequency of peak: 180 MHz. Frequency of oscillation: 150 MHz.
Bandwidth: 56 MHz. Output resistance $r_{out} = 10.9$ Ω.

Chapter 6:

6.1: $\beta = \left(\frac{R_3 \parallel (R_2 + R_1)}{R_3 \parallel (R_2 + R_1) + R_4}\right)\left(\frac{R_1}{R_1 + R_2}\right)$. Ideal opamp: $A_{CL} = 1/\beta = 63$ V/V.
Opamp with $A = 100$ V/V: $A_{CL} = A/(1 + A\beta) = 38.65$ V/V.

6.2: $A_{\min} = 3087$ V/V.

6.3: Single stage: $A_{CL} = 100$ V/V: Tolerance: 3%. $A_{CL} = 10$ V/V: Tolerance: 0.3%.
Two-stage amplifier: Total gain $A = 100$ V/V. Tolerance: $\sim 0.6\%$

6.4: $A_{CL\max} = 20$ V/V; $\beta = 0.045$ V/V.

6.5: $A_{CL} = 18.18$ V/V; $r_{out\,CL} = 0.909$ kΩ.

6.6: For $r_{out} = 0$: $A_{CL} = -R(A_v/(1+A_v))$; $r_{inCL} = R/(1+A_v)$; $A_{CL} = -9.524$ kΩ; $r_{inCL} = 476$ Ω.
For $r_{out} = 100$ Ω: $A_{CL} = -R(A_v - r_{out}/R)/(1+A_v))$; $r_{inCL} = (R+r_{out})/(1+A_v)$;
$A_{CL} = -9.519$ kΩ; $r_{inCL} = 481$ Ω.

6.7: $A\beta = 50$; $\beta = 5$ kΩ; $i_o/v_{in} = 0.196$ mA/V.

6.8: $L(s) = (sRC/(1+sRC))(A_0/(1+sA_0/(2\pi\,\text{GBW})))$; PM = 90.6°.

6.9: $L(s) = (sRC/(1+sRC))(A_0/(1+sA_0/(2\pi\,\text{GBW})))(1/(1+s/(2\pi f_{p2})))$; PM = 52.7°.

6.10: PM = 55°. $f'_{p1} = 145$ kHz.

6.11: PM = 36° for $f_{p1} = 270$ kHz; PM = 54° for $f'_{p1} = 145$ kHz; PM = 65.5° for $f_{p1} = 93$ kHz.

6.12: Simulated values: $f_{p1} = 96$ kHz; $f_{p2} = 72.9$ MHz. $f_z = 22.5$ MHz.
Calculated values from node equations: $f_{p1} = 95.3$ kHz; $f_{p2} = 72.6$ MHz; $f_z = 22.4$ MHz.
Approximate values from (6.62), (6.59) and (6.54): $f_{p1} = 100$ kHz; $f_{p2} = 61.7$ MHz;
$f_z = 22.4$ MHz.
Approximate values from (6.60), (6.59) and (6.54): $f_{p1} = 112$ kHz; $f_{p2} = 61.7$ MHz;
$f_z = 22.4$ MHz.
The values from simulation and node equations match perfectly. The approximate values match within 15%.

Chapter 7:

7.1: $g_{m2} = 10\,g_{m1}$; $C_c = \left(\frac{g_{m1}}{g_{m2}}\right)\left(C_L + C_1 + \sqrt{(C_1+C_L)^2 + 2\left(\frac{g_{m2}}{g_{m1}}\right)C_1C_L}\right)$.
$g_{m2} = 5.0$ mA/V. $C_c = 2.14$ pF. Phase margin: 61°. Bandwidth: 57 MHz.

7.2: $C_c = 3.35$ pF. Bandwidth: 36 MHz.

7.3: $g_{m2} = 7.9$ mA/V. Bandwidth: 54 MHz.

7.4: $W_1 = W_2 = W_3 = W_4 = 13.9$ µm.

7.5: $W_7 = 290$ µm. $I_{BN6} = 1.043$ mA. $C_{gs7} \simeq 1.7$ pF. $C_c = 2.14$ pF.

7.6: Phase margin: 56°. Bandwidth: 53 MHz.

7.7: Method 1: Increase C_c to increase ω_{p2}/ω_{tl}.
New value: $C_c = 3.25$ pF. Bandwidth: 37 MHz.
Method 2: Increase W_7 and I_{BN6} by the same factor to increase g_{m7}, hence ω_z/ω_{tl} and ω_{p2}/ω_{tl}.
New values: $W_7 = 661$ µm; $I_{BN6} = 2.38$ mA. Bandwidth: 47 MHz.
Method 3: Insert a resistor R_c in series with C_c to move the zero.
$R_c = 514$ Ω. Bandwidth: 49 MHz.

7.8: $v_{IN\,\min} = 0.85$ V; $v_{IN\,\max} = 1.52$ V.

7.9: SR $\simeq 40$ V/μs (both rising and falling output signal).

Chapter 8:

8.1: $W_1 = 25$ μm. $R_w = 216\ \Omega$. $I_{O1} = 90$ μA. $I_{O2} = 128$ μA.

8.2: $W_1 = 25$ μm. $W_3 = 100$ μm. $I_{O1} = 90$ μA. $I_{O2} = 90$ μA.

8.3: $I_{IN} = 100$ μA: Maximum relative error: 4.8%. $I_{IN} = 400$ μA: Maximum relative error: 2.4%.

8.4: $R_B = 1.53$ kΩ. $V_{O\,\max} = 0.98$ V.

8.5: $W_5/L_5 = 6.25$. $V_{O\,\min} = 0.4$ V.
With the channel-length modulation included: $W_5/L_5 = 5.9$ and $r_o \simeq 12.8$ MΩ.
With bulk effect: $W_5/L_5 = 4.5$, $V_{O\,\min} \simeq 0.4$ V and $r_o \simeq 16.1$ MΩ.

8.6: $R_B = 7.22$ kΩ. $V_{O\,\max} = 1.40$ V.
With bulk effect and channel-length modulation: $R_B = 6.65$ kΩ. $r_o \simeq 89.4$ kΩ.
$V_{O\,\max} = 1.42$ V.

8.7: $R_B = 34.4$ kΩ. $g_{m2} = 98.2$ μA/V. $g_{mb2} = 26.3$ μA/V. $r_{ds2} = 2.15$ MΩ.
Calculated output resistance: $r_o = 11.4$ MΩ. Simulated output resistance: $r_o = 11.4$ MΩ.

8.8: $V_{D1} = 536.0$ mV. $\partial V_{D1}/\partial T = -2.17$ mV/K.
$V_{DM} = 476.4$ mV. $\partial(V_{D1} - V_{DM})/\partial T = 0.1984$ mV/K. $G = 10.94$. $V_{\text{ref}} = 1.19$ V.

8.9: $L_0 \simeq A\, g_{m1}\, R_1 = 600$ V/V. Open loop: $\partial V_{\text{ref}}/\partial V_{DD} \simeq g_{m1}\,(R_2 + r_{d1}) = 6.76$ V/V.
Closed loop: $\partial V_{\text{ref}}/\partial V_{DD} \simeq ((R_2 + r_{d1})/R_1)/A = 11.3$ mV/V.

8.10: $L_0 \simeq A_v\,(g_{m1}/(g_{m1} + g_{mb1}))\,(R_1/(R_1 + R_2)) = 694$ V/V.
Load regulation $\partial V_O/\partial I_L \simeq -(1/(A\, g_{m1}))\,(1 + R_2/R_1) = -12$ μV/mA $= -12$ mΩ.

8.11: $V_O = 1.5$ V, ripple: 2.64 mV.
Load regulation $\partial V_O/\partial I_L \simeq -10.9$ μV/mA $= -10.9$ mΩ.
Line regulation $\partial V_O/\partial V_{IN} \simeq 13.2$ mV/V.

Appendix C – Transistor Models

The parameters from Table 3.1 are used in several problems. For convenience, Table 3.1 is shown below.

Parameter:	μC_{ox}	V_{to}	$\lambda' = \lambda L$	γ	$\lvert 2\Phi_F \rvert$
NMOS:	180 µA/V^2	0.40 V	0.10 µm/V	0.5 $\sqrt{\text{V}}$	0.7 V
PMOS:	45 µA/V^2	−0.42 V	0.14 µm/V	0.5 $\sqrt{\text{V}}$	0.7 V

Table 3.1: Shichman-Hodges transistor parameters for a generic 0.18 µm CMOS process.

A corresponding LTspice model also including transistor parameters for capacitive effects is shown below in Fig. 3.35 where the values of '`Lambda`' correspond to a channel length of 1 µm.

```
.model NMOS-SH nmos (Kp=180u Vto=0.40 Lambda=0.10 Gamma=0.5 Phi=0.7
+TOX=4.0n CGSO=0.29n CGBO=0 CGDO=0.29n CJ=3.65m CJSW=0.79n)

.model PMOS-SH pmos (Kp=45u Vto=-0.42 Lambda=0.14 Gamma=0.5 Phi=0.7
+TOX=4.2n CGSO=0.28n CGBO=0 CGDO=0.28n CJ=1.38m CJSW=1.44n)
```

Figure 3.35: LTspice transistor models including parameters for calculating small-signal capacitances.

Also the BSIM3 transistor models from Fig. 3.44 are used in some problems. For convenience, Fig. 3.44 is shown below.

Generic BSIM3 model for 0.18 µm CMOS process. Adapted from 'http://ptm.asu.edu/modelcard/180nm_bulk.txt '.

```
* Predictive Technology Model Beta Version
* 180nm NMOS SPICE Parametersv (normal one)
.model NMOS-BSIM NMOS
+Level=49
+Lint=4.e-08              Tox=4.e-09
+Vth0=0.3999              Rdsw=250
+Tref=27.0                version=3.1
+Xj=6.0000000E-08         Nch=5.9500000E+17
+lln=1.0000000            lwn=1.0000000          wln=0.00
+wwn=0.00                 ll=0.00
+lw=0.00                  lwl=0.00               wint=0.00
+wl=0.00                  ww=0.00                wwl=0.00
+Mobmod=1                 binunit=2
+Dwg=0.00                 Dwb=0.00
+K1=0.5613000             K2=1.0000000E-02
+K3=0.00                  Dvt0=8.0000000         Dvt1=0.7500000
+Dvt2=8.0000000E-03       Dvt0w=0.00             Dvt1w=0.00
+Dvt2w=0.00               Nlx=1.6500000E-07      W0=0.00
+K3b=0.00                 Ngate=5.0000000E+20
+Vsat=1.3800000E+05       Ua=-7.0000000E-10      Ub=3.5000000E-18
+Uc=-5.2500000E-11        Prwb=0.00
+Prwg=0.00                Wr=1.0000000           U0=3.5000000E-02
+A0=1.1000000             Keta=4.0000000E-02     A1=0.00
+A2=1.0000000             Ags=-1.0000000E-02     B0=0.00
+B1=0.00
+Voff=-0.12350000         NFactor=0.9000000      Cit=0.00
+Cdsc=0.00                Cdscb=0.00             Cdscd=0.00
+Eta0=0.2200000           Etab=0.00              Dsub=0.8000000
+Pclm=5.0000000E-02       Pdiblc1=1.2000000E-02  Pdiblc2=7.50E-03
+Pdiblcb=-1.3500000E-02   Drout=1.7999999E-02    Pscbe1=8.66E+08
+Pscbe2=1.0000000E-20     Pvag=-0.2800000        Delta=1.00E-02
+Alpha0=0.00              Beta0=30.0000000
+kt1=-0.3700000           kt2=-4.0000000E-02     At=5.5000000E+04
+Ute=-1.4800000           Ua1=9.5829000E-10      Ub1=-3.3473E-19
+Uc1=0.00                 Kt1l=4.0000000E-09     Prt=0.00
+Cj=0.00365               Mj=0.54                Pb=0.982
+Cjsw=7.9E-10             Mjsw=0.31              Php=0.841
+JS=1.50E-08              JSW=2.50E-13
+N=1.0                    Xti=3.0                Cgdo=2.786E-10
+Cgso=2.786E-10           Cgbo=0.0E+00           Capmod= 2
+NQSMOD=0                 Elm=5                  Xpart=1
+Cgsl=1.6E-10             Cgdl=1.6E-10           Ckappa=2.886
+Cf=1.069e-10             Clc=0.0000001          Cle=0.6
+Dlc=4E-08                Dwc=0                  Vfbcv=-1
```

```
* Predictive Technology Model Beta Version
* 180nm PMOS SPICE Parametersv (normal one)
.model PMOS-BSIM PMOS
+Level=49
+Lint=3.e-08              Tox=4.2e-09
+Vth0=-0.42               Rdsw=450
+Tref=27.0                version=3.1
+Xj=7.0000000E-08         Nch=5.9200000E+17
+lln=1.0000000            lwn=1.0000000          wln=0.00
+wwn=0.00                 ll=0.00
+lw=0.00                  lwl=0.00               wint=0.00
+wl=0.00                  ww=0.00                wwl=0.00
+Mobmod= 1                binunit=2
+Dwg=0.00                 Dwb=0.00
+K1=0.5560000             K2=0.00
+K3=0.00                  Dvt0=11.2000000        Dvt1=0.7200000
+Dvt2=-1.0000000E-02       Dvt0w=0.00             Dvt1w=0.00
+Dvt2w=0.00               Nlx=9.5000000E-08      W0=0.00
+K3b=0.00                 Ngate=5.0000000E+20
+Vsat=1.0500000E+05       Ua=-1.2000000E-10      Ub=1.0000000E-18
+Uc=-2.9999999E-11        Prwb=0.00
+Prwg=0.00                Wr=1.0000000           U0=8.0000000E-03
+A0=2.1199999             Keta=2.9999999E-02     A1=0.00
+A2=0.4000000             Ags=-0.1000000         B0=0.00
+B1=0.00
+Voff=-6.40000000E-02      NFactor=1.4000000      Cit=0.00
+Cdsc=0.00                Cdscb=0.00             Cdscd=0.00
+Eta0=8.5000000           Etab=0.00              Dsub=2.8000000
+Pclm=2.0000000           Pdiblc1=0.1200000      Pdiblc2=8.00E-05
+Pdiblcb=0.1450000        Drout=5.0000000E-02    Pscbe1=1.00E-20
+Pscbe2=1.0000000E-20     Pvag=-6.0000000E-02    Delta=1.00E-02
+Alpha0=0.00              Beta0=30.0000000
+kt1=-0.3700000           kt2=-4.0000000E-02     At=5.5000000E+04
+Ute=-1.4800000           Ua1=9.5829000E-10      Ub1=-3.3473E-19
+Uc1=0.00                 Kt1l=4.0000000E-09     Prt=0.00
+Cj=0.00138               Mj=1.05                Pb=1.24
+Cjsw=1.44E-09            Mjsw=0.43              Php=0.841
+JS=1.50E-08              JSW=2.50E-13
+N=1.0                    Xti=3.0                Cgdo=2.786E-10
+Cgso=2.786E-10           Cgbo=0.0E+00           Capmod=2
+NQSMOD=0                 Elm=5                  Xpart=1
+Cgsl=1.6E-10             Cgdl=1.6E-10           Ckappa=2.886
+Cf=1.058e-10             Clc=0.0000001          Cle=0.6
+Dlc=3E-08                Dwc=0                  Vfbcv=-1
```

Figure 3.44: Library file 'BSIM3_018.lib' with BSIM3 models for a generic 0.18 µm CMOS process, adapted from Predictive Technology Model Website, http://ptm.asu.edu/.

INDEX