SIGNALS AND SYSTEMS

THIRD EDITION

Chi-Tsong Chen
State University of New York at Stony Brook

New York Oxford
OXFORD UNIVERSITY PRESS
2004

Oxford University Press

Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016
www.oup.com

Library of Congress Cataloging-in-Publication Data
Chen, Chi-Tsong.
Signals and systems / Chi-Tsong Chen.—3rd ed.
p. cm.
Rev. ed. of: System and signal analysis. 2nd ed. c1994.
Includes bibliographical references and index.
ISBN 0-19-515661-7 (acid-free paper)
1. System analysis. 2. Signal theory (Telecommunication) I. Chen, Chi-Tsong. System and signal analysis. II. Title.

QA402.C4423 2004
003—dc22

2003064970

Printing number: 9 8 7 6 5 4 3 2 1

Printed in the United States of America
on acid-free paper

To my grandchildren

Jordan, Lauren, Leo,

and those yet to come

PREFACE

This text introduces basic concepts in signals and systems and their associated mathematical and computational tools. Its purpose is to provide a common background for subsequent courses in electrical engineering: communications, control, electronic circuits, filter design, and digital signal processing. It is also useful in the study of vibrations and control in mechanical engineering. The reader is assumed to have knowledge of general physics (including simple circuit analysis), simple matrix operations, and basic calculus. Some knowledge of differential equations is helpful but not essential.

This text is a revision of *System and Signal Analysis, 2nd edition,* which was published in 1994. It differs considerably in presentation and emphasis from the second edition and from other texts on signals and systems:

1. Most texts on signals and systems prepare students for subsequent courses in communications, control, and signal processing. Some lean more heavily on communications, some more on control. The course on signals and systems at Stony Brook was changed in the mid-1990s from an elective to a required course for all students in the electrical engineering and computer engineering programs. To reflect this change, this text uses (in addition to passive RLC networks and simple mechanical systems) operational amplifiers (op amps) as examples to discuss system concepts, introduces op-amp circuit implementation, and discusses feedback design. Thus, this text also prepares students for subsequent electronics courses.
2. Most texts on signals and systems give a fairly complete treatment of transform theory and then discuss engineering applications. This revision, in response to feedback from alumni working in industry, introduces engineering concepts at the earliest possible stage and then introduces transform theory only as needed. Thus, the discussion of transform theory is not exhaustive. We skip many topics that are only of academic interest or rarely arise in practice. Instead, we focus on concepts (such as speeds of response and operational frequency ranges of devices) that are relevant to engineering. We also include engineering constraints when discussing mathematics, such as in the feedback implementation of inverse systems. Otherwise, a design may become simply a mathematical exercise with no engineering bearing.
3. This revision stresses computer computation. It introduces the fast Fourier transform (FFT) to compute frequency spectra. We discuss its actual employment rather than its internal

programming. Because most, if not all, system responses in MATLAB[1] are computed from state-space (ss) equations, we develop ss equations and discuss their computer computation and their op-amp circuit implementations.

4. Mathematical training is important in engineering, we believe, not in its abstract concepts and theoretical results but in its rigorous methodology. Thus, we define every term carefully and develop every topic logically. For example, we define λ as a *zero* of the rational function $H(s) = N(s)/D(s)$ if $H(\lambda) = 0$. Many engineering texts define λ as a zero of $H(s)$ if $N(\lambda) = 0$. This is correct only if $D(\lambda) \neq 0$ or $N(s)$ and $D(s)$ have no common factor at λ. This leads to the concepts of *coprimeness* and *degree*. These concepts are important in engineering, because they are used in *minimal realizations* and in optimal and pole-placement designs in control, yet they are not discussed in most other texts. We also distinguish between *magnitudes* and *amplitudes* and between *lumped systems* and *distributed systems*. Note that the results in lumped systems are not necessarily applicable to distributed systems. For every positive result, we show a negative result to illustrate the importance of the condition involved.
5. This revision tries to present every topic in the simplest possible way. For example, the discussion leading to the FFT is self-contained even without developing the discrete Fourier transform (DFT). The presentation of the Routh test for checking stability is believed to be simpler than the conventional cross-product method. We also simplify the concepts of controllability and observability in ss equations to the concepts of coprimeness and degree in transfer functions and use them to discuss the redundancy of systems.

The following is a brief description of each chapter. It also provides a global view and introduces key terminology of the subject area:

1. *Signals:* We introduce *continuous-time* (CT), *discrete-time* (DT), and *digital signals*. We give reasons for not studying digital signals even though they are used in all computer computation and digital signal processing. We then introduce some simple CT signals, discuss their various manipulations, and give reasons for using *complex exponential functions* to define frequency components of CT sinusoids. Unlike CT signals, DT signals can be plotted against *time instants* or *time index* (without specifying the *sampling period*). The latter is used in their manipulations,[2] and the former is used in discussing their frequency content. We show that the frequency of DT sinusoids cannot be directly defined. We then define it *formally* from the frequency of CT sinusoids and justify it physically. This leads to the concepts of *Nyquist frequency range* for DT signals and *frequency aliasing* due to time sampling.
2. *Systems:* We model a system as a black box whose every input, a signal, excites a unique output, another signal. We then introduce systems with and without memory. For systems with memory, we introduce the concepts of *causality, state* (set of initial conditions), *lumpedness, zero-state (forced) responses,* and *zero-input (natural) responses*. We then

[1]MATLAB is a registered trademark of The MathWorks, Inc.

[2]They involve manipulations of streams of numbers. The procedures are the same for any sampling period T and, consequently, are independent of T.

introduce the concepts of *linearity* (L) and *time invariance* (TI). We discuss their implications and explain why the study of LTI systems is relatively simple. We then use examples, including op amps, to show that LTI lumped systems are obtained in practice by modeling, approximation, and simplification and are valid only for limited input signal ranges.

3. *Mathematical descriptions of systems:* We use the concepts of linearity and time invariance to develop *discrete* and *integral convolutions* for LTI systems and *difference* and *differential equations* for LTI and lumped systems. The development is generic and is applicable to any system, be it electrical, mechanical, chemical, or biomedical.
4. *CT signal analysis:* We use the Fourier series and Fourier transform to develop *frequency spectra* for CT signals. We discuss several sufficient conditions for signals to have frequency spectra and argue that frequency spectra of most practical signals are well-defined, bounded, and continuous. We show that frequency spectra of signals reveal explicitly the distribution of energy in frequencies. The concept is essential in discussing many topics in later chapters.
5. *The sampling theorem and spectral computation:* Frequency spectra of most, if not all, practical signals cannot be expressed in closed form and cannot be computed analytically. The only way to compute them is numerically from their time samples. Thus, we establish the relationship between the frequency spectra of a CT signal and its sampled DT sequence and then present the *Nyquist sampling theorem*. We then introduce the *fast Fourier transform* (FFT) to compute frequency spectra of DT and CT signals.
6. *CT system qualitative analysis:* We introduce *transfer functions* through the Laplace transform and give reasons for studying only *proper rational* transfer functions. We introduce *poles* and *zeros* and use them to develop general responses of systems. We then introduce the concepts of *frequency response* and *stability*[3] for systems and establish the equation

$$Y(j\omega) = H(j\omega)U(j\omega) \tag{P.1}$$

where $Y(j\omega)$ is the output's frequency spectrum, $U(j\omega)$ is the input's frequency spectrum, and $H(j\omega)$ is a system's frequency response. We show that the equation has physical meaning only if the system is stable. We discuss how to compute the Fourier transform using the Laplace transform and show that the phasor analysis discussed in most network texts is applicable only to stable systems. We give reasons for *not* discussing in this text the Fourier analysis of systems even though it is discussed in most other texts. We also give reasons for *not* using transfer functions in computer computation.
7. *CT system quantitative analysis:* For computer computation, we transform transfer functions into *state-space* (ss) equations, called the *realization* problem. The name "realization" is justified by the fact that every ss equation can be readily simulated on a computer and implemented using an op-amp circuit. Thus ss equations are more convenient for computer computation and synthesis, whereas transfer functions are more convenient for qualitative

[3] Every passive RLC network is stable, and its stability study is unnecessary. However, a circuit that contains active elements such as op amps or a computer program can easily become unstable. Thus its stability study is imperative.

analysis and design. The analytical study of ss equations is not discussed because it plays no role in the applications mentioned above. We also compare the two descriptions and justify the use of transfer functions in disregarding zero-input responses of systems. We then discuss an identification scheme to develop a more realistic model for op amps and caution the use of *linear sweep sinusoids* in identification.

8. *Applications:* This chapter introduces three independent topics. The first topic is basic in all engineering disciplines, the second topic is basic in control and electronics, and the last topic is basic in communication:
   - **(a)** *Model reductions:* Amplifiers, seismometers, and accelerometers are all based on reduced models. Thus, model reduction is widely used in practice. We use (P.1) to develop *operational frequency ranges* for devices and demonstrate that a device will yield the intended result only if the frequency spectra of input signals lie inside the operational frequency range of the device. Note that seismometers and accelerometers have transfer functions of the same form, but they have different operational frequency ranges and, consequently, different reduced models and different design requirements. We also show that the accelerometers used in automobiles to trigger airbags are much more complex than the simple model discussed in this or any other similar text and cannot be easily analyzed. Thus engineering is more than mathematics and physics; it needs innovation, the construction of prototypes and their repetitive testings and improvements, and years of development.
   - **(b)** *Feedback:* We discuss the *loading problem* in connecting two systems. We then use an example to demonstrate the main reason for using feedback: reduction of the effects of parameter variations. Feedback, however, introduces the stability problem. We show that the stability of a negative or positive feedback system is independent of the stability of its subsystems. We also use feedback to implement *inverse systems* and discuss its limitations. We then design Wien-bridge oscillators—directly and by using a feedback model—and relate the Barkhausen criterion to the pole condition.
   - **(c)** *Modulations:* We introduce two modulation schemes and show the role of (P.1) in their demodulations.
9. *DT system qualitative analysis:* We introduce the z-transform, DT transfer functions, stability, and frequency responses.
10. *DT system quantitative analysis:* We develop DT state-space equations directly from high-order difference equations. The procedure has no CT counterpart. The remainder of this chapter closely follows the CT case.

Most results in this text can be obtained by typing a small number of MATLAB functions.[4] Thus, it is more important than ever to understand basic ideas and procedures. Most exercises in the text are designed to check understanding of the topic involved and require minimal numerical computation. Thus the reader should solve the exercises before proceeding to the next topic. We also suggest that the reader solve the problems at the end of each chapter by hand and then

[4] All results in this text are obtained using MATLAB 5.3 Student Version.

verify the results using a computer. In addition, we recommend that the reader repeat all of the programs in the text. The programs will yield only essential results because the programs skip nonessential functions such as the sizing of figures and the drawing of coordinates.

The logical sequence of the chapters is as follows:

$$\text{Chapters 1--4} \Rightarrow \begin{cases} \text{Sections 5.1--5.3} \Rightarrow \text{Sections 5.4--5.6} \\ \text{Chapter 6} \Rightarrow \begin{cases} \text{Chapter 7} \\ \text{Sections 8.1--8.3} \\ \left.\begin{array}{l}\text{Section 8.4}\\ \text{Section 8.5}\end{array}\right\} \Rightarrow \text{Section 8.6} \\ \text{Section 8.7--8.8} \end{cases} \\ \text{Chapter 9} \Rightarrow \text{Chapter 10} \end{cases}$$

This text contains more material than can be covered in one semester. A one-semester sophomore/junior course at Stony Brook covers (skipping the asterisked sections) Chapters 1 through 4, Sections 5.1 through 5.3, Chapters 6 and 7, and parts of Chapter 8. Clearly, other arrangements are also possible. A solutions manual is available from the publisher.

Many people helped me in writing this text. Mr. Anthony Olivo performed many op-amp circuit experiments for me. I consulted Professors Armen Zemanian and John Murray whenever I had any questions or doubts. The first draft of this revision was reviewed by a number of reviewers. Their comments prompted me to rearrange and rewrite many sections. Many people at Oxford University Press, including Peter C. Gordon, Danielle Christensen, Barbara Brown, Trent Haywood, Mary Hopkins, and Brian Kinsey, were most helpful in this undertaking. I thank them all. Finally, I thank my wife, Bih-Jau, for her support.

A NOTE TO THE READER

When I was an undergraduate student about forty-five years ago, I did every single one of the assigned problems and was an "A" student. I believed that I understood most subjects well. This belief was reinforced by my passing a competitive entrance exam to a master's-degree program in Taiwan. Again I completed the degree with flying colors and was confident for my next challenge.

My confidence was completely shattered when I started to do research under Professor Charles A. Desoer at the University of California, Berkeley. Under his critical and constant questioning, I realized that I did not understand the subject of study at all. More important, I discovered that my method of studying had been incorrect: learning only the mechanics of solving problems without learning the underlying concepts. From that time on, whenever I studied a topic, I would ponder every statement carefully and then study its implications. Are the implications still valid if some word in the statement is missing? Why? After some thought, I would re-read the topic or article. It often took me several iterations of pondering and re-reading to fully grasp certain ideas and results. I also learned to construct simple examples to gain insight and, by keeping in mind the goals of a study, to differentiate between what is essential and what is secondary or not important. *It takes a great deal of time and thought to really understand a subject*.

As a consequence of my Ph.D. training, I became fascinated by mathematics for its absolute correctness and rigorous development—no ambiguity and no approximation. Thus in the early part of my teaching career, I tended to teach more mathematics than engineering, as is evident in the fact that I expanded the book *Linear System Theory and Design* from 431 to 662 pages between its first and second editions. With the information explosion, a widening gap between theory and practice, less-prepared students, and our limited time and energy, I realized in the middle of my teaching career that it is unnecessary to burden engineering students, especially non-Ph.D. students, with too much mathematics. Thus I cut the third edition of the aforementioned book in half, from 662 to 334 pages, by deleting those topics that were mostly of academic interest and seemed to have no practical application. In the same spirit, my intention with *Signals and Systems, 3rd edition,* has been to develop a focused and concise text on signals and systems that still contains all material that an undergraduate student or a practicing engineer needs to know. I hope I have succeeded in this endeavor.

Students taking a course on signals and systems usually take three or four other courses at the same time. They may also have many distractions: part-time jobs, relationships, or the Internet. They simply do not have the time to really ponder a topic. Thus I fully sympathize with their lack of understanding. When students come to my office to ask questions, I always insist that they try to solve the problems themselves by going back to the original definitions and then by developing the answers step by step. Most of the time, the students discover that the questions were not difficult at all. Thus if the reader finds a topic difficult, he or she should go back to think about the basic definitions and then follow the steps logically. Do not get discouraged. We suggest that the reader refer occasionally to the preface to keep in mind the goals of the study. It is hoped that after finishing this book, the reader will be comfortable with all the italicized terms in the preface.

Chi-Tsong Chen
October 2003

CONTENTS

May be omitted without loss of continuity.

CHAPTER 1 Signals

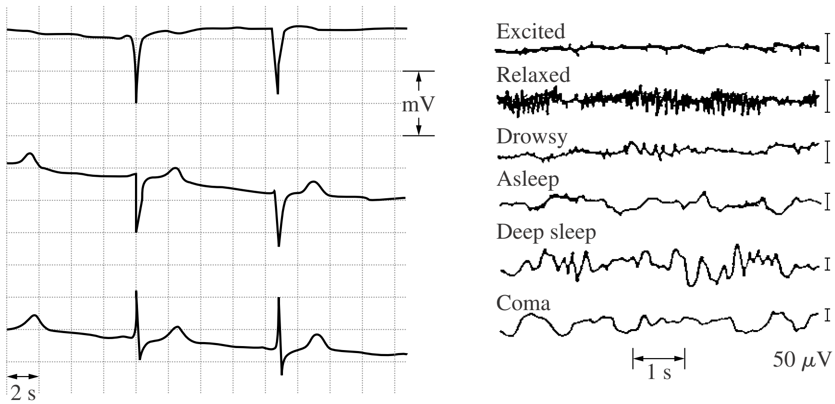
1.1 INTRODUCTION

This text studies signals and systems. We encounter both of them daily almost everywhere. When using a telephone, your voice, an acoustic signal, is transformed by the microphone, a system, into an electrical signal. That electrical signal is transmitted, maybe through a pair of copper wires or a satellite circulating around the earth, to the other party and then transformed back, using a loudspeaker, another system, into your voice. On its way, the signal may have been processed many times by many different systems. In addition to their role in communications, signals are used in medical diagnoses and in detecting an object, such as an airplane or a submarine. For example, the electrocardiogram (EKG) shown in Figure 1.1(a) and the brain waves (EEG) shown in Figure 1.1(b) can be used to determine the heart condition and the state of mind of the patient. Figures 1.2(a) and 1.2(b) show another type of signal: the total number of shares traded and the closing price of a stock at the end of each trading day in the New York Stock Exchange. Other examples of signals are the U.S. gross domestic product (GDP), consumer price index, and unemployment rate. We also use signals to control our environment and to transfer energy. To control the temperature of a home, we may set the thermostat at 20°C in the daytime and, to save energy, 15°C in the evening. Such a control signal is shown in Figure 1.3(a). Electricity is delivered to our home in the sinusoidal waveform shown in Figure 1.3(b), which in the United States has peak magnitude of $110 \times \sqrt{2}$ volts and frequency of 60 hertz (Hz, cycles per second).

All aforementioned signals depend on one independent variable—namely, time—and are called one-dimensional signals. Pictures are signals with two independent variables; they depend on the horizontal and vertical positions and are called two-dimensional signals. Temperature, wind speed, and air pressure are four-dimensional signals because they depend on the geographical location (longitude and latitude), altitude, and time. If we study the temperature at a fixed location, then the temperature becomes a one-dimensional signal, a function of time. This text studies only one-dimensional signals, and the independent variable is time. No complex-valued signals can arise in the real world; thus we study only real-valued signals.

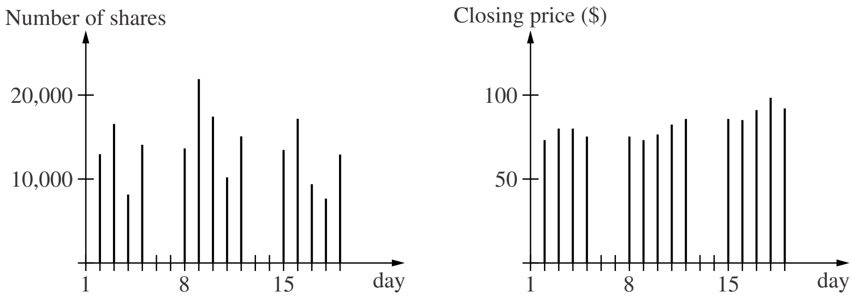
1.2 CONTINUOUS-TIME (CT), DISCRETE-TIME (DT), AND DIGITAL SIGNALS

A signal is called a continuous-time (CT) signal if it is defined at every time instant in a time interval of interest and its amplitude can assume any value in a continuous



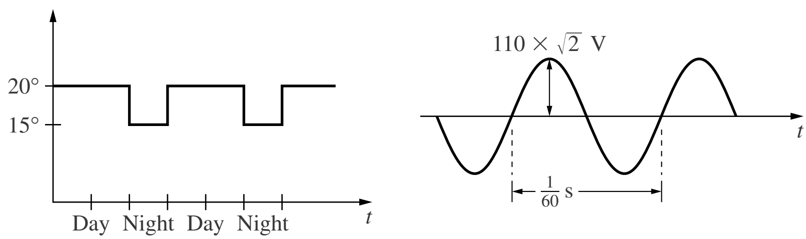
(a) (b)

Figure 1.1 Continuous-time signals.



(a) (b)

Figure 1.2 Digital signals.



(a) (b)

Figure 1.3 (**a**) Control signal. (**b**) Household electric voltage.

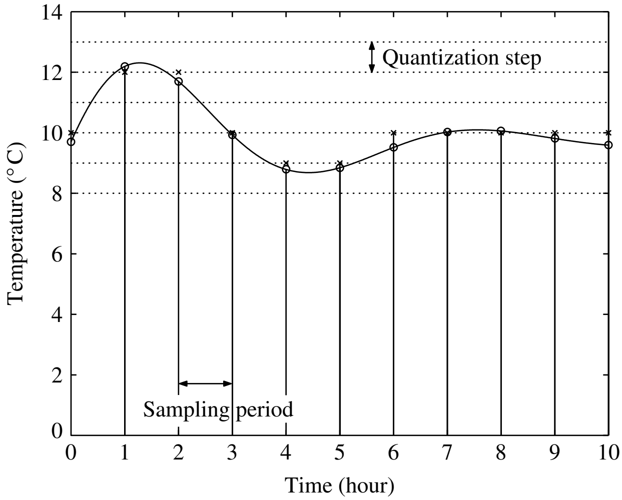


Figure 1.4 Sampling and quantization.

range.[1] All signals in Figures 1.1 and 1.3 are CT signals. A CT signal is also called an *analog* signal because its waveform is often analogous to that of the physical variable. Note that a CT signal is not necessarily a continuous function of time as shown in Figure 1.3(a); it can be discontinuous.

A signal is called a discrete-time (DT) signal if it is defined only at discrete time instants, and its amplitude can assume any value in a continuous range. It is called a *digital* signal if its amplitude can assume a value only from a given finite set. The signals in Figure 1.2 are digital signals because they are defined at the closing time of each trading day, and the total number of shares and price are limited to integers and cents, respectively. Although some signals are inherently digital, most DT and digital signals are obtained from CT signals by *sampling* or *discretization in time* and *quantization in amplitude* as we discuss next.

Our world is an analog world, and most signals are analog signals. Suppose the temperature at Stony Brook is plotted in Figure 1.4 with a solid line. It is obtained by using a thermometer with analog readout. Clearly it is a CT signal. Now if we read and record the temperature on the hour, then the signal is as shown with small circles. It is called a sampled signal. The time instants at which we record the temperature are called sampling instants. The sampling instants need not be equally spaced. If they are, the time interval between two immediate sampling instants is called the *sampling period*. It is called the *sample time* in MATLAB. Thus *sampling a CT signal yields a DT signal*. We give one more example.

[1]There are infinitely many numbers in any nonzero range—for example, in $[-1, 1]$ or $[0.999, 1]$.

EXAMPLE 1.2.1

Consider the signal $x(t) = 2e^{-0.3t} \sin t$, for all $t \geq 0$. The signal is a CT signal as shown in Figures 1.5(a) and 1.5(b) with solid lines. If we take its samples with sampling period T, then the signal becomes[2]

$$x[n] := x(nT) = x(t)|_{t=nT} = 2e^{-0.3nT} \sin nT$$

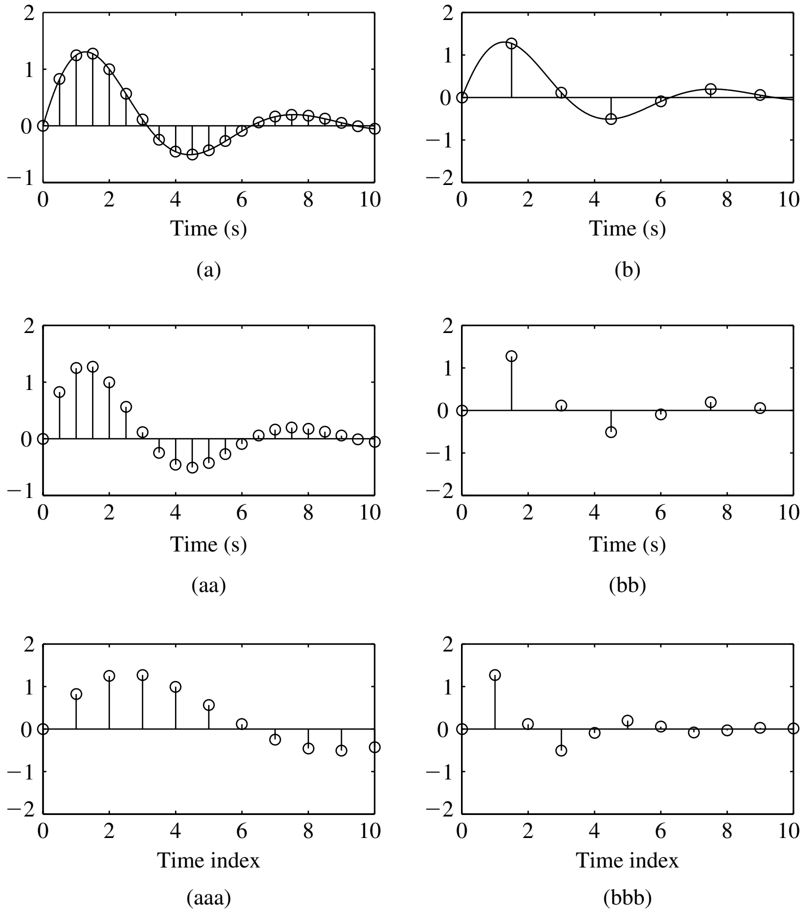


Figure 1.5 (**a**) CT signal and its sampled signal with sampling period $T = 0.5$. (**aa**) Plot of the sampled sequence in (a) against time. (**aaa**) Plot of (aa) against time index. (**b**), (**bb**), and (**bbb**) Repetitions of (a), (aa), and (aaa) with $T = 1.5$.

[2]We use $A := B$ to denote that A, by definition, equals B. We use $A =: B$ to denote that B, by definition, equals A.

where n is called the *time index* and can assume only integers. We use the convention that the variable inside a pair of brackets must be an integer. The signal $x[n] = x(nT)$ is defined only at discrete time instants nT and is a DT signal. For example, if $T_1 = 0.5$, then the DT signal is

$$x_1[n] = x(nT_1) = 2e^{-0.3nT_1} \sin nT_1 = 2e^{-0.15n} \sin 0.5n$$

for $n = 0, 1, 2, \ldots$. By direct substitution, we can compute $x_1[0] = 0$, $x_1[1] = 0.08253$, and so forth. The DT signal is plotted in Figure 1.5(aa) with time as its horizontal ordinate and in Figure 1.5(aaa) with time index as its horizontal ordinate. The signal consists of the sequence of numbers

$$\begin{array}{cccccc} 0 & 0.08253 & 1.2468 & 1.2721 & 0.9981 & \cdots \\ \uparrow & & & & & \end{array}$$

that appears at $n = 0, 1, 2, 3, 4, \ldots$, or at the discrete time instants $t = 0, 0.5, 1, 1.5, 2, \ldots$. Thus a DT signal is also called a time sequence. Note the use of an uparrow to indicate the time index $n = 0$. An uparrow will not be shown if the first entry occurs at $n = 0$.

For a different sampling period, we will obtain a different DT signal. For example, if $T_2 = 1.5$, then we have

$$x_2[n] = x(nT_2) = 2e^{-0.3nT_2} \sin nT_2 = 2e^{-0.45n} \sin 1.5n$$

This DT signal, as shown in Figures 1.5(b), 1.5(bb), and 1.5(bbb), consists of the sequence of numbers

$$0 \quad 1.2721 \quad 0.1148 \quad -0.5068 \quad -0.0924 \quad \cdots$$

that appears at $n = 0, 1, 2, 3, 4, \ldots$, or at $t = 0, 1.5, 3, 4.5, 5, \ldots$. This is a different time sequence.

If a DT signal is plotted with respect to time, then we can detect roughly the waveform of the original CT signal as in Figures 1.5(aa) and 1.5(bb). If it is plotted with respect to time index, it will be difficult to visualize the original CT signal as shown in Figures 1.5(aaa) and 1.5(bbb). It is important to remember that associated with every time sequence, there is an underlying sampling period.

If the temperature at Stony Brook is measured with a digital thermometer with integer readout, then the measured temperature can assume only integers as denoted by the horizontal dotted lines in Figure 1.4. These lines are called the *quantization levels*. The value between immediate quantization levels is called the *quantization step*. If a measured temperature does not fall exactly on a level, then it must be rounded to the nearest level as shown with crosses in Figure 1.4. This is called *quantization*. Thus *quantizing a DT signal yields a digital signal*. A digital computer can take only sequences of numbers; their values are limited by the number of bits used. Thus all signals processed on digital computers are digital signals. Although the digital signals in Figure 1.2 are not sampled or quantized from CT signals, we still can say that their sampling period is 24 hours and their quantization levels are respectively 1 and 0.01.

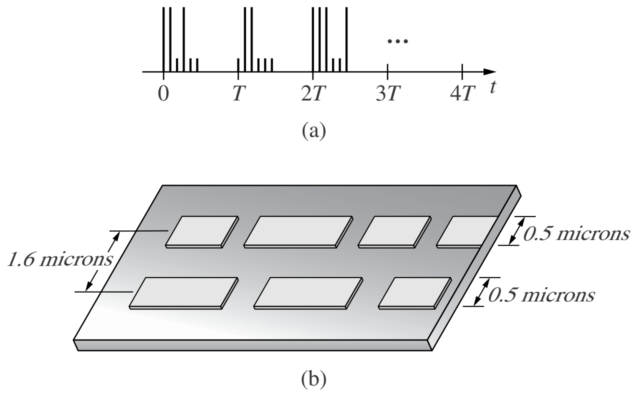


Figure 1.6 Representations of digital signals.

In processing and transmission, digital signals are often coded using a string of ones and zeros as shown in Figure 1.6(a) or

$$\begin{array}{cccccc} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 \\ \text{sign bit} & 2^1 & 2^0 & 2^{-1} & 2^{-2} & 2^{-3} \end{array}$$

where a_i can assume only 1 or 0 and is called a binary digit or bit. The set has six bits and a total of $2^6 = 64$ quantization levels. The quantization step is $2^{-3} = 0.125$. On a digital computer with a 5.5-volt power supply, one may be represented by a voltage ranging from 3 to 5 volts and zero from 0 to 1 volt. A compact disc (CD) has a spiral track separated by 1.6 microns (1.6×10^{-6} meters). On the track, zeros are represented by elongated bumps of width 0.5 micron and height 0.125 micron as shown in Figure 1.6(b). Sounds or data are encoded with extra bits (called eight-fourteen modulation), so each gap or one is separated by at least two zeros. Digital signals so encoded are clearly much less susceptible to noise and power variation than are CT signals. Coupled with the flexibility and versatility of digital circuits and digital computers, digital signal processing (DSP) has become increasingly popular and important. For example, although human voices are CT signals, they are now being transmitted digitally as shown in Figure 1.7(a) to avoid cross-talk, distortion, and transmission noise. More generally, a CT signal can be processed digitally as shown in Figure 1.7(b). The *transducer* is a device that transforms a signal from one form to another. For example, a microphone is a transducer that transforms voice, an acoustic signal, into an electrical signal. The transformed signal may require some conditioning, such as amplification (if the signal level is too small to drive the next device) or filtering to eliminate the noise introduced by the transducer or to reduce the effect of the so-called frequency aliasing. The conditioned analog signal is then transformed into a digital signal using an analog-to-digital converter (ADC). After digital processing using a personal computer or DSP processor, the processed digital signal can be transformed back to an analog signal using a digital-to-analog converter (DAC). The analog signal can be displayed on a monitor, smoothed by passing through an analog lowpass filter, or amplified to drive a system such as a speaker. ADC and DAC are now widely available commercially. Thus the conversion of analog signals to digital signals and vice versa can now be easily carried out.

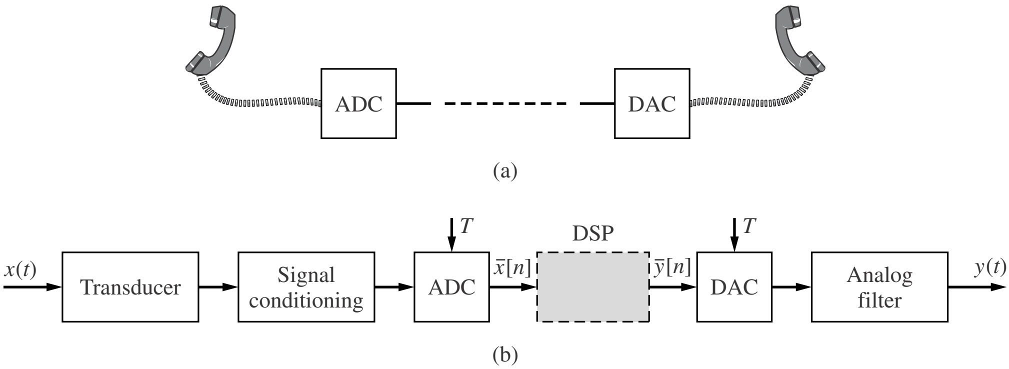


Figure 1.7 (**a**) Digital voice communication. (**b**) Digital processing of analog signals.

Although signals processed in computers, specialized digital hardware, and DSP processors are all digital signals, their analytical study is complicated. Clearly the operations (addition, subtraction, multiplication, and division) of binary-coded digital signals is more complex or at least less familiar to us than the operations of uncoded CT and DT signals. But the main reason is that quantization is not a linear operation. To simplify the discussion, we use decimal numbers to illustrate this point. Suppose every number is to be rounded to its nearest integer (that is, the quantization step is 1); then we have

$$Q(2.7 + 3.7) = Q(6.4) = 6 \neq Q(2.7) + Q(3.7) = 3 + 4 = 7$$

and

$$Q(2.7 \times 3.7) = Q(9.99) = 10 \neq Q(2.7) \times Q(3.7) = 3 \times 4 = 12$$

where Q stands for quantization. Because of these nonlinear phenomena, analytical study of digital signals is difficult. There are, however, no such problems in studying DT signals. For this reason, *in analysis and design, all digital signals will be considered as DT signals*. On the other hand, *all DT signals must be quantized to become digital signals in digital processing*. As we can see from Figure 1.4, there are errors between a DT signal and its corresponding digital signal. However, if the quantization step is small, such as 10^{-12} in general-purpose computers, the errors are very small and can often be ignored. If the quantization errors are not negligible such as in using specialized hardware, they are studied using statistical methods and its study is outside the scope of this text. In conclusion, this text studies only CT and DT signals as in every other text on signals and systems.

1.3 ELEMENTARY CT SIGNALS

A one-dimensional signal can be represented by $x(t)$. At every t, the signal is required to assume a unique value; otherwise the signal is not well defined. In mathematics, $x(t)$ is called a *function*. Thus there is no difference between signals and functions and they will be used interchangeably.

A signal $x(t)$ is called a *positive-time* signal if $x(t) = 0$ for $t < 0$, and it is called a *negative-time* signal if $x(t) = 0$ for $t > 0$. In other words, a positive-time signal can have nonzero values

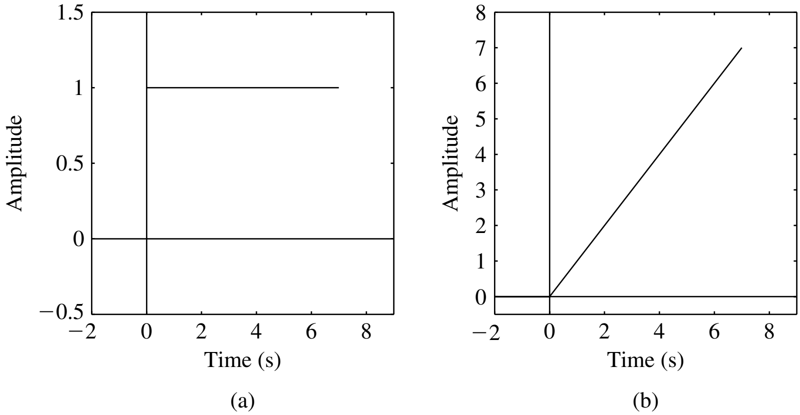


Figure 1.8 (**a**) Step function. (**b**) Ramp function.

only for $t \geq 0$, and a negative-time signal can have nonzero values only for $t \leq 0$.[3] If a signal has nonzero values for some $t < 0$ and some $t \geq 0$, then it is called a two-sided signal. Whether a signal is positive-time or negative-time depends on where $t = 0$ is. In practice, $t = 0$ can be selected as the time instant we start to generate, measure, or study a signal. Thus we encounter mostly positive-time signals.

Step Function The (unit) step function is defined as

$$q(t) = \begin{cases} 1 & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \tag{1.1}$$

and shown in Figure 1.8(a). The value of $q(t)$ at $t = 0$ is defined as 1. In fact, it can be defined as 0, 0.5, -10, 10^{10}, or any other value and will not affect our discussion. The information or energy of CT signals depends on not only amplitude but also *time duration*. An isolated t has zero width (zero time duration), and its value will not contribute any energy. Thus the value of $q(t)$ at $t = 0$ is immaterial.

The step function defined in (1.1) has amplitude 1 for $t \geq 0$. If it has amplitude -2, we write $-2q(t)$. We make distinction between *amplitude* and *magnitude*. Both are required to be real-valued. An amplitude can be negative, zero, or positive. A magnitude can be zero or positive but not negative.

Ramp Function The ramp function is defined as

$$r(t) = \begin{cases} t & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \tag{1.2}$$

---

[3]Positive-time signals are called *causal signals* in some texts. Causality is, as we will discuss in the next chapter, a property of systems and is used in this text to describe only systems.

and shown in Figure 1.8(b). It has slope 1 for $t \geq 0$. To generate a ramp function with slope a, all we do is to write $ar(t)$. The ramp function can be obtained from the step function as

$$r(t) = \int_{-\infty}^{t} q(\tau)\,d\tau = \int_{0_-}^{t} q(\tau)\,d\tau$$
$$= \int_{0}^{t} q(\tau)\,d\tau = \int_{0_+}^{t} q(\tau)\,d\tau \qquad (1.3)$$

where 0_- denotes a time instant infinitesimally smaller than 0, and 0_+ is infinitesimally larger than 0. We will obtain the same $r(t)$ no matter what value $q(0)$ is. This confirms our assertion that the value of $q(0)$ is immaterial.

Real Exponential Function—Time Constant Consider the exponential function

$$x(t) = e^{-at} \qquad (1.4)$$

for $t \geq 0$, where a is real and nonnegative (zero or positive). If a is negative, the function grows exponentially to infinity. No device can generate such a signal. If $a = 0$, the function reduces to the step function in (1.1). If $a > 0$, the function decreases exponentially to zero as t approaches ∞. We plot in Figure 1.9(a) two real exponential functions with $a = 2$ (solid line) and 0.1 (dotted line). The larger a is, the faster e^{-at} vanishes. In order to see e^{-2t} better, we replot them in Figure 1.9(b) using a different time scale.

A question we often ask in engineering is, At what time will e^{-at} become zero? Mathematically speaking, it becomes zero only at $t = \infty$. However, in engineering, the function can be considered to become zero when its magnitude is less than 1% of its peak magnitude. For example, in a scale with its reading from 0 to 100, a reading of 1 or less can be considered to be 0. We give an estimate for e^{-at}, with $a > 0$, to reach roughly zero. Let us define $t_c := 1/a$. It is called the *time constant*. Because

$$\frac{x(t+t_c)}{x(t)} = \frac{e^{-a(t+t_c)}}{e^{-at}} = e^{-at_c} = e^{-1} = 0.37$$

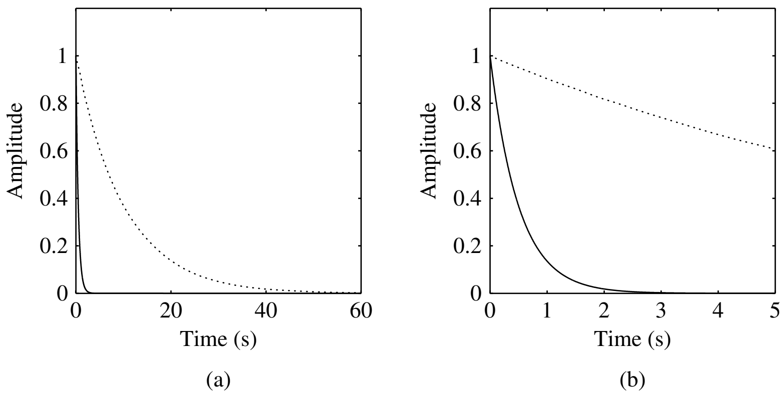


Figure 1.9 (**a**) The functions e^{-at} with $a = 0.1$ (dotted line) and $a = 2$ (solid line) for t in $[0, 60]$. (**b**) Replot of (a) for t in $[0, 5]$.

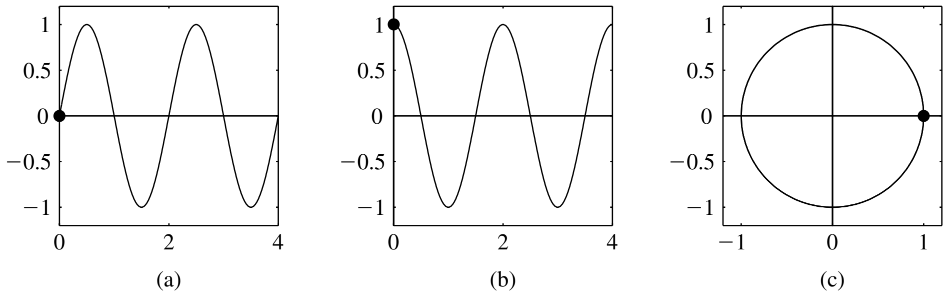


Figure 1.10 (**a**) $\sin \omega_0 t$. (**b**) $\cos \omega_0 t$. (**c**) $e^{j\omega_0 t}$.

the magnitude of e^{-at} decreases to 37% of its original magnitude whenever the time increases by one time constant $t_c = 1/a$. Because $(0.37)^5 = 0.007 = 0.7\%$, the magnitude of e^{-at} decreases to less than 1% of its original magnitude in five time constants. Thus we often consider e^{-at} to have reached zero in five time constants. For example, the time constant of $e^{-0.1t}$ is $1/0.1 = 10$ and $e^{-0.1t}$ reaches zero in $5 \times 10 = 50$ seconds as shown in Figure 1.9(a). The signal e^{-2t} has time constant $1/2 = 0.5$ and takes $5 \times 0.5 = 2.5$ seconds to reach zero as shown in Figure 1.9(b).

Sinusoidal Functions Consider the sinusoidal functions

$$\sin \omega_0 t \quad \text{and} \quad \cos \omega_0 t \tag{1.5}$$

shown in Figures 1.10(a) and 1.10(b). They can be considered as the projection respectively on the vertical and horizontal axes of the point A rotating, with a constant speed, counterclockwise around a unit circle shown in Figure 1.10(c). The time P, in seconds, in which they complete one cycle is called the *fundamental period* or, simply, period. The number of cycles in one second is called the *frequency,* expressed in hertz (Hz), denoted by $f_0 = 1/P$. Because one cycle has 2π radians, the frequency expressed in rad/s is $\omega_0 = 2\pi f_0 = 2\pi/P$.

1.4 MANIPULATIONS OF CT SIGNALS

This section discusses some manipulations of signals. The procedure is very simple. We select a number of t and then do direct substitution as we will illustrate with examples.

1.4.1 Shifting and Flipping

Given the signal $x(t)$ shown in Figure 1.11(a), we plot the signal defined by $x(t-1)$. To plot $x(t-1)$, we select a number of t and then find its values from the given $x(t)$. For example, if $t = -1$, then $x(t-1) = x(-2)$, which equals 0 as we can read from Figure 1.11(a). If $t = 0$, then $x(t-1) = x(-1)$, which equals 0. If $t = 1$, then $x(t-1) = x(0)$. Because $x(0_-) = 0$ and $x(0_+) = 1$, we consider separately $t = 1_-$ and $t = 1_+$, where a_- denotes a number infinitesimally smaller than a while a_+ denotes a number infinitesimally larger than a. Proceeding forward, we have the following:

t	-1	0	1_-	1_+	2	3	4
$x(t-1)$	0	0	0	1	1	0	0

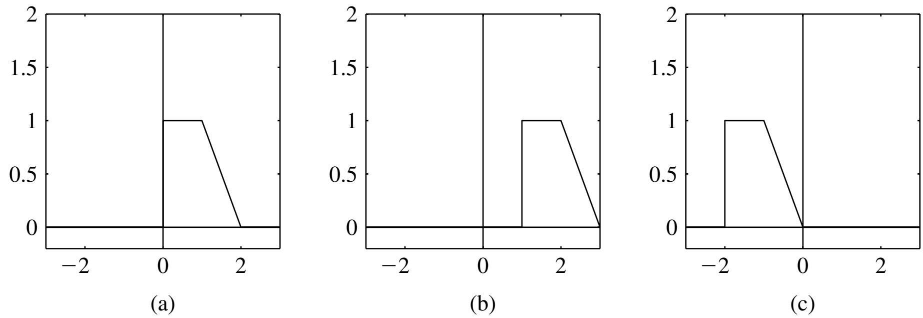


Figure 1.11 (**a**) A given function $x(t)$. (**b**) $x(t-1)$. (**c**) $x(t+2)$.

From these, we can plot in Figure 1.11(b) $x(t-1)$. Using the same procedure, we can plot $x(t+2)$ in Figure 1.11(c). We see that $x(t-t_0)$ simply shifts $x(t)$ from $t=0$ to t_0. That is, $x(t-t_0)$ shifts $x(t)$ to the right by t_0 seconds if $t_0>0$, and it shifts $x(t)$ to the left by $|t_0|$ seconds if $t_0<0$. Note that right shifting is time delay, whereas left shifting is time advance.

Time shifting, especially delay, is common in practice. Music from a compact disc (CD) or an audio tape is a delayed signal. It is not possible to advance a signal in *real time*. For example, no sound will appear before it is uttered. Taped signals, however, can be advanced or delayed with respect to a selected reference time $t=0$. Operations from tapes may be called non-real-time operations. Strictly speaking, advanced signals from tapes are still delayed signals from the time they were recorded.

Using the same procedure, we can plot $x(-t)$, $x(-t-1)$, and $x(-t+2)=x(2-t)$ in Figures 1.12(a) through 1.12(c) for the given $x(t)$ in Figure 1.11(a). We see that $x(-t)$ is the *flipping* of $x(t)$ with respect to $t=0$ and $x(-t+t_0)$ is the shifting of $x(-t)$ to t_0. That is, $x(-t+t_0)$ shifts $x(-t)$ to the right by t_0 seconds if $t_0>0$, and it shifts $x(-t)$ to the left by $|t_0|$ seconds if $t_0<0$. For example, $x(t+2)$ shifts $x(t)$ to the left by two seconds and $x(-t+2)$ shifts $x(-t)$ to the right by two seconds. Note also that if $x(t)$ is positive time, then $x(-t)$ is a negative-time signal.

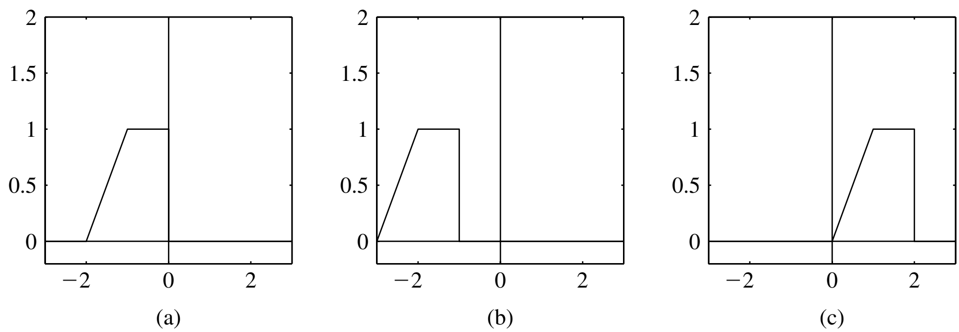


Figure 1.12 (**a**) $x(-t)$ of Figure 1.11(a). (**b**) $x(-t-1)$. (**c**) $x(-t+2)$.

It is important to mention that if we are not sure whether a signal is shifted to the right or left, we can always select a number of t to verify it. After all, *the rules of shifting and flipping are developed from substituting*.

1.4.2 Multiplication and Addition

Consider two signals. Their multiplication and addition form two new signals. The multiplication and addition are to be carried out at every time instant. However, if the two given signals have some patterns, some computation can be skipped as the next example illustrates.

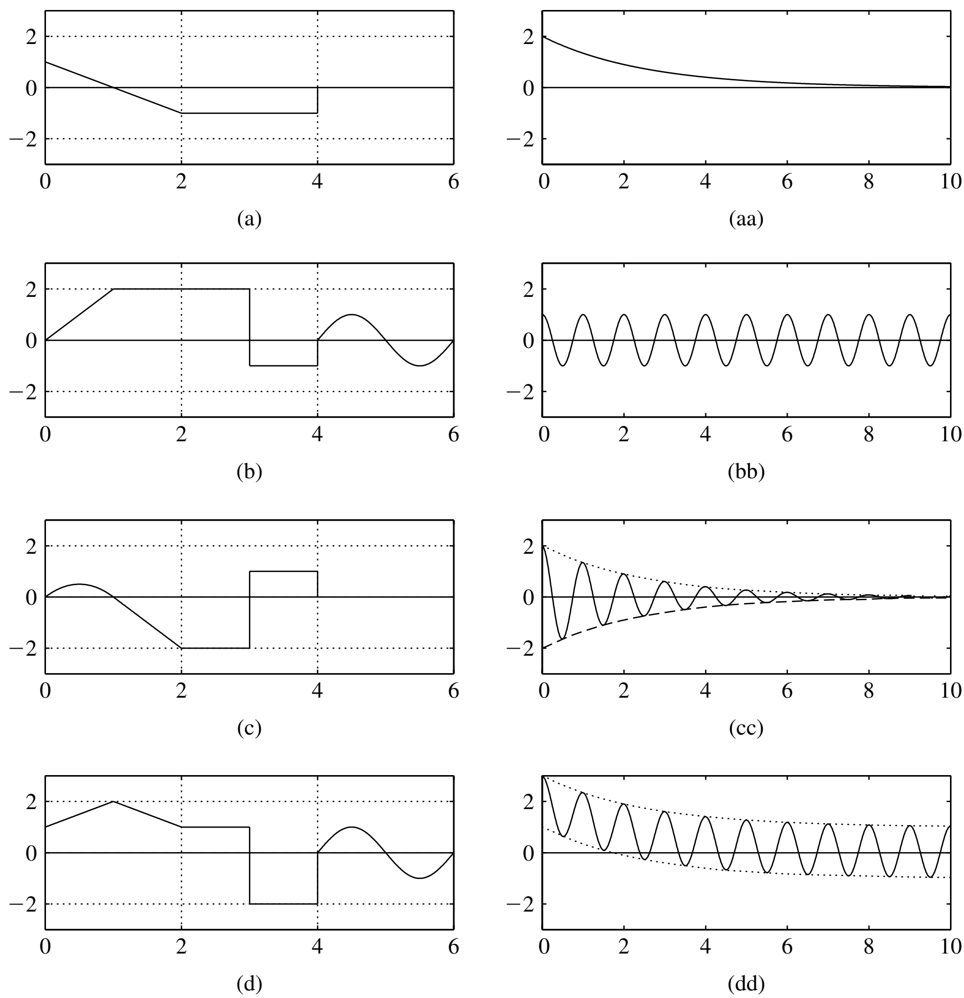


Figure 1.13 (**a**) and (**b**) Two given signals. (**c**) Their product. (**d**) Their sum. (**aa**) $x_1(t) = 2e^{-0.4t}$. (**bb**) $x_2(t) = \cos 2\pi t$. (**cc**) Their product. (**dd**) Their sum.

Consider the two signals $x_1(t)$ and $x_2(t)$ shown in Figures 1.13(a) and 1.13(b). We first list their values at $t = 0, 1, \ldots, 6$ in the following:

t	0	1	2	3_-	3_+	4_-	4_+	5	6
$x_1(t)$	1	0	−1	−1	−1	−1	0	0	0
$x_2(t)$	0	2	2	2	−1	−1	0	0	0
$x_1(t)x_2(t)$	0	0	−2	−2	1	1	0	0	0
$x_1(t)+x_2(t)$	1	2	1	1	−2	−2	0	0	0

Note that $x_1(t)$ is discontinuous at $t = 4$ and $x_2(t)$ is discontinuous at $t = 3$ and 4. Thus their values at 3_-, 3_+, 4_-, and 4_+ are specified. The products and sums of $x_1(t)$ and $x_2(t)$ are computed in the fourth and fifth rows. Next we consider the functions between these points. Consider the time interval [0, 1]. Both $x_1(t)$ and $x_2(t)$ are straight lines (polynomials of t with degree 1), thus their sum is also a straight line as shown in Figure 1.13(d) that is obtained by connecting 1 at $t = 0$ and 2 at $t = 1$. Their product, however, is a polynomial of t^2. Thus it is not a straight line; it is a curve connecting 0 at $t = 0$ to 0 at $t = 1$. For this example, we can obtain the curve (a polynomial of t with degree 2) analytically; but it is simpler to compute it directly at a number of points. For example, we compute $x_1(0.5)x_2(0.5) = 0.5 \times 1 = 0.5$ and then connect them as shown in Figure 1.13(c). Clearly the more points computed, the more accurate the plot. Next we consider the time interval [4, 6]. The function $x_1(t)x_2(t)$ is identically zero and the function $x_1(t) + x_2(t)$ equals $x_2(t)$ because $x_1(t)$ is identically zero in the interval. The rest can be similarly verified.

We plot in Figures 1.13(aa) and 1.13(bb) $x_1(t) = 2e^{-0.4t}$ and $x_2(t) = \cos 2\pi t$, and we plot their product in Figure 1.13(cc). The product equals zero where $x_2(t)$ equals zero and equals $x_1(t)$ where $x_2(t) = \cos 2\pi t = 1$. Its envelopes, because $|x_1(t)\cos 2\pi t| \le |x_1(t)|$, are governed by $\pm x_1(t)$ as plotted in Figure 1.13(cc) with dotted and dashed lines. Figure 1.13(dd) shows the addition of $2e^{-0.4t}$ and $\cos 2\pi t$. It simply shifts $\cos 2\pi t$ up by $2e^{-0.4t}$ at each t.

EXERCISE 1.4.1

Let $x(t)$ be a two-sided signal. Show that $x(t)q(t)$ is a positive-time signal and that $x(t)q(-t)$ is a negative-time signal. Does the equation

$$x(t)q(t) + x(t)q(-t) = x(t) \tag{1.6}$$

hold for all t? Can the equation be used?

Answers

No. It does not hold at $t = 0$. Yes.

1.4.3 Modulation

Consider a signal $x(t)$. Its multiplication by $\cos \omega_c t$ is called *modulation*. The signal $x(t)$ is called the *modulating signal;* $x(t)\cos \omega_c t$, *modulated signal;* $\cos \omega_c t$, *carrier signal;* and ω_c, *carrier frequency*. Signals sent out by radio or television stations are all modulated signals.

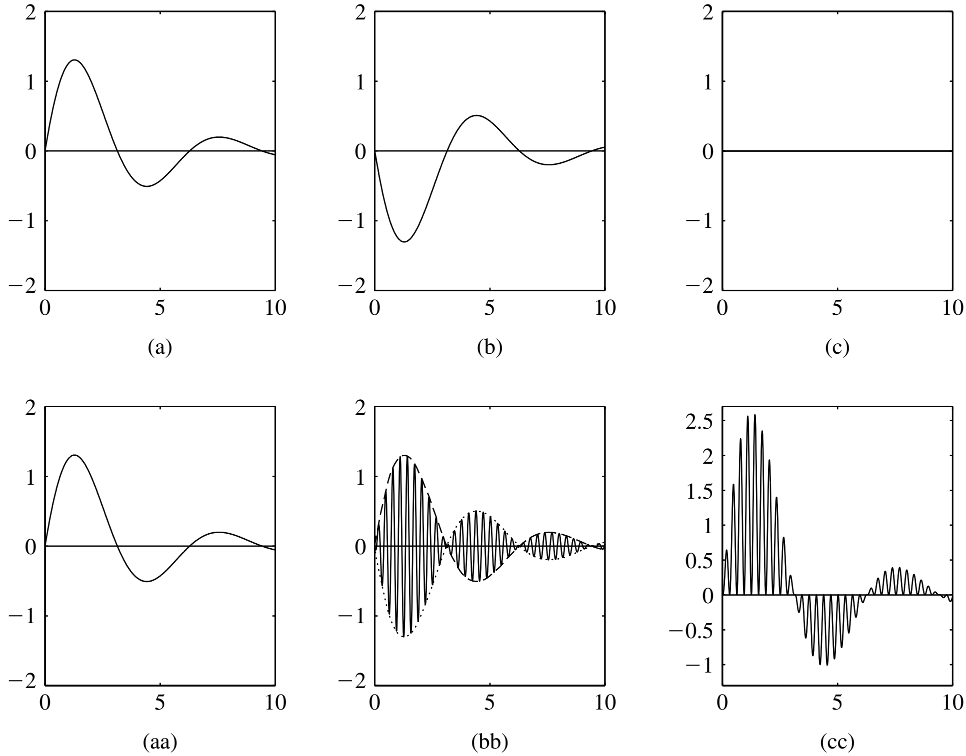


Figure 1.14 (**a**) $x_1(t) = 2e^{-0.3t}\sin t$. (**b**) $x_2(t) = -x_1(t)$. (**c**) $x_1(t) + x_2(t)$. (**aa**) $x_1(t)$. (**bb**) $x_{2m} = x_2(t)\cos 20t$. (**cc**) $x_1(t) + x_{2m}$.

Otherwise they will interfere with each other and cannot be recovered. We use a simple example to illustrate this fact. Consider the signal $x_1(t) = 2e^{-0.3t}\sin t$ shown in Figure 1.14(a). The signal $x_2(t)$ in Figure 1.14(b) equals $-x_1(t)$. Their sum is identically zero as shown in Figure 1.14(c), and there is no way to recover $x_1(t)$ and $x_2(t)$ from their sum. Figure 1.14(aa) shows $x_1(t)$ and Figure 1.14(bb) shows the modulated signal $x_2(t)\cos\omega_c t$ with $\omega_c = 20$. Their sum is shown in Figure 1.14(cc). If the carrier frequency ω_c is selected properly, it is possible to recover both $x_1(t)$ and $x_2(t)$ from their sum as we will demonstrate in Chapter 8.

1.4.4 Windows and Pulses

Consider the functions shown in Figures 1.15(a) and 1.15(b). The former is called a rectangular window $w_a(t)$ of length $2a$, and the latter a shifted window $w_L(t)$ of length L. The rectangular window $w_a(t)$ is a two-sided signal and is symmetric or even $[w_a(t) = w_a(-t)]$. The shifted window is a positive-time signal. Windows can be used to truncate signals.

We now express the rectangular window using the step function $q(t)$ defined in (1.1). The functions $q(t+a)$ and $q(t-a)$ shift $q(t)$ to $-a$ and a, respectively. Thus we have

$$w_a(t) = q(t+a) - q(t-a)$$

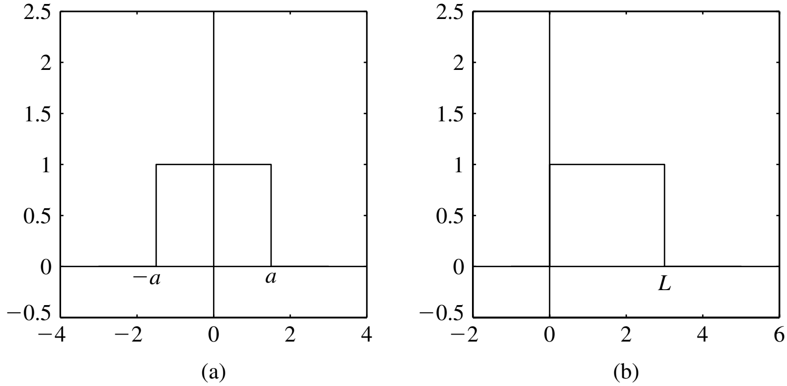


Figure 1.15 (**a**) Rectangular window of length $2a$. (**b**) Shifted window of length L.

It can also be expressed as

$$w_a(t) = q(t+a)q(-t+a)$$

Note that $q(t+a) = 0$ for $t < -a$, and $q(-t+a) = 0$ for $t > a$. Thus their product yields $\omega_a(t)$.

EXERCISE 1.4.2

Express the shifted window $w_L(t)$ in Figure 1.15(b) in terms of the window $w_a(t)$ in Figure 1.15(a). Also express it using $q(t)$.

Answers

$w_L(t) = w_a(t-a) = q(t) - q(t-L) = q(t)q(-t+L)$ with $a = L/2$

We show in Figure 1.16(a) a rectangular pulse with width a and height $1/a$, and we show in Figures 1.16(b) and 1.16(c) two triangular pulses with base width $2a$ and height $1/a$. They all have area 1 for all $a > 0$.

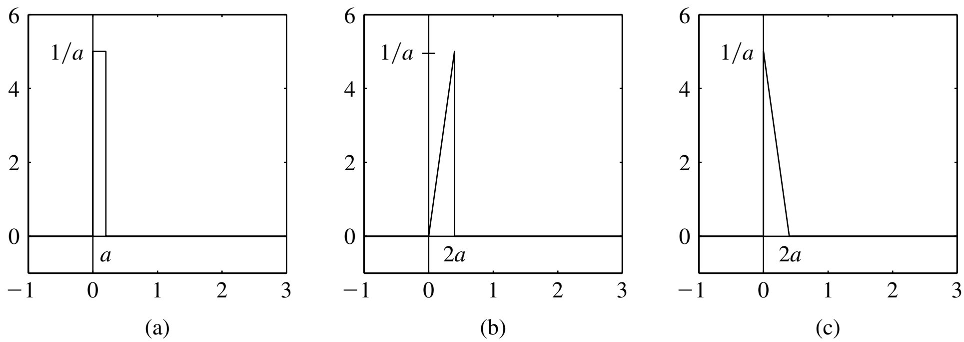


Figure 1.16 (**a**) Pulse with area 1. (**b**) and (**c**) Triangular pulses with area 1.

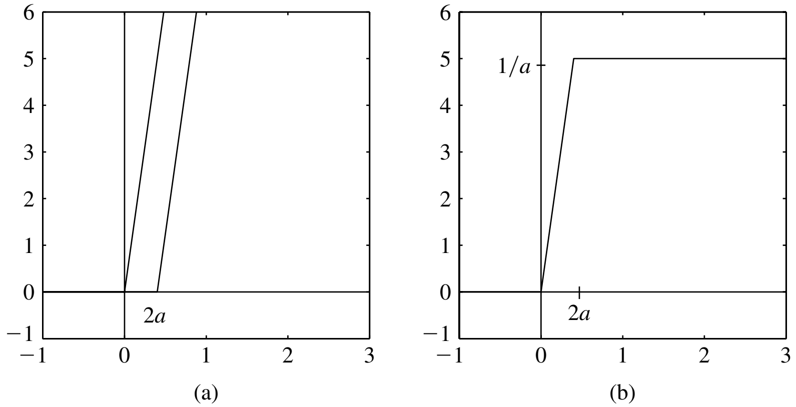


Figure 1.17 (**a**) Two ramp functions with slope $1/2a^2$. (**b**) Their difference.

In computer simulation such as in using SIMULINK,[4] we may have to generate the three pulses using the step and ramp functions defined in (1.1) and (1.2). The rectangular pulse can be easily decomposed as

$$\delta_a(t) = \frac{1}{a}[q(t) - q(t-a)] = \begin{cases} 1/a & \text{for } 0 \le t \le a \\ 0 & \text{for } t < 0 \text{ and } t > a \end{cases} \tag{1.7}$$

To generate the triangular pulse in Figure 1.16(b), we plot in Figure 1.17(a) $r(t)/2a^2$ and $r(t-2a)/2a^2$; they are ramp functions with slope $1/2a^2$. Their difference is shown in Figure 1.17(b); it consists of the triangular pulse in Figure 1.16(b) and a step function starting from $t = 2a$ and with amplitude $1/a$. Thus the triangular pulse in Figure 1.16(b) can be expressed as

$$\frac{r(t)}{2a^2} - \frac{r(t-2a)}{2a^2} - \frac{q(t-2a)}{a}$$

EXERCISE 1.4.3

Show that the triangular pulse in Figure 1.16(c) can be expressed as

$$\frac{1}{a}\left[q(t) - \frac{r(t)}{2a} + \frac{r(t-2a)}{2a}\right]$$

1.5 IMPULSE

Consider the pulse defined in Figure 1.16(a). It has width a and height $1/a$. As a decreases, the pulse becomes higher and narrower but its area remains to be 1. We define

$$\delta(t) := \lim_{a\to 0} \delta_a(t) \tag{1.8}$$

[4]SIMULINK is a registered trademark of The MathWorks, Inc.

It is called the Dirac delta function, δ-function, impulse function, or, simply, *impulse*. The impulse occurs at $t = 0$ and has the property

$$\delta(t) = \begin{cases} \infty & \text{if } t = 0 \\ 0 & \text{if } t \neq 0 \end{cases} \tag{1.9}$$

and

$$\int_{-\infty}^{\infty} \delta(t)\,dt = \int_{0}^{0_+} \delta(t)\,dt = 1 \tag{1.10}$$

Because $\delta(t)$ is zero everywhere except at $t = 0$, the integration interval from $-\infty$ to ∞ in (1.10) can be reduced to the immediate neighborhood of $t = 0$. The width from 0 to 0_+ is practically zero but still includes the whole impulse at $t = 0$. If an integration interval does not include the impulse, then the integration is 0 such as

$$\int_{-\infty}^{-1} \delta(t)\,dt = 0, \qquad \int_{1}^{10} \delta(t)\,dt = 0$$

If integration intervals touch the impulse such as

$$\int_{-1}^{0} \delta(t)\,dt, \qquad \int_{0}^{3} \delta(t)\,dt$$

then ambiguity may occur. If the impulse is defined from the pulse in Figure 1.16(a), then the former is 0 and the latter is 1. However, it is possible to define the impulse so that the former is 1 and the latter 0 or both 0.5. To avoid these situations, we assume that *whenever an integration touches an impulse, the integration covers the whole impulse and its integration equals 1.*

The impulse as defined is not an ordinary function such as the step function defined in (1.1). The value at any isolated t of an ordinary function does not contribute anything to an integration. For example, no matter what value $q(t)$ in (1.1) assumes at $t = 0$ such as 10^{10}, we still have

$$\int_{0_-}^{0_+} q(t)\,dt = 0$$

This is in contrast to (1.10). Thus the impulse is not an ordinary function; it is called a *generalized function*.

The impulse $\delta(t) = 1 \times \delta(t - 0)$ occurs at $t = 0$ and has area or *weight* 1. The impulse $\delta(t - t_0)$ is the impulse $\delta(t)$ shifted from $t = 0$ to $t = t_0$. Thus the impulse $a\delta(t - t_0)$, where a is a real constant, occurs at $t = t_0$ and has weight a. Note that a can be negative. The impulse has an important property that we will constantly use in the text. Let $x(t)$ be an ordinary function and continuous at t_0. Then we have

$$x(t)\delta(t - t_0) = x(t_0)\delta(t - t_0) \tag{1.11}$$

and

$$\begin{aligned} \int_{t=-\infty}^{\infty} x(t)\delta(t - t_0)\,dt &= x(t_0)\int_{-\infty}^{\infty} \delta(t - t_0)\,dt = x(t_0)\cdot 1 \\ &= x(t)|_{t=t_0} = x(t_0) \end{aligned} \tag{1.12}$$

Equation (1.11) follows directly from the property that $\delta(t - t_0)$ is 0 everywhere except at t_0, and (1.12) follows immediately from (1.11) and (1.10). The integration in (1.12) is called the *sifting* or *sampling* property of the impulse. Whenever an integrand is the product of a function and an impulse, we can move the function out of the integration and then replace the integration variable by the number or variable that makes the argument of the impulse 0. For example, we have

$$\begin{aligned}\int_0^{\infty} \sin(t-2)\delta(t-3)\,dt &= \sin(t-2)|_{t-3=0}\\ &= \sin(3-2) = \sin 1 = 0.84\end{aligned}$$

Thus $\delta(t-3)$ picks out or sifts the value of $\sin(t-2)$ at $t=3$. We also have

$$\int_{-\infty}^{\infty} x(t-\tau)\delta(\tau-3)\,d\tau = x(t-\tau)|_{\tau=3} = x(t-3)$$

and

$$\int_{-\infty}^{\infty} x(\tau)\delta(t-\tau-3)\,d\tau = x(\tau)|_{\tau=t-3} = x(t-3)$$

They shift the function $x(t)$ to $t=3$. Thus the sifting property can be used to achieve time shifting.

We discuss a relationship between impulses and step functions. Consider the function shown in Figure 1.18(a) and its derivative shown in Figure 1.18(b). Note that the function increases from 0 at time 0 to 1 at time $t=a$. Thus the slanted line has slope $1/a$ and its derivative is $1/a$. As $a \to 0$, the function in Figure 1.18(a) becomes the step function defined in (1.1) and the function in Figure 1.18(b) becomes an impulse. Thus we have

$$\delta(t) = \frac{dq(t)}{dt} \tag{1.13}$$

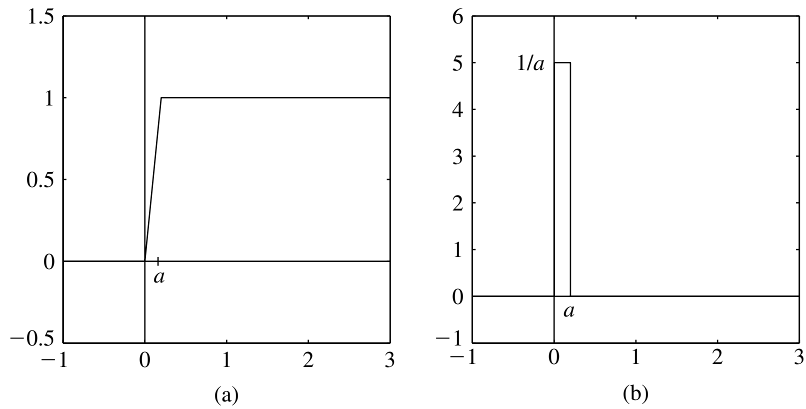


Figure 1.18 (**a**) A function. (**b**) Its derivative.

EXAMPLE 1.5.1

Consider the function

$$x(t) = q(t) + 0.5q(t-1) - 1.8q(t-2.5)$$

plotted in Figure 1.19(a). Its derivative is

$$\dot{x}(t) := \frac{dx(t)}{dt} = \delta(t) + 0.5\delta(t-1) - 1.8\delta(t-2.5)$$

and plotted in Figure 1.19(b). Each impulse is denoted by a hollow arrow with height equal to its weight. Note that its actual height is either ∞ or $-\infty$.

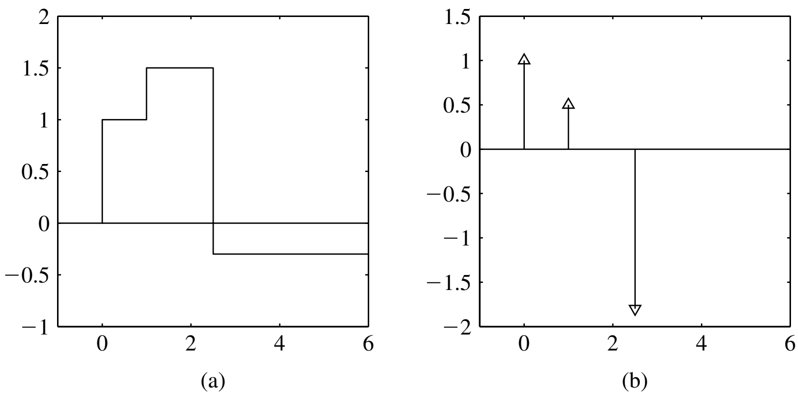


Figure 1.19 (**a**) Stepwise function. (**b**) Its derivative.

EXERCISE 1.5.1

Verify

$$q(t) = \int_{\tau=0}^{\infty} \delta(t-\tau)\,d\tau$$

and

$$x(t) = \int_{\tau=-\infty}^{\infty} \delta(t-\tau)\,d\tau = 1$$

for all t.

Mathematically speaking, we should not define the impulse as in (1.8). It should be defined using the two properties in (1.10) and (1.12). The two triangular pulses in Figures 1.16(b) and 1.16(c) have the two properties as $a \to 0$, and they can also be used to define the impulse. Thus the exact waveform of the impulse is immaterial so long as it has a very small time duration and a nonzero area. Thus impulses can be used to approximate an impact or collision. Striking a hammer generates an impulse with some nonzero weight.

1.5.1 Piecewise-Constant Approximation of CT Signals

A CT signal is often approximated by a piecewise-constant or stair-step function as shown in Figure 1.20. Let $x(t)$ be a CT signal as shown in Figure 1.10(a) with a dotted line, and let $x_a(t)$ be the stair-step function as shown with a solid line. We develop a mathematical expression for $x_a(t)$. First we modify (1.7) as $\delta_T(t - nT)$; this pulse occurs at $t = nT$ and has width T and height $1/T$. Clearly, $\delta_T(t - nT) \times T$ has height 1. The pulse starting at $t = nT$ in Figure 1.20(a) has width T and height $x(nT)$. Thus it can be expressed as $x(nT)\delta_T(t - nT)T$. This is a general expression applicable to every pulse in Figure 1.20(a). Thus the stair-step function in Figure 1.20(a) can be expressed as

$$x_a(t) = \sum_{n=0}^{\infty} x(nT)[\delta_T(t - nT)]T \tag{1.14}$$

This is a general expression and is valid for any $T > 0$.

We plot $x_a(t)$ in Figures 1.20(a) and 1.20(b) for $T = 0.3$ and 0.05. We see that the smaller T is, the better $x_a(t)$ approximates $x(t)$. We show that the approximation becomes exact as $T \to 0$. Let us define $\tau = nT$. As $T \to 0$, the variable τ becomes a continuum, T can be written as $d\tau$, the summation in (1.14) becomes an integration, and the pulse δ_T becomes an impulse. Thus as $T \to 0$, the expression in (1.14) becomes

$$x_a(t) = \int_0^{\infty} x(\tau)\delta(t - \tau)\, d\tau = x(t) \tag{1.15}$$

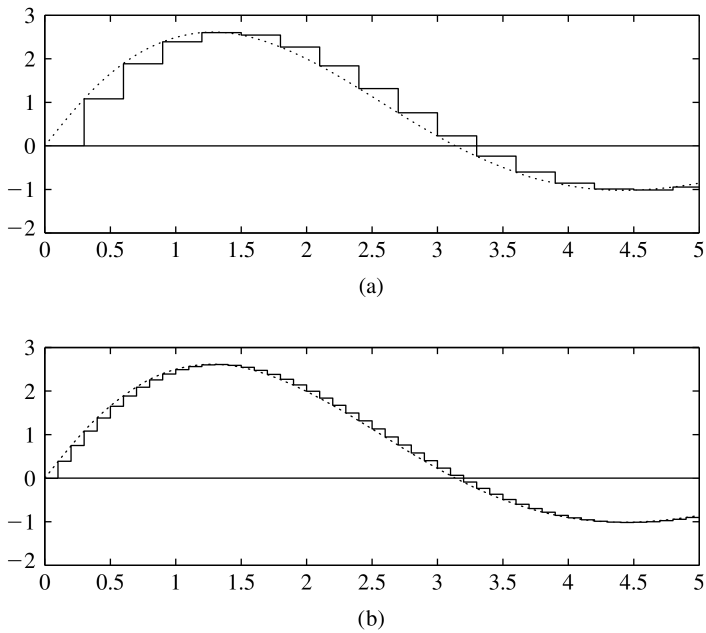


Figure 1.20 (**a**) Stair-step function with $T = 0.3$. (**b**) With $T = 0.05$.

The integration reduces to $x(t)$ by using the sifting property of impulses. Thus the stair-step function is a good and legitimate approximation. The expression in (1.14) will be used in Section 3.7.

A CT signal $x(t)$ cannot be stored in a computer because it has infinitely many points in any nonzero time interval. However, once it is approximated by a stair-step function, then the stair-step function can be easily stored. For example, for the stair-step function in Figure 1.20(a) with $T = 0.3$, the signal can be represented as

$$t = [0 \quad 0.3 \quad 0.6 \quad 0.9 \quad 1.2 \quad \cdots]$$

and

$$x = [0 \quad 1.08 \quad 1.89 \quad 2.39 \quad 2.60 \quad \cdots]$$

It is important that the lengths of the two arrays or vectors must be the same. This essentially stores the samples of $x(t)$ at sampling instants nT.

If time instants are equally spaced, then they can be more easily expressed in MATLAB as `t=0:T:tf`, which means that the time instants start from 0 with increment T up the increment after which the next increment is larger than the final time t_f. If the second argument T is missing, then the increment is 1. For example, $n = 0 : 5$ and $n = 0 : 5.3$ both mean $n = 0, 1, 2, 3, 4, 5$. This convention will be used throughout this text.

If a signal can be expressed in closed form such as $x(t) = 4e^{-0.3t} \sin t$ and if we select the sampling period as T and take N samples, then the signal can be stored in a computer as `n=0:N-1;t=n*T;x=4*exp(-0.3*t).*sin(t)`.

1.6 ELEMENTARY DT SIGNALS AND THEIR MANIPULATION

This section discusses the discrete-time (DT) counterparts of Sections 1.3 through 1.5. Consider the DT signal

$$x[n] := x(nT)$$

where T is the sampling period and n is the time index. As shown in Figure 1.5, the signal can be plotted against time or time index. The waveform and the frequency content of a DT signal clearly depend on T. However, in its processing or manipulation, the sampling period does not play any role. For example, when we process a DT signal (a sequence of numbers) on a computer, the processed signal will appear immediately on the monitor once the processing is completed. It clearly has nothing to do with the sampling period. In a real-time processing, once a value at time instant nT is received, the computer will start to compute using the values received at nT and earlier. Depending on the speed of the computer and the complexity of the program, it takes some finite time to complete the computation. The result is then stored in a memory location. A control logic will send it out at time instant $(n + 1)T$. In real-time processing, the only restriction is that the sampling period T be larger than the time needed for computation. Thus in processing or manipulation, we can suppress the sampling period and consider $x[n]$ without specifying T. For this reason, all DT signals in this section will be plotted with respect to time index n.

A DT signal $x[n]$ is a positive-time sequence if $x[n] = 0$ for $n < 0$, and it is a negative-time sequence if $x[n] = 0$ for $n > 0$. Otherwise it is called a two-sided sequence. In practice, $n = 0$ can be selected as the time instant we start to generate, measure, or study a DT signal. Thus we encounter in practice mostly positive-time sequences.

The operations of flipping, shifting, multiplication, and addition for CT signals are directly applicable to DT signals. Thus the discussion here will be brief. When we add or multiply two DT signals, their sampling periods are implicitly assumed to be the same. Otherwise, the operation is not defined.

Step Sequence The (unit) step sequence is defined as

$$q[n] = \begin{cases} 1 & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases} \tag{1.16}$$

as shown in Figure 1.21(a). Then $q[-n]$ is the flipping of $q[n]$ with respect to $n = 0$ as shown in Figure 1.21(b). The sequences $q[n - n_0]$ is the shifting of $q[n]$ to n_0, and $q[-n - n_0]$ is the shifting of $q[-n]$ to $-n_0$. They can be verified by direct substitutions.

Impulse Sequence The impulse sequence is defined as

$$\delta[n] = \begin{cases} 1 & \text{for } n = 0 \\ 0 & \text{for } n \neq 0 \end{cases}$$

It has only one nonzero entry at $n = 0$ and is zero elsewhere. It is also called the *Kronecker delta sequence* and is the DT counterpart of the CT impulse defined in Section 1.5. The sequence $\delta[n - k]$ is the shifting of $\delta[n]$ to $n = k$. Thus we have

$$\delta[n - k] = \begin{cases} 1 & \text{for } n = k \\ 0 & \text{for } n \neq k \end{cases} \tag{1.17}$$

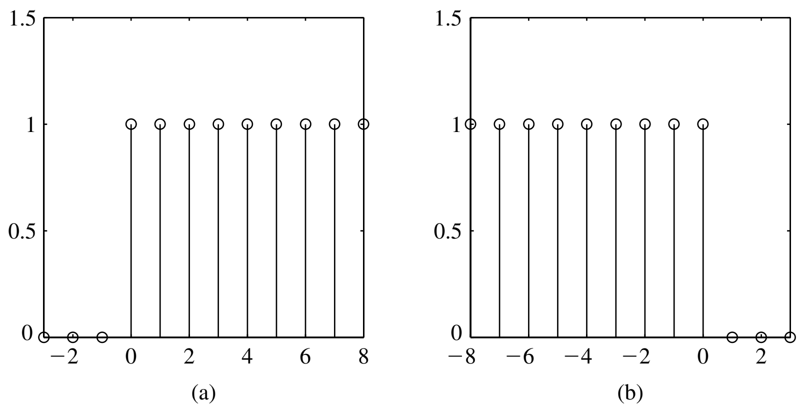


Figure 1.21 (**a**) Step sequence $q[n]$. (**b**) Flipped step sequence $q[-n]$.

Unlike the CT impulse which cannot be generated in practice, the impulse sequence can be easily generated.

Let $x[n]$ be a two-sided sequence. Then $x[n]q[n]$ is a positive-time sequence and $x[n]q[-n]$ is a negative-time sequence. Consider

$$x[n]q[n] + x[n]q[-n] = x[n] \tag{1.18}$$

The equality holds for all nonzero integers. At $n = 0$, we have $x[0]+x[0] \neq x[0]$. Equation (1.18) is the counterpart of (1.6) for CT signals. For CT signals, the value at any isolated time instant is immaterial and (1.6) can be used. For DT signals, the value at every sampling instant counts. Thus (1.18) cannot be used and must be modified as

$$x[n]q[n] + x[n]q[-n] - x[n]\delta[n] = x[n] \tag{1.19}$$

This equation holds for all integers n.

EXERCISE 1.6.1

Let N be a positive integer. Then the sequence

$$w_N[n] = \begin{cases} 1 & \text{for } -N \leq n \leq N \\ 0 & \text{for } n < -N \text{ and } n > N \end{cases} \tag{1.20}$$

is called a *rectangular window* of length $2N + 1$. Show

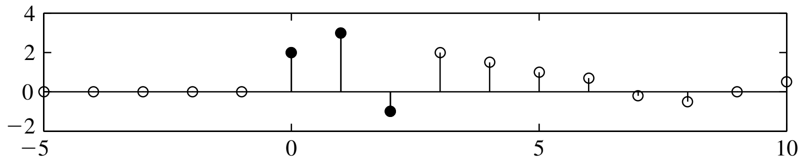
$$w_N[n] = q[n + N] - q[n - N - 1] = q[n + N]q[-n + N]$$

We next develop the DT counterpart of (1.15). Let $x[n]$ be a DT positive-time signal as shown in Figure 1.22(a). We decompose it as shown in Figures 1.22(b), 1.22(c), 1.22(d), and so forth. The sequence in Figure 1.22(b) is the impulse sequence $\delta[n]$ with height $x[0]$ or $x[0]\delta[n-0]$. The sequence in Figure 1.22(c) is the sequence $\delta[n]$ shifted to $n = 1$ with height $x[1]$ or $x[1]\delta[n-1]$. The sequence in Figure 1.22(d) is $x[2]\delta[n - 2]$. Thus the DT signal in Figure 1.22(a) can be expressed as

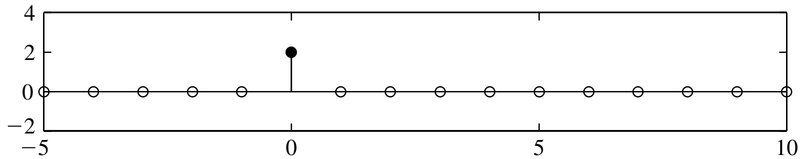
$$\begin{aligned} x[n] &= x[0]\delta[n-0] + x[1]\delta[n-1] + x[2]\delta[n-2] + \cdots \\ &= \sum_{k=0}^{\infty} x[k]\delta[n-k] \end{aligned} \tag{1.21}$$

This holds for every n. Note that in the summation, n is fixed and k ranges from 0 to ∞. For example, if $n = 10$, then (1.21) becomes

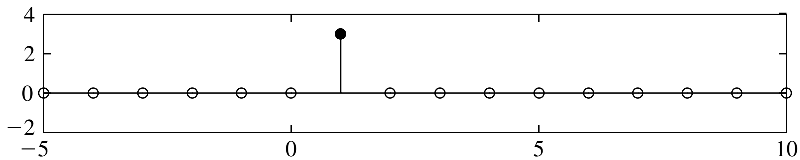
$$x[10] = \sum_{k=0}^{\infty} x[k]\delta[10 - k]$$



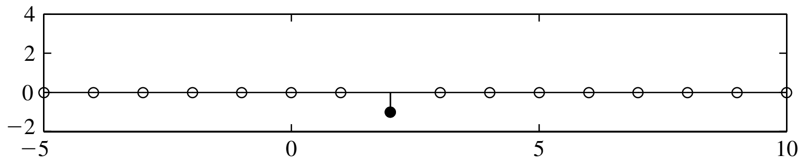
(a)



(b)



(c)



(d)

Figure 1.22 (**a**) DT signal $x[n]$. (**b**) $x[0]\delta[n]$. (**c**) $x[1]\delta[n-1]$. (**d**) $x[2]\delta[n-2]$.

As k ranges from 0 to ∞, every $\delta[10-k]$ is zero except $k = 10$. Thus the infinite summation reduces simply to $x[10]$ and the equality holds. We see that the derivation of (1.21) is much simpler than its CT counterpart in (1.15).

EXAMPLE 1.6.1

Consider a sequence defined by $x[0] = 1$, $x[1] = 0$, $x[2] = -2$, $x[3] = 1$, $x[4] = -3$. The sequence can be represented using two arrays or two vectors such as $n = [0\ \ 1\ \ 2\ \ 3\ \ 4]$, $x = [1\ \ 0\ \ -2\ \ 1\ \ -3]$. It can also be expressed as

$$x[n] = \delta[n] - 2\delta[n-2] + \delta[n-3] - 3\delta[n-4]$$

The former can be used as an input in a computer program, and the latter is a mathematical expression.

EXERCISE 1.6.2

Express the following sequences using two arrays and using impulse sequences:

(a) $x[-5] = 2, x[-3] = -4, x[0] = 1, x[1] = 2, x[3] = 4$.

(b) $x[n] = n + 2$, for $n = 0, 1, 3$, and $x[n] = 0$ otherwise.

Answers

(a) $n = -5 : 3$, $x = [2\ 0\ -4\ 0\ 0\ 1\ 2\ 0\ 4]$, or $n = [-5\ -3\ 0\ 1\ 3]$, $x = [2\ -4\ 1\ 2\ 4]$, $x[n] = 2\delta[n+5] - 4\delta[n+3] + \delta[n] + 2\delta[n-1] + 4\delta[n-3]$

(b) $n = [0\ 1\ 2\ 3]$, $x = [2\ 3\ 0\ 5]$, $x[n] = 2\delta[n] + 3\delta[n-1] + 5\delta[n-3]$

EXERCISE 1.6.3

Verify

$$q[n] = \sum_{k=0}^{\infty} \delta[n-k] \quad \text{and} \quad x[n] = \sum_{k=-\infty}^{\infty} \delta[n-k] = 1$$

for all n. They are the counterparts of the equations in Exercise 1.5.1.

Real Exponential Sequence Sampling the CT real exponential function $x(t) = e^{-at}$ with sampling period T yields $x(nT) = e^{-anT} =: b^n$ with $b := e^{-aT}$. Thus the real exponential sequence can be defined as

$$x[n] = b^n \tag{1.22}$$

for some real b. If $|b| > 1$, the sequence grows unbounded as $n \to \infty$. The sequence is a step sequence if $b = 1$, and it assumes 1 and -1 alternatively if $b = -1$. If $|b| < 1$, the sequence approaches zero *exponentially*.

For an exponentially decreasing sequence, we may define, following the CT case, its *time constant* as $t_c = -1/\ln|b|$, where ln stands for natural logarithm. Because $|b| < 1$, its natural logarithm is negative and the time constant is positive. Let us define

$$n_c := \text{round}(-5/\ln|b|)$$

where "round" rounds a number to its nearest integer. We claim that the sequence b^n, with $|b| < 1$, takes n_c samples for its magnitude to decrease to less than 1% of its original magnitude. We verify this by direct substitution. We compute b^{n_c} for various b:

b	0.99	0.95	0.9	0.8	0.5	0.3	0.1	0.05
n_c	497	97	47	22	7	4	2	2
b^{n_c}	0.0068	0.0069	0.0071	0.0074	0.0078	0.008	0.01	0.0025

We see that all b^{n_c} are less than 1% of its original magnitude $b^0 = 1$. Thus, as in the CT case, the magnitude of an exponentially decreasing sequence decreases to less than 1% of its peak magnitude in n_c samples or $n_c T$ seconds, roughly five time constants.

1.7 CT SINUSOIDAL SIGNALS

A CT signal[5] $x(t)$ is said to be periodic with period P if

$$x(t) = x(t+P) \tag{1.23}$$

for all t. It means that the signal repeats itself every P seconds. If (1.23) holds, then

$$x(t) = x(t+P) = x(t+2P) = \cdots = x(t+kP)$$

for every positive integer k. Thus if $x(t)$ is periodic with period P, it is also periodic with period $2P, 3P, \ldots$. The smallest such P is called the *fundamental period* or, simply, the *period* unless stated otherwise. The frequency $2\pi/P$ (in rad/s) or $1/P$ (in Hz) is called the *fundamental frequency* of the periodic signal. The simplest CT periodic signals are sine and cosine functions. The sinusoids $\sin\omega_0 t$ and $\cos\omega_0 t$ have fundamental frequency ω_0 and fundamental period $2\pi/\omega_0$.

Consider two periodic signals $x_i(t)$ with period P_i, for $i = 1, 2$. Is their linear combination $x(t) = a_1 x_1(t) + a_2 x_2(t)$, for any constants a_i, periodic? The condition for $x(t)$ to be periodic is that $x_1(t)$ and $x_2(t)$ have a common period or there exist integers k_1 and k_2 such that

$$k_1 P_1 = k_2 P_2 \quad \text{or} \quad \frac{P_1}{P_2} = \frac{k_2}{k_1}$$

Thus $x(t)$ is periodic if and only if P_1/P_2 is a rational number, a ratio of two integers.

EXAMPLE 1.7.1

The sinusoid $\sin 2t$ is periodic with period $P_1 = 2\pi/2 = \pi$. The sinusoid $\sin\pi t$ is periodic with period $P_2 = 2\pi/\pi = 2$. Because $P_1/P_2 = \pi/2$ is not a rational number, $\sin 2t + \sin\pi t$ is not a periodic signal. In other words, the signal will never repeat itself or its period is infinity.

Consider the function

$$x(t) = -2 + 3\sin 0.6t - 4\cos 0.6t - 2.4\cos 2.7t \tag{1.24}$$

The sinusoids $\sin 0.6t$ and $\cos 0.6t$ are periodic with period $P_1 = 2\pi/0.6$. Their linear combination $x_1(t) := 3\sin 0.6t - 4\cos 0.6t$ is clearly also periodic with period P_1. The sinusoid $x_2(t) := 2.4\cos 2.7t$ is periodic with period $P_2 = 2\pi/2.7$. Because

$$\frac{P_1}{P_2} = \frac{2\pi/0.6}{2\pi/2.7} = \frac{2.7}{0.6} = \frac{27}{6} = \frac{9}{2} =: \frac{\bar{k}_2}{\bar{k}_1}$$

[5]The study of this section and its subsections may be postponed until Chapter 4.

is a rational number, $x_1(t) - x_2(t)$ is periodic with fundamental period $2P_1 = 9P_2 = 4\pi/0.6 = 20\pi/3$. Note that to find the fundamental period, the two integers $\bar{k}_1$ and $\bar{k}_2$ must be *coprime* or have no common integer factor other than 1. Because

$$-2 = -2\sin(0 \cdot t + \pi/2) = 2\sin(0 \cdot t - \pi/2) = -2\cos(0 \cdot t) = 2\cos(0 \cdot t \pm \pi)$$

the constant -2 has zero frequency. However, it is periodic with any period. Thus we conclude that the signal in (1.24) is periodic with fundamental period $20\pi/3$ and fundamental frequency $2\pi/(20\pi/3) = 6/20 = 0.3$ rad/s.

The fundamental frequency computed above was computed through the fundamental period. In fact we can compute it directly without computing first the fundamental period. Given a number a, a number d is called a *divisor* of a if a/d is an integer. For example, the number 0.6 has divisors $0.6, 0.3, 0.2, 0.1, 0.06, \ldots$. The number 0.4 is not a divisor of 0.6 because $0.6/0.4 = 1.5$ is not an integer.

Consider sinusoidal signals $x_i(t)$ with frequency ω_i for $i = 1, 2$. If ω_1 and ω_2 have no common divisor, then $x_1(t) + x_2(t)$ is not periodic. For example, $\sin 2t + \sin \pi t$ is not periodic because 2 and π have no common divisor. However, if ω_1 and ω_2 have common divisors, then any linear combination of $x_1(t)$ and $x_2(t)$ is periodic. Its fundamental frequency is their *greatest common divisor (gcd)*. For example, consider the signal in (1.24). The frequency 0.6 has divisors $0.6, 0.3, 0.2, 0.1, \ldots,$ and 2.7 has divisors $2.7, 0.9, 0.3, 0.1, \ldots$. They have common divisors $0.3, 0.1, \ldots$. Their gcd is 0.3. Thus the signal in (1.24) is periodic with fundamental frequency 0.3 rad/s.

1.7.1 Frequency Components

Consider a time signal $x(t)$. An important question we may ask is its frequency content. For the signal in (1.24), the answer is simple. It consists of one sine function with frequency 0.6 rad/s and amplitude 3, and three cosine functions with frequencies 0, 0.6, and 2.7 and amplitudes -2, -4, and -2.4, respectively. This is one way of specifying the frequency content of the signal and is called a frequency-domain description.

There are, however, other ways of specifying the frequency content as we discuss next. Using $\cos\omega t = -\sin(\omega t - \pi/2)$ and

$$a\sin\omega t + b\cos\omega t = A\sin(\omega t + \theta)$$

where $A = \sqrt{a^2 + b^2} > 0$ and $\theta = \tan^{-1}(b/a)$, we can write (1.24) as

$$x(t) = 2\sin(0 \cdot t - \pi/2) + 5\sin(0.6t - 0.93) + 2.4\sin(2.7t - \pi/2) \tag{1.25}$$

Thus the periodic signal can also be specified as consisting of three sine functions of frequencies 0, 0.6, and 2.7 with corresponding magnitudes 2, 5, and 2.4 and phases $-\pi/2$, -0.93, and $-\pi/2$. Using

$$\sin(\omega t + \phi) = \cos(\omega t + \phi - \pi/2)$$

we can write (1.24) or (1.25) as

$$x(t) = 2\cos(0 \cdot t - \pi) + 5\cos(0.6t - 2.5) + 2.4\cos(2.7t - \pi) \tag{1.26}$$

Thus the periodic signal can also be specified as consisting of three cosine functions of frequencies 0, 0.6, and 2.7 with corresponding magnitudes 2, 5, and 2.4 and phases $-\pi$, -2.5, and $-\pi$.

In conclusion, the frequency content of the periodic signal in (1.24) can be expressed using the set of sine and cosine functions, the set of sine functions, or the set of cosine functions. Each set is called a *basis*. Thus in discussing the frequency content of a signal, we must specify the basis.

1.7.2 Complex Exponentials—Positive and Negative Frequencies

We discussed in the preceding section several ways of describing the frequency content of periodic signals. Because signals encountered in practice are mostly nonperiodic or aperiodic, the descriptions must be extended to aperiodic signals. However, those descriptions are not as easily extendible as the basis to be introduced next.

Consider the function $e^{j\omega_0 t}$, with $j := \sqrt{-1}$, or, using Euler's identity,

$$e^{j\omega_0 t} = \cos \omega_0 t + j \sin \omega_0 t \tag{1.27}$$

It is a complex-valued function with real part $\cos \omega_0 t$ and imaginary part $\sin \omega_0 t$, and it is called the *complex exponential function* with frequency ω_0. If we plot $e^{j\omega_0 t}$ against time as shown in Figure 1.23(a), then the plot will be three-dimensional and is difficult to visualize. Thus it is customary to plot $e^{j\omega_0 t}$ on a complex plane, as shown in Figure 1.23(b), with time as the parameter on the plot. Because

$$|e^{j\omega_0 t}| = \sqrt{(\cos \omega_0 t)^2 + (\sin \omega_0 t)^2} = 1$$

for all t, the plot of $e^{j\omega_0 t}$ on a complex plane is very simple, a unit circle as shown.

Using (1.27) and $e^{-j\omega_0 t} = \cos \omega_0 t - j \sin \omega_0 t$, we can readily develop

$$\sin \omega_0 t = \frac{e^{j\omega_0 t} - e^{-j\omega_0 t}}{2j} \quad \text{and} \quad \cos \omega_0 t = \frac{e^{j\omega_0 t} + e^{-j\omega_0 t}}{2} \tag{1.28}$$

Substituting these into (1.24) yields

$$x(t) = -2 + 3 \times \frac{e^{j0.6t} - e^{-j0.6t}}{2j} - 4 \times \frac{e^{j0.6t} + e^{-j0.6t}}{2} - 2.4 \times \frac{e^{j2.7t} + e^{-j2.7t}}{2}$$

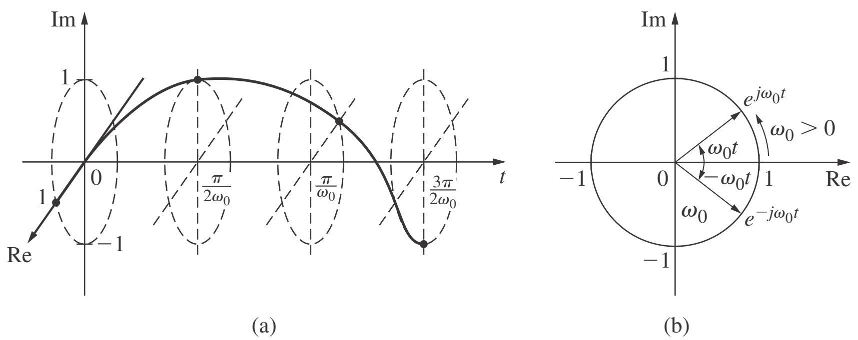


Figure 1.23 (**a**) Three-dimensional plot of $e^{j\omega_0 t}$. (**b**) Two-dimensional plot of $e^{j\omega_0 t}$.

which can be simplified as

$$x(t) = -2e^{j0\cdot t} + (-2 - 1.5j)e^{j0.6t} + (-2 + 1.5j)e^{-j0.6t} - 1.2e^{j2.7t} - 1.2e^{-j2.7t} \quad (1.29)$$

In terms of this expression, the periodic signal $x(t)$ is said to have frequency components at 0, ± 0.6, and ± 2.7 with corresponding coefficients -2, $-2 \mp 1.5j$, -1.2, and -1.2.

There are two problems in using (1.29). In (1.24) through (1.26), we encountered only positive frequencies and real numbers. Here we encounter complex numbers and both positive and negative frequencies. In spite of these two problems, the expression can be more easily extended to aperiodic signals. Thus complex exponentials are used extensively, if not exclusively, in discussing frequency content of signals.

Before proceeding, we give an interpretation of negative frequency. Consider the plot of $e^{j\omega_0 t}$ in Figure 1.23(b). By convention, if $\omega_0 > 0$, the function $e^{j\omega_0 t}$ rotates around the unit circle *counterclockwise* as t increases; whereas $e^{-j\omega_0 t}$ rotates *clockwise*. Thus positive and negative frequencies indicate only different directions of rotation. There is no way to avoid the use of complex coefficients, thus we discuss some of their operations in the next subsection.

1.7.3 Magnitudes and Phases; Even and Odd

We encountered complex coefficients in (1.29). Complex numbers can be expressed in polar form as

$$\alpha + j\beta = Ae^{j\theta}$$

where α, β, and θ are real and can be positive or negative. The number A, however, must be real and positive (can be zero, but not negative); it equals $\sqrt{\alpha^2 + \beta^2}$ and is called the *magnitude*. The number θ equals $\tan^{-1}(\beta/\alpha)$ with unit in radians and is called the *phase* or *angle*. For example, we have

$$\begin{aligned} -2 - 1.5j &= \sqrt{(-2)^2 + (-1.5)^2}e^{j\tan^{-1}((-1.5)/(-2))} \\ &= 2.5e^{-j2.5} = 2.5e^{-j143^\circ} = 2.5e^{j217^\circ} \end{aligned}$$

and

$$\begin{aligned} -2 + 1.5j &= \sqrt{(-2)^2 + (1.5)^2}e^{j\tan^{-1}(1.5/(-2))} \\ &= 2.5e^{j2.5} = 2.5e^{j143^\circ} = 2.5e^{-j217^\circ} \end{aligned}$$

In order to see better their magnitudes and phases, we plot them on complex planes as shown in Figures 1.24(a) and 1.24(b). They are vectors emitting from the origin as shown. The magnitude is simply the length of the vector or the distance from the origin and equals 2.5. The phase is the angle measured from the real axis. By convention, an angle is positive if measured counterclockwise and is negative if measured clockwise. The vector $-2 - 1.5j$ is in the third quadrant and has phase -2.5 rad or $-2.5 \times 180/\pi = -143^\circ$ if measured clockwise; it has phase $2\pi - 2.5 = 3.78$ rad or $360 - 143 = 217^\circ$ if measured counterclockwise. The vector $-2 + 1.5j$ is in the second quadrant and has angle 143° or -217°. When we compute the phase of a complex number, it is helpful to draw it on a complex plane; otherwise we may obtain an incorrect angle especially in using a calculator.

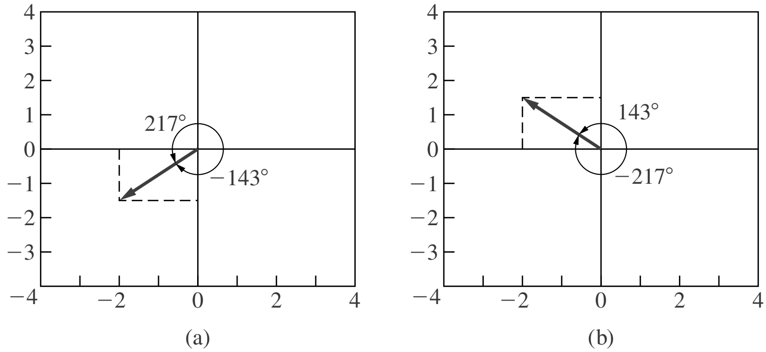


Figure 1.24 (**a**) Vector $-2 - 1.5j$. (**b**) Vector $-2 + 1.5j$.

EXERCISE 1.7.1

Express 2, $2 + 2j$, $2 - 2j$, $-2 + 2j$, $-2 - 2j$, and -2 in polar form.

Answers

$2e^{j0}$, $2.83e^{j\pi/4} = 2.83e^{0.785j}$, $2.83e^{-j\pi/4}$, $2.83e^{j3\pi/4}$, $2.83e^{-j3\pi/4}$, $2e^{j\pi}$

EXERCISE 1.7.2

Plot $e^{j\pi/2}$, $e^{j\pi}$, $e^{j3\pi/2}$, $e^{j2\pi}$, $e^{-j\pi/2}$, $e^{-j\pi}$, $e^{-j3\pi/2}$, and $e^{-j2\pi}$ on the complex plane and find their values.

Answers

$j, -1, -j, 1, -j, -1, j, 1$

EXERCISE 1.7.3

Show $e^{j\theta} = e^{j(\theta+2k\pi)}$ for any integer k.

Two angles are considered the same if they differ by 360° or its integer multiple, written as

$$\theta_1 = \theta_2 \qquad (\text{mod } 360 \text{ or } 2\pi) \tag{1.30}$$

where mod stands for modulo. For example, we have

$$217 = -143, \quad 143 = -217 \qquad (\text{mod } 360)$$

as shown in Figure 1.24, and

$$25 = 385 = -335 = 745 \qquad (\text{mod } 360)$$

In view of the above, the expression of an angle is not unique. In order to have a unique expression, we express every angle inside the range $-180^\circ < \theta \leq 180^\circ$ or $-\pi < \theta \leq \pi$, denoted as $(-180^\circ, 180^\circ]$ or $(-\pi, \pi]$. Note the use of left parenthesis and right bracket. If we adopt the range $[-180^\circ, 180^\circ]$, then the angle 180° can also be expressed as -180° and its expression is not unique.

Now we express the coefficients of (1.29) in polar form as

$$\begin{aligned} x(t) = {} & 2e^{j\pi}e^{j0\cdot t} + 2.5e^{-j2.5}e^{j0.6t} + 2.5e^{j2.5}e^{-j0.6t} \\ & + 1.2e^{j\pi}e^{j2.7t} + 1.2e^{j\pi}e^{-j2.7t} \end{aligned} \tag{1.31}$$

Note that -2 and -1.2 are not in polar form; their polar forms are $2e^{j\pi}$ and $1.2e^{j\pi}$. This is the standard expression used in discussing frequency content of periodic signals. It is called the *complex Fourier series* and will be introduced formally in Chapter 4. The coefficients are called *Fourier series coefficients*.

In terms of the expression in (1.31), the signal can also be specified by the five frequencies 0, ± 0.6, and ± 2.7 with corresponding coefficients $2e^{j\pi}$, $2.5e^{\mp j2.5}$, and $1.2e^{j\pi}$. This is called a frequency-domain description, and the coefficients are plotted against frequencies as shown in Figure 1.25. Because the coefficients are generally complex, we plot in Figure 1.25(a) their magnitudes against frequencies and plot in Figure 1.25(b) their phases. It is also possible to plot their real and imaginary parts against frequencies. However, such plots have no physical meaning; whereas the magnitude plot reveals, as we will discuss in Chapter 4, the distribution of the power of $x(t)$ in frequencies.

We discuss an important property of Fourier series coefficients. Let us use an asterisk to note the complex conjugate of a number. For example, we have $(2 - j3)^* = 2 + j3$ and

$$(e^{j2.7t})^* = (\cos 2.7t + j \sin 2.7t)^* = \cos 2.7t - j \sin 2.7t = e^{-j2.7t}$$

Thus *complex conjugation simply changes j into $-j$ or $-j$ into j*. Clearly, a number a is real-valued if and only if $a^* = a$. Likewise, a signal $x(t)$ is real-valued if and only if $x^*(t) = x(t)$, for all t.

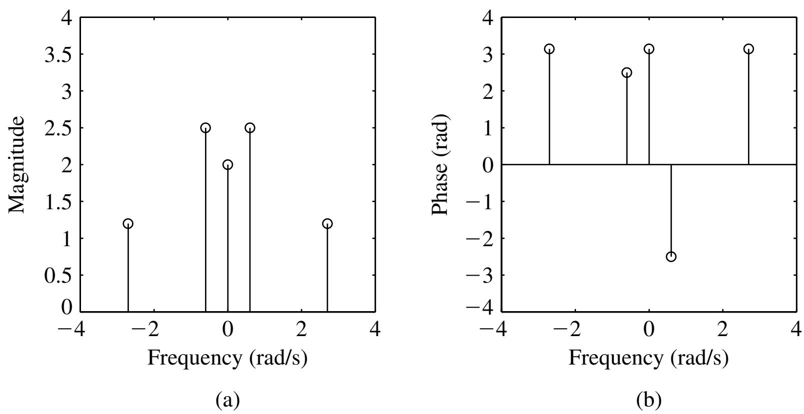


Figure 1.25 (**a**) Magnitude plot. (**b**) Phase plot.

A function $f(t)$ is *even* or *symmetric* with respect to t if $f(t) = f(-t)$, for all t in $(-\infty, \infty)$. It is *odd* or *antisymmetric* if $f(t) = -f(-t)$. It is clear that a positive-time signal $x(t)$ can never be even or odd. The preceding definitions still hold if t is replaced by other variables. For example, $f(\omega)$ $[f(m)]$ is even (with respect to ω $[m]$) if $f(\omega) = f(-\omega)$ $[f(m) = f(-m)]$ and is odd if $f(\omega) = -f(-\omega)$ $[f(m) = -f(-m)]$.

We now express $x(t)$ in (1.31) more generally as

$$x(t) = c_0 + \sum_m \left[c_m e^{j\omega_m t} + c_{-m} e^{-j\omega_m t}\right]$$

where c_0 is the coefficient associated with $\omega = 0$, c_m is associated with positive frequency ω_m, and c_{-m} is associated with negative frequency $-\omega_m$. Taking the complex conjugate of $x(t)$ yields

$$x^*(t) = c_0^* + \sum_m \left[c_m e^{j\omega_m t} + c_{-m} e^{-j\omega_m t}\right]^* = c_0^* + \sum_m \left[c_m^* e^{-j\omega_m t} + c_{-m}^* e^{j\omega_m t}\right]$$

If $x(t)$ is real, then $x(t) = x^*(t)$. Thus we have

$$c_0 + \sum_m \left[c_m e^{j\omega_m t} + c_{-m} e^{-j\omega_m t}\right] = c_0^* + \sum_m \left[c_{-m}^* e^{j\omega_m t} + c_m^* e^{-j\omega_m t}\right]$$

which implies $c_0 = c_0^*$, $c_m = c_{-m}^*$, and $c_{-m} = c_m^*$ for all m. Thus c_0 must be real and c_m and c_{-m} must be *conjugate symmetric* in the sense $c_m = c_{-m}^*$. In other words, if $c_m = \alpha + j\beta$, then we have $c_{-m} = \alpha - j\beta$, which implies

$$|c_m| = \sqrt{\alpha^2 + \beta^2} = \sqrt{\alpha^2 + (-\beta)^2} = |c_{-m}|$$

and

$$\measuredangle c_m = \tan^{-1}(\beta/\alpha) = -\tan^{-1}((-\beta)/\alpha) = -\measuredangle c_{-m}$$

In conclusion, the Fourier series coefficients of real $x(t)$ must be conjugate symmetric with respect to ω. Their magnitudes are even and their phases are odd.

Let us check the magnitude plot in Figure 1.25(a). It is indeed an even function of ω. The phase plot in Figure 1.25(b), however, is not odd in the usual sense. If $f(\omega)$ is odd, then $f(\omega) = -f(-\omega)$, for all ω. In general, $f(0) = -f(0)$ implies $f(0) = 0$. This, however, may not be so if f is an angle. For example, $\measuredangle c_0 = \pi$ at $\omega = 0$ in Figure 1.25(b) is different from zero and still meets $\measuredangle c_0 = \pi = -\measuredangle c_0 = -\pi$ (mod 2π). The phases at $\omega = 2.7$ and -2.7 are mathematically odd if both are selected as π; they are odd in the usual sense if one is selected as π and the other as $-\pi$. In conclusion, the phase plot in Figure 1.25(b) is an odd function of ω.

EXERCISE 1.7.4

What are the fundamental frequency and fundamental period of the signal

$$x(t) = 2\cos 4t + 2\sin 4t$$

Express it in complex Fourier series. Plot the magnitudes and phases of its coefficients against frequency. Is the magnitude plot even? Is the phase plot odd?

Answers

4, 0.5π, $1.4e^{-j\pi/4}e^{4t} + 1.4e^{j\pi/4}e^{-4t}$. Yes. Yes.

EXERCISE 1.7.5

Repeat Exercise 1.7.4 for the signal

$$x(t) = 3 + \cos 2t - 2\sin 6t$$

Answers

$2, \pi, 3 + 0.5e^{2t} + 0.5e^{-2t} + e^{j\pi/2}e^{6t} + e^{-j\pi/2}e^{-6t}$. Yes. Yes.

EXERCISE 1.7.6

Is the CT signal

$$x(t) = 1.5 - 2\sin t + 3\cos t + 1.2\sin \pi t$$

periodic? Verify that the magnitudes and phases of its Fourier series coefficients are as shown in Figure 1.26. Is the magnitude plot even? Is the phase plot odd?

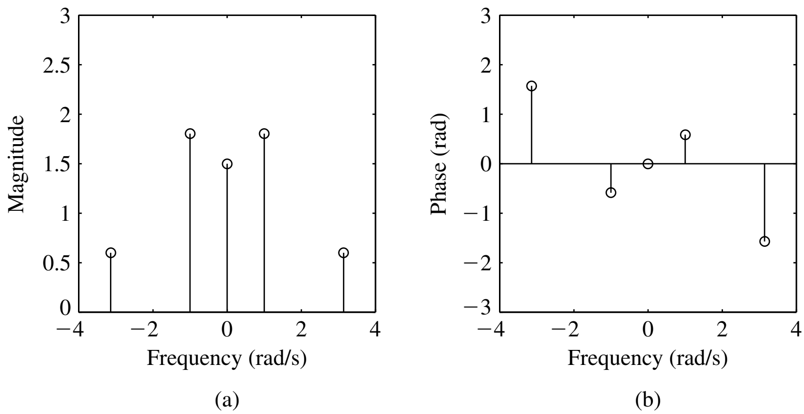


Figure 1.26 (**a**) Magnitude plot. (**b**) Phase plot.

Answers

No. Yes. Yes.

To conclude this section, we mention that the *frequency spectrum* is the standard terminology used to describe the frequency content of signals. As we will discuss in Chapter 4, the frequency spectrum of (1.24) or (1.31) is

$$\begin{aligned} X(\omega) = 2\pi[&2e^{j\pi}\delta(\omega) + 2.5e^{-j2.5}\delta(\omega - 0.6) + 2.5e^{j2.5}\delta(\omega + 0.6) \\ &+ 1.2e^{j\pi}\delta(\omega - 2.7) + 1.2e^{j\pi}\delta(\omega + 2.7)] \end{aligned} \tag{1.32}$$

which is obtained from (1.31) by replacing every $e^{\pm j\omega_0 t}$ by $\delta(\omega \mp \omega_0)$ and then multiplying the whole equation by 2π. Clearly, (1.32) is different from (1.31). The spectrum consists of five impulses at $\omega = 0, \pm 0.6$, and ± 2.7. Although some texts call the coefficients of (1.31) the discrete frequency spectrum of (1.24), we will not use the terminology to avoid its possible confusion with the spectrum defined in (1.32).

1.8 DT SINUSOIDAL SEQUENCES AND NYQUIST FREQUENCY RANGE

The CT sinusoid $\sin \omega_0 t$ is periodic for every ω_0 in $(-\infty, \infty)$ and will repeat itself every $P = 2\pi/\omega_0$ seconds. Its counterpart in the DT case is much more complex as we discuss in this section.[6]

Consider a DT sequence $x[n] = x(nT)$, where n is an integer ranging from $-\infty$ to ∞ and $T > 0$ is the sampling period. The sequence is said to be periodic with period N samples or NT seconds if there exists an integer N such that

$$x[n] = x[n + N]$$

for all integer n. It means that the sequence repeats itself *after* every N samples. If $x[n]$ is periodic with period N, then it is also periodic with period $2N, 3N, \ldots$. The smallest such N is called the *fundamental period* or, simply, the period, unless stated otherwise.

Every CT sinusoid $\sin \omega_0 t$ is periodic for every ω_0. This is, however, not the case for every DT sinusoid $\sin \omega_0 nT$. Indeed, if $\sin \omega_0 nT$ is periodic for the given ω_0 and T, then there exists an integer N such that

$$\sin \omega_0 nT = \sin \omega_0 (n + N)T = \sin \omega_0 nT \cos \omega_0 NT + \cos \omega_0 nT \sin \omega_0 NT$$

This holds for every integer n if and only if

$$\cos \omega_0 NT = 1 \quad \text{and} \quad \sin \omega_0 NT = 0$$

In order to meet these conditions, we must have

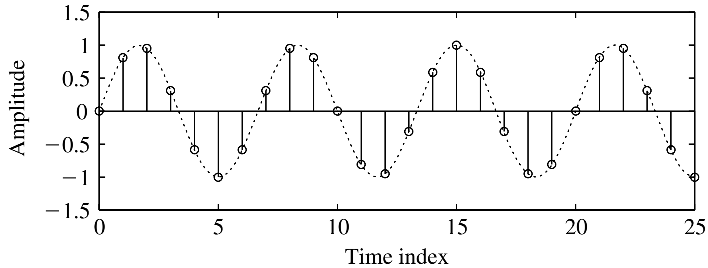
$$\omega_0 NT = k2\pi \quad \text{or} \quad N = \frac{2k\pi}{\omega_0 T} \tag{1.33}$$

for some positive integer k. In other words, the DT sinusoid $\sin \omega_0 nT$ is periodic if and only if there exists a positive integer k to make $2k\pi/\omega_0 T$ an integer. This is not always possible for every $\omega_0 T$. For example, consider $\sin \omega_0 T n = \sin 2n$ with $\omega_0 T = 2$. Because $N = 2k\pi/\omega_0 T = 2k\pi/2 = k\pi$ and because π is an irrational number (cannot be expressed as a ratio of two integers), there exists no integer k to make $N = k\pi$ an integer. Thus $\sin 2n$ is not periodic. In order for such k to exit in (1.33), $\omega_0 T$ must be a rational-number multiple of π. For example, if $\omega_0 T = 0.3\pi$, then we have

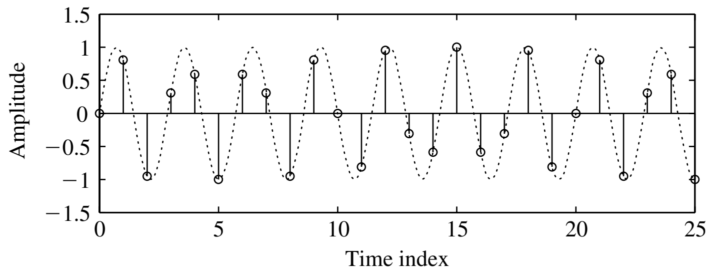
$$N = \frac{2k\pi}{0.3\pi} = \frac{2k}{0.3} = \frac{20k}{3}$$

---

[6]The study of this section may be postponed until Chapter 5.



(a)



(b)

Figure 1.27 (**a**) $\sin 0.3\pi n$. (**b**) $\sin 0.7\pi n$.

If $k = 3$, then $N = 20k/3 = 20$. Thus $\sin 0.3\pi n$, for any T, is periodic with period 20 samples or $20T$ seconds. Note that $k = 6, 9, 12$ also make N an integer. The smallest such positive k yields the fundamental period of the sequence. We plot in Figure 1.27(a) $\sin 0.3\pi n$ for $T = 1$ and $n = 0 : 25$. The sequence indeed repeats itself after 20 samples.

EXERCISE 1.8.1

Verify that $\sin 0.7\pi n$ with $T = 1$ shown in Figure 1.27(b) is periodic with period 20 samples.

A sinusoid $\sin \omega_0 nT$, as discussed above, may or may not be periodic. If it is periodic, how do we define its frequency? The frequency of a CT sinusoid with fundamental period P (in seconds) is defined as $2\pi/P$ (in rad/s). If we follow this definition, then the frequency of $\sin \omega_0 nT$ that is periodic with fundamental period N samples or NT seconds should be $2\pi/NT$. If we accept this definition, then the two sinusoidal sequences in Figure 1.27 should have, because of having the same fundamental period, the same frequency. This is clearly not acceptable. Thus we cannot define the frequency of $\sin \omega_0 nT$ as in the CT case. In conclusion, $\sin \omega_0 nT$ *may not be periodic. Even if it is periodic, its frequency should not be defined from its period as in the CT case.*

If we cannot use the fundamental period to define the frequency of DT sinusoids, how should we define it? The most natural way is to use their envelopes as shown in Figure 1.27 with dotted lines. More specifically, a CT sinusoid $\sin \bar{\omega} t$ is called an *envelope* of $\sin \omega_0 nT$ if its sample with sampling period T equals $\sin \omega_0 nT$, that is,

$$\sin \omega_0 nT = \sin \bar{\omega} t|_{t=nT}$$

for all n. Clearly, $\sin \omega_0 t$ is an envelope of $\sin \omega_0 nT$. However, there are, as we will show shortly, other envelopes. The envelope $\sin \bar{\omega} t$ is a CT sinusoid with well-defined frequency $\bar{\omega}$. Thus it is possible to use the frequency of CT sinusoids to define the frequency of DT sinusoidal sequences. Before doing so, we discuss a difference between CT and DT sinusoids.

For CT sinusoids, if $\omega_1 \neq \omega_2$, then $\sin \omega_1 t \neq \sin \omega_2 t$. This, however, is not the case for DT sinusoids. We show that if

$$\omega_1 = \omega_2 \qquad (\text{mod } 2\pi/T) \tag{1.34}$$

then

$$\sin \omega_1 nT = \sin \omega_2 nT$$

for all n. The condition in (1.34) means that even though ω_1 and ω_2 are different, if their difference is $k2\pi/T$, for some integer k (zero, negative, or positive), or

$$\omega_1 - \omega_2 = k\frac{2\pi}{T}$$

then we still have $\sin \omega_1 nT = \sin \omega_2 nT$. Indeed, we have

$$\begin{aligned}\sin \omega_1 nT &= \sin((\omega_2 + k2\pi/T)nT) = \sin(\omega_2 nT + 2kn\pi) \\ &= \sin \omega_2 nT \cos 2kn\pi + \cos \omega_2 nT \sin 2kn\pi = \sin \omega_2 nT\end{aligned}$$

where we have used $\cos 2kn\pi = 1$ and $\sin 2kn\pi = 0$ for all integers n and k. For example, if $T = 5$, then $2\pi/T = 2 \times 3.14/5 = 1.256$ and

$$0.4 = 0.4 + 1.256 = 0.4 - 1.256 = 0.4 + 2 \times 1.256 \qquad (\text{mod } 2\pi/T = 1.256)$$

or

$$0.4 = 1.656 = -0.856 = 2.912 \qquad (\text{mod } 2\pi/T = 1.256)$$

Thus we have, with $T = 5$,

$$\sin 0.4nT = \sin 1.656nT = \sin(-0.856)nT = \sin 2.912nT \tag{1.35}$$

for all n. In other words, we can use any one in (1.35) to plot the sinusoidal sequence. Or, equivalently, all the solid dots in Figure 1.28 denote the same sinusoidal sequence. Because of this fact, the ω_0 in $\sin \omega_0 nT$ is often said to be periodic with period $2\pi/T$. Thus the representation ω_0 in $\sin \omega_0 nT$ is not unique.

Because $\sin \omega_0 t$ is an envelope of $\sin \omega_0 nT$, the sinusoids $\sin 0.4t$, $\sin 1.656t$, $\sin(-0.856)t$, and $\sin 2.912t$, following (1.35) and shown in Figure 1.29, are all envelopes of any sinusoidal sequence in (1.35). Thus, every DT sinusoid $\sin \omega_0 nT$ has infinitely many CT sinusoidal envelopes.

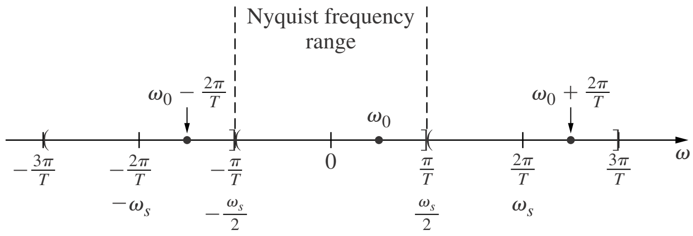


Figure 1.28 Periodicity of ω_0 in $\sin \omega_0 nT$.

In order to define the frequency of $\sin \omega_0 nT$ through its envelopes, we must use an envelope that is unique in some sense. The envelope $\sin \bar{\omega} t$ with the smallest $|\bar{\omega}|$ will be called the *primary envelope*. The frequency of the primary envelope will then be defined as the frequency of $\sin \omega_0 nT$. This is stated as a definition.

DEFINITION 1.1 The frequency of DT sinusoidal sequence $\sin \omega_0 nT$, periodic or not, is defined as the frequency of CT sinusoidal function $\sin \bar{\omega} t$ with the smallest $|\bar{\omega}|$ such that the sample of $\sin \bar{\omega} t$, with sampling period T, equals $\sin \omega_0 nT$.

From Figure 1.28, we see that the $\bar{\omega}$ that has the smallest magnitude must lie in the range $-\pi/T \leq \bar{\omega} \leq \pi/T$. If $\bar{\omega} = \pi/T$ and $\bar{\omega} = -\pi/T$, then the two $\bar{\omega}$ have the same magnitude. In this case, we adopt the former. Thus the frequency range of DT sinusoids becomes

$$-\frac{\pi}{T} < \omega \leq \frac{\pi}{T} \qquad \text{(in rad/s)}$$

denoted as $(-\pi/T, \pi/T]$ (note the left parenthesis and right bracket). This will be called the *Nyquist frequency range* (NFR). We call π/T (rad/s), following Reference 27, the *Nyquist frequency*. It is the *highest possible frequency* that a DT sinusoid with sampling period T can assume.

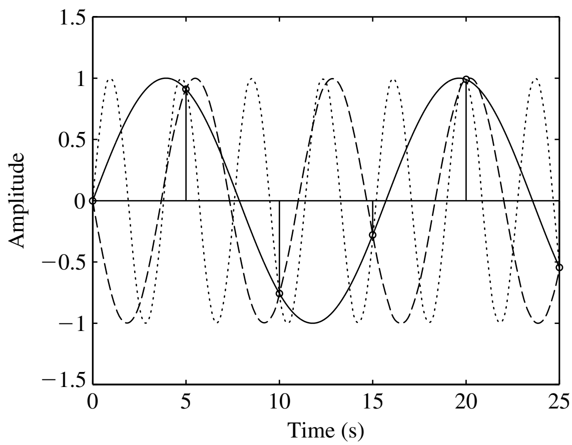


Figure 1.29 Envelopes of $\sin 0.4nT$ with $T = 5$.

Let us define $f_s = 1/T$ (in Hz) and $\omega_s = 2\pi/T$ (in rad/s). They are called the *sampling frequency*. If we define $f = \omega/2\pi$, then the Nyquist frequency range is

$$-0.5\omega_s = -\frac{\pi}{T} < \omega \leq \frac{\pi}{T} = 0.5\omega_s \qquad \text{(in rad/s)} \tag{1.36}$$

or

$$-0.5 f_s < f \leq 0.5 f_s \qquad \text{(in Hz)} \tag{1.37}$$

and the Nyquist frequency is $0.5\omega_s$ (rad/s) or $0.5 f_s$ (Hz), half of the sampling frequency. If $T = 1$, the NFR becomes $(-\pi, \pi]$ (in rad/s) or $(-0.5, 0.5]$ (in Hz). We see that the frequency range of DT sinusoid sequences is determined entirely by the sampling frequency or sampling period.

Now we discuss how to determine the frequency of $\sin \omega_0 nT$. The procedure is very simple. First we compute the NFR for the given T. If ω_0 lies inside the range, then the frequency of $\sin \omega_0 nT$ is ω_0. If ω_0 lies outside the range, we find the $\bar{\omega}$ that lies inside the range and equals ω_0 modulo $\omega_s = 2\pi/T$, and then the frequency of $\sin \omega_0 nT$ is $\bar{\omega}$.

EXAMPLE 1.8.1

Consider $\sin \omega_0 nT = \sin 5.15\pi n$. Because $\omega_0 T = 5.15\pi$, ω_0 is not defined without specifying T. If the sampling period is $T = 0.1$, then its NFR is $(-\pi/0.1, \pi/0.1] = (-10\pi, 10\pi]$. First we write $\sin 5.15\pi n = \sin 51.5\pi nT$. Because 51.5π is outside the range, it must be reduced to lie inside the range by subtracting $2\pi/T = 20\pi$ or its multiple. Clearly, we have

$$51.5\pi = 31.5\pi = 11.5\pi = -8.5\pi \qquad (\text{mod } \omega_s = 20\pi)$$

Thus the frequency of $\sin 5.15\pi n$ with sampling period $T = 0.1$ is -8.5π rad/s or -4.25 Hz.

EXAMPLE 1.8.2

Consider $\sin 6.5n = \sin 65nT$ with sampling period $T = 0.1$. Its frequency must lie inside $(-10\pi, 10\pi] = (-31.4, 31.4]$. Because

$$65 = 65 - 62.8 = 2.2 \qquad (\text{mod } 2\pi/0.1 = 62.8)$$

the frequency of $\sin 6.5n$ with $T = 0.1$ is 2.2 rad/s. Note that $\sin 6.5n$ is not periodic, but its frequency is still defined.

We recapitulate what has been discussed so far. The DT sinusoid $\sin \omega_0 nT$ may or may not be periodic. Even if it is periodic, its frequency should not be defined from its fundamental period as in CT sinusoids. The frequency of $\sin \omega_0 nT$ is defined as the frequency of CT sinusoid $\sin \bar{\omega} t$, where $\bar{\omega}$ equals ω_0 modulo $2\pi/T$ and lies inside $(-\pi/T, \pi/T]$. This implies that if $|\omega_0| < \pi/T$ or $T < \pi/|\omega_0|$, then

$$\text{Frequency of DT } \sin \omega_0 nT = \text{Frequency of CT } \sin \omega_0 t = \omega_0$$

Thus, the same notation ω has been used to denote the frequencies of CT and DT sinusoids. The frequency range of CT sinusoids is $(-\infty, \infty)$. The frequency range of DT sinusoids is $(-\pi/T, \pi/T]$. The frequency of a CT sinusoid can be as large as desired. The frequency of a DT sinusoid, however, is limited by the Nyquist frequency π/T.

We mention that what has been discussed for $\sin \omega_0 nT$ is also applicable to the cosine sequence $\cos \omega_0 nT$ and the complex exponential sequence $e^{j\omega_0 nT}$. For example, we have

$$e^{j\omega_0 nT} = e^{j(\omega_0 \pm 2\pi/T)nT} = e^{j(\omega_0 \pm 4\pi/T)nT} = \cdots \tag{1.38}$$

and the frequency of $e^{j\omega_0 nT}$ equals $\bar{\omega}$ rad/s, where $\bar{\omega}$ lies inside the NFR $(-\pi/T, \pi/T] = (-\omega_s/2, \omega_s/2]$ and equals ω_0 modulo $\omega_s = 2\pi/T$. If we use $f_0 = \omega_0/2\pi$ (in Hz) and $f_s = 1/T$, then the NFR is $(-f_s/2, f_s/2]$ (in Hz), and the frequency of $e^{j2\pi f_0 nT}$ equals $\bar{f}$ (Hz), where $\bar{f}$ lies inside $(-f_s/2, f_s/2]$ and equals f_0 modulo f_s.

We give a physical justification of the frequency of the complex exponential sequence $x[n] = x(nT) = e^{j2\pi f_0 nT}$. First we assume $T = 1$ or $f_s = 1$; that is, the sampling period is one second. We use solid dots to plot in Figure 1.30 $x[n] = x(nT) = e^{j2\pi f_0 n}$, for $n = 0 : 4$. Let $f_0 = \pm 1, \pm 2, \ldots$; that is, the dot rotates, in every second, complete cycles (no matter how many) in either direction. However, the dot appears to be stationary as shown in Figure 1.30(a). Thus the sequence $e^{j2\pi f_0 n}$ for any integer f_0 has, to our perception, frequency zero. This is indeed the frequency given by Definition 1.1 because $\pm 1 = \pm 2 = 0$ (modulo $f_s = 1$).

If the dot rotates, in every second, $5/4, 9/4, \ldots$ cycles (counterclockwise) or $-3/4, -7/4, \ldots$ cycles (clockwise), then the dot appears as shown in Figure 1.30(b). The sequence rotates, to our perception, $1/4$ cycle per second or has frequency $1/4$ Hz. Note that $5/4 = -3/4 = 1/4$ (modulo $f_s = 1$).

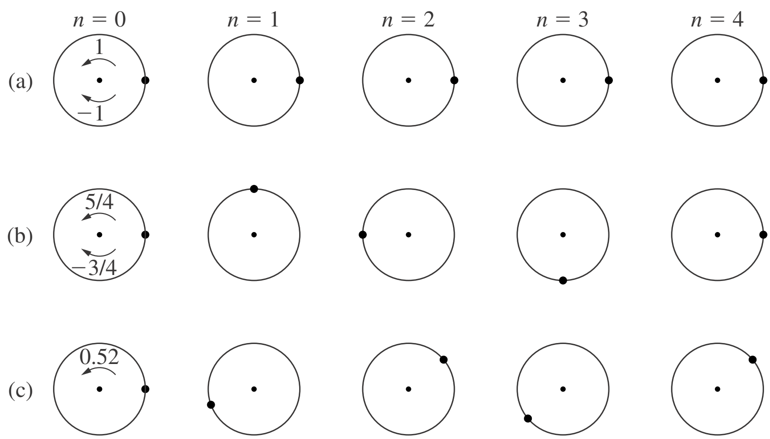


Figure 1.30 (**a**) The dot rotates, in every second, complete cycles (no matter how many) in either direction. However, it appears to be stationary or has frequency 0. (**b**) The dot rotates, in every second, $5/4, 9/4, \ldots$ cycles (counterclockwise) or $-3/4, -7/4, \ldots$ cycles (clockwise). However it appears to rotate $1/4$ cycle per second. (**c**) The dot rotates 0.52 cycles per second (counterclockwise). However, it appears to rotate -0.48 cycles per second (clockwise). Thus its frequency cannot be larger than $0.5 f_s = 0.5$ (Hz).

Next we show that if $T = f_s = 1$, then the highest frequency of $e^{j2\pi f_0 n}$ is $0.5f_s = 0.5$. Indeed, if $f_0 = 0.52$ or the dot rotates, in every second, 0.52 cycle (counterclockwise), it however will appear as rotating -0.48 (clockwise) as shown in Figure 1.30(c). Thus the frequency of $e^{j2\pi f_0 nT}$ is, to our perception, -0.48 Hz. In other words, the frequency of $e^{j2\pi f_0 n}$ cannot be larger than $0.5f_s = 0.5$.

To conclude this section, we mention that the NFR can be selected as $(-\pi/T, \pi/T]$ or $[-\pi/T, \pi/T)$ in rad/s, but not $[-\pi/T, \pi/T]$. If we select the range $[-\pi/T, \pi/T]$, then the frequency at $-\pi/T$ is the same as the frequency at π/T, and there is a redundancy. If we use the unit of Hz, then the NFR can be selected as $(-0.5f_s, 0.5f_s]$ or $[-0.5f_s, 0.5f_s)$ but not $[-0.5f_s, 0.5f_s]$.

1.9 SAMPLING AND FREQUENCY ALIASING

Because of many advantages of digital techniques, CT signals are now widely processed digitally.[7] To do so, we must first select a sampling period T and then sample the CT signal into a DT signal. We first use an example to discuss the effect of T on the frequency of the sampled sequence of $\sin \omega_0 t$. Note that the discussion is equally applicable to $\cos \omega_0 t$ and $e^{j\omega_0 t}$.

EXAMPLE 1.9.1

Consider the CT sinusoid $x(t) = \sin 3t$. It has frequency 3 rad/s. Its sampled sequence with sampling period T is

$$x[n] = x(nT) = \sin 3nT$$

The frequency of $\sin 3nT$, as discussed in the preceding section, depends on T. For any T, the NFR is $(-\pi/T, \pi/T]$. If 3 lies inside the range, then the frequency of $\sin 3nT$ is 3 rad/s. If not, we must reduce 3 to lie inside the range by subtracting $2\pi/T$ or its multiple. This is illustrated in the following:

T	$(-\pi/T, \pi/T]$	Frequency of $\sin 3nT$
0.1	$(-31.4, 31.4]$	3
0.5	$(-15.7, 15.7]$	3
1	$(-3.14, 3.14]$	3
1.8	$(-1.74, 1.74]$	$3 - 3.48 = -0.48$
2.6	$(-1.2, 1.2]$	$3 - 2.4 = 0.6$

We see that for $T = 0.1, 0.5$, and 1, the frequency of $\sin 3nT$ equals the frequency of the original sinusoid. For $T = 1.8$, the frequency 3 lies outside the NFR $(-1.74, 1.74]$; thus the frequency of $\sin 3nT$ is not 3 rad/s. Its frequency should be -0.48, which lies inside the range and equals 3 modulo $\omega_s = 2\pi/T = 3.48$. This is reasonable in view

[7] The study of this section may be postponed until Chapter 5.

of Figure 1.31(a). We call -0.48 an *aliased frequency*. Similarly, if $T = 2.6$, then the frequency of $\sin 3nT$ is 0.6 rad/s. It is different from the frequency of $\sin 3t$ and is an aliased frequency. Thus we often say that *time sampling may introduce frequency aliasing*.

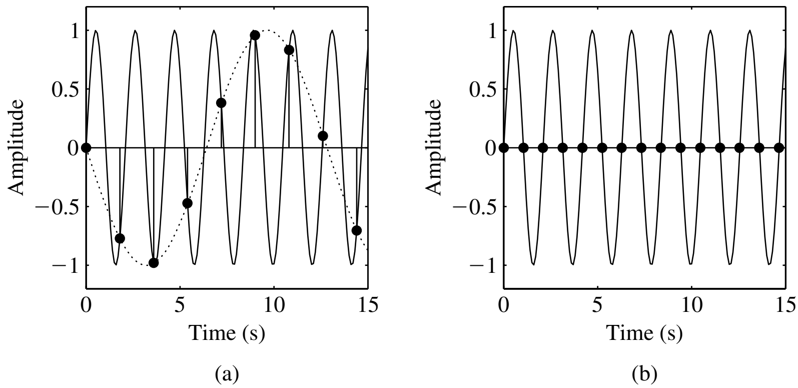


Figure 1.31 (**a**) Sample of $\sin 3t$ with sample period $T = 1.8$. (**b**) With $T = \pi/3$.

Using the procedure in the preceding example, we can show that the frequency of a CT sinusoid and the frequency of its sampled sequence are related as shown in Figure 1.32. The relationship in Figure 1.32 has a very important implication. The sampled sequence, with sampling period $T = 1/f_s$, of the sinusoid $\sin \omega_0 t = \sin 2\pi f_0 t$, $\cos \omega_0 t = \cos 2\pi f_0 t$, or $e^{j\omega_0 t} = e^{j2\pi f_0 t}$ with frequency $\omega_0 > 0$ and fundamental period $P = 2\pi/\omega_0$ has frequency $\omega_0 = 2\pi f_0$ if

$$T < \frac{\pi}{\omega_0} = \frac{P}{2} \quad \text{or} \quad f_s > 2f_0 \tag{1.39}$$

On the other hand, if (1.39) is not met, then the sampled sequence has a frequency, called an *aliased frequency,* which is different from ω_0. Note that if $T = \pi/\omega_0$, the sampled sequence still yields an aliased frequency. For example, for $x(t) = \sin 3t$ in Example 1.9.1, if $T = \pi/3$, then $x(nT) = \sin 3nT = \sin \pi n = 0$, for all n as shown in Figure 1.31(b). The sampled sequence has frequency 0, instead of 3.

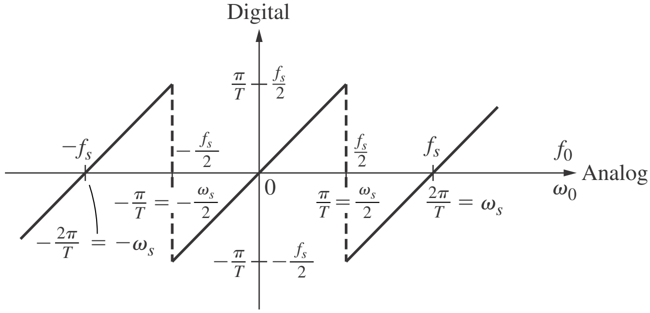


Figure 1.32 Relationship between the frequency of a CT sinusoid and that of its sampled sequence.

We discuss an application of Figure 1.32. Consider $e^{j\omega_0 t} = e^{j2\pi f_0 t}$. If $f_s = 1/T = f_0$ or $Tf_0 = 1$, then its sampled sequence $e^{j2\pi f_0 nT} = e^{j2\pi n}$ is 1 for for all n and thus has frequency 0. This zero frequency can also be read out from Figure 1.32. This fact has an important application. A stroboscope is a device that periodically emits flashes of light and acts as a sampler. The frequency of emitting flashes can be varied. We aim a stroboscope at an object that rotates at a constant speed. We increase the strobe frequency and observe the object. The object will *appear* to increase its rotational speed and then suddenly to reverse its rotational direction.[8] If we continue to increase the strobe frequency, the speed of the object will appear to slow down and then to stop rotating. The strobe frequency at which the object appears stationary is the rotational speed of the object. Thus a stroboscope can be used to measure rotational velocities. It can also be used to study vibrations of mechanical systems.

We give one more example to discuss the effect of frequency aliasing due to sampling.

EXAMPLE 1.9.2

Consider the CT signal

$$x(t) = \cos 50t + 2\cos 70t \tag{1.40}$$

It consists of two frequencies 50 and 70 rad/s and is plotted in Figure 1.33 with solid lines. Its sampled sequence with sampling period T is

$$x(nT) = \cos 50nT + 2\cos 70nT \tag{1.41}$$

Let us select $T = \pi/60$ or, equivalently, $\omega_s = 2\pi f_s = 120$ (rad/s). Its sampled sequence is plotted in Figure 1.33(a) with solid dots. For this T, the Nyquist frequency range (NFR) is $(-60, 60]$. Because 70 is outside the range, the frequency of $\cos 70nT$ is not 70. Because $70 = -50$ (mod $2\pi/T = 120$), we have

$$\cos 70nT = \cos(-50)nT$$

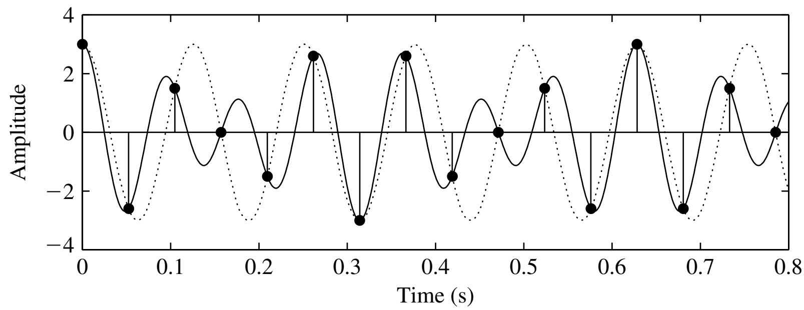
Thus the frequency of $\cos 70nT$ is -50 rad/s, which is an aliased frequency. Thus, for $T = \pi/60$, the sampled sequence in (1.41) becomes

$$x(nT) = \cos 50nT + 2\cos(-50)nT = 3\cos 50nT \tag{1.42}$$

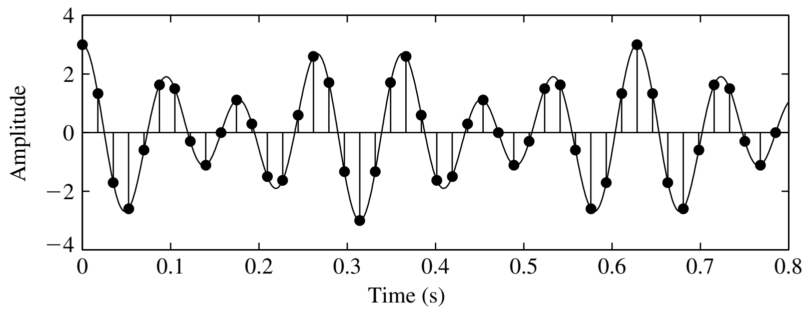
Its primary envelope $3\cos 50t$ is plotted in Figure 1.33(a) with a dotted line. We see that from the sampled sequence in (1.42), it is not possible to detect the existence of $\cos 70t$ in the original CT signal in (1.40). Thus if we sample (1.40) with sampling period $T = \pi/60$, some information of $x(t)$ will be lost in its sample sequence.

Next we select T as $\pi/180$ or, equivalently, $\omega_s = 360$ (rad/s). Its sampled sequence is plotted in Figure 1.33(b) with solid dots. For this T, the NFR is $(-180, 180]$. Because both 50 and 70 lie inside the range, the sampled sequence in (1.41) cannot be simplified. In other words, the frequency components of the CT signal are retained in its sampled sequence.

[8]This phenomenon may appear in a wagon's wheel in old western movies.



(a)



(b)

Figure 1.33 (**a**) Sample of $\cos 50t + 2\cos 70t$ with sample period $T = \pi/60$. (**b**) With $T = \pi/180$.

EXERCISE 1.9.1

Verify that the sampled sequence of (1.40) with sampling period $T = \pi/45$ is given by

$$x(nT) = \cos 40nT + 2\cos 20nT \tag{1.43}$$

Can the frequency components of $x(t)$ in (1.40) be determined from its sampled sequence in (1.43)?

Answer

No.

If we want to process a CT signal $x(t)$ digitally using its sampled sequence $x(nT)$, the sampling period T must be selected so that all information—in particular, all frequency components of $x(t)$—is retained in $x(nT)$. From the preceding discussion, we can now develop the condition. Consider a CT signal $x(t)$ that contains a number of frequency components. Let $\omega_{max} = 2\pi f_{max}$

be the largest frequency. If the sampling period $T = 1/f_s$ is selected so that

$$T < \frac{\pi}{\omega_{max}} \quad \text{or} \quad f_s > 2f_{max} \tag{1.44}$$

then its NFR contains all frequency components of $x(t)$. Under this condition, no frequency aliasing will occur in sampling $x(t)$. Thus $x(nT)$ will contain all frequency components of $x(t)$. This is in fact the *sampling theorem* which states that under the condition in (1.44), $x(t)$ can be recovered from $x(nT)$. This will be established in Chapter 5.

PROBLEMS

1.1 Consider the signal $x(t) = 1 + 2\cos\pi t$. Plot roughly the signal for t in $[0, 5]$. Compute its sampled sequence with sampling period $T = 0.5$; that is, compute the values of $x(nT) = x(0.5n)$, for $n = 0, 1, \ldots, 9, 10$ denoted as $n = 0 : 10$. Plot $x(nT)$ with respect to time t and with respect to time index n.

1.2 Repeat Problem 1.1 with sampling period $T = 1$ and for $n = 0 : 5$.

1.3 Consider the signal $x_1(t)$ shown in Figure 1.34(a). Plot $x_1(t-1)$, $x_1(-t+2)$, $x_1(t-1) + x_1(-t+2)$, $x_1(t-1) - x_1(-t+2)$, and $x_1(t-1)x_1(-t+2)$.

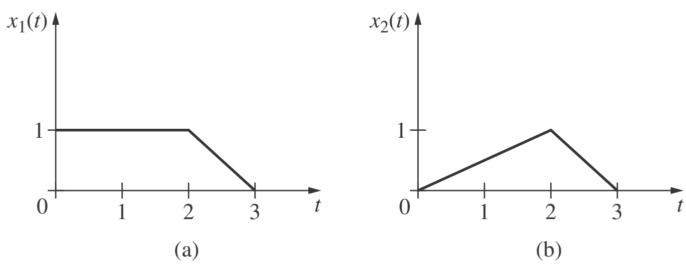


Figure 1.34

1.4 Repeat Problem 1.3 for the signal shown in Figure 1.34(b).

1.5 Express the signals in Figure 1.34 using the step and ramp functions defined in (1.1) and (1.3).

1.6 Express the signals in Figure 1.35 in terms of step and ramp functions.

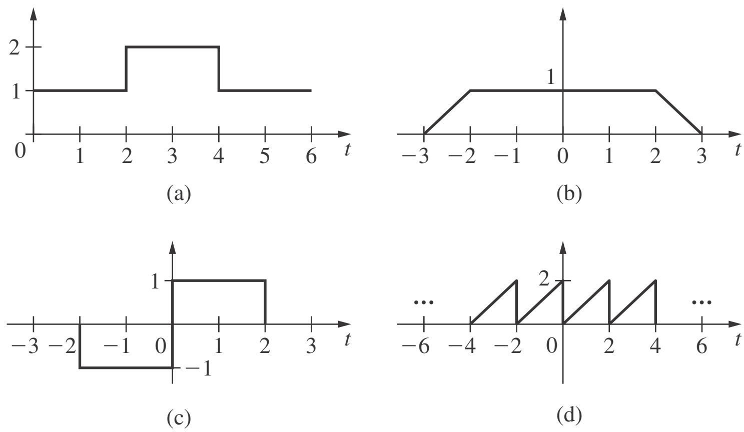


Figure 1.35

1.7 Consider the signal in Figure 1.11(a). It starts from $t = 0$ and ends at $t = 2$ and is said to have time duration 2.

(a) Plot $x(2t)$. What is its time duration?

(b) Plot $x(0.5t)$, What is its time duration?

(c) Show that if $a > 1$, then the time duration of $x(at)$ is smaller than that of $x(t)$. This speeds up the signal and is called *time compression*.

(d) Show that if $0 < a < 1$, then the time duration of $x(at)$ is larger than that of $x(t)$. This slows down the signal and is called *time expansion*.

1.8 A signal $x(t)$ is even if $x(t) = x(-t)$. It is odd if $x(t) = -x(-t)$. Which signal in Figure 1.35 is even? Which is odd?

1.9 Consider a signal $x(t)$. Define

$$x_e(t) = \frac{x(t) + x(-t)}{2} \quad \text{and} \quad x_o(t) = \frac{x(t) - x(-t)}{2}$$

Show that $x_e(t)$ is even and $x_o(t)$ is odd.

1.10 Problem 1.9 implies that every signal can be expressed as the sum of an even function and an odd function. Find the even and odd functions for the signals in Figures 1.34(a) and 1.35(a).

1.11 Use the time constant and period to sketch roughly the following signals for $t \geq 0$:

(a) $x_1(t) = e^{-0.2t}$

(b) $x_2(t) = \sin \pi t$

(c) $x_1(t)x_2(t)$

1.12 Find the derivatives of the signals in Figures 1.35(a) and 1.35(c).

1.13 Sketch the signal in Figure 1.35(a) modulated by $\cos 2\pi t$.

1.14 Compute

(a) $\int_0^9 [\cos \pi\tau]\delta(\tau - 3)\, d\tau$

(b) $\int_5^9 [\cos \pi\tau]\delta(\tau - 3)\, d\tau$

(c) $\int_{-\infty}^{\infty} [\cos(t - \tau)]\delta(\tau + 3)\, d\tau$

(d) $\int_0^{\infty} [\cos(t - \tau)]\delta(\tau + 3)\, d\tau$

(e) $\int_{-\infty}^{0} [\cos(t - \tau)]\delta(\tau + 3)\, d\tau$

1.15 Consider the sequence $x_1[n]$ shown in Figure 1.36(a). Plot $x_1[n+1]$, $x_1[-n+2]$, $x_1[n+1] + x_1[-n+2]$, and $x_1[n+1]x_1[-n+2]$.

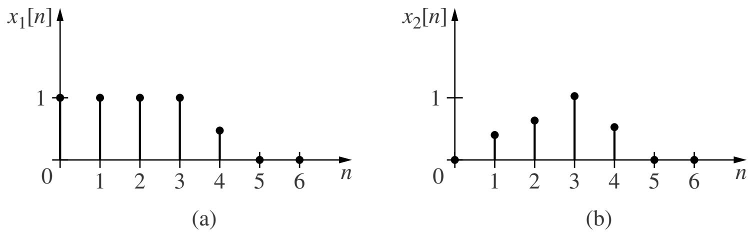


Figure 1.36

1.16 Repeat Problem 1.15 for the sequence $x_2[n]$ shown in Figure 1.36(b).

1.17 Express the sequences in Figures 1.36(a) and 1.36(b) in terms of the step and ramp sequences. The step sequence is defined in (1.16). The ramp sequence can be defined as $r[n] = n$ for $n \geq 0$ and as $r[n] = 0$ for $n < 0$.

1.18 Express the sequences in Figures 1.36(a) and 1.36(b) in terms of the impulse sequence defined in (1.17).

1.19 Consider a DT signal $x[n]$. Show that

$$x_e[n] = \frac{x[n] + x[-n]}{2} \quad \text{and} \quad x_o[n] = \frac{x[n] - x[-n]}{2}$$

are respectively even and odd.

1.20 Show

$$x[n]\delta[n - n_0] = x[n_0]\delta[n - n_0]$$

where n_0 is a fixed integer, and

$$\sum_{n=-\infty}^{\infty} x[n]\delta[n-n_0] = x[n_0]$$

They are the DT counterparts of (1.11) and (1.12). The latter is called the *sifting property* of the impulse sequence.

1.21 Consider $w_2[n]$ as defined in (1.20) with $N = 2$. Compute

$$y_1[n] = \sum_{k=-\infty}^{n} w_2[k]$$

and

$$y_2[n] = \sum_{k=-\infty}^{\infty} w_2[k]$$

for $n = -5 : 5$.

1.22 Consider the CT signal

$$x(t) = 2 + \sin 1.4t - 4\cos 2.1t$$

Is it periodic? If yes, find its fundamental period. (a) Express it using exclusively sine functions. What are their frequencies, corresponding magnitudes, and phases? (b) Express it using exclusively cosine functions. What are their frequencies, corresponding magnitudes, and phases?

1.23 Express the signal in Problem 1.22 using complex exponentials. Plot its magnitudes with respect to frequency. Is it even? Plot its phases. Is it odd?

1.24 Is the signal

$$x(t) = 2 + \sin 2t - 3\cos \pi t$$

periodic? Can it be expressed using complex exponentials?

1.25 Consider the signal

$$x(t) = -2.5 + 2\sin 2.8t - 4\cos 2.8t + 1.2\cos 4.9t$$

Is it periodic? If yes, find its fundamental period and fundamental frequency. Express it using complex exponentials. What are its Fourier series coefficients. Plot their magnitudes and phases. Are they respectively even and odd?

1.26 Which of $\sin 6.8\pi n$, $\cos 0.2n$, $\sin 4.9n$, $\sin 1.6\pi n$, and $-\sin 1.1n$ are periodic sequences? If they are, find their fundamental periods in samples. Are their frequencies defined without specifying the sampling period?

1.27 Are the sequences $\sin 6.9\pi n$, $\sin 4.9\pi n$, $\sin 2.9\pi n$, $\sin 0.9\pi n$, $\sin(-1.1)\pi n$, and $\sin(-3.1)\pi n$, for $T = 1$ the same sequence? If not, how many different sequences are there? Find their frequencies.

1.28 What are the frequencies of the sequences in Problem 1.26 if $T = 1$?

1.29 Find the frequencies of $\sin 1.2\pi n$ for $T = 0.5$ and $T = 0.1$.

1.30 Consider the sequences shown in Figure 1.37 with sampling period $T = 0.5$. What are their frequencies?

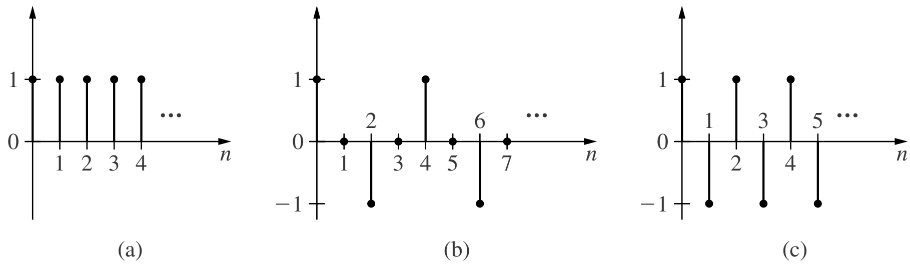


Figure 1.37

1.31 Consider $\cos \omega_k n$ with $T = 1$ and

$$\omega_k = \frac{2\pi k}{N}$$

where N is a given positive integer and k is an integer ranging from $-\infty$ to ∞. How many different $\cos \omega_k n$ are there? What are their frequencies?

1.32 Consider the CT sinusoid $x(t) = \sin 4t$. What are the frequencies of its sampled sequences if the sampling period T are 0.3, 0.6, 0.9, and 1.2? Do they equal the frequency of $\sin 4t$? Under what condition on T will the frequency of $\sin 4nT$ equal 4 rad/s?

1.33 Consider the CT sinusoid $x(t) = \sin \pi t$. What are its sampled sequences $x(nT)$ if the sampling period T are 0.5, 0.9, 1, and 1.5? What are their frequencies. Under what condition on T will the frequency of its sampled sequence be the same as the frequency of $\sin \pi t$?

1.34 Consider the CT signal $x(t) = \sin 50t + 2\sin 70t$. Verify that its sampled sequence $x(nT)$ is $-\sin 50nT$ if T is selected as $T = \pi/60$? Can you determine the two frequencies of $x(t)$ from $x(nT)$? What is the sampled sequence of $x(t)$ if $T = \pi/45$? Can you determine the two frequencies of $x(t)$ from this sampled sequence? Repeat the question for $T = \pi/180$.

1.35 What are the sampled sequences of $x(t) = \cos 50t + 2\sin 70t$ for $T = \pi/45, \pi/60$, and $\pi/180$? Under what condition on T will all frequencies of $x(t)$ be retained in $x(nT)$?

1.36 Consider the CT signal in Problem 1.25. If it is sampled with sampling period $T = 1$, what is its sampled signal? Does the sampled sequence contain all frequency components of the CT signal?

1.37 Repeat Problem 1.36 with $T = 0.5$.

CHAPTER 2 Systems

2.1 INTRODUCTION

Just like signals, systems exist everywhere. There are many types of systems. A system that transforms a signal from one form to another is called a *transducer*. There are two transducers in every telephone set: a microphone that transforms the voice into an electrical signal and a loudspeaker that transforms the electrical signal into the voice. Other examples of transducers are strain gauges, flow-meters, thermocouples, accelerometers, and seismometers. Signals acquired by transducers often are corrupted by noise. The noise must be suppressed or, if possible, eliminated. This can be achieved by designing a system called *filter*. If the signal level is too small or too large to be processed by the next system, the signal must be amplified or attenuated by designing an *amplifier*. Motors are systems that are used to drive compact discs, audio tapes, or huge satellite dishes. Generators are also systems—commonly used to generate electricity. They can also be used as power amplifiers in control systems.

In this text, a system will be modeled as a *black box* with at least one input terminal and one output terminal, as shown in Figure 2.1. We may or may not know the content or internal structure of the box; therefore, we call it a black box. Note that a terminal does not necessarily mean a physical terminal, such as a wire sticking out of the box. It merely indicates that a signal may be applied or measured from the terminal. We assume that if an excitation or input signal is applied to the input terminal, a *unique* response or output signal will be measurable or observable at the output terminal. This unique relationship between the excitation and response, input and output, or cause and effect is essential in defining a system.

A system with one input terminal and one output terminal is called a single-input single-output (SISO) system. A system with two or more input terminals and two or more output terminals is called a multi-input multi-output (MIMO) system. Clearly we may have multi-input single-output (MISO) or single-input multi-output (SIMO) systems. This text studies mostly SISO systems. Most concepts and results, however, are applicable and extendible to MIMO systems.

A system is called a *continuous-time* (CT) *system* if it accepts CT signals as its input and generates CT signals as its output. Likewise, a system is a *discrete-time* (DT) *system* if it accepts DT signals as its input and generates DT signals as its output. All DT signals are assumed to be equally spaced in time with sampling period T. We use $u(t)$ or $u[n] = u(nT)$ to denote the input and use $y(t)$ or $y[n] = y(nT)$ to denote the output. A system can be classified dichotomously as causal or noncausal, lumped or distributed, linear or nonlinear, time-invariant or time-varying. We introduce these concepts first for CT systems and then for DT systems.

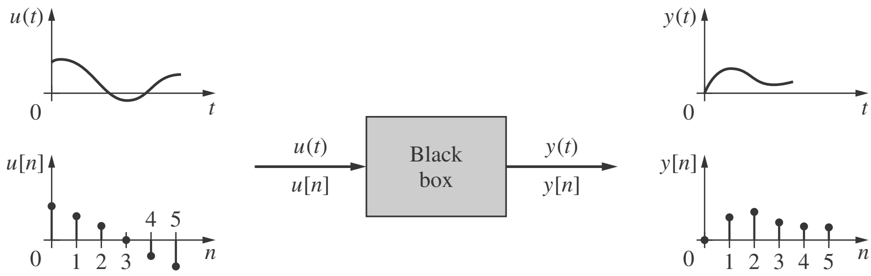


Figure 2.1 Black box.

2.2 CT SYSTEMS WITH AND WITHOUT MEMORY

A system is called a *memoryless system* if its output $y(t_0)$ depends only on the input $u(t)$ applied at time t_0; it is independent of the input applied before t_0 or after t_0. In other words, the current output of a memoryless system depends only on the current input; it does not depend on past or future inputs. An amplifier with gain a is a memoryless system because its input and output are related by $y(t) = au(t)$. Consider the resistive network shown in Figure 2.2. Its input voltage $u(t)$ and output voltage $y(t)$ are related by

$$y(t) = \frac{R_2}{R_1 + R_2} u(t)$$

It is a memoryless system. Many practical systems such as operational amplifiers and diodes are often modeled as memoryless systems as we will discuss in later sections.

If the current output of a system depends on past or future input, the system is said to have memory. For example, the system described by

$$y(t) = u(t-1) + 2u(t) - 3u(t+2)$$

has memory because the current output $y(t)$ depends on the past input $u(t-1)$, current input $u(t)$, and future input $u(t+2)$.

A system is called a *causal* or *nonanticipatory* system if its current output depends on past and current inputs but not on future input. An integrator is described by

$$y(t) = \int_{\infty}^{t} u(\tau)\, d\tau \tag{2.1}$$

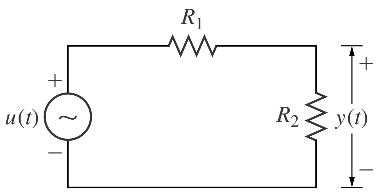


Figure 2.2 Resistive network.

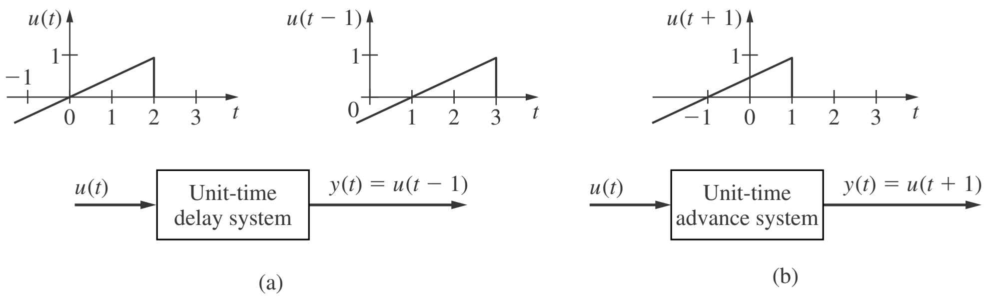


Figure 2.3 (**a**) Unit-time delay system. (**b**) Unit-time advance system.

Its output does not depend on future input. Thus it is a causal system. A *unit-time delay system* is defined by

$$y(t) = u(t-1) \tag{2.2}$$

Its output equals the input delayed by one second, as shown in Figure 2.3(a). It is also a causal system.

If a system is noncausal or anticipatory, then its current output depends on future input. For example, consider the *unit-time advance system* defined by

$$y(t) = u(t+1) \tag{2.3}$$

Its output equals the input advanced by one second as shown in Figure 2.3(b). The output at time t depends on the input at $t+1$, a future time. Thus it is a noncausal system. A noncausal system can *predict* what input will be applied in the future. No physical system has such a capability. Thus every physical system is causal, and *causality is a necessary condition for a system to be built or realized in the real world.*

The concept of causality is important in filters design. An *ideal filter* is noncausal and thus cannot be built in practice. Therefore, the filter design problem becomes finding a causal system to approximate as closely as possible the ideal filter. The concept of causality is implicitly associated with time (past, current, and future), and it is important in communication and control systems that operate in real time. In picture processing, we can develop an algorithm (a DT system) that utilizes all information surrounding a pixel to improve the quality of the picture. This is, in some sense, a noncausal system. This text studies only causal systems.

EXERCISE 2.2.1

Are the following systems with or without memory, causal or noncausal?

(a) $y(t) = 5u(t)$

(b) $y(t) = \sin u(t)$

(c) $y(t) = \sin u(t) + \sin u(t-1)$

(d) $y(t) = \sin u(t) + \sin u(t+1)$

(e) $y(t) = (\sin t)u(t)$

(f) $y(t) = (\sin t)u(t-1)$

Answers

(a) Without, causal. (b) Without, causal. (c) With, causal. (d) With, noncausal. (e) Without, causal. (f) With, causal.

2.3 THE CONCEPT OF STATE—SET OF INITIAL CONDITIONS

Current output of a causal system is affected by past input. How far back in time will the past input affect the current output? Strictly speaking, the time should go all the way back when the system was first built. To simplify the discussion, we assume the time to be $-\infty$. In other words, the input from $(-\infty, t]$ affects the output at t. Tracking the input $u(t)$ from $-\infty$ onward is, if not impossible, very difficult. The concept of state can resolve this problem.

Consider the block with mass m sliding on a frictionless floor shown in Figure 2.4(a). The applied force $u(t)$ is considered as the input and the displacement $y(t)$ the output. Using Newton's law, we have

$$m\frac{d^2y(t)}{dt^2} = u(t) \tag{2.4}$$

Thus the velocity of the mass is

$$v(t) := \frac{dy(t)}{dt} = \int_{t_0}^{t} \frac{u(\tau)}{m}\,d\tau + v(t_0) \tag{2.5}$$

and the displacement is

$$y(t) = \int_{t_0}^{t} v(\tau)\,d\tau + y(t_0) \tag{2.6}$$

We see that $y(t)$ depends on $y(t_0)$ and $v(t)$ for t in $[t_0, t]$, which in turn depends on $v(t_0)$ and $u(t)$ for t in $[t_0, t]$. Thus if we know $y(t_0)$ and $v(t_0)$, then we can determine uniquely $y(t)$ from $u(t)$ for $t \geq t_0$. The input applied before t_0 is no longer needed. In some sense, the effect of the input applied before t_0 on future output is summarized in $y(t_0)$ and $v(t_0)$. The set of $y(t)$ and

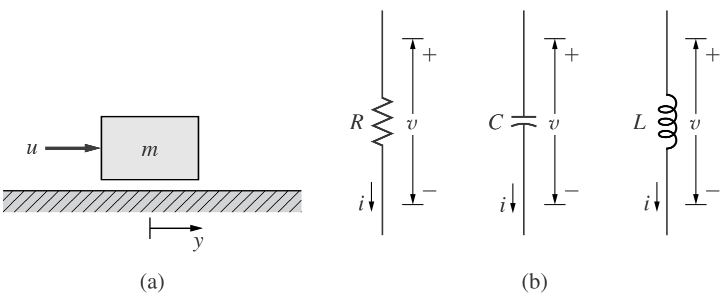


Figure 2.4 (**a**) Mechanical system. (**b**) Electrical systems.

$v(t)$ expressed as a column vector, denoted by a boldface x,

$$\mathbf{x}(t) := \begin{bmatrix} y(t) \\ v(t) \end{bmatrix}$$

is called the *state,* and its components are called the *state variables*. If we know the initial state at t_0 or, equivalently, the initial position $y(t_0)$ and the initial velocity $v(t_0) = dy(t_0)/dt$, then there is no more need to know the input applied before t_0.

We next discuss the state of RLC networks. Consider the electrical components shown in Figure 2.4(b). The voltage $v(t)$ across the resistor and the current through it are related by $v(t) = Ri(t)$. It is a memoryless element. The current and voltage of the capacitor with capacitance C are related by

$$i(t) = \frac{d}{dt}[Cv(t)] = C\frac{dv(t)}{dt}$$

where C has been assumed to be a constant. Its integration yields

$$v(t) = \frac{1}{C}\int_{t_0}^{t} i(\tau)\,d\tau + v(t_0) \tag{2.7}$$

In order to determine $v(t)$ from $i(t)$ for $t \geq t_0$, we need the information of $v(t)$ at $t = t_0$. Thus $v(t_0)$ is the initial state of the capacitor. Similarly, the voltage and current of an inductor with inductance L are related by

$$v(t) = \frac{d}{dt}[Li(t)] = L\frac{di(t)}{dt}$$

or

$$i(t) = \frac{1}{L}\int_{t_0}^{t} v(\tau)\,d\tau + i(t_0) \tag{2.8}$$

Thus the inductor current $i(t_0)$ is the initial state of the inductor. In general,[1] all capacitor voltages and all inductor currents can be selected as state variables. For the network shown in Figure 2.5, if we assign the capacitor voltages $v_1(t)$ and $v_2(t)$ and the inductor current $i(t)$ as shown, then the state of the network is

$$\mathbf{x}(t) := \begin{bmatrix} v_1(t) \\ v_2(t) \\ i(t) \end{bmatrix} \tag{2.9}$$

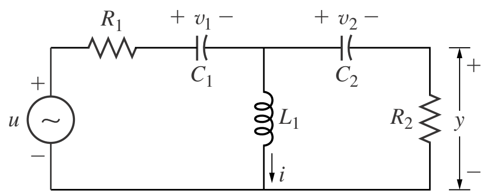


Figure 2.5 Network with three state variables.

[1]There are some exceptions as we will discuss in Chapter 7.

that consists of three state variables. It is important to mention that in assigning the voltages and current, we must also assign their polarities; otherwise, they are not well defined. For easy reference, we give a formal definition of the state.

DEFINITION 2.1 The state of a system at time t_0 is the information at t_0 that, together with the input $u(t)$, for $t \geq t_0$, determine uniquely the output $y(t)$ for $t \geq t_0$.

A system is called a *lumped system* if its number of state variables is finite. It is called a *distributed system* if its number of state variables is infinity. The mechanical system in Figure 2.4(a) has two state variables and the electrical system in Figure 2.5 has three state variables, thus they are lumped systems. Every memoryless system is lumped because it has no state variable or its number of state variable is zero.

EXAMPLE 2.3.1

Consider the unit-time delay system shown in Figure 2.3(a). In order to determine $y(t)$, for $t \geq 0$, in addition to the input $u(t)$, for $t \geq 0$, we need the information of $u(t)$ for all t in $[-1, 0)$. There are infinitely many points between -1 and 0; thus the state consists of infinitely many state variables and the system is distributed.

EXAMPLE 2.3.2

Consider the unit-time advance system shown in Figure 2.3(b). The output $y(t)$, for $t \geq 0$, can be determined uniquely from $u(t)$ for $t \geq 0$, without any other information. Thus the system has no state variable and is lumped. Note that the system, as mentioned earlier, is not causal.

EXERCISE 2.3.1

Which of the systems in Exercise 2.2.1 are lumped.

Answers

(a), (b), (d), and (e).

In conclusion, a system is lumped if the effect of past input on future output can be summarized in a finite number of initial conditions. Otherwise, it is distributed. CT systems that consist of transmission lines, waveguides, or time-delay elements are distributed systems.

2.3.1 Zero-Input Response and Zero-State Response

Using the concept of states, we now can express the output of a system as

$$\left.\begin{array}{l}\mathbf{x}(t_0) \\ u(t), \quad t \geq t_0\end{array}\right\} \rightarrow y(t), \quad t \geq t_0 \tag{2.10}$$

There is no more need to track past input all the way to $-\infty$ or to the time the system was first built. We call (2.10) a state input–output pair of the system.

If $u(t) = 0$ for $t \geq t_0$, the output is excited exclusively by initial conditions and is called the *zero-input response*. If $\mathbf{x}(t_0) = \mathbf{0}$, the output is excited exclusively by the input applied from t_0 onward and is called the *zero-state response*. The zero-input response is excited exclusively by nonzero initial conditions (the input is identically zero), thus the response is determined entirely by the *structure* or *nature* of the system. Consequently, the zero-input response is also called the *natural response* or *unforced response*. The zero-state response is excited exclusively by the applied input. Different inputs clearly excite different zero-state responses. Thus the zero-state response is also called the *forced response*.

The zero-input or natural response can be expressed as, using (2.10),

$$\left.\begin{array}{l}\mathbf{x}(t_0) \neq \mathbf{0} \\ u(t) = 0, \quad t \geq t_0\end{array}\right\} \rightarrow y_{zi}(t), \quad t \geq t_0 \qquad \text{(zero-input response)} \tag{2.11}$$

and the zero-state or forced response can be expressed as

$$\left.\begin{array}{l}\mathbf{x}(t_0) = \mathbf{0} \\ u(t), \quad t \geq t_0\end{array}\right\} \rightarrow y_{zs}(t), \quad t \geq t_0 \qquad \text{(zero-state response)} \tag{2.12}$$

where the subscripts *zi* and *zs* stand for zero input and zero state, respectively. The zero-state response can also be expressed as $\{u(t) \rightarrow y(t)\}$, without specifying $\mathbf{x}(t_0) = \mathbf{0}$.

A system is said to be *initially relaxed* at t_0 if its initial state at t_0 is zero. In this case, the input applied before t_0 has no more effect on $y(t)$ for $t \geq t_0$. For example, consider the unit-time delay system shown in Figure 2.3(a). If no input is applied after $t_0 - 1$, then the system is initially relaxed at t_0. In practice, if a system contains no time delay element or is lumped, and if the output of the system is identically zero for a short period of time before applying any input, the system is initially relaxed at the end of the period. *If a system is initially relaxed at t_0, there is no loss of generality to assume $u(t) = 0$ and $y(t) = 0$ for all $t < t_0$.*

2.4 LINEARITY OF MEMORYLESS SYSTEMS

This section introduces the concept of linearity. First we discuss memoryless systems. For memoryless systems, there is no initial state and the state input–output pair in (2.10) can be reduced to $\{u(t) \rightarrow y(t)\}$. A memoryless system is linear if for *any* two pairs $\{u_i(t) \rightarrow y_i(t)\}$, for $i = 1, 2$, we have

$$u_1(t) + u_2(t) \rightarrow y_1(t) + y_2(t) \qquad \text{(additivity)} \tag{2.13}$$

and, for any real constant α,

$$\alpha u_1(t) \rightarrow \alpha y_1(t) \qquad \text{(homogeneity)} \tag{2.14}$$

Otherwise it is a nonlinear system. The first condition is called the *additivity property,* and the second one is called the *homogeneity property*. We use examples to illustrate these two properties. For memoryless systems, the output y and the input u can be related by a curve or straight lines as shown in Figure 2.6. For the memoryless system described by Figure 2.6(a), we can read from the plot $\{u_1 = 1 \rightarrow y_1 = 2\}$, $\{u_2 = 2 \rightarrow y_2 = 2\}$, and $\{u_3 = 3 \rightarrow y_3 = 2\}$. If the system has the additivity property, then the output of $(u_1 + u_2 = u_3)$ should be $y_1 + y_2 = 4$

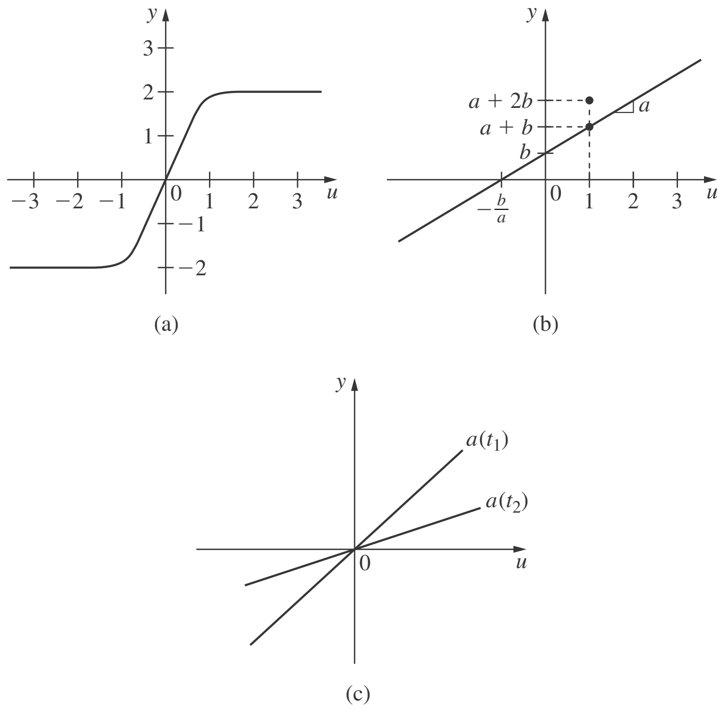


Figure 2.6 (**a**) Nonlinear. (**b**) Nonlinear. (**c**) Linear.

instead of $y_3 = 2$. If the system has the homogeneity property, the output of $3u_1 = 3 = u_3$ should be $3y_1 = 3 \times 2 = 6$. This is not the case. Thus the memoryless system is not linear. The nonlinearity is called *saturation* and often arises in practice.

Consider a memoryless system with input and output related by a straight line $y = au + b$, for some constants a and b, as shown in Figure 2.6(b). From the plot, we can read out $\{u_1 = 0 \rightarrow y_1 = b\}$ and $\{u_2 = 1 \rightarrow y_2 = a + b\}$. If the system has the additivity property, we should have $\{u_1 + u_2 = 1 \rightarrow y_1 + y_2 = b + a + b = a + 2b\}$. This is not the case. Thus the system is not linear. In fact, *a memoryless system is linear if and only if its input and output can be related by a straight line passing through the origin* as shown in Figure 2.6(c) such as

$$y(t) = a(t)u(t)$$

for some time function $a(t)$.

Before proceeding, we mention that the additivity property *almost* implies the homogeneity property but not conversely. See Problems 2.3 and 2.4. We also mention that the two properties in (2.13) and (2.14) are jointly called the *principle of superposition* and can be combined as, for any constants α_1 and α_2,

$$\alpha_1 u_1(t) + \alpha_2 u_2(t) \rightarrow \alpha_1 y_1(t) + \alpha_2 y_2(t) \qquad \text{(superposition)} \qquad (2.15)$$

EXAMPLE 2.4.1

Consider a memoryless systems with input u and output y related by $y(t) = \cos u(t)$. Because

$$\cos(u_1(t) + u_2(t)) \neq \cos u_1(t) + \cos u_2(t)$$

the system is not linear.

EXAMPLE 2.4.2 (Modulation)

Consider a memoryless systems with input u and output y related by $y(t) = (\cos 20t)u(t)$. Because

$$(\cos 20t)(\alpha_1 u_1(t) + \alpha_2 u_2(t)) = \alpha_1(\cos 20t)u_1(t) + \alpha_2(\cos 20t)u_2(t)$$

the system is linear. This system is actually the modulation discussed in Section 1.4.3. Thus *the modulation is a linear process*.

EXERCISE 2.4.1

Show that $y(t) = u^2(t)$ is nonlinear and that $y(t) = t^2 u(t)$ is linear.

2.4.1 Linearity of Systems with Memory

This subsection extends the linearity concept discussed for memoryless systems to systems with memory. The response of such a system depends not only on the input but also on the initial state. Thus the additivity and homogeneity properties should also apply to the initial state. Let

$$\left.\begin{array}{ll} \mathbf{x}_i(t_0) & \\ u_i(t), & t \geq t_0 \end{array}\right\} \rightarrow y_i(t), \quad t \geq t_0$$

for $i = 1, 2$, be *any* two state input–output pairs of a system. The system is linear if the following properties hold:

$$\left.\begin{array}{ll} \mathbf{x}_1(t_0) + \mathbf{x}_2(t_0) & \\ u_1(t) + u_2(t), & t \geq t_0 \end{array}\right\} \rightarrow y_1(t) + y_2(t), \quad t \geq t_0 \qquad \text{(additivity)} \tag{2.16}$$

and

$$\left.\begin{array}{ll} \alpha\mathbf{x}_1(t_0) & \\ \alpha u_1(t), & t \geq t_0 \end{array}\right\} \rightarrow \alpha y_1(t), \quad t \geq t_0 \qquad \text{(homogeneity)} \tag{2.17}$$

for any constant α. Otherwise, the system is said to be *nonlinear*. As for memoryless systems, the first condition is called the additivity property and the second one is called the homogeneity property. Jointly they are called the superposition property and can be combined into one equation as in (2.15).

We discuss some important implications of linearity. If $\alpha = 0$ in (2.17), then we have

$$\left.\begin{matrix} \mathbf{0} \\ 0, \quad t \geq t_0 \end{matrix}\right\} \rightarrow 0, \quad t \geq t_0$$

It means that if $\mathbf{x}(t_0) = \mathbf{0}$ and if $u(t) = 0$, for $t \geq t_0$, then $y(t) = 0$, for $t \geq t_0$. Thus a necessary condition for a system to be linear is that when the system is initially relaxed, the output must be identically zero if no input is applied. For example, consider the memoryless system in Figure 2.6(b). Its output is $b \neq 0$ even though no input is applied ($u = 0$). Thus the memoryless system violates this necessary condition and is, therefore, not linear.

Equation (2.16) implies

$$\text{Response of } \begin{cases} \mathbf{x}(t_0) \\ u(t), \quad t \geq t_0 \end{cases}$$
$$= \text{Response of } \begin{cases} \mathbf{x}(t_0) \\ u(t) = 0, \quad t \geq t_0 \end{cases} + \text{Response of } \begin{cases} \mathbf{x}(t_0) = \mathbf{0} \\ u(t), \quad t \geq t_0 \end{cases}$$

or

$$\begin{aligned} \text{Total response} &= \text{Zero-input response} + \text{Zero-state response} \\ &= \text{Natural response} + \text{Forced response} \end{aligned} \tag{2.18}$$

It means that for a linear system, the total response can always be decomposed as the zero-state (forced) and zero-input (natural) responses. Because of this property, the zero-state response and the zero-input response of a linear system can be computed separately. Their sum then yields the total response. If we are asked to compute the total response of a nonlinear system, it is useless to compute separately the zero-state and zero-input responses, because the total response can be very different from their sum.

If u_1 and u_2 in (2.16) and (2.17) are identically zero, then the additivity and homogeneity properties apply to zero-input responses. If initial states in (2.16) and (2.17) are zeros, then the two properties apply to zero-state responses. Thus for a linear system, its zero-input responses have the additivity and homogeneity properties. So have its zero-state responses. For example, consider the system shown in Figure 2.7(a). The block denoted by $\mathcal{S}$ is a linear subsystem whose description will be discussed in Chapter 7. The output y of the subsystem is negatively fed back to the input as shown. The small circle, called an *adder* or a *summer,* has two inputs u and y, as indicated by the two entering arrows, and one output e as indicated by the departing arrow. The

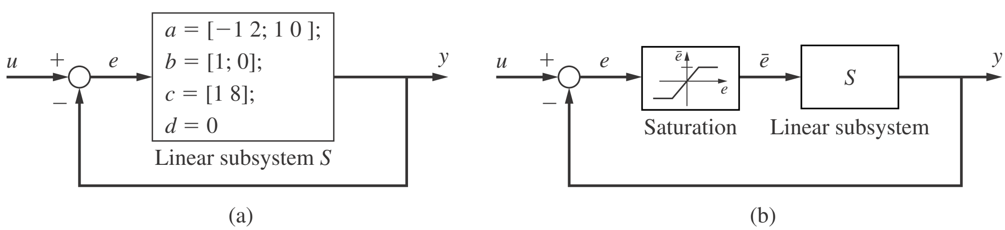


Figure 2.7 (**a**) Linear system. (**b**) Nonlinear system.

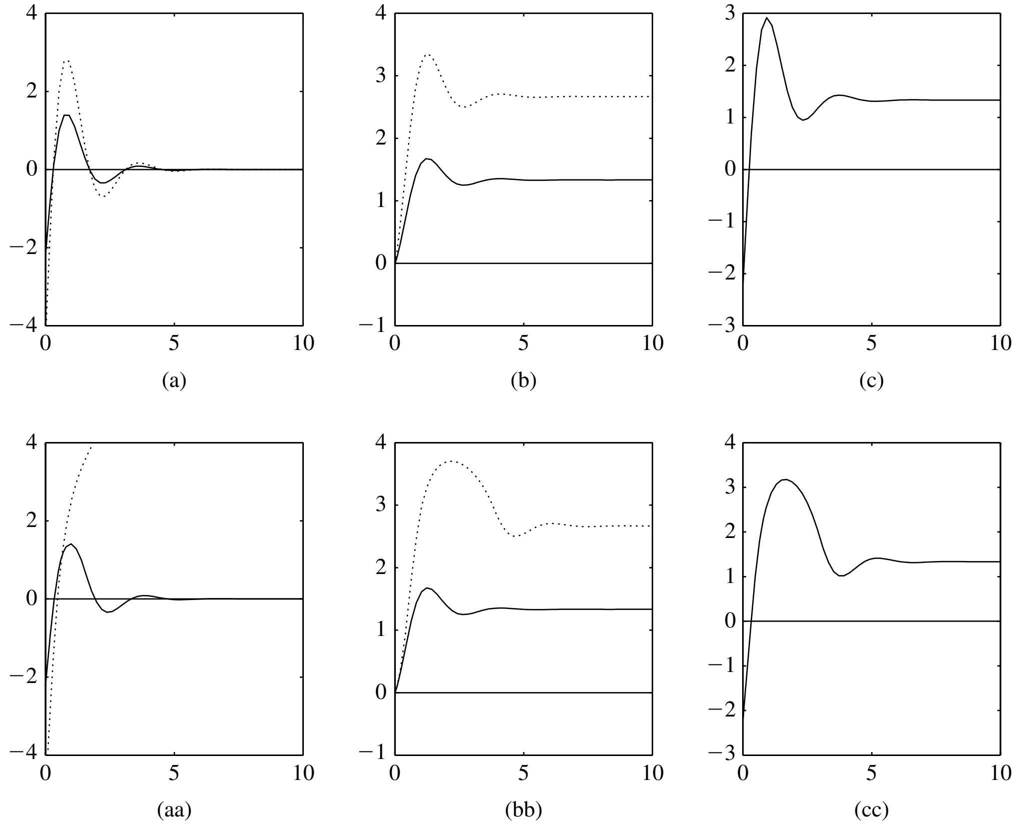


Figure 2.8 (**a**) Zero-input responses of the linear system in Figure 2.7(a) excited by $[1\ \ -0.4]'$ (solid line) and $[2\ \ -0.8]'$ (dotted line). (**b**) Zero-state responses of the linear system excited by step inputs with amplitude 1 (solid line) and amplitude 2 (dotted line). (**c**) Total response of the linear system excited by $u(t) = 1$ and $\mathbf{x}(0) = [1\ \ -0.4]'$. (**aa**), (**bb**), and (**cc**) Corresponding responses of the nonlinear system in Figure 2.7(b). These plots are obtained using SIMULINK.

arrow associated with u has a positive sign, and the arrow associated with y has a negative sign. Thus the output e equals $+u - y$ or $e = u - y$. The resulting feedback system is still linear. Its zero-input response y_{1zi} excited by the initial states $\mathbf{x}_1(0) = [1\ \ -0.4]'$, where the prime denotes the transpose, is plotted in Figure 2.8(a) with a solid line. Its zero-input response y_{2zi} excited by $\mathbf{x}_2(0) = [2\ \ -0.8]'$ is shown in Figure 2.8(a) with a dotted line. Because $\mathbf{x}_2(0) = 2\mathbf{x}_1(0)$, we have $y_{2zi}(t) = 2y_{1zi}(t)$ as shown. Figure 2.8(b) shows with a solid line the zero-state response $y_{1zs}(t)$ excited by $u_1(t) = 1$, for $t \geq 0$, and shows with a dotted line the zero-state response $y_{2zs}(t)$ excited by $u_2(t) = 2$. Because the system is linear and $u_2(t) = 2u_1(t)$, we have $y_{2zs}(t) = 2y_{1zs}$ as expected. Figure 2.8(c) plots the response excited by the input $u_1(t) = 1$ and the initial state $\mathbf{x}_1(0)$. The response is indeed the sum of $y_{1zi}(t)$ and $y_{1zs}(t)$.

Next we consider the feedback system in Figure 2.7(b). It contains a nonlinear element, thus the feedback system is nonlinear. We plot in Figure 2.8(aa) with a solid line its zero-input response $y_{1zi}(t)$ excited by $\mathbf{x}_1(0) = [1\ -0.4]'$ and plot with a dotted line its zero-input response $y_{2zi}(t)$ excited by $\mathbf{x}_2(0) = [2\ -0.8]'$. Although $\mathbf{x}_2(0) = 2\mathbf{x}_1(0)$, we do not have $y_{2zi} = 2y_{1zi}$. In fact, $y_{2zi}(t)$ grows to infinity. Figure 2.8(bb) shows the zero-state responses $y_{1zs}(t)$ (solid line) and $y_{2zs}(t)$ (dotted line) excited by $u_1(t) = 1$ and $u_2(t) = 2$, respectively. Again we do not have $y_{2zs}(t) = 2y_{1zs}(t)$. For the nonlinear system, the total response in Figure 2.8(cc) is not the sum of $y_{1zi}(t)$ and $y_{1zs}(t)$.

In conclusion, the characteristics of a linear system is independent of initial conditions and applied inputs. This is not the case for nonlinear systems. Thus the study of nonlinear systems is complicated and must be carried out with respect to specific initial conditions and inputs.

2.5 TIME INVARIANCE AND ITS IMPLICATION

If the characteristics of a system do not change with time, the system is said to be time-invariant. Otherwise, it is time-varying. If the resistances, capacitances, and inductance of the network elements in Figure 2.5 are constants, independent of time, then the network is a time-invariant system. A burning rocket is a time-varying system because its mass decreases rapidly with time. Strictly speaking, most physical systems are time-varying. For example, an automobile or a television set cannot last forever; it may break down after ten years. However, it can be modeled in the first few years as a time-invariant system. In fact, most physical systems can be so modeled over a limited time interval. Just like causality, time invariance is defined for systems, not for signals. Signals are mostly time-varying. If a signal is time-invariant, such as $u(t) = 1$ or $q(t) = 1$, then it is a very simple or a trivial signal.

If the characteristics of a system do not change with time, the output waveform excited by some initial state and input will always be the same no matter when the initial time is, as shown in Figure 2.9. Mathematically, this is stated as follows. For any state input–output pair of a time-invariant system

$$\left.\begin{array}{ll}\mathbf{x}(t_0) & \\ u(t), & t \geq t_0\end{array}\right\} \rightarrow y(t), \quad t \geq t_0$$

and any t_1, we have

$$\left.\begin{array}{ll}\mathbf{x}(t_0+t_1) & \\ u(t-t_1), & t \geq t_0+t_1\end{array}\right\} \rightarrow y(t-t_1), \quad t \geq t_0+t_1 \qquad \text{(time shifting)} \tag{2.19}$$

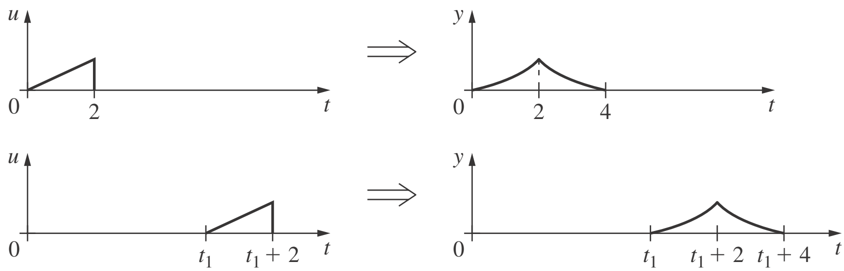


Figure 2.9 Time shifting property.

This is called the time shifting property. If a system is time-invariant, then its zero-input responses have the time shifting property. So have its zero-state responses. For time-invariant systems, the initial time is not critical. Thus we may assume, without loss of generality, that $t_0 = 0$. Note that $t_0 = 0$ is a relative one; it is the time instant we start to study or to apply an input to a system.

EXAMPLE 2.5.1

Consider a linear system. Suppose we are given the two pairs of zero-state responses shown in Figures 2.10(a) and 2.10(b). We pose the following questions: (i) Is the system time-invariant or time-varying? (ii) Can we compute the outputs excited by the inputs u_3 and u_4 shown in Figures 2.10(c) and 2.10(d)?

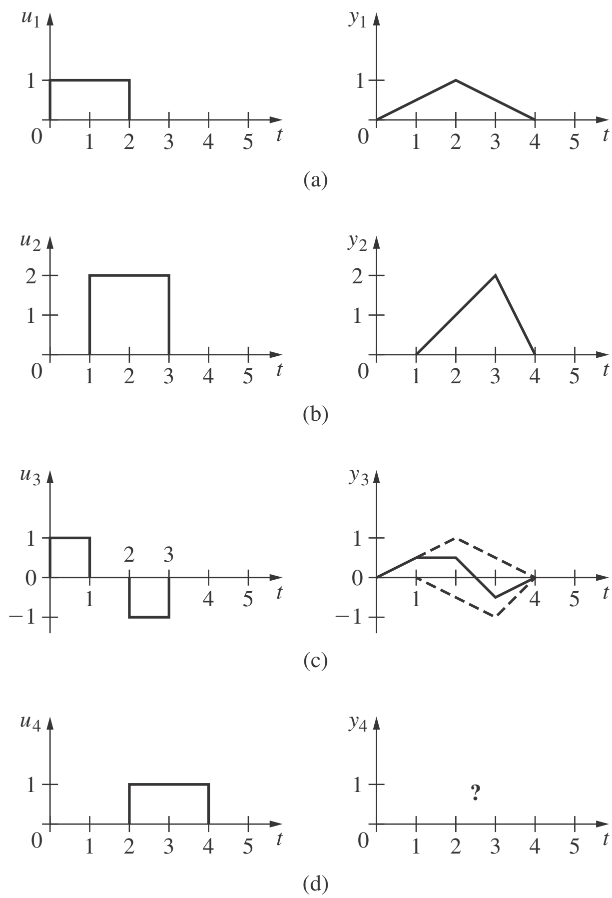


Figure 2.10 (**a**) and (**b**) Two given input–output pairs. (**c**) The output of u_3 can be obtained from (a) and (b) because $u_3(t) = u_1(t) - 0.5u_2(t)$ which involves no time shifting. (**d**) The output of u_4 cannot be obtained from (a) and (b) because the system is time-varying.

The system is linear; therefore if we double the input u_1 in Figure 2.10(a), then the output will be doubled. The input u_2 in Figure 2.10(b) is clearly the shifting of $2u_1$ by one second. If the system were time-invariant, then y_2 should equal $2y_1$ shifted by one second. This is not the case. Thus the linear system is a time-varying system.

The input u_3 in Figure 2.10(c) equals $u_1 - 0.5u_2$. Therefore the output y_3 should equal $y_1 - 0.5y_2$. The manipulation is carried out point by point (check the values at $t = 0, 1, 2, 3$, and 4). In computing the output in Figure 2.10(c), we use only the superposition property; no time shifting is used. Although the input u_4 is the shifting of u_1 by two seconds or the shifting of $0.5u_2$ by one second, because the system is time-varying, there is no way to deduce from Figures 2.10(a) and 2.10(b) the output excited by u_4.

EXERCISE 2.5.1

Which of the systems in Exercise 2.2.1 are time-invariant?

Answers

(a), (b), (c), and (d).

2.6 IMPLICATIONS OF LINEARITY AND TIME INVARIANCE—ZERO-STATE RESPONSES

This section studies some important implications of linearity and time invariance. Responses of LTI systems can always be decomposed into zero-input (natural) responses and zero-state (forced) responses. Recall that zero-input responses are responses excited by nonzero initial conditions at time $t = 0$. One may wonder, Where do the initial conditions come from? Although they may be caused by external disturbances or noise, mostly they are the residue of the input applied before $t = 0$. Thus in some sense, zero-input responses can be considered to be parts of zero-state responses. This is indeed the case under some minor conditions as we will discuss in Chapter 7. Thus we study in this section only zero-state responses. In other words, every system will be assumed to be initially relaxed at $t_0 = 0$. In practice, if the output of an LTI lumped system is identically zero before applying any input, then the system is initially relaxed.

For zero-state responses, because all initial states are zero, state input–output pairs can be reduced to input–output pairs. See Problems 1.1 and 1.2. The most important implication of linearity and time invariance is as follows: The characteristics of such a system can be completely determined from any one input–output pair.[2] In other words, from any input–output pair, we can deduce everything we want to know about the system. In particular, we can compute or predict from the known input–output pair the output excited by any input.

We discuss first memoryless systems. Consider the network with input $u(t)$ shown in Figure 2.11, taking from Reference 12. Suppose we are interested in the current $y(t)$ passing through

[2]There is a mathematically contrived exception. See Problems 2.18 and 2.19.

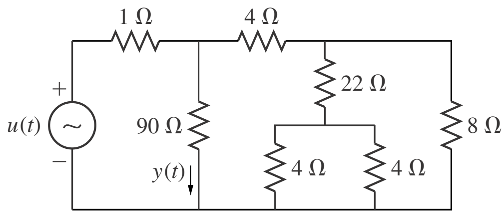


Figure 2.11 Resistive network.

the 90-Ω resistor. The network consists of only resistors with constant resistances and is therefore a memoryless LTI system. Such a system can be described by $y(t) = au(t)$ for some constant a. One way to find a is to carry out analysis. Alternatively, we may find a by measurement. If we apply a step input ($u(t) = 1$ for $t \geq 0$) and the current is measured as 0.01 A for $t \geq 0$ ($y(t) = 0.01$), then we have $a = 0.01/1 = 0.01$ and the input and output of the network is completely described by $y(t) = 0.01u(t)$. In conclusion, if an LTI system is memoryless, its input–output relationship can be obtained by one measurement.

For LTI systems with memory, the situation is more complex. We use an example to establish intuitively the assertion. We show that if we know the input–output pair of an LTI system shown in Figure 2.12(a), then we can compute or predict the output of the system excited by any input.

Let us denote the input and output in Figure 2.12(a) by $u_1(t)$ and $y_1(t)$. This is in fact the input–output pair shown in Figure 2.10(a). Because the system in Figure 2.10 is time-varying, there is no way to know the output excited by the u_4 shown in Figure 2.10(d). However, if the system is time-invariant, then its output is simply the shifting of $y_1(t)$ by two seconds as shown in Figure 2.12(b). That is, $u_1(t-2)$ excites $y_1(t-2)$. Similarly, the input $u_1(t-4)$ excites $y_1(t-4)$ as shown in Figure 2.12(c). Because

$$q(t) = u_1(t) + u_1(t-2) + u_1(t-4) + \cdots$$

the output of the system excited by the step input is given by

$$y_q(t) = y_1(t) + y_1(t-2) + y_1(t-4) + \cdots$$

as shown in Figure 2.12(d). The output $y_q(t)$ is called the *step response*.

Now suppose we are given an arbitrary input $u(t)$ as shown in Figure 2.13. The input signal can be approximated by a sequence of pulses as shown. See Section 1.5.1. Clearly, the narrower the pulses, the more accurate the approximation. Suppose we select the pulse width as 0.5. We first generate the response excited by the pulse at $t = 0$ with height 1 and width 0.5 from the step response shown in Figure 2.12(d). We shift the step function $q(t)$ to $t = 0.5$ and multiply it by -1 to yield the input $-q(t-0.5)$ as shown in Figure 2.12(e). Then the output equals $-y_q(t-0.5)$ as shown. Adding $q(t)$ and $-q(t-0.5)$ yields the pulse $p(t)$ shown in Figure 2.12(f). Its output $y_p(t)$ is the sum of $y_q(t)$ and $-y_q(t-0.5)$ as shown. Using $y_p(t)$ and the homogeneity, additivity, and time shifting properties, we can obtain approximately the output excited by the input shown in Figure 2.13. This justifies the assertion that the output excited by any input can be computed from the given input–output pair shown in Figure 2.12(a). This will be formally established in Chapters 3 and 6.

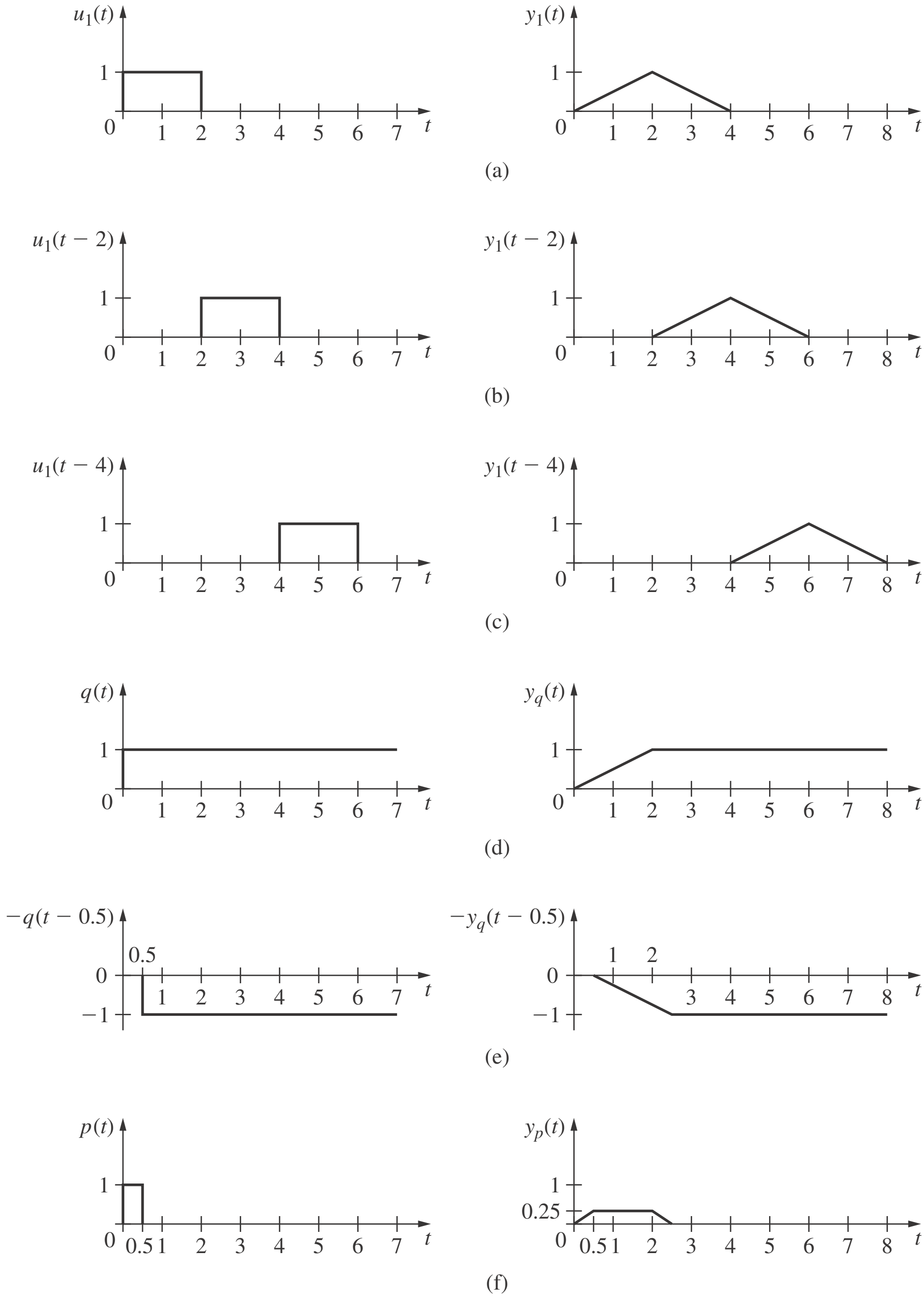


Figure 2.12 (**a**) A given input–output pair. (**b**)–(**f**) Their outputs all can be obtained from (a).

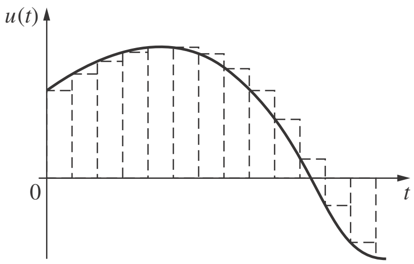


Figure 2.13 Input signal.

EXERCISE 2.6.1

Consider a linear time-invariant (LTI) system with its step response shown in Figure 2.14(a). Verify the input–output pairs shown in Figures 2.14(b) and 2.14(c).

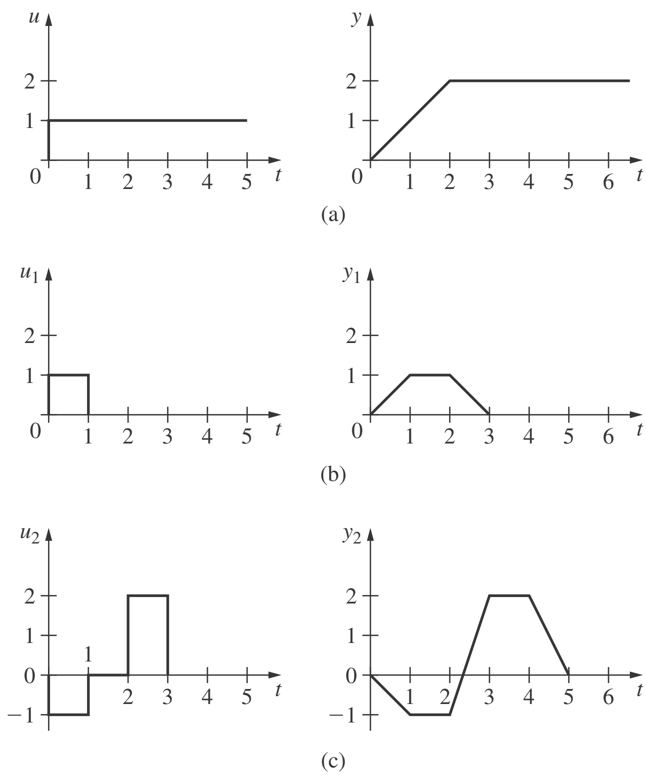


Figure 2.14 (**a**) A given input–output pair. (**b**) and (**c**) Develop their outputs from (a).

2.7 MODELING CT LTI LUMPED SYSTEMS

The concepts of linearity and time invariance are defined from state input–output pairs without referring to internal structures of systems. Thus the concepts are very powerful and applicable to any system, be it an electrical, mechanical, chemical, aeronautical, electromechanical, or biomedical system. If a system is known to be linear and time-invariant, then its characteristics can be determined by applying an arbitrary input and measuring its output. From the input–output pair, we can deduce everything we want to know about the system. Now the question is: How do we know a system is linear and time-invariant?

Linearity is defined from state input–output pairs. If we find some pairs that do not meet the additivity or homogeneity property, then we can conclude that the system is not linear. However, if the pairs meet the two properties, we still cannot conclude the system to be linear because the two properties must apply to all possible state input–output pairs. There are infinitely many of them. It is not possible to check them all. Thus there is no way to check linearity from the input and output terminals. Similarly, it is not possible to conclude the time invariance of a system because the shifting property must hold forever (from now to time infinity). Fortunately, linear and time-invariant systems are obtained by *modeling* as we discuss next. Furthermore, we develop only lumped models because distributed models are difficult to develop.

Most systems are built by interconnecting a number of subsystems or components. If every component is linear and time-invariant, then the overall system is linear and time-invariant. However, no physical component is linear and time invariant in the mathematical sense. For example, a resistor with resistance R is not linear because if the applied voltage is very large, the resistor may burn out. However, within its power limitation, the resistor can very well be considered to be a linear element. Likewise, disregarding saturation, the charge Q stored in a capacitor is related to the applied voltage v by $Q = Cv$, where C is the capacitance; the flux F generated by an inductor is related to its current i by $F = Li$, where L is the inductance. They are linear elements. The resistance R, capacitance C, and inductance L probably will change after 100 years, thus they are not time-invariant. However, their values will remain constant within a short time interval. Thus they may be considered as linear and time-invariant. If all resistors, capacitors, and inductors of RLC networks such as the ones in Figures 2.5 and 2.11 are so modeled, then the networks are linear and time-invariant.

We give a different example. Consider the mechanical system shown in Figure 2.15. It consists of a block with mass m connected to a wall by a spring. The input $u(t)$ is the force applied to

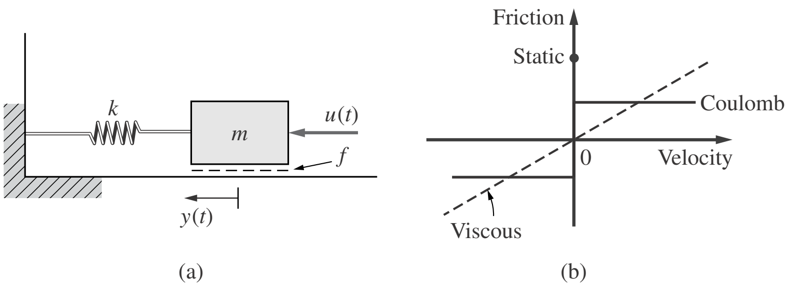


Figure 2.15 (**a**) Mechanical system. (**b**) Frictions.

the block, and the output $y(t)$ is the displacement measured from the equilibrium position. The spring is a nonlinear element because it will break if it is stretched over its elastic limit. However, within some limit, the spring can be described by Hooke's law as

$$\text{Spring force} = ky(t) \tag{2.20}$$

where k is called the *spring constant*. This is a linear element. The friction between the block and the floor is very complex and may consist of three parts: static, Coulomb, and viscous frictions as shown in Figure 2.15(b). Note that the coordinates are friction versus velocity. When the block is stationary (its velocity is zero), we need a certain amount of force to overcome its static friction to start its movement. Once the block is moving, there is a constant friction, called the Coulomb friction, to resist its movement. In addition, there is a friction, called the viscous friction, which is proportional to the velocity as shown with a dashed line in Figure 2.15(b) or

$$\text{Viscous friction} = f \times \text{Velocity} = f\dot{y}(t) \tag{2.21}$$

where $\dot{y}(t) = dy(t)/dt$ and f is called the *viscous friction coefficient* or *damping coefficient*. It is a linear relationship. Because of the viscous friction between its body and the air, a sky diver may reach a constant falling speed. If we disregard the static and Coulomb frictions and consider only the viscous friction, then the mechanical system can be modeled as a linear system within the elastic limit of the spring.

In conclusion, most physical systems are nonlinear and time-varying. However, within the time interval of interest and a limited operational range, we can model many physical systems as linear and time-invariant. Thus the systems studied in this text, in fact in most texts, are actually models of physical systems. Modeling is an important problem. If a physical system is properly modeled, we can predict the behavior of the physical system from its model. Otherwise, the behavior of the model may differ appreciably from that of the physical system.

2.8 IDEAL OPERATIONAL AMPLIFIERS

The operational amplifier, or op amp for short, is one of the most important circuit elements. It is built in integrated circuit (IC) form; it is small in size, inexpensive, and widely available. An op amp is usually represented as shown in Figure 2.16(a) and modeled as shown in Figure 2.16(b). It is a simplified model of the operational amplifier. In reality, the op amp is very complex and may contain over twenty transistors. It has more terminals than the three shown. There are terminals to be connected to power supplies—for example, with 15 and −15 volts—and terminals for balancing. These terminals are usually not shown. The op amp in Figure 2.16(a) has two input terminals and one output terminal. The one with a minus sign is called the *inverting terminal,* and the one with a plus sign is called the *noninverting terminal*. Let e_- and e_+ be the voltages at the inverting and noninverting terminals. Then the output voltage v_o is a function of $e_d := e_+ - e_-$, called the *differential input voltage*. If an op amp is modeled as a memoryless system, then v_o and e_d can be related as shown in Figure 2.16(c). For $|e_d|$ small, typically less than 0.1 mV (millivolt $= 10^{-3}$ volt), the output is a linear function of the differential input voltage. For $|e_d|$ large, the output voltages saturate at levels roughly one or two volts below the supply voltages. Clearly the op amp is a nonlinear system. We mention that the current i_o at the output terminal is generally limited—for example, 25 mA—and is provided by the power supplies.

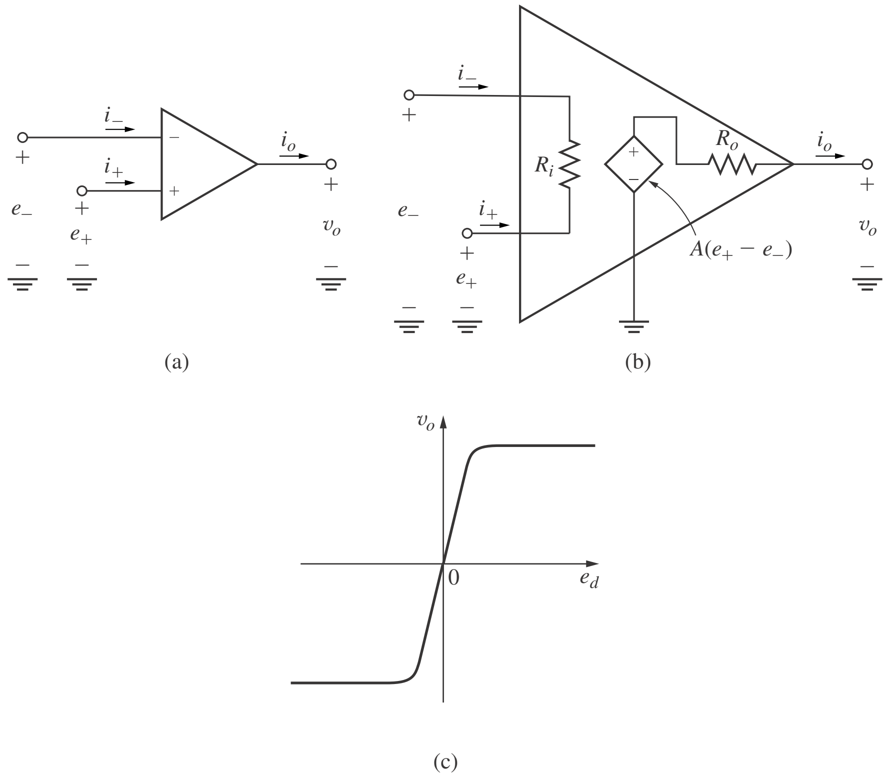


Figure 2.16 (**a**) Op amp. (**b**) Its model. (**c**) Its characteristic.

Some op amps are designed to operate in the saturation regions, such as comparators. If $e_d > \epsilon$, where ϵ is on the order of mV, then the output will remain in the positive saturation region. If $e_d < \epsilon$, then the output will remain in the negative saturation region. Using this property, an op amp can be used as a *comparator* to turn on or off a device. It can also be used to design many useful circuits. See References 8, 11, and 23.

Most op amps encountered in this text operate only in linear region. In this case, the output voltage and the differential input voltage are related by

$$v_o(t) = A(e_+(t) - e_-(t)) = Ae_d(t) \tag{2.22}$$

where A is called the *open-loop gain*. The resistor R_i in Figure 2.16(b) is the input resistance and R_o the output resistance. The resistance R_i is generally very large, greater than 10^4 Ω, and R_o is very small, less than 50 Ω. The open-loop gain A is very large, usually over 10^5.

In order for an op amp to operate in its linear region, we must introduce feedback from the output terminal to input terminals either directly or through resistors or capacitors.[3] We consider the simplest connection in Figure 2.17(a), in which the output terminal is connected to the

[3]Inductors are bulky and difficult to fabricate in integrated circuit form. Thus they are not used.

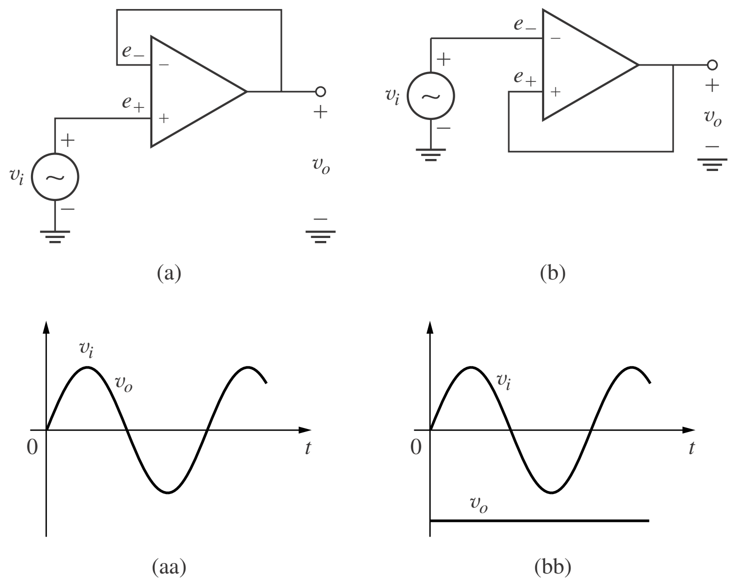


Figure 2.17 (**a**) Voltage follower. (**b**) Unstable system. (**aa**) Input and output of (a). (**bb**) Input and output of (b).

inverting input terminal and the input signal $v_i(t)$ is applied to the noninverting terminal, that is, $e_-(t) = v_o(t)$ and $v_i(t) = e_+(t)$. Using (2.22), we have immediately

$$v_o(t) = A(v_i(t) - v_o(t)) \quad \text{or} \quad (1 + A)v_o(t) = Av_i(t)$$

which implies, if $A = 10^5$,

$$v_o(t) = \frac{A}{1 + A} v_i(t) = 0.99999 v_i(t)$$

In practice, we can consider $v_o(t) = v_i(t)$ for all t. Thus the op-amp circuit is called a *voltage follower*. See References 6, 8, 11, and 23. For example, if we apply $v_i(t) = \sin 2t$, the output will be $v_o(t) = \sin 2t$ as shown in Figure 2.17(aa).

An inquisitive reader may ask: If $v_o(t) = v_i(t)$, why not connect $v_i(t)$ directly to the output instead of going through a voltage follower? In practice, connecting two devices or systems may affect each other and alter their descriptions. We give a simple example. Consider a 10-V voltage supply source with an internal resistance of 1 Ω. If no load is connected to the source, the terminal voltage is 10 V. If a load with resistance 4 Ω is connected to the terminal, the voltage source supplies a terminal voltage of $10 \times 4/(4 + 1) = 8$ V across the load. If the load resistance is 10 Ω, then the voltage source supplies a terminal voltage of $10 \times 10/(10 + 1) = 9.1$ V. Thus for different loads, the same voltage source yields different terminal voltages. This is called the *loading* effect. The effect is not desirable in connecting two devices and should be eliminated. Because an op amp has a large gain, large input resistance, and small output resistance, inserting a voltage follower between two devices can reduce or eliminate the loading effect. In other words, a voltage follower can act as a shield or a buffer. Thus voltage followers are widely used in practice. They are also called *buffers* or *isolating amplifiers*.

An op amp is called ideal if $A = \infty$, $R_i = \infty$, and $R_o = 0$. If the output voltage is limited to the supply voltages, say ± 15, then (2.22) implies

$$e_+ = e_- \tag{2.23}$$

Thus the two input terminals are *virtually short*. If $R_i = \infty$, then we have

$$i_- = -i_+ = \frac{e_- - e_+}{R_i} = 0 \tag{2.24}$$

Thus the two input terminals are *virtually open*. The op amp is a unique device whose input terminals are virtually short and open at the same time. It is important to mention that the equalities in (2.23) and (2.24) are not in the mathematical sense. They are only roughly equal. For example, for a typical op amp with $A = 10^5$ and $R_i = 10^4$. If the supply voltages are ± 15 V, then all variables including the output are limited to ± 15. Thus (2.22) implies $|e_+ - e_-| \leq 15/10^5 = 0.00015$. If $R_i = 10^4$, then the current $i_- = -i_+$ is less than $15/(10^5 \times 10^4)$. Thus the two properties in (2.23) and (2.24) roughly hold in practice. Using the two properties, analysis and design of op-amp circuits can be greatly simplified. For example, for the voltage follower in Figure 2.17(a), because $v_i = e_+$ and $v_o = e_-$, (2.23) implies $v_o(t) = v_i(t)$. There is no need to carry out any computation.

2.8.1 DAC and ADC

This subsection uses the ideal op-amp model to build a digital-to-analog converter (DAC).[4] Consider the op-amp circuit shown in Figure 2.18(a). It consists of four latches, four transistor switches, five resistors with resistances shown, an op amp, and a precision voltage reference E. The 4-bit input, denoted by $\{x_1\ x_2\ x_3\ x_4\}$, latched in the register, controls the closing and opening of the switches. A switch is closed if $x_i = 1$, open if $x_i = 0$. The output voltage is denoted by v_o. Because the noninverting terminal is grounded, we have $e_+ = 0$ and, consequently, $e_- = 0$. Thus we have $i_f = v_o/R$ and $i_k = x_k E/(2^k R)$, for $k = 1:4$. Kirchhoff's current law implies that the algebraic sum of all currents at the inverting terminal is 0. Thus we have

$$i_f = i_1 + i_2 + i_3 + i_4$$

or

$$\frac{v_o}{R} = \left(\frac{x_1 E}{2R} + \frac{x_2 E}{4R} + \frac{x_3 E}{8R} + \frac{x_4 E}{16R}\right)$$

which implies

$$v_0 = \left(2^{-1}x_1 + 2^{-2}x_2 + 2^{-3}x_3 + 2^{-4}x_4\right) E$$

For example, if $\{x_1\ x_2\ x_3\ x_4\} = \{1\ \ 1\ \ 0\ \ 1\}$ and $E = 10$, then

$$\begin{aligned} v_o &= (1 \times 2^{-1} + 1 \times 2^{-2} + 0 \times 2^{-3} + 1 \times 2^{-4}) \times 10 \\ &= (0.5 + 0.25 + 0.0625) \times 10 = 8.125 \end{aligned}$$

---

[4]This subsection may be skipped without loss of continuity.

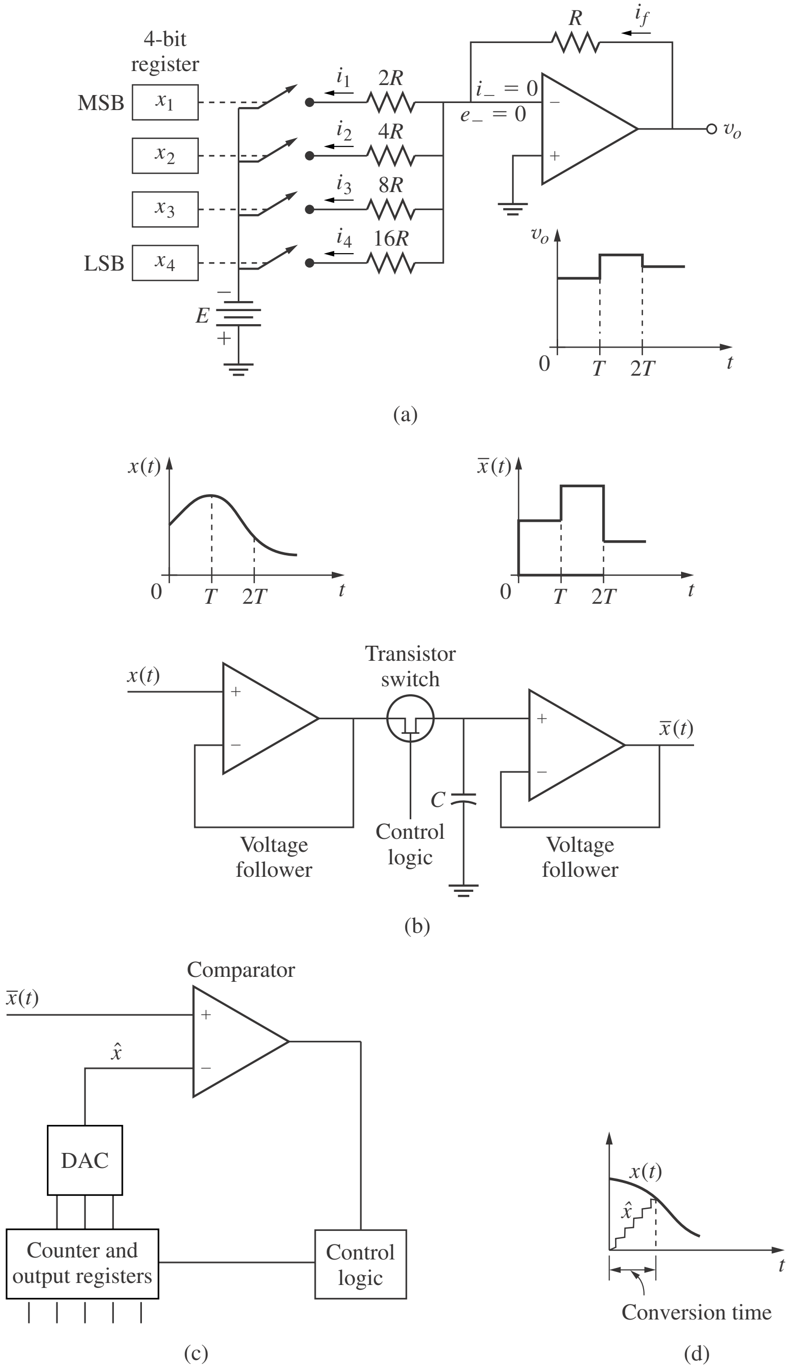


Figure 2.18 (**a**) 4-bit DAC. (**b**) Sample-and-hold circuit. (**c**) ADC. (**d**) Conversion time.

The output of the op amp will hold this value until the next set of binary input arrives. Thus the output of the DAC is stepwise as shown in Figure 2.18(a). The discontinuities of the output can be smoothed by passing it through an analog lowpass filter.

We next discuss the conversion of analog signals into digital signals. An analog signal $x(t)$ is first sampled at sampling instants as shown in Figure 2.18(b). It is carried out using a transistor switch. The voltage $\bar{x}(t)$ at each sampling instant is held in the capacitor C shown. To eliminate the loading problem, two voltage followers are used to shield the capacitor as shown. The capacitor will hold the voltage $\bar{x}(t)$ until the next sampling instant. Thus the circuit is called a *sample-and-hold circuit*. The conversion of the voltage into digital form is carried out using the arrangement shown in Figure 2.18(c). It consists of an op amp that acts as a comparator, a DAC, a counter and output registers, and a control logic. At each sampling instant, the counter starts to drive the DAC. Its output $\hat{x}$ is compared with the signal $\bar{x}$ to be converted. The counter stops as soon as the DAC output $\hat{x}$ equals or exceeds $\bar{x}$. The value of the counter is then transferred to the output register and is the digital representation of the analog signal $\bar{x}$. Because it takes time to carry out each conversion, if the analog signal $x(t)$ changes rapidly, the value converted may not be the value at the sampling instant as shown in Figure 2.18(d). For this reason, we need the sample-and-hold circuit, which is often built as a part of ADC. In the ADC, we see the use of three op amps; one is used as a nonlinear element and two are used as linear elements.

2.8.2 A More Realistic Model

Consider the voltage follower shown in Figure 2.17(a). Using (2.23), we readily obtain $v_o(t) = v_i(t)$. Now we consider the op-amp circuit shown in Figure 2.17(b) in which the input signal $v_i(t)$ is applied to the inverting terminal and the output $v_o(t)$ is connected to the noninverting terminal. Using (2.23), we also obtain $v_o(t) = v_i(t)$. However, if we apply $v_i(t) = \sin 2t$, the output $v_o(t)$ is not $\sin 2t$. Instead the output goes to the negative saturation region as shown in Figure 2.17(bb). Thus the circuit cannot be used as a voltage follower.

Now the question is: Why can the op-amp circuit in Figure 2.17(a) be used, but not the one in Figure 2.17(b)? To answer the question, we need the concepts of stability and model reduction which we will introduce in Chapters 6 and 8. As a preview, we discuss briefly the stability issue. The output in Figure 2.17(a) is fed back to the inverting terminal (the terminal with a negative sign), thus it is called negative feedback. The output in Figure 2.17(b) is fed back to the noninverting terminal (the terminal with a positive sign), thus it is called positive feedback. It is often said that negative feedback stabilizes a system and positive feedback destabilizes a system. This happens to be true for the circuits in Figure 2.17. Thus the op-amp circuit in Figure 2.17(a) is stable and can be used in practice, whereas the op-amp circuit in Figure 2.17(b) is not stable and cannot be used in practice. However, in general, a negative feedback system can be unstable and a positive feedback system can be stable. See Problems 6.17 and 8.17 through 8.19. Thus it is important to analyze the stability of the circuits in Figure 2.17. In order to do so, we need a more realistic model for the op amp as we discuss next.

Consider (2.22) and the characteristic shown in Figure 2.16(c). There we have modeled the op amp to be memoryless. In reality, the op amp has memory. By this, we mean that the output $v_o(t_0)$ depends on $e_d(t) = e_+(t) - e_-(t)$ not only at $t = t_0$ but also for $t < t_0$. In this case, (2.22) is no longer valid. Instead, (2.22) must be replaced by an integration equation such as

$$v_o(t) = \int_{\tau=0}^{t} h(t-\tau)e_d(\tau)\,d\tau \tag{2.25}$$

where $h(t)$ is some function, or by a differential equation such as

$$\frac{d^2v_o(t)}{dt^2} + a_2\frac{dv_o(t)}{dt} + a_3v_o(t) = b_1e_d(t) \tag{2.26}$$

for some constants a_2, a_3, and b_1, or by an algebraic equation such as

$$V_o(s) = A(s)E_d(s) \tag{2.27}$$

Equations (2.25) and (2.26) will be developed in the next chapter, and (2.27) will be developed in Chapter 6. Using a more realistic model, we will formally answer in Chapter 8 the posed question.

2.9 IDEAL DIODES AND RECTIFIERS

This section discusses a memoryless time-invariant nonlinear element that will be used in Section 8.3.[5] Consider the *diode* with voltage $v(t)$ and current $i(t)$ shown in Figure 2.19(a). Its ideal v–i characteristic is shown in Figure 2.19(b). If the diode is *forward biased* or $v(t) > 0$, the diode acts as short circuit and the current $i(t)$ will flow. If the diode is *reverse biased* or $v(t) < 0$, the diode acts as open circuit and no current can flow. Thus the current of a diode can flow only in one direction. A more realistic v–i characteristic is shown in Figure 2.19(c). A practical diode will conduct only if $v(t)$ is larger than a *threshold voltage* v_c and has a voltage drop across the diode when it is conducting. It will also break down; that is, the current also flows in the reverse direction if $v(t) < -M$, for some large M. Using the ideal model, analysis can be simplified.

We use ideal diodes to discuss half-wave and full-wave rectifiers. Consider the circuit shown in Figure 2.20(a). If the input voltage v_i is positive, the diode conducts and acts as a short circuit. Thus we have $v_o = v_i$. If v_i is negative, the diode is reverse biased and acts as an open circuit. Thus we have $v_o = 0$. Therefore v_i and v_o can be related as shown in Figure 2.20(b).

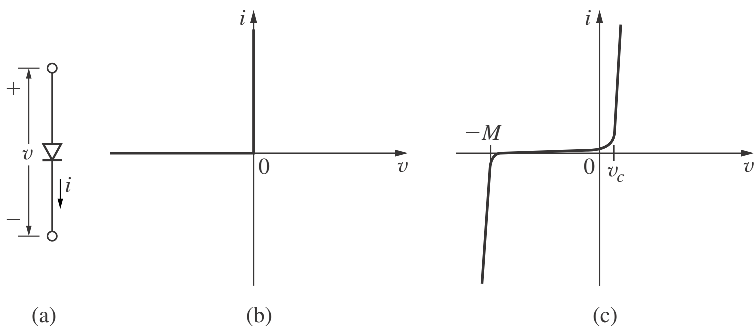


Figure 2.19 (**a**) Diode. (**b**) Ideal characteristic. (**c**) Realistic characteristic.

[5]This section may be skipped without loss of continuity.

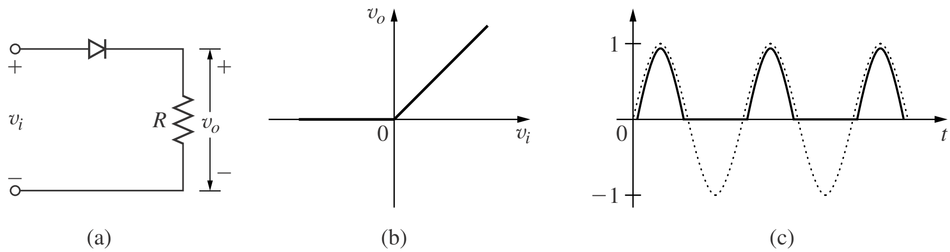


Figure 2.20 (**a**) Half-wave rectifier. (**b**) Its characteristic. (**c**) Input $v_i(t) = \sin \omega_0 t$ (dotted line) and output (solid line).

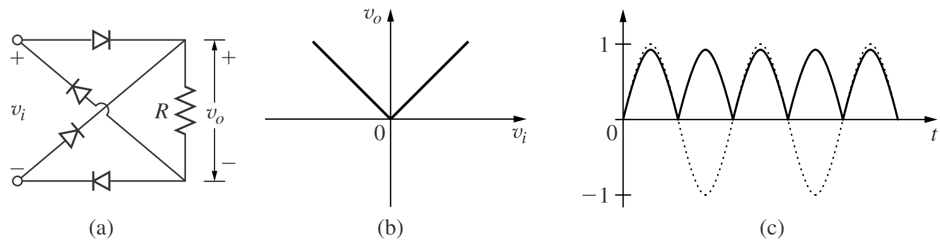


Figure 2.21 (**a**) Full-wave rectifier. (**b**) Its characteristic. (**c**) Input $v_i(t) = \sin \omega_0 t$ (dotted line) and output $v_o(t) = |\sin \omega_0 t|$ (solid line).

If $v_i(t) = \sin \omega_0 t$ as shown in Figure 2.20(c) with a dotted line, then the output $v_o(t)$ is as shown in Figure 1.20(c) with a solid line. The circuit is called a *half-wave rectifier* and can be used to transform an ac voltage into a dc voltage. It can also function as a switch, as we will discuss in Example 4.2.5.

Consider now the circuit shown in Figure 2.21(a). If $v_i > 0$, the diode conducts along the outer loop and $v_o = v_i$. If $v_i(t) < 0$, the diode conducts through the two diagonal branches and $v_o = -v_i > 0$. Thus the input and output can be related as shown in Figure 2.21(b). If $v_i(t) = \sin \omega_0 t$ as shown in Figure 2.21(c) with a dotted line, then the output is $v_o(t) = |\sin \omega_0 t|$ as shown in Figure 2.21(c) with a solid line. The circuit is called a *full-wave rectifier* and is more efficient than the half-wave rectifier in converting an ac voltage into a dc voltage.

2.10 DISCRETE-TIME LTI SYSTEMS

In the remainder of this chapter, we develop the discrete-time counterpart of what has been discussed for CT systems. Because most concepts are directly applicable, the discussion will be brief.

A system is called a discrete-time (DT) system if an input sequence $u[n]$ excites a unique output sequence $y[n]$. A DT system is called a memoryless system if its output $y[n_0]$ depends only on the input $u[n_0]$ applied at the same sampling instant. Otherwise it is a DT system with memory. Thus the output $y[n]$ of a DT system with memory depends on past, current, and future inputs. It is causal or nonanticipatory if the output does not depend on future input. Thus the output of a causal system depends only on past and current inputs. The past input should go all the way back to $n = -\infty$. If we introduce the state $\mathbf{x}[n_0]$ at time instant n_0, then its knowledge and the input $u[n]$, for $n \geq n_0$, determine uniquely the output $y[n]$, for $n \geq n_0$, denoted as

$$\left.\begin{array}{l}\mathbf{x}[n_0] \\ u[n], \quad n \geq n_0\end{array}\right\} \rightarrow y[n], \quad n \geq n_0 \tag{2.28}$$

The vector $\mathbf{x}$ is called the state, and its components are called state variables. Equation (2.28) is called a state input–output pair. If the number of state variables is finite, the DT system is a lumped system. Otherwise, it is a distributed system.

A DT system is linear if for *any* two state input–output pairs

$$\left.\begin{array}{l}\mathbf{x}_i[n_0] \\ u_i[n], \quad n \geq n_0\end{array}\right\} \rightarrow y_i[n], \quad n \geq n_0 \tag{2.29}$$

for $i = 1, 2$, we have

$$\left.\begin{array}{l}\mathbf{x}_1[n_0] + \mathbf{x}_2[n_0] \\ u_1[n] + u_2[n], \quad n \geq n_0\end{array}\right\} \rightarrow y_1[n] + y_2[n], \quad n \geq n_0 \qquad \text{(additivity)} \tag{2.30}$$

and

$$\left.\begin{array}{l}\alpha\mathbf{x}_1[n_0] \\ \alpha u_1[n], \quad n \geq n_0\end{array}\right\} \rightarrow \alpha y_1[n], \quad n \geq n_0 \qquad \text{(homogeneity)} \tag{2.31}$$

for *any* constant α. These two conditions are similar to (2.16) and (2.17) in the CT case. Thus the concept of linearity in DT systems is identical to the concept in the CT systems and all discussion for CT systems are directly applicable to DT systems. For example, the response of every DT linear system can be decomposed as

$$\text{Total response} = \text{Zero-state (forced) response} + \text{Zero-input (natural) response} \tag{2.32}$$

and zero-state responses have the additivity and homogeneity properties. So have zero-input responses.

A DT system is time invariant if its characteristics do not change with time. Mathematically, it can be stated as follows. For any state input–output pair

$$\left.\begin{array}{l}\mathbf{x}_i[n_0] = \mathbf{x}_0 \\ u_i[n], \quad n \geq n_0\end{array}\right\} \rightarrow y_i[n], \quad n \geq n_0 \tag{2.33}$$

and for any integer n_1, we have

$$\left.\begin{array}{l}\mathbf{x}_i[n_0 + n_1] = \mathbf{x}_0 \\ u_i[n - n_1], \quad n \geq n_0 + n_1\end{array}\right\} \rightarrow y_i[n - n_1], \quad n \geq n_0 + n_1 \qquad \text{(time shifting)} \tag{2.34}$$

This is called the time shifting property. Thus, for a time-invariant DT system, the response will always be the same no matter when the input sequence and initial state are applied. Thus we may assume, without loss of generality, that $n_0 = 0$ for time-invariant DT systems.

EXAMPLE 2.10.1

Consider a savings account in a bank. Let $u[n]$ be the amount of money deposited on the nth day and let $y[n]$ be the total amount of money in the account at the end of the nth day. Strictly speaking, $u[n]$ and $y[n]$ are digital signals. If we consider them as DT signals, then the savings account can be considered as a DT system. If the interest rate changes with time, then the system is a time-varying system. If the interest rate depends on the total amount of money in the account, then the system is a nonlinear system. If the interest rate is fixed and is the same no matter what $y[n]$ is, then the savings account is an LTI system.

As in the CT case, under some minor conditions, zero-input responses of DT LTI lumped systems can be considered as parts of zero-state responses. Thus there is no need to study zero-input responses and we may concentrate on zero-state responses. In this case, the output is excited exclusively by the applied input and a state input–output pair reduces to an input–output pair. The zero-state response of a DT LTI system excited by any input can be obtained from any single input–output pair. The situation is identical to the CT case.

EXAMPLE 2.10.2

Consider a DT LTI system. Suppose the zero-state response of the system excited by $u[n] = \delta[n]$ is as shown in Figure 2.22(a). Because the input is an impulse sequence, the output is called the *impulse response* and will be denoted by $h[n]$. We show that the output of the system excited by any input sequence can be computed from the impulse response. For example, consider the input sequence shown in Figure 2.22(b). The input

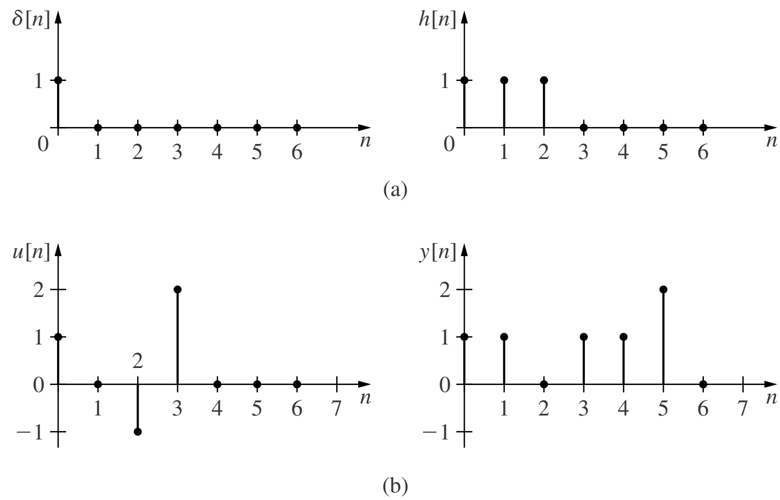


Figure 2.22 (**a**) Impulse response. (**b**) Output excited by $u[n]$.

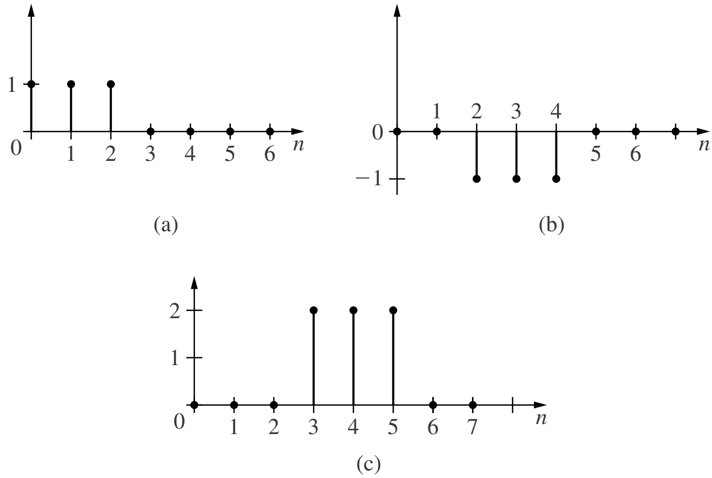


Figure 2.23 Outputs excited by (**a**) $\delta[n]$, (**b**) $-\delta[n-2]$, and (**c**) $2\delta[n-3]$.

can be expressed as

$$u[n] = \delta[n] - \delta[n-2] + 2\delta[n-3] \tag{2.35}$$

The output excited by $\delta[n]$ is repeated in Figure 2.23(a). The output excited by $\delta[n-2]$ is the shifting of the impulse response by two samples and then multiplied by -1 as shown in Figure 2.23(b). The output excited by $2\delta[n-3]$ is the shifting of the impulse response by three samples and then multiplied by 2 as shown in Figure 2.23(c). Their sum is shown in Figure 2.22(b) and is the output excited by the input in (2.35). This is a graphical method of computing the output and reveals the basic ideas involved. They can also be expressed mathematically as follows:

$$\begin{aligned}
\delta[n] &\rightarrow \delta[n] + \delta[n-1] + \delta[n-2] \\
-\delta[n-2] &\rightarrow -\delta[n-2] - \delta[n-3] - \delta[n-4] \\
2\delta[n-3] &\rightarrow 2\delta[n-3] + 2\delta[n-4] + 2\delta[n-5]
\end{aligned}$$

The sum of the three left-hand-side terms is the input in (2.35). Thus its output is the sum of the three right-hand-side terms and is given by

$$y[n] = \delta[n] + \delta[n-1] + \delta[n-3] + \delta[n-4] + 2\delta[n-5] \tag{2.36}$$

as shown in Figure 2.22(b).

2.11 CONCLUSION

Systems can be classified dichotomously as shown in Figure 2.24. We use LTI systems to denote linear time-invariant systems and LTIL systems to denote LTI lumped systems. Most problems in LTIL systems have been solved, and various analysis and design methods are now available. This text studies mainly the analysis part of LTIL systems and touches only slightly on LTI distributed systems.

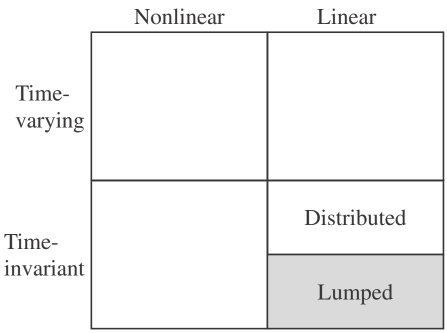


Figure 2.24 Classification of systems.

The class of systems studied in this text constitute only a small part of systems. However, many physical systems can be so modeled under limited time intervals and operational ranges. Thus its study is of fundamental importance. Its study is also a prerequisite for studying nonlinear and/or time varying systems.

PROBLEMS

2.1 Consider a CT LTI system. Let y_i be the output excited by the input u_i and initial state $\mathbf{x}_i(0)$, for $i = 1, 2, 3$. If $\mathbf{x}_1(0) = \mathbf{x}_2(0) = \mathbf{x}_3(0) = \mathbf{a} \neq \mathbf{0}$, which of the following statements are correct?

(a) If $u_3 = u_1 + u_2$, then $y_3 = y_1 + y_2$.

(b) If $u_3 = 0.5(u_1 + u_2)$, then $y_3 = 0.5(y_1 + y_2)$.

(c) If $u_3 = u_1 - u_2$, then $y_3 = y_1 - y_2$.

2.2 In Problem 2.1, if $\mathbf{x}_1(0) = \mathbf{x}_2(0) = \mathbf{x}_3(0) = \mathbf{0}$, which of the three statements are correct? Give your reasons.

2.3 Consider a CT system whose input and output are related by

$$y(t) = \begin{cases} u^2(t)/u(t-1) & \text{if } u(t-1) \neq 0 \\ 0 & \text{if } u(t-1) = 0 \end{cases}$$

Show that the system satisfies the homogeneity property but not the additivity property. Thus the homogeneity property does not imply the additivity property.

2.4 Show that if $\{u_1 + u_2\} \rightarrow \{y_1 + y_2\}$, then

$$\left\{\frac{a}{b}u_1\right\} \rightarrow \left\{\frac{a}{b}y_1\right\}$$

for any integers a and b. Thus the additivity property implies $\{\alpha u_1\} \rightarrow \{\alpha y_1\}$, for any rational number α. In other words, the additivity property *almost* implies the homogeneity property.

2.5 Discuss whether or not each of the following equations is memoryless, linear, time-invariant, and causal:

(a) $y(t) = -2 + 3u(t)$

(b) $y(t) = \sqrt{u(t)}$

(c) $y(t) = u(t)u(t-1)$

(d) $y(t) = tu(t)$

(e) $y(t) = \int_{t_0}^{t} u(\tau)\,d\tau + y(t_0)$

(f) $y(t) = \int_{t_0}^{t} \tau u(\tau)\,d\tau + y(t_0)$

2.6 Consider a CT system whose input and output are related by

$$y(t) = \begin{cases} u(t) & \text{for } t \le b \\ 0 & \text{for } t > b \end{cases}$$

for a fixed b. This is called a *truncation* operator because it truncates the input after time b. Is it memoryless? Linear? Time-invariant? Lumped?

2.7 Consider a CT linear system. Its zero-state responses excited by u_1 and u_2 are shown in Figure 2.25(a). Is the system time-varying or time-invariant? Can you find the zero-state responses excited by the inputs u_3, u_4, and u_5 shown in Figure 2.25(b)?

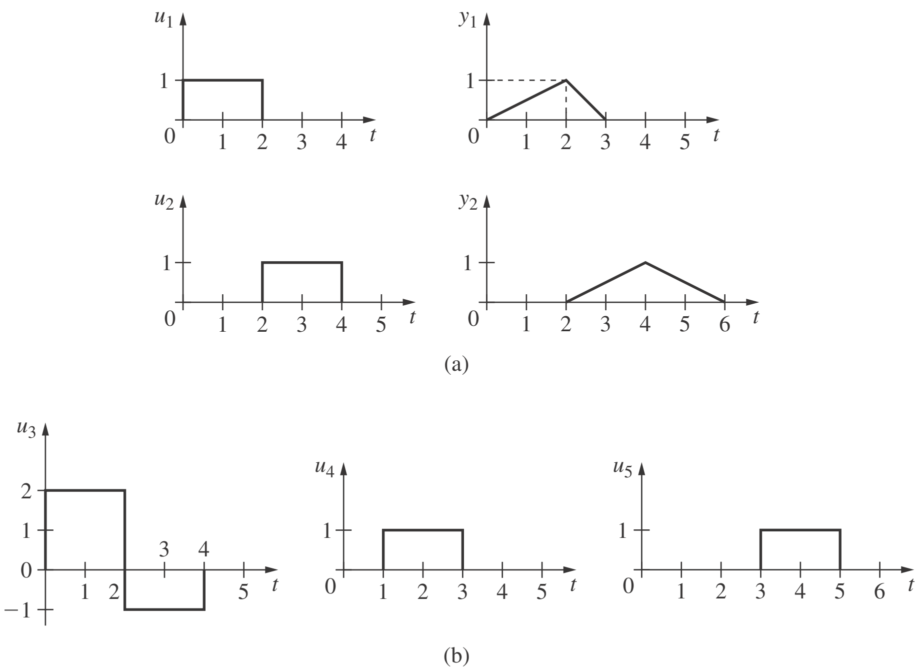


Figure 2.25

2.8 Consider a CT LTI system. Given the input and output pair in Figures 2.26(a) and 2.26(b). Find the output excited by the input shown in Figure 2.26(c).

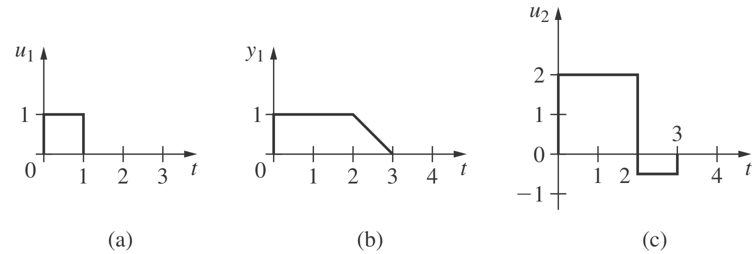


Figure 2.26

2.9 Consider a CT LTI system. Suppose its step response (the zero-state response excited by a step input or $u(t) = q(t)$) is as shown in Figure 2.27(a). Find the outputs excited by the inputs shown in Figures 2.27(b) through 2.27(d).

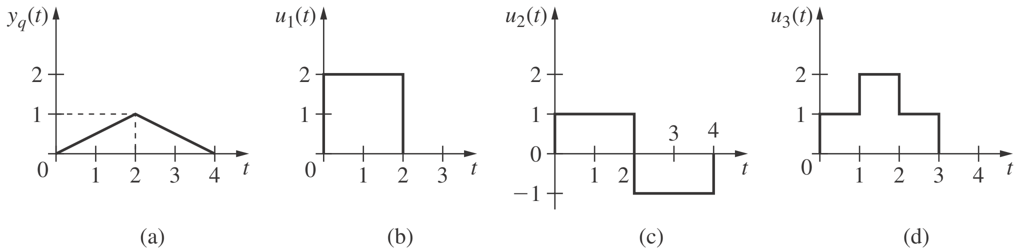


Figure 2.27

2.10 Consider the system shown in Figure 2.28 in which the diodes are assumed to be ideal. Are u and y linearly related? Other than this type of purposely contrived example, generally, a system is linear and time-invariant if and only if all its subsystems are linear and time-invariant.

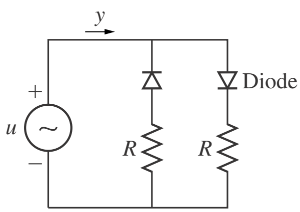


Figure 2.28

2.11 Use (2.22) with $A = 10^5$ to show that the output $v_o(t)$ of the op-amp circuit in Figure 1.17(b) roughly equal the input $v_i(t)$.

2.12 Consider the op-amp circuit shown in Figure 2.29(a). Show $v_o = -(R_2/R_1)v_i$. It is called an *inverting amplifier*. If $R_1 = R_1 = R$ as shown in Figure 2.29(b), then $v_o(t) = -v_i(t)$ and the circuit is called an *inverter*.

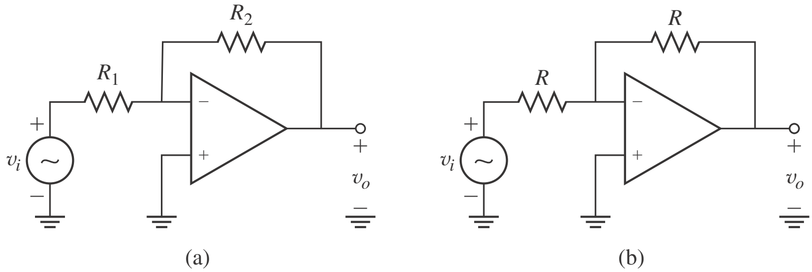


Figure 2.29

2.13 Verify $v_o = -(R_2/R_1)v_i$ for the op-amp circuit shown in Figure 2.30(a).[6] The circuit however cannot be used as an inverting amplifier. See Problem 8.2. Verify that the input and output of the circuit in Figure 2.30(b) are related by

$$v_o(t) = \left(1 + \frac{R_2}{R_1}\right) v_i(t)$$

It is a noninverting amplifier and can be used in practice. See Problem 8.3.

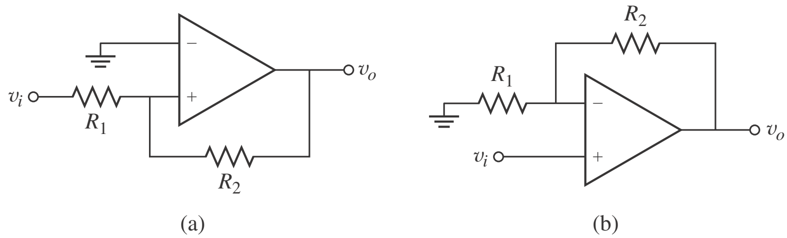


Figure 2.30

2.14 Consider the op-amp circuit shown in Figure 2.31. Verify that the voltage $v(t)$ and current $i(t)$ shown are related by $v(t) = -Ri(t)$. Thus the circuit can be used to implement a resistor with a negative resistance.

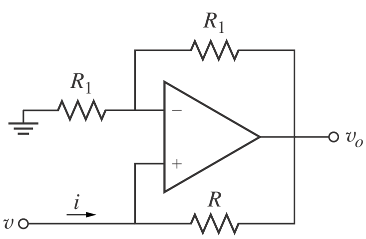


Figure 2.31

[6]All voltages in op-amp circuits are with respect to ground as in Figure 2.29. For simplicity, they will not be so indicated in the remainder of this text.

2.15 Suppose the impulse response of a DT LTI system is as shown in Figure 2.22(a). What are the outputs excited by the input sequences $u_1[n]$ and $u_2[n]$ shown in Figure 2.32?

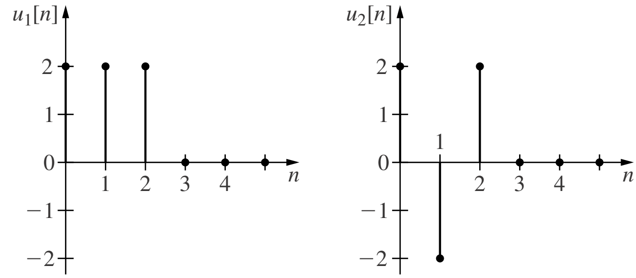


Figure 2.32

2.16 Consider a DT LTI system with the input-output pair shown in Figure 2.33. What is its step response (the output excited by the step sequence $q[n]$)? What is its impulse response (the output excited by the impulse sequence $\delta[n]$)?

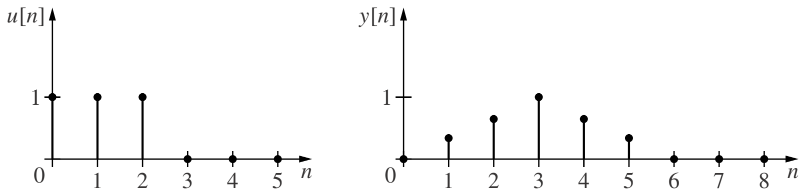


Figure 2.33

2.17 Consider the memoryless system shown in Figure 2.6(b) whose input u and output y are not linearly related. Now let us introduce a new output $\bar{y}$ defined as $\bar{y} = y - b$. What is the equation relating u and $\bar{y}$? Is it linear? This technique is often used to model nonlinear systems. See Problems 7.22 and 7.23.

2.18 Consider a "system" whose output at t is the algebraic sum of the amounts of all continuities of the input from $-\infty$ to t. See Reference 14 (p. 3). For example, if $u_1(t) = \sin 2t$, then $y_1(t) = 0$ because $\sin 2t$ has no discontinuity. If $u_2(t) = q(t)$, then $y_2(t) = q(t)$. Show that the "system" is linear and time-invariant. However, for this LTI "system," it is not possible to obtain $\{u_2 \rightarrow y_2\}$ from $\{u_1 \rightarrow y_1\}$.

2.19 Consider the "system" defined in Problem 2.18. Show that its impulse response is $\delta(t)$ if we define the impulse as the limit of the pulse in Figure 1.16(a) as $a \rightarrow 0$. Show that its impulse response is $-q(t - 2a)/a$ as $a \rightarrow 0$ if we use the triangular pulse in Figure 1.16(b) to define the impulse or $q(t - 2a)/a$ if we use the triangular pulse in Figure 1.16(c). In order to qualify as a system, every input must excite a unique output. The "system" does not have a unique impulse response and is, therefore, not a system as defined in this text.

CHAPTER 3

Convolutions, Difference, and Differential Equations

3.1 INTRODUCTION

The study of systems consists of four steps: modeling, developing mathematical equations, analysis, and design. We developed in Chapter 2 the class of linear time-invariant (LTI) causal systems to be used in modeling. Now we shall develop mathematical equations to describe them. The response of such a system can always be decomposed as

$$\text{Total response} = \text{Zero-state response} + \text{Zero-input response}$$

Furthermore, the zero-state response and the zero-input response can be studied separately. We first develop an equation, called a convolution, that describes only zero-state responses and then develop a difference or differential equation that describes both zero-state and zero-input responses. Because the mathematics involved in DT systems is much simpler, we first develop equations for DT systems and then develop equations for CT systems. As discussed in Section 1.6, in manipulation of DT signals, the sampling period does not play any role. Thus we suppress the sampling period and use only the time index in this chapter.

3.1.1 Preliminary

Every DT LTI system without memory can be described by $y[n] = au[n]$ for some constant a. The equation is applicable at every sampling instant and can be simply written as $y = au$ without specifying the sampling instant. Such a relationship can be described by a straight line with slope a passing through the origin of the input–output plane as shown in Figure 2.6(c). Thus the study of such a system is very simple.

We now consider a DT system that is causal and has memory. The current output $y[n]$ of such a system depends on not only the current input $u[n]$ but also past inputs $u[n-k]$, for $k = 1, 2, \ldots,$ for example, such as

$$y[n] = 2u[n] + 3u[n-1] - 1.5u[n-3] \tag{3.1}$$

for all $n \geq 0$. Here we assume that the system is initially relaxed at $n = 0$ or, equivalently, $u[n] = 0$ and $y[n] = 0$ for all $n < 0$. Thus the equation describes only the zero-state or forced response of the system. We claim that the system is linear. Indeed, for any two input–output pairs $\{u_1[n] \to y_1[n]\}$ and $\{u_2[n] \to y_2[n]\}$, we can verify, by direct substitution,

$\{u_1[n] + u_2[n] \rightarrow y_1[n] + y_2[n]\}$ and $\{\alpha u_1[n] \rightarrow \alpha y_1[n]\}$. Thus the system has the additivity and homogeneity properties and is therefore linear. The system is characterized by the three parameters 2, 3, and -1.5. They are constant (independent of time), thus the system is time-invariant. Thus the system described by (3.1) is linear and time-invariant (LTI). The expression is called a *linear combination* of $u[k]$ with constant parameters. Unlike a memoryless system where we can suppress the sampling instant, for a system with memory the segment of past inputs will be different for different n. Thus the sampling instant n must appear explicitly in the equation. Such a relationship can no longer be described by a straight line as in Figure 2.6(c) or any graph. Thus the only way to describe LTI systems is to use mathematical equations.

We show in this chapter that every DT system that is LTI, causal, and initially relaxed at $n = 0$ can be described by an equation such as the one in (3.1). In order to do so, we introduce first the concept of impulse responses.

3.2 DT IMPULSE RESPONSES

Consider a DT LTI system that is initially relaxed at $n = 0$. Let $h[n]$ be the output excited by the input $u[n] = \delta[n]$. Because the input is an impulse sequence, this particular output is called the *impulse response,* that is,

$$h[n] = \text{Impulse response} = \text{Output excited by the input } \delta[n]$$

Clearly, different DT systems have different impulse responses.

EXAMPLE 3.2.1

Consider the savings account discussed in Example 2.10.1. We assume that the interest rate is fixed and is the same no matter how much money is in the account. Under this assumption, the account is an LTI system. To simply the discussion, we assume that the interest rate is 0.01% per day and compounded daily. If we deposit one dollar the first day ($n = 0$) and none thereafter (the input $u[n]$ is an impulse sequence), then we have

$$y[0] = 1$$
$$y[1] = y[0] + y[0] \times 0.0001 = 1 + 0.0001 = 1.0001$$

and

$$\begin{aligned} y[2] &= y[1] + y[1] \times 0.0001 = 1.0001 + 1.0001 \times 0.0001 \\ &= 1.0001(1 + 0.0001) = 1.0001^2 \end{aligned}$$

In general, we have $y[n] = (1.0001)^n$. This particular output is called the impulse response, that is,

$$h[n] = (1.0001)^n \tag{3.2}$$

for $n = 0, 1, 2, \ldots$.

EXAMPLE 3.2.2

Consider the DT system described by (3.1). Its impulse response can be computed by direct substitution. If $u[n] = \delta[n]$, then $u[0] = 1$ and $u[n] = 0$ for $n \neq 0$. From (3.1), we have

$$\begin{aligned}
n = 0: &\quad y[0] = 2u[0] + 3[-1] - 1.5u[-3] = 2 \cdot 1 + 3 \cdot 0 - 1.5 \cdot 0 = 2 \\
n = 1: &\quad y[1] = 2u[1] + 3u[0] - 1.5u[-2] = 2 \cdot 0 + 3 \cdot 1 - 1.5 \cdot 0 = 3 \\
n = 2: &\quad y[2] = 2u[2] + 3u[1] - 1.5u[-1] = 2 \cdot 0 + 3 \cdot 0 - 1.5 \cdot 0 = 0 \\
n = 3: &\quad y[3] = 2u[3] + 3u[2] - 1.5u[0] = 2 \cdot 0 + 3 \cdot 0 - 1.5 \cdot 1 = -1.5 \\
n = 4: &\quad y[4] = 2u[4] + 3u[3] - 1.5u[1] = 2 \cdot 0 + 3 \cdot 0 - 1.5 \cdot 0 = 0
\end{aligned}$$

and $y[n] = 0$ for $n \geq 4$. This output is, by definition, the impulse response of the system. Thus the impulse response is

$$h[0] = 2, \qquad h[1] = 3, \qquad h[2] = 0, \qquad h[3] = -1.5,$$

and $h[n] = 0$ for $n \geq 4$.

EXAMPLE 3.2.3

The impulse response of a moving average filter of length 5 is defined by

$$h[n] = \begin{cases} 1/5 = 0.2 & \text{for } n = 0, 1, 2, 3, 4 \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

The reason for calling it a moving average filter will be given later.

EXERCISE 3.2.1

What is the impulse response of a multiplier with gain a?

Answer

$a\delta[n]$

EXERCISE 3.2.2

A DT system is called a *unit-sample delay element* or, simply, a *unit delay element* if its output equals the input delayed by one sample or $y[n] = u[n-1]$. What is its impulse response?

Answer

$\delta[n-1]$

EXAMPLE 3.2.4[1]

Consider the systems shown in Figure 3.1. Let $r[n]$ and $y[n]$ be respectively the input and output of the overall systems. Each overall system consists of a multiplier with gain a and a unit delay element as shown. Let the input and output of the multiplier be denoted by $e[n]$ and $u[n]$. Then we have $u[n] = ae[n]$ and $y[n] = u[n-1]$. The output in Figure 3.1(a) is fed positively back into the input to yield $e[n] = r[n] + y[n]$; therefore, the system is called a positive feedback system. The system in Figure 3.1(b) is called a negative feedback system because the output is fed negatively back into the input to yield $e[n] = r[n] - y[n]$.

Now let us compute the impulse response of the positive feedback system. The impulse response describes only the zero-state response; therefore, in its computation, the system is implicitly assumed to be initially relaxed or $r[n] = e[n] = u[n] = y[n] = 0$ for $n < 0$. Now, if $r[n] = \delta[n]$, then the response of the system is

n	0	1	2	3	4	5	$\cdots$
$r[n]$	1	0	0	0	0	0	
$y[n] = u[n-1]$	0	a	a^2	a^3	a^4	a^5	
$e[n] = r[n] + y[n]$	1	a	a^2	a^3	a^4	a^5	
$u[n] = ae[n]$	a	a^2	a^3	a^4	a^5	a^6	

It is obtained as follows: The input $r[n]$ is an impulse sequence; therefore, its values at $n = 0, 1, 2, 3, \ldots,$ are $1, 0, 0, 0, \ldots$. Because the system is initially relaxed, we have $u[-1] = 0$ and, consequently, $y[0] = u[-1] = 0$. The sum of $r[0]$ and $y[0]$ yields $e[0] = 1$. Because $u[n] = ae[n]$, we have $u[0] = a \times 1 = a$. Once $u[0]$ is computed, we have $y[1] = u[0] = a$. We then compute $e[1]$, $u[1]$, and then $y[2]$. Proceeding forward, we can complete the table. The output is, by definition, the impulse response $h[n]$ of the system. Thus, we have

$$h[0] = 0, \qquad h[1] = a, \qquad h[2] = a^2,$$

and, in general,

$$h[n] = a^n \tag{3.4}$$

for $n = 1, 2, \ldots$.

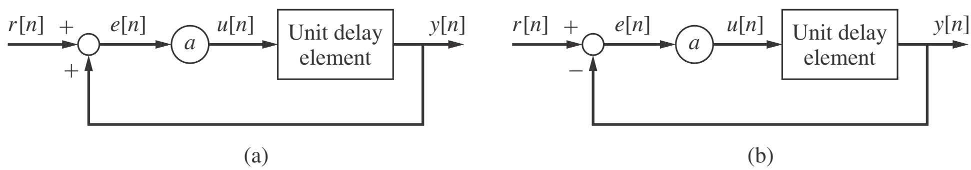


Figure 3.1 (**a**) Positive feedback system. (**b**) Negative feedback system.

[1]This example may be skipped without loss of continuity.

EXERCISE 3.2.3

Show that the impulse response of the negative feedback system in Figure 3.1(b) is $h[0] = 0$, and

$$h[n] = (-1)^{n-1}a^n \tag{3.5}$$

for $n = 1, 2, \ldots$.

If a system is causal and initially relaxed, no output will appear before applying an input. Thus for any causal system, we have

$$h[n] = 0 \qquad \text{for } n < 0 \tag{3.6}$$

In fact, (3.6) is the necessary and sufficient condition for a system to be causal. Clearly, the preceding examples are all causal systems.

3.2.1 FIR and IIR Systems

A DT LTI system is called an IIR (infinite impulse response) system if its impulse response has infinitely many nonzero entries. The savings account and the two systems in Figure 3.1 are all IIR systems. A DT LTI system is called an FIR (finite impulse response) system if its impulse response has a finite number of nonzero entries. An FIR system of length N has impulse response $h[n]$, for $n = 0, 1, 2, \ldots, N-1$ with $h[N-1] \neq 0$, and $h[n] = 0$ for $n \geq N$. The system described by (3.1) has, as computed in Example 3.2.2, $h[3] \neq 0$ and $h[n] = 0$ for $n \geq 4$. Thus it is FIR with length 4. The moving average filter discussed in Example 3.2.3 is also an FIR system; it has length 5.

The impulse response of a memoryless system with gain a is $h[0] = a$ and $h[n] = 0$ for $n \geq 1$. Thus *a memoryless system is an FIR system of length one*. For the FIR system of length 4 described by (3.1), its current output $y[n]$ depends on the current input $u[n]$ and past three inputs $u[n-1]$, $u[n-2]$, $u[n-3]$. Generally, the current output of an FIR system of length N depends on the current input and past $(N-1)$ inputs. Thus the system must remember the past $(N-1)$ inputs or has a memory of $(N-1)$ samples. For an IIR system, its current output depends on the current and *all* past inputs. Thus an IIR system has infinite memory.

3.3 DT LTI SYSTEMS—DISCRETE CONVOLUTIONS

Consider a discrete-time LTI system with impulse response $h[n]$. We show that the zero-state response of the system excited by any input sequence can be computed from $h[n]$. Let $u[n]$, $n = 0, 1, 2, \ldots$, be an arbitrary input sequence. Then it can be expressed as

$$u[n] = \sum_{k=0}^{\infty} u[k]\delta[n-k] \tag{3.7}$$

See (1.21). Note that for each fixed k, $u[k]\delta[n-k]$ is a sequence of n and is zero everywhere except at $n = k$, where its amplitude is $u[k]$. If the system is linear and time-invariant, then we have

$$\delta[n] \rightarrow h[n] \qquad \text{(definition)}$$

$$\delta[n-k] \rightarrow h[n-k] \qquad \text{(time shifting)}$$

$$u[k]\delta[n-k] \rightarrow u[k]h[n-k] \qquad \text{(homogeneity)}$$

$$\sum_{k=0}^{\infty} u[k]\delta[n-k] \rightarrow \sum_{k=0}^{\infty} u[k]h[n-k] \qquad \text{(additivity)}$$

Thus the output excited by the input $u[n]$, for $n \geq 0$, is given by[2]

$$y[n] = \sum_{k=0}^{\infty} h[n-k]u[k] \tag{3.8}$$

If a system is causal, then $h[n] = 0$ for $n < 0$, which implies $h[n-k] = 0$ for $k > n$. For each fixed n, the integer variable k in (3.8) ranges from 0 to ∞. However, for $k > n$, we have $h[n-k] = 0$. Thus, for a causal system, the upper summation limit ∞ in (3.8) can be replaced by n as

$$y[n] = \sum_{k=0}^{n} h[n-k]u[k] \tag{3.9}$$

for $n = 0, 1, 2, \ldots$. We see that the current output $y[n]$ depends on the input $u[k]$ for $0 \leq k \leq n$. The summation is called a *discrete convolution*. Because it relates the input and output, it is also called an *input–output description* of the system. The description is developed without using any information of the system other than the properties of linearity, time invariance, and causality. Thus it is a general formula applicable to *any* system so long as it is linear, time-invariant, causal, and initially relaxed at $n = 0$. It is applicable to lumped as well as distributed systems.

We develop an alternative form of (3.9). Define $\bar{k} := n - k$. Then $k = n - \bar{k}$ and (3.9) can be written as

$$y[n] = \sum_{\bar{k}=n}^{0} h[\bar{k}]u[n-\bar{k}] = \sum_{\bar{k}=0}^{n} u[n-\bar{k}]h[\bar{k}]$$

Because $\bar{k}$ is a dummy variable, it can be replaced by any symbol, in particular, by k. Thus we have

$$y[n] = \sum_{k=0}^{n} h[n-k]u[k] = \sum_{k=0}^{n} u[n-k]h[k] \tag{3.10}$$

for $n = 0, 1, 2, \ldots$. In other words, the discrete convolution has two equivalent forms. Either form can be used. Because h and u can be interchanged, (3.10) is said to have the *commutative* property. Note that (3.8) does not have the property.

[2]Wherever the sign of infinity appears in an equation, the questions of convergence and the validity of its operations arise. See References 13, 20, 33, and the footnote on page 152. In this text, we will not be concerned with these mathematical subtleties and will proceed intuitively.

The discrete convolution in (3.10) is an algebraic equation, and its computation is simple and straightforward. By direct substitution, we have

$$\begin{aligned}
y[0] &= h[0]u[0] \\
y[1] &= h[1]u[0] + h[0]u[1] \\
y[2] &= h[2]u[0] + h[1]u[1] + h[0]u[2] \\
&\;\;\vdots
\end{aligned}$$

Thus, if the impulse response of a system is known, its zero-state response excited by any input can be computed from (3.10).

EXAMPLE 3.3.1

Consider the savings account discussed in Example 3.2.1. If we deposit $u[0] = \$100.00$, $u[1] = -\$50.00$ (withdraw), $u[2] = \$200.00$, and $u[10] = \$50.00$, what is the total amount of money in the account at the end of the eleventh day? Note that $n = 0$ is the first day. Thus the eleventh day is $n = 10$.

The impulse response of the system was computed in Example 3.2.1 as $h[n] = (1.0001)^n$. Thus the account can be described by

$$y[n] = \sum_{k=0}^{n} h[n-k]u[k] = \sum_{k=0}^{n} (1.0001)^{n-k}u[k]$$

and, for $n = 10$,

$$y[10] = \sum_{k=0}^{10} (1.0001)^{10-k}u[k]$$

Because $u[k]$ is nonzero only at $k = 0, 1, 2,$ and 10, the equation becomes

$$\begin{aligned}
y[10] = {} & (1.0001)^{(10-0)}u[0] + (1.0001)^{(10-1)}u[1] + (1.0001)^{(10-2)}u[2] \\
& + (1.0001)^{(10-10)}u[10] = (1.0001)^{10} \times 100 + (1.0001)^9 \times (-50) \\
& + (1.0001)^8 \times 200 + 1 \times 50 = 300.2151
\end{aligned}$$

This is the amount of money in the account at the end of the eleventh day. From the equation, we see that $u[0]$ is multiplied by $(1.0001)^{10}$ because there are 10 days between the first day and the eleventh day. The amount $u[1]$ is multiplied by $(1.0001)^9$ because it earns 9 days' interest. The amount $u[10]$ is multiplied by 1 because it does not earn any interest. Thus the equation is correct.

EXAMPLE 3.3.2

Consider a memoryless LTI system such as a multiplier with gain a. Its impulse response is $h[n] = a\delta[n]$ or $h[0] = a$ and $h[n] = 0$ for $n > 0$. In this case, either form of (3.10) reduces to $y[n] = au[n]$. Thus, $y[n] = au[n]$ is a special case of convolutions.

For an FIR system of length N, we have $h[n] = 0$ for $n \geq N$, which implies $h[n-k] = 0$ for $n - k \geq N$ or $k \leq n - N$. Thus (3.10) can be modified as

$$y[n] = \sum_{k=n-N+1}^{n} h[n-k]u[k] = \sum_{k=0}^{N-1} h[k]u[n-k] \tag{3.11}$$

For example, for the FIR filter of length 5 in Example 3.2.3, (3.11) becomes

$$\begin{aligned} y[n] &= \sum_{k=0}^{4} h[k]u[n-k] \\ &= 0.2\,(u[n] + u[n-1] + u[n-2] + u[n-3] + u[n-4]) \end{aligned}$$

The output $y[n]$ is the sum of the five inputs $u[n-k]$, for $k = 0:4$, divided by 5. In other words, $y[n]$ is the average of its current and past four inputs. Thus the system is called a *moving-average filter*.

Before proceeding, we mention that if a DT LTI system is initially relaxed at $n_0 = -\infty$ and the input is applied from $-\infty$, then (3.8) must be modified as

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[n-k]u[k] =: h[n] * u[n] \\ &= \sum_{k=-\infty}^{\infty} u[n-k]h[k] =: u[n] * h[n] \end{aligned} \tag{3.12}$$

This equation is widely used and is often denoted by an asterisk. It has the commutative property as shown. The equation reduces to (3.10) if the system is causal and the input is applied from $n = 0$ onward, or $u[n] = h[n] = 0$ for $n < 0$. Although the use of (3.12) and its CT counterpart can simplify some mathematical derivation, it will suppress an important engineering phenomenon. See Section 6.8.3. Furthermore, no input can be applied from $-\infty$, thus we prefer the use of (3.10) and (3.8).

3.3.1 Underlying Procedure of Discrete Convolutions

Discrete convolutions are algebraic equations and can be computed, as illustrated in the preceding section, by direct substitutions. In this subsection, we use an example to discuss how to compute them graphically. This will reveal the procedure involved in convolutions. Consider

$$h[n] = \begin{cases} n-1 & \text{for } n = 0, 1, 2, 3 \\ 0 & \text{for } n > 3 \text{ and } n < 0 \end{cases}$$

and

$$u[n] = \begin{cases} n-2 & \text{for } n = 0, 1, 2, 3, 4 \\ 0 & \text{for } n > 4 \text{ and } n < 0 \end{cases}$$

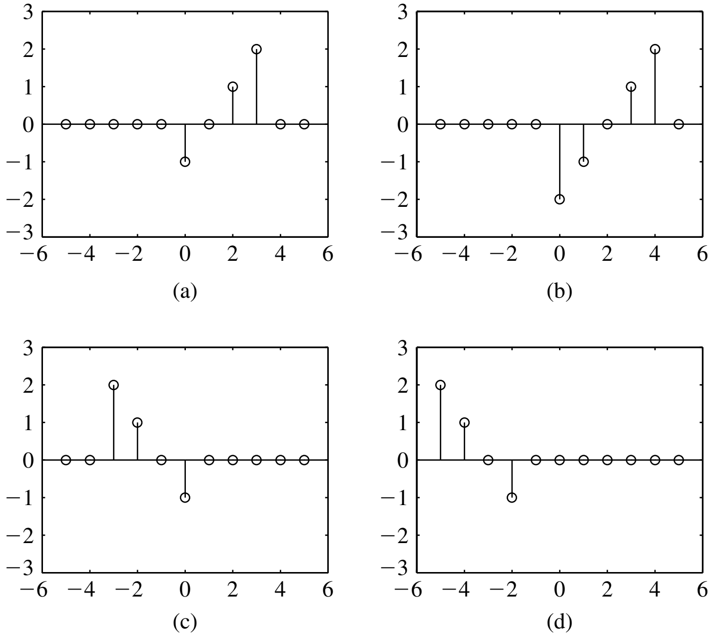


Figure 3.2 (**a**) Given $h[k]$. (**b**) Given $u[k]$. (**c**) Flipping $h[k]$ to yield $h[-k]$. (**d**) Shifting $h[-k]$ to $n = -2$ to yield $h[-2-k]$.

as shown in Figures 3.2(a) and 3.2(b). Note that the integer variable n can be changed to any symbol, and it has been changed to k in Figures 3.2(a) and 3.2(b). Instead of discussing (3.10), we discuss the computation of

$$y[n] = \sum_{k=-\infty}^{\infty} h[n-k]u[k] \tag{3.13}$$

for all n. The equation consists of two integer variables n and k both ranging from $-\infty$ to ∞. However, we cannot vary them at the same time. To compute it, we must select first an n and then carry out the summation over k. As discussed in Section 1.4.1, $h[n-k]$ is the flipping of $h[k]$ with respect to $k = 0$ and then shifting the flipped sequence $h[-k]$ in Figure 3.2(c) to n. For example, if $n = -2$, then $h[-2-k]$ is as shown in Figure 3.2(d). The product of $h[-2-k]$ and $u[k]$ is 0 for all k because their nonzero entries do not overlap. Their sum is 0; thus we have $y[-2] = 0$. Thus computing $y[n]$, for each n, requires four steps:

1. Flipping $h[k]$ with respect to $k = 0$ to yield $h[-k]$.
2. Shifting $h[-k]$ to n to yield $h[n-k]$.
3. Multiplying $h[n-k]$ and $u[k]$ for all k.
4. Summing up all nonzero $h[n-k]u[k]$ to yield $y[n]$.

The procedure can also be carried out in a table form as

k	$\cdots$	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	$\cdots$
$u[k]$		0	0	0	0	0	0	-2	-1	0	1	2	0	0	
$h[-2-k]$		0	2	1	0	-1	0	0	0	0	0	0	0	0	
$u \times h$		0	0	0	0	0	0	0	0	0	0	0	0	0	

The first row is the time index k. The second row is the given $u[k]$. The third row is the shifting of the flipped $h[-k]$ to $n = -2$. The fourth row is the products of $u[k]$ and $h[-2-k]$ for all k; they are all zeros. The sum of the last row yields $y[-2]$. To find $y[3]$, we shift the third row to $n = 3$ to yield

k	$\cdots$	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	$\cdots$
$u[k]$		0	0	0	0	0	0	-2	-1	0	1	2	0	0	
$h[3-k]$		0	0	0	0	0	0	2	1	0	-1	0	0	0	
$u \times h$		0	0	0	0	0	0	-4	-1	0	-1	0	0	0	

The products of $u[k]$ and $h[3-k]$ are listed in the fourth row. Their sum is -6. Thus we have $y[3] = -6$.

The preceding procedure can be carried out systematically as follows. We write the flipping of $h[n]$ on a paper strip. To find the output $y[n]$ at $n = -2$, we shift the strip to $n = -2$ and carry out multiplications and summation. Once $y[-2]$ is computed, we slide the strip one position to the right or to $n = -1$ and repeat the process. Because there are no overlaps of nonzero entries, we have $y[-1] = 0$. Next we slide the strip to $n = 0$. There is only one overlap of -2 and -1. Their product is 2. Thus we have $y[0] = 2$. We then slide the strip to $n = 1$. There are two overlaps: -1 and -1 at $k = 1$ and -2 and 0 at $k = 0$. Their products are 1 and 0. Their sum yields $y[1] = 1$. Proceeding forward, we can obtain

n	$\cdots$	-2	-1	0	1	2	3	4	5	6	7	8	9	$\cdot$
$y[n]$		0	0	2	1	-2	-6	-4	1	4	4	0	0	

Note that for $n \geq 8$, we have $y[n] = 0$ because there are no overlaps of nonzero entries.

We mention that if $h[n]$ is a finite sequence of length N and $u[n]$ a finite sequence of length P, then their convolution has at most $N + P - 1$ nonzero entries. In MATLAB, typing

```
u=[-2 -1 0 1 2];
h=[-1 0 1 2];
conv(u,h)
```

will yield `2 1 -2 -6 -4 1 4 4`. The convolution in MATLAB is not computed as discussed above. It is computed using a DT state-space equation as we will discuss in Section 10.4.1.

3.4 DT LTI LUMPED SYSTEMS—DIFFERENCE EQUATIONS

Every discrete-time linear time-invariant system can, as shown in the preceding sections, be described by a DT convolution. In this section, we show that the savings account studied in Example 3.2.1 can also be described by a different type of equation.

EXAMPLE 3.4.1

Consider the savings account discussed in Example 3.2.1. Its output $y[n]$ (total amount of money in the account at the end of the nth day) and input $u[n]$ (the amount of money deposited in the nth day) can be described by the convolution

$$y[n] = \sum_{k=0}^{n} (1.0001)^{n-k} u[k] \tag{3.14}$$

Now we will use (3.14) to develop a different type of equation to describe the system.

Equation (3.14) holds for every positive integer n. Replacing n in (3.14) by $n+1$ yields

$$y[n+1] = \sum_{k=0}^{n+1} (1.0001)^{n+1-k} u[k]$$

which can be expanded as

$$\begin{aligned} y[n+1] &= \sum_{k=0}^{n} (1.0001)^{n+1-k} u[k] + (1.0001)^{n+1-(n+1)} u[n+1] \\ &= 1.0001 \sum_{k=0}^{n} (1.0001)^{n-k} u[k] + u[n+1] \end{aligned} \tag{3.15}$$

Substituting (3.14) into (3.15) yields

$$y[n+1] = 1.0001 y[n] + u[n+1] \tag{3.16}$$

or

$$y[n+1] - 1.0001 y[n] = u[n+1] \tag{3.17}$$

This is called a first-order linear difference equation with constant coefficients or LTI difference equation. Thus the savings account can be described by a convolution and a difference equation.

This example shows that some convolution can be transformed into a difference equation. This is not necessarily possible for every convolution. For example, the system with the impulse response

$$h[n] = \begin{cases} 0 & \text{for } n \le 0 \\ 1/n & \text{for } n \ge 1 \end{cases}$$

or

$$h[n] = \begin{cases} 0 & \text{for } n = 0, 2, 4, \ldots \\ 1/n! & \text{for } n = 1, 3, 5, \ldots \end{cases}$$

cannot be transformed into an LTI difference equation. Such a system is a distributed system. If a system is lumped, then its convolution can always be transformed into a difference equation.

3.4.1 Setting Up Difference Equations

As discussed in Section 2.7, CT linear time-invariant systems are obtained by modeling. We have the same situation for DT systems. If a DT system is modeled as linear, time-invariant, and lumped, then we can develop its difference equation directly from the model without computing first its convolution. For example, for the savings account, if the amount of money at the end of the nth day is $y[n]$, then $y[n+1]$ equals the sum of the principle $y[n]$, its interest $0.0001y[n]$, and the amount of money deposited in the $(n+1)$th day, that is,

$$y[n+1] = y[n] + 0.0001y[n] + u[n+1] = 1.0001y[n] + u[n+1]$$

This is (3.17). This derivation is considerably simpler. We give more examples.

EXAMPLE 3.4.2

Let $y[n]$ be the U.S. population in year n and let $b[n]$ and $d[n]$ be, respectively, the birth rate and death rate. If $u[n]$ is the number of immigrants entering the United States, then we have

$$y[n+1] = y[n] + b[n]y[n] - d[n]y[n] + u[n] = (1 + b[n] - d[n])y[n] + u[n]$$

This is a first-order difference equation. It is a time-varying equation if its coefficients change year by year. If $b[n]$ and $d[n]$ are constants, then the equation reduces to

$$y[n+1] = (1 + b - d)y[n] + u[n]$$

This is an LTI difference equation. It is clear that after the equation is solved, $y[n]$ must be rounded or truncated to an integer. Furthermore, the model does not include illegal immigrants.

EXAMPLE 3.4.3 (Amortization)

In purchasing an automobile or a house, we may assume some debt and then pay it off by monthly installment. This is known as amortization. Let $y[0]$ be the initial total amount of money borrowed. We assume that we pay back $u[n]$ at month n. The unpaid debt will carry a monthly charge of $100r\%$. If $y[n]$ is the amount of debt at month n, then we have

$$y[n+1] = y[n] + ry[n] - u[n+1] = (1+r)y[n] - u[n+1]$$

This is an LTI difference equation.

3.4.2 From Difference Equations to Convolutions

Not every convolution can be transformed into a difference equation. However, if the difference-equation description of a system is known, then the impulse response of the system can be readily computed. For example, suppose a system is described by the difference equation

$$y[n+2] - 0.1y[n+1] - 0.06y[n] = u[n+1] + 2u[n] \tag{3.18}$$

By definition, if a system is initially relaxed or, equivalently, $y[n] = u[n] = 0$, for $n < 0$, then the output excited by the input $u[n] = \delta[n]$ is the impulse response. We substitute $u[n] = \delta[n]$ into (3.18) and rewrite it as

$$y[n+2] = 0.1y[n+1] + 0.06y[n] + \delta[n+1] + 2\delta[n] \tag{3.19}$$

For $n = -2$, the equation becomes

$$y[0] = 0.1y[-1] + 0.06y[-2] + \delta[-1] + 2\delta[-2]$$

which yields $y[0] = 0$. Substituting $n = -1 : 2$ recursively into (3.19) yields

$$\begin{aligned}
y[1] &= 0.1y[0] + 0.06y[-1] + \delta[0] + 2\delta[-1] = 1 \\
y[2] &= 0.1y[1] + 0.06y[0] + \delta[1] + 2\delta[0] = 0.1 \cdot 1 + 0 + 0 + 2 = 2.1 \\
y[3] &= 0.1y[2] + 0.06y[1] + \delta[2] + 2\delta[1] = 0.1 \cdot 2.1 + 0.06 \cdot 1 + 0 + 0 = 0.27 \\
y[4] &= 0.1y[3] + 0.06y[2] + \delta[3] + 2\delta[2] = 0.1 \cdot 0.27 + 0.06 \cdot 2.1 + 0 + 0 = 0.153
\end{aligned}$$

Thus we have $h[0] = 0$, $h[1] = 1$, $h[2] = 2.1$, $h[3] = 0.27$, $h[4] = 0.153, \ldots$. In conclusion, the impulse response of any difference equation can be obtained by direct substitution. Once the impulse response is computed, the zero-state response of the system can also be described by a convolution.

EXERCISE 3.4.1

Compute the impulse response sequence of the difference equation

$$y[n] - y[n-1] = 0.2(u[n] - u[n-3])$$

Is it an FIR or IIR system?

Answers

$h[n] = 0.2$ for $n = 0 : 2$ and $h[n] = 0$, for $n \geq 3$. FIR.

EXERCISE 3.4.2

Compute the impulse response sequence of the difference equation

$$y[n+1] - 0.5y[n] = 0.2u[n]$$

Is it an FIR or IIR filter?

Answers

$h[0] = 0$ and $h[n] = 0.2 \times (0.5)^{n-1}$ for $n \geq 1$. IIR.

3.5 COMPARISON OF DISCRETE CONVOLUTIONS AND DIFFERENCE EQUATIONS

Convolutions can be used to describe LTI distributed and lumped systems, whereas difference equations describe only lumped systems. In this sense, convolutions are more general than difference equations. For DT LTI lumped systems, we can use either convolutions or difference equations. It is natural to ask at this point which description is simpler to use for analysis and design. To answer this question, we compare them in the following:

1. *Difference equations require less computation than convolutions.* We use the savings account discussed in Example 3.4.1 to demonstrate this fact. Consider the convolution in (3.14). Because there are $(n+1)$ terms in the summation, it requires $n+1$ multiplications and n additions to compute each $y[n]$. Thus computing $y[n]$ for $n = 0:9$ requires $1+2+\cdots+10 = 55$ multiplications and $0+1+\cdots+9 = 45$ additions.[3] If we use the difference equation in (3.16), we need one addition and one multiplication to compute each $y[n]$. Thus computing $y[n]$ for $n = 0 : 9$ requires only 10 multiplications and 10 additions. They are considerably less than those for the convolution in (3.14).
2. *Difference equations requires less memory than convolutions*. Consider the difference equation in (3.18). The equation can be stored in a computer as $[1\ \ -0.1\ \ -0.06]$ and $[0\ \ 1\ \ 2]$, which are the coefficients of the equation. If we use its convolution equation, then we need to store $h[n]$, for all $n \geq 0$. There are infinitely many of them.
3. *Convolutions describe only zero-state responses*. Whenever we use a convolution, all initial conditions are implicitly assumed to be zero. Difference equations can be used even if initial conditions are not zero. Thus difference equations describe both zero-state (forced) responses and zero-input (natural) responses.

Because difference equations have many advantages over convolutions, we use mainly difference equations in studying LTI lumped systems. For distributed systems, we have no choice but to use convolutions.

3.6 GENERAL FORMS OF DIFFERENCE EQUATIONS

This section introduces two different but equivalent forms of difference equations.[4] Consider

$$\begin{aligned} &a_1 y[n+N] + a_2 y[n+N-1] + \cdots + a_{N+1} y[n] \\ &\quad = b_1 u[n+M] + b_2 u[n+M-1] + \cdots + b_{M+1} u[n] \end{aligned} \tag{3.20}$$

where a_i and b_i are real constants, and M and N are positive integers.[5] In this equation, we always assume $a_1 \neq 0$ and $b_1 \neq 0$. We also assume either $a_{N+1} \neq 0$ or $b_{M+1} \neq 0$. The rest of

[3] The fast Fourier transform (FFT), which will be introduced in Chapter 5, can also be used to compute convolutions. See Section 10.4.1.

[4] The study of this section and its subsection may be postponed until Chapter 9.

[5] There is no standard way of assigning coefficients a_i and b_i. We assign them to agree with those used in MATLAB.

a_i and b_i, however, can be zero or nonzero. If both a_{N+1} and b_{M+1} are zero such as in

$$2y[n+3] + 3y[n+2] = u[n+2] - u[n+1]$$

with $N = 3$ and $M = 2$, then the equation can be written as, reducing all indices by 1,

$$2y[n+2] + 3y[n+1] = u[n+1] - u[n] \tag{3.21}$$

and we have $N = 2$ and $M = 1$. Thus if at least one of a_{N+1} and b_{M+1} is nonzero, then N and M will be unique. The difference equation in (3.20) is said to be in *advanced form* and of order $\max(N, M)$. The difference equation in (3.21) clearly has order $\max(2, 1) = 2$.

Subtracting 2 from all indices of (3.21) yields

$$2y[n] + 3y[n-1] = u[n-1] - u[n-2] \tag{3.22}$$

This is called a *delayed-form* difference equation. Its general form is

$$\begin{aligned} &\bar{a}_1 y[n] + \bar{a}_2 y[n-1] + \cdots + \bar{a}_{\bar{N}+1} y[n-\bar{N}] \\ &\quad = \bar{b}_1 u[n] + \bar{b}_2 u[n-1] + \cdots + \bar{b}_{\bar{M}+1} u[n-\bar{M}] \end{aligned} \tag{3.23}$$

where $\bar{a}_i$ and $\bar{b}_i$ are real constants, and $\bar{N}$ and $\bar{M}$ are positive integers. In this equation, we always assume $\bar{a}_{\bar{N}+1} \neq 0$ and $\bar{b}_{\bar{M}+1} \neq 0$. We also assume either $\bar{a}_1 \neq 0$ or $\bar{b}_1 \neq 0$. The rest of $\bar{a}_i$ and $\bar{b}_i$, however, can be zero or nonzero. If at least one of $\bar{a}_1$ and $\bar{b}_1$ is nonzero, then $\bar{N}$ and $\bar{M}$ will be unique. The delayed-form difference equation in (3.23) is said to have order $\max(\bar{N}, \bar{M})$.

The difference equations in (3.20) and (3.23) can describe both causal and noncausal systems. We discuss the conditions for them to be causal. For the advanced-form difference equation in (3.20), if $M > N$, then $n + M > n + N$ and $n + M$ is a future sampling instant with respect to $n + N$. Thus $y[n + N]$ depends on the future input $u[n + M]$ and the equation describes a noncausal system. Thus (3.20) describes a causal system *if and only if* $N \geq M$. For (3.23), if $\bar{a}_1 = 0$, then we require $\bar{b}_1 \neq 0$ in order to have unique $\bar{N}$ and $\bar{M}$. In this case, the output $y[n-1]$ depends on $u[n]$, a future input. Thus the delayed-form difference equation in (3.23) is causal *if and only if* $\bar{a}_1 \neq 0$; there is no condition imposed on $\bar{N}$ and $\bar{M}$. The integer $\bar{N}$ can be larger than, equal to, or smaller than $\bar{M}$.

As we saw in (3.21) and (3.22), delayed-form difference equations can be easily obtained from advanced-form and vice versa. We use both forms in this text.

3.6.1 Recursive and Nonrecursive Difference Equations

Consider the difference equation in (3.23). For convenience, we normalize $\bar{a}_1$ to 1. If $\bar{a}_i = 0$ for all $i > 1$, then the equation reduces to

$$y[n] = \bar{b}_1 u[n] + \bar{b}_2 u[n-1] + \cdots + \bar{b}_{\bar{M}+1} u[n-\bar{M}] \tag{3.24}$$

This is called a *nonrecursive* difference equation. Its output depends only on current and past inputs and is independent of *past outputs*. If $\bar{a}_1 = 1$ and $\bar{a}_i \neq 0$ for some $i > 1$, then (3.23) is called a *recursive* difference equation and can be written as

$$\begin{aligned} y[n] = &-\bar{a}_2 y[n-1] - \cdots - \bar{a}_{\bar{N}+1} y[n-\bar{N}] \\ &+ \bar{b}_1 u[n] + \bar{b}_2 u[n-1] + \cdots + \bar{b}_{\bar{M}+1} u[n-\bar{M}] \end{aligned} \tag{3.25}$$

Its current output depends not only on current and past inputs but also on *past outputs*. In this equation, because past outputs are needed in computing current output, the equation must be computed in the order of $y[0]$, $y[1]$, $y[2]$, and so forth, or must be computed recursively from $n = 0$ onward. There is no such restriction in computing the nonrecursive equation in (3.24).

We discuss the impulse response of the nonrecursive difference equation in (3.24). By definition, if a system is initially relaxed or, equivalently, $u[n] = 0$ and $y[n] = 0$ for all $n < 0$, then the output excited by $u[n] = \delta[n]$ is the impulse response of the equation. Direct substitution yields

$$h[0] = y[0] = \bar{b}_1, \quad h[1] = y[1] = \bar{b}_2, \ldots, \quad h[\bar{M}] = y[\bar{M}] = \bar{b}_{\bar{M}+1}$$

and $h[n] = y[n] = 0$ for all $n > \bar{M} + 1$. This is an FIR filter of length $(\bar{M} + 1)$. Thus *every nonrecursive difference equation describes an FIR system.* Because the difference equation in (3.24) has order $\bar{M}$, it is called an $\bar{M}$th-order FIR filter. Thus an FIR filter of order $\bar{M}$ has length $\bar{M} + 1$.

In fact, the convolution description of an FIR filter is a nonrecursive difference equation. For example, consider an FIR filter of length 4—that is, $h[n] = 0$ for $n < 0$ and $n > 3$. Then its input and output are related by

$$y[n] = \sum_{k=0}^{n} h[k]u[n-k] = \sum_{k=0}^{3} h[k]u[n-k]$$

or

$$y[n] = h[0]u[n] + h[1]u[n-1] + h[2]u[n-2] + h[3]u[n-3]$$

This is a third-order nonrecursive difference equation. Thus *there is no difference between convolutions and nonrecursive difference equations in describing FIR filters.*

Impulse responses of most recursive difference equations have infinitely many nonzero entries. The reason is that the impulse input will yield some nonzero outputs. This nonzero outputs will then propagate to all outputs because current output depends on past output. Thus *difference equations describing IIR filters must be recursive*. Although every FIR filter can be described by a nonrecursive difference equation, in some special situation we may develop a recursive difference equation to describe it to save the number of operations as the next example illustrates.

EXAMPLE 3.6.1

Consider the 30-day moving average filter or the filter with impulse response

$$h[n] = \begin{cases} 1/30 = 0.033 & \text{for } n = 0:29 \\ 0 & \text{for } n < 0 \text{ and } n > 29 \end{cases}$$

Its convolution and nonrecursive difference equation both equal

$$y[n] = 0.033\,(u[n] + u[n-1] + u[n-2] + \cdots + u[n-28] + u[n-29]) \quad (3.26)$$

Its computation requires 29 additions and one multiplication.

Let us reduce the indices of (3.26) by one to yield

$$y[n-1] = 0.033\,(u[n-1] + u[n-2] + u[n-3] + \cdots + u[n-29] + u[n-30])$$

Subtracting this equation from (3.26) yields

$$y[n] - y[n-1] = 0.033\,(u[n] - u[n-30]) \tag{3.27}$$

This is a recursive difference equation. To compute $y[n]$, we rewrite it as

$$y[n] = y[n-1] + 0.033\,(u[n] - u[n-30])$$

Computing each $y[n]$ requires one multiplication and two additions; thus the recursive equation in (3.27) requires less computation than the nonrecursive difference equation in (3.26).

3.7 CT LTI SYSTEMS—INTEGRAL CONVOLUTIONS

We discuss in the remainder of this chapter the continuous-time counterpart of what has been discussed for DT systems. We show that if a CT system is linear, time-invariant, and initially relaxed at $t = 0$, then the output $y(t)$ excited by any input $u(t)$ can be described by the integral convolution

$$y(t) = \int_{\tau=0}^{\infty} h(t-\tau)u(\tau)\,d\tau \tag{3.28}$$

where $h(t)$ is the output excited by the impulse $u(t) = \delta(t)$ and is called the *impulse response*. This is the CT counterpart of (3.8). Its derivation is similar to the DT case but requires a limiting process. Let $h_a(t)$ be the output of the system excited by the input $u(t) = \delta_a(t)$ defined in (1.7) and plotted in Figure 1.16(a). As discussed in Section 1.5.1, any input signal $u(t)$ can be approximated by

$$u(t) \approx \sum_{k=0}^{\infty} u(ka)\delta_a(t-ka)a$$

This is the same as (1.14) except that T is replaced by a and n by k. Now if the system is linear and time invariant, then we have

$$\begin{aligned}
\delta_a(t) &\to h_a(t) && \text{(definition)}\\
\delta_a(t-ka) &\to h_a(t-ka) && \text{(time shifting)}\\
u(ka)\delta_a(t-ka)a &\to u(ka)h_a(t-ka)a && \text{(homogeneity)}\\
\sum_{k=0}^{\infty} u(ka)\delta_a(t-ka)a &\to \sum_{k=0}^{\infty} u(ka)h_a(t-ka)a && \text{(additivity)}
\end{aligned}$$

The left-hand side of the last row equals roughly the input, thus the output of the system is given by

$$y(t) \approx \sum_{k=0}^{\infty} u(ka)h_a(t-ka)a \tag{3.29}$$

We define $\tau = ka$. If $a \to 0$, the pulse $\delta_a(t)$ becomes the impulse $\delta(t)$, τ becomes a continuum, a can be written as $d\tau$, the summation becomes an integration, and the approximation becomes an equality. Thus we have, as $a \to 0$,

$$y(t) = \int_0^{\infty} h(t-\tau)u(\tau)\, d\tau \tag{3.30}$$

where $h(t)$ is the output excited by the input $u(t) = \delta(t)$ and is called the impulse response of the system.

The system is causal if and only if $h(t) = 0$ for $t < 0$. This follows from the fact that if a system is causal and initially relaxed at $t = 0$, then no input will appear before applying an input. If $h(t) = 0$ for $t < 0$, then $h(t-\tau) = 0$ for $t < \tau$ and (3.30) reduces to

$$y(t) = \int_0^t h(t-\tau)u(\tau)\, d\tau = \int_0^t h(\tau)u(t-\tau)\, d\tau \tag{3.31}$$

The second equality can be obtained from the first by changing variables. Equation (3.31) is the CT counterpart of (3.10). Note that (3.31) has the commutative property, but (3.30) does not have.

We study mostly linear, time-invariant, and lumped systems. Such DT systems are classified as FIR and IIR systems. For CT systems, except for memoryless systems such as amplifiers, all such CT systems have impulse responses of infinite duration. In other words, we encounter almost exclusively CT IIR systems. Thus we make no such classification for CT systems.

3.7.1 Impulse Responses and Step Responses

In theory, the impulse response of a CT LTI system can be obtained by measurement from the input and output terminals without knowing the internal structure of the system. We apply at the input terminal an impulse, then the signal measured at the output terminal is the impulse response. However, because an impulse has zero width and infinite height, it cannot be generated in practice. If no impulse can be generated, there is no impulse response to be measured. We now discuss a method of obtaining the impulse response without generating an impulse.

Unlike impulses, step functions can be easily generated in practice. Let $y_q(t)$ be the output of a system excited by a (unit) step function. Such output is called the *step response*. If $u(t) = q(t) = 1$ for $t \geq 0$, then $u(t-\tau) = 1$ for all τ in $[0, t]$. Thus we have

$$y_q(t) = \int_0^t h(\tau)u(t-\tau)\, d\tau = \int_0^t h(\tau)\, d\tau$$

Its differentiation with respect to t yields

$$\frac{dy_q(t)}{dt} = \frac{d}{dt}\left(\int_0^t h(\tau)\, d\tau\right) = h(t) \tag{3.32}$$

Thus the impulse response can be obtained by differentiating the step response.

EXERCISE 3.7.1

Consider an LTI system with its step response shown in Figure 2.12(d). What is its impulse response?

Answer

$h(t) = 0.5[q(t) - q(t-2)]$

Even though impulse responses can be obtained by differentiating step responses, differentiation will amplify high-frequency noise that often exists in electrical systems. Thus this approach of obtaining impulse responses is rarely used in practice. This is illustrated in the next example.

EXAMPLE 3.7.1

Consider the network shown in Figure 3.3(a). Let us apply a step input, that is, $u(t) = 1$, for $t \geq 0$. Suppose its output (step response) is measured as $y_q(t) = 1 - e^{-0.5t}$ for $t \geq 0$ and shown in Figure 3.4(a). Then the impulse response of the network is

$$h(t) = \frac{y_q(t)}{dt} = 0.5e^{-0.5t}$$

and is plotted in Figure 3.3(b).

Differentiation can be approximated on a computer by

$$\frac{dy_q(t)}{dt} = \frac{y_q(t+\Delta) - y_q(t)}{\Delta}$$

with a small Δ. For the step response shown in Figure 3.4(a), if we select $\Delta = 0.01$, then the result of its differentiation is shown in Figure 3.4(b). It is very close to the impulse response shown in Figure 3.3(b).

Now suppose the measured step response is corrupted by uniformly distributed random noise in $(-0.0025, 0.0025)$.[6] Because the noise is very small in magnitude,

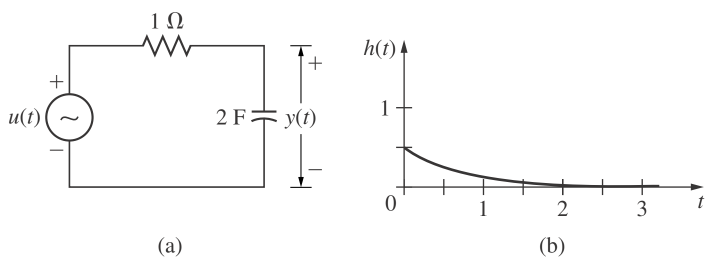


Figure 3.3 (**a**) Network. (**b**) Its impulse response.

[6]Such a noise can be generated in MATLAB as `0.005*(rand(1,N)-0.5)`, where `rand(1,N)` generates a string of N random numbers uniformly distributed in (0, 1).

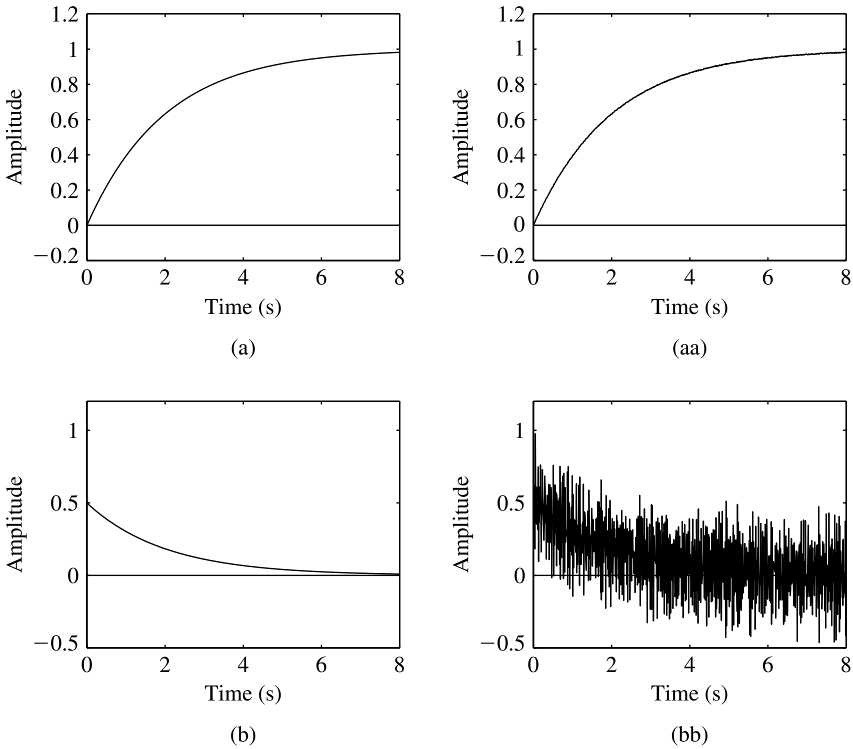


Figure 3.4 (**a**) Step response. (**b**) Its differentiation. (**aa**) Step response corrupted by random noise. (**bb**) Its differentiation.

the corrupted step response shown in Figure 3.4(aa) is indistinguishable from the one in Figure 3.4(a). Its differentiation, as shown in Figure 3.4(bb), is very different from $h(t)$. Thus (3.32) cannot be used. Because high-frequency noise often exists in measured signals, differentiation should be avoided whenever possible.

In summary, although impulse responses can be measured in theory, it is difficult to achieve in practice. However there are analytical methods of computing impulse responses as we will discuss in Chapter 6.

3.7.2 Graphical Computation of Convolutions

We discuss in this section graphical computation of CT convolutions.[7] We use a simple example to illustrate the procedure. Consider the $h(t)$ and $u(t)$ shown in Figures 3.5(a) and 3.5(b). The impulse response $h(t)$ is zero for $t < 0$ and $t > 2$, and the input $u(t)$ is zero for $t < 0$ and $t > 4$.

---

[7]This subsection may be skipped without loss of continuity.

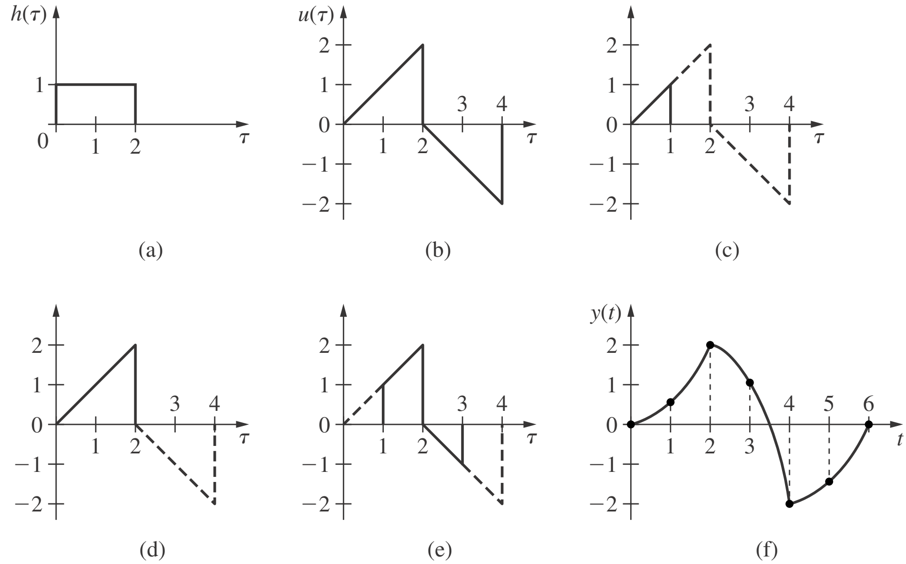


Figure 3.5 (**a**) $h(\tau)$. (**b**) $u(\tau)$. (**c**) $h(1-\tau)u(\tau)$. (**d**) $h(2-\tau)u(\tau)$. (**e**) $h(3-\tau)u(\tau)$. (**f**) $y(t)$.

Note that they are plotted with respect to τ. For the two signals, the convolution in (3.30) can be written as, similar to (3.12),

$$y(t) = \int_{\tau=-\infty}^{\infty} h(t-\tau)u(\tau)\,d\tau \tag{3.33}$$

To compute graphically $y(t)$ at $t = t_0$ requires four steps: (1) Flip $h(\tau)$ to yield $h(-\tau)$. (2) Shift $h(-\tau)$ to t_0 to yield $h(t_0-\tau)$. (3) Multiply $h(t_0-\tau)$ and $u(\tau)$ for all τ. (4) Compute its total area which yields $y(t_0)$. For example, if $t_0 < 0$, the nonzero parts of $h(t_0-\tau)$ and $u(\tau)$ do not overlap. Thus their product is identically zero and its total area is zero. Thus we have $y(t) = 0$ for $t < 0$. For $t_0 = 1$, the product $h(1-\tau)u(\tau)$ is shown in Figure 3.5(c) with a solid line. The triangle has area $(1 \times 1)/2 = 0.5$. Thus we have $y(1) = 0.5$. For $t = 2$, their product is shown in Figure 3.5(d) and its total area is $(2 \times 2)/2 = 2$. Thus we have $y(2) = 2$. For $t = 3$, their product is shown in Figure 3.5(e). There are positive and negative areas. The positive area is $2 - 0.5 = 1.5$ and the negative area is -0.5. Their sum is $1.5 - 0.5 = 1$. Thus we have $y(3) = 1$. Proceeding forward, we can obtain $y(4) = -2$, $y(5) = -1.5$, and $y(t) = 0$ for $t \geq 6$. The computed $y(t)$ at $t = 1, 2, 3, 4,$ and 5 are plotted in Figure 3.5(f) with solid dots. The rest of $y(t)$ is obtained by interpolation as shown.[8] Clearly, if we compute $y(t)$ at more points, then the result will be more accurate.

[8]For t in $[0, 2]$, the area is $0.5t^2$. Thus the interpolation should be polynomials of degree 2 instead of straight lines (polynomials of degree 1).

EXERCISE 3.7.2

Verify that the convolution of the two signals in Figures 3.6(a) and (b) is as shown in Figure 3.6(c).

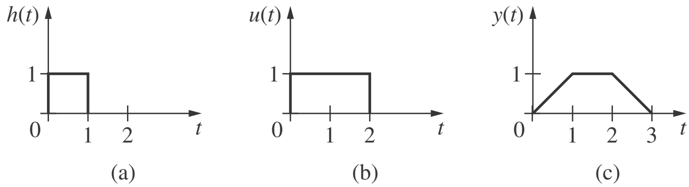


Figure 3.6 (**a**) $h(t)$. (**b**) $u(t)$. (**c**) Their convolution.

3.8 CT LTI LUMPED SYSTEMS—DIFFERENTIAL EQUATIONS

Every CT linear time-invariant system can be described by an integral convolution. In this section we use an example to show that some convolution can be transformed into a differential equation.

EXAMPLE 3.8.1

Consider the network shown in Figure 3.3(a) and discussed in Example 3.7.1. Its impulse response is $h(t) = 0.5e^{-0.5t}$, and its output $y(t)$ excited by any input $u(t)$ can be described by

$$y(t) = \int_0^t h(t-\tau)u(\tau)\,d\tau = \int_0^t 0.5e^{-0.5(t-\tau)}u(\tau)\,d\tau$$
$$= 0.5e^{-0.5t}\int_0^t e^{0.5\tau}u(\tau)\,d\tau$$

Let us differentiate the output:

$$\dot{y}(t) := \frac{dy(t)}{dt} = \frac{d}{dt}[0.5e^{-0.5t}] \times \left[\int_0^t e^{0.5\tau}u(\tau)\,d\tau\right]$$
$$+0.5e^{-0.5t}\frac{d}{dt}\left[\int_0^t e^{0.5\tau}u(\tau)\,d\tau\right] = -0.5 \times 0.5e^{-0.5t}\left[\int_0^t e^{0.5\tau}u(\tau)\,d\tau\right]$$
$$+0.5e^{-0.5t}e^{0.5t}u(t) = -0.5y(t) + 0.5u(t)$$

Thus we have

$$\dot{y}(t) + 0.5y(t) = 0.5u(t)$$

This is a first-order linear differential equation with constant coefficients or a first-order LTI differential equation. Thus the network can be described by a convolution and a differential equation.

Not every integral convolution can be transformed into a differential equation. If a system is distributed, then such a transformation is not possible. The network in Figure 3.3(a) has one state variable. Thus it is a lumped system, and its convolution can be so transformed as shown. This is true for every LTI lumped system.

If a DT system is linear and time-invariant, then it can be described by a discrete convolution. If it is lumped as well, then it can also be described by a difference equation of the advanced form shown in (3.20) or the delayed form shown in (3.23). Difference equations are preferable to discrete convolutions because they require less computation and memory and they describe both zero-state and zero-input responses. We have similar situations for CT systems. To be more specific, if a CT system is linear and time-invariant, then it can be described by an integral convolution. If it is lumped as well, then it can also be described by a differential equation of the form

$$\begin{aligned} a_1 y^{(N)}(t) + a_2 y^{(N-1)}(t) + \cdots + a_{N+1} y(t) \\ = b_1 u^{(M)}(t) + b_2 y^{(M-1)}(t) + \cdots + b_{M+1} u(t) \end{aligned} \tag{3.34}$$

where $y^{(k)}(t) = d^k y(t)/dt^k$, and a_i and b_i are real constants. As in (3.20), we assume $a_1 \neq 0$ and $b_1 \neq 0$ and assume at least one of a_{N+1} and b_{M+1} to be nonzero; otherwise, N and M can be reduced. Unlike difference equations that have both advanced and delayed forms, we use only the form in (3.34).[9] The differential equation is said to have order $\max(N, M)$.

In the difference equation in (3.20), we require $N \geq M$ in order for the equation to describe a DT causal system. In (3.34), we require $N \geq M$ for different reasons. Consider $y(t) = du(t)/dt$, a pure differentiator. It has $N = 0$ and $M = 1$. Is it a causal system? The answer depends on how we define the differentiation. If we define it as, with $\Delta > 0$,

$$y(t) = \lim_{\Delta \to 0} \frac{u(t + \Delta) - u(t)}{\Delta}$$

then the output at t depends on the future input $u(t + \Delta)$ and the differentiator is not causal. However, if we adopt

$$y(t) = \lim_{\Delta \to 0} \frac{u(t) - u(t - \Delta)}{\Delta}$$

then the current output depends only on the current and past input and the differentiator is causal. Thus whether a differentiator is causal or not is open to argument.

Even so, differentiators are rarely used if a signal contains high-frequency noise as demonstrated in Figure 3.4. We give one more example.

EXAMPLE 3.8.2

Consider

$$u(t) = \sin 0.1t + 0.01 \sin 1000t$$

where $\sin 0.1t$ is a desired signal and $0.01 \sin 1000t$ is noise. Let us define its *signal-to-noise ratio (SNR)* as the ratio of their peak magnitudes. Then its SNR is $1/0.01 = 100$.

---

[9] The form in (3.34) is equivalent to the advanced-form difference equation in (3.20). The form corresponding to the delayed form in (3.23) is an integral equation, which is not used in studying CT systems.

If the input $u(t)$ goes through a differentiator, then its output is

$$y(t) = \frac{du(t)}{dt} = 0.1\cos 0.1t + 0.01 \times 1000 \cos 1000t$$

Now its SNR becomes $0.1/10 = 0.01$. Thus the output is completely dominated by the noise and is useless. See Problem 4.8 using a different definition of SNR.

In conclusion, differentiators and, more generally, differential equations of the form shown in (3.34) with $N < M$ are not used in practice if signals contain high-frequency noise. Thus we study mainly differential equations with $N \geq M$.

3.8.1 Setting Up Differential Equations

As discussed in Section 2.7, linear, time-invariant, and lumped systems are obtained in practice by modeling. Once such a model is available, we can develop its differential equation directly without computing first its impulse response. This is illustrated with examples.

EXAMPLE 3.8.3

Consider the network shown in Figure 3.3(a). It is a CT LTIL system. If the voltage across the capacitor is $y(t)$, then its current is $2\dot{y}(t)$. This is the current flowing around the loop. Thus the voltage drop across the 1-Ω resistor is $2\dot{y}(t)$. Applying Kirchhoff's voltage law around the loop yields

$$2\dot{y}(t) + y(t) = u(t)$$

This is the differential equation developed in Example 3.8.1. This procedure is much simpler than the one in Example 3.8.1.

EXAMPLE 3.8.4

Consider the mechanical system shown in Figure 2.15. For a limited range, the spring can be modeled as linear with spring force $ky(t)$, where k is the spring constant. If we consider only viscous friction and disregard the static and Coulomb frictions, then the friction is linear and equals $f\dot{y}(t)$, where f is the viscous friction coefficient and $\dot{y}(t) := dy(t)/dt$. The applied force $u(t)$ must overcome the spring force and friction, and the remainder is to accelerate the mass. Thus Newton's law implies

$$u(t) - f\dot{y}(t) - ky(t) = m\ddot{y}(t)$$

with $\ddot{y}(t) := d^2y(t)/dt^2$, or

$$m\ddot{y}(t) + f\dot{y}(t) + ky(t) = u(t) \tag{3.35}$$

This is a second-order differential equation that describes the mechanical system.

EXERCISE 3.8.1

The suspension system of an automobile shown in Figure 3.7(a) can be modeled as shown in Figure 3.7(b). The model consists of a block with mass m, which denotes the weight of the automobile. When the automobile hits a pothole or curb, a vertical force $u(t)$ is applied to the mass and causes the automobile to oscillate. The suspension system consists of a spring with spring constant k and a dashpot which represents the shock absorber. The spring generates the force $ky(t)$, where $y(t)$ is the vertical displacement measured from equilibrium. The dashpot is modeled to generate the viscous friction $f\dot{y}(t)$. Find a differential equation to describe the system.

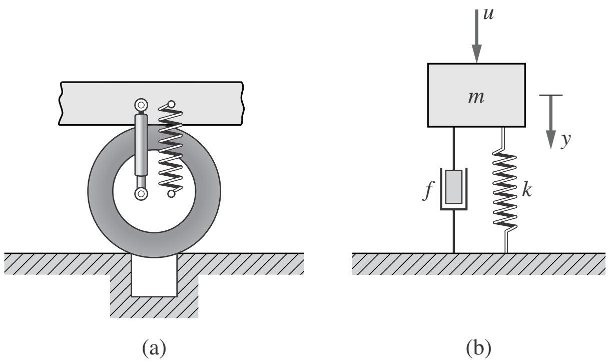


Figure 3.7 (**a**) Automobile suspension system. (**b**) Its model.

Answer

$m\ddot{y}(t) + f\dot{y}(t) + ky(t) = u(t)$

EXAMPLE 3.8.5

Consider the network shown in Figure 3.8. It consists of two resistors with resistances 3 and 4 Ω, one inductor with inductance 1 henry (H), and one capacitor with capacitance 2 farad (F). The input $u(t)$ is a voltage source, and the output $y(t)$ is the voltage across the capacitor as shown. The current passing through the capacitor is $2\dot{y}(t)$, and the current passing through the 4-Ω resistor is $y(t)/4 = 0.25y(t)$. Thus the current

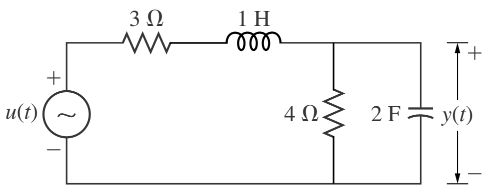


Figure 3.8 Network.

passing through the 3-Ω resistor and 1-H inductor is $0.25y(t) + 2\dot{y}(t)$. Their voltages are $3(0.25y(t) + 2\dot{y}(t))$ and $0.25\dot{y}(t) + 2\ddot{y}(t)$, respectively. Applying Kirchhoff's voltage law around the outer loop yields

$$3(0.25y(t) + 2\dot{y}(t)) + 0.25\dot{y}(t) + 2\ddot{y}(t) + y(t) = u(t)$$

or

$$2\ddot{y}(t) + 6.25\dot{y}(t) + 1.75y(t) = u(t) \tag{3.36}$$

This is a second-order differential equation that describes the network.

EXERCISE 3.8.2

Find a differential equation to describe the network shown in Figure 3.3(a) if the 1-Ω resistor is replaced by a 1-H inductor.

Answer

$2\ddot{y}(t) + y(t) = u(t)$

EXAMPLE 3.8.6 (Electromechanical System)[10]

Motors are widely used in homes and factories. There are ac, dc, hydraulic, and other types of motors. We consider only dc motors that can be modeled as shown in Figure 3.9. Such motors, with power ranging from a fraction of a watt to over 10,000 hp, are used in steel rolling mills, electric locomotives, cranes, golf carts, and many control systems. The model in Figure 3.9 consists of a field circuit and an armature circuit. The field circuit is connected to a constant voltage supply and thus provides a constant magnetic field; the circuit can also be replaced by a permanent magnet. The input $u(t)$ is applied to the armature circuit. Thus it is called an armature-controlled dc motor. The motor

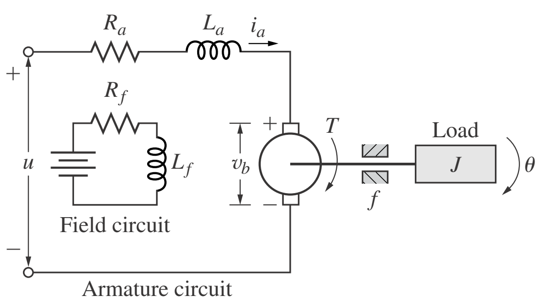


Figure 3.9 Armature-controlled dc motor.

[10]This example may be skipped without loss of continuity.

generates a torque $T(t)$ to drive a load with moment of inertia J. If we identify the following between translational and rotational movements

Translation		Rotational
Force	$\longleftrightarrow$	Torque
Mass	$\longleftrightarrow$	Moment of inertia
Linear displacement $y(t)$	$\longleftrightarrow$	Angular displacement θ

then we have, as in (3.35),

$$J\ddot{\theta}(t) + f\dot{\theta}(t) = T(t) \tag{3.37}$$

where f is the viscous friction coefficient. Here we have assumed the motor shaft to be rigid, thus there is no torsional spring force.

The torque $T(t)$ generated is modeled to be linearly proportional to the armature current $i_a(t)$ as $k_t i_a(t)$, where k_t is a constant. In reality, the circuit may saturate and the model holds only for a limited range of $u(t)$. A rotating motor shaft will generate a back electromagnetic force (emf) voltage $v_b(t)$ as shown. The voltage $v_b(t)$ is modeled as $v_b(t) = k_b\dot{\theta}(t)$, where $\dot{\theta} = d\theta/dt$ is the shaft angular velocity and k_b is a constant. Substituting $T(t) = k_t i_a(t)$ into (3.37) yields

$$J\ddot{\theta}(t) + f\dot{\theta}(t) = k_t i_a(t) \tag{3.38}$$

From the armature circuit we have

$$u(t) = R_a i_a(t) + L_a \dot{i}_a(t) + v_b(t) = R_a i_a(t) + L_a \dot{i}_a(t) + k_b\dot{\theta}(t) \tag{3.39}$$

where $\dot{i}_a(t) = di_a(t)/dt$. Substituting the $i_a(t)$ in (3.38) into (3.39), we will finally obtain

$$JL_a\theta^{(3)}(t) + (R_aJ + L_af)\ddot{\theta}(t) + (R_af + k_vk_t)\dot{\theta}(t) = k_tu(t) \tag{3.40}$$

This is a third-order differential equation that describes the motor.

EXAMPLE 3.8.7 (Hydrautic Tanks)[11]

In chemical plants, it is often necessary to maintain liquid levels. A simplified model is shown in Figure 3.10 in which

q_i, q_1, q_o = rates of liquid flow

A_1, A_2 = areas of the cross section of tanks

h_1, h_2 = liquid levels

R_1, R_2 = flow resistances, controlled by valves

[11] This example may be skipped without loss of continuity.

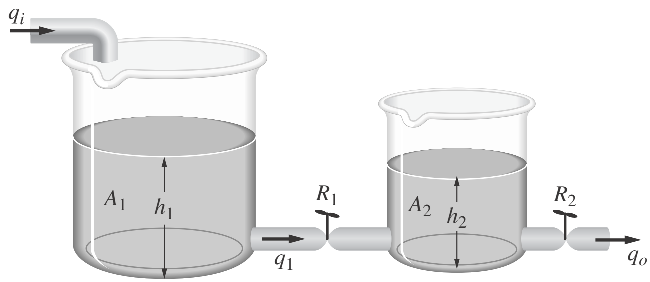


Figure 3.10 Control of liquid levels.

It is assumed that q_1 and q_o are governed by

$$q_1(t) = \frac{h_1(t) - h_2(t)}{R_1} \quad \text{and} \quad q_o(t) = \frac{h_2(t)}{R_2} \tag{3.41}$$

They are proportional to relative liquid levels and inversely proportional to flow resistances. The changes of liquid levels are governed by

$$A_1 dh_1(t) = (q_i(t) - q_1(t))\,dt \quad \text{and} \quad A_2 dh_2(t) = (q_1(t) - q_o(t))\,dt$$

or

$$A_1 \dot{h}_1(t) = q_i(t) - q_1(t) \quad \text{and} \quad A_2 \dot{h}_2(t) = q_1(t) - q_o(t) \tag{3.42}$$

where $\dot{h}_k(t) = dh_k(t)/dt$ for $k = 1, 2$. Let q_i and q_o be the input and output. Now we develop a differential equation to describe the system. Differentiating (3.41) yields

$$\dot{h}_2(t) = R_2 \dot{q}_o(t) \quad \text{and} \quad \dot{h}_1(t) = R_1 \dot{q}_1(t) + \dot{h}_2(t) = R_1 \dot{q}_1(t) + R_2 \dot{q}_o(t)$$

Substituting them into (3.42) yields

$$q_i(t) - q_1(t) = A_1(R_1 \dot{q}_1(t) + R_2 \dot{q}_o(t)) \tag{3.43}$$

$$q_1(t) - q_o(t) = A_2 R_2 \dot{q}_o(t) \tag{3.44}$$

Substituting $q_1(t)$ in (3.44) and its derivative into (3.43), we obtain

$$q_i(t) - (q_o(t) + A_2 R_2 \dot{q}_o(t)) = A_1 R_1(\dot{q}_o(t) + A_2 R_2 \ddot{q}_o(t)) + A_1 R_2 \dot{q}_o(t)$$

which can be simplified as

$$A_1 A_2 R_1 R_2 \ddot{q}_o(t) + (A_1 R_1 + A_2 R_2 + A_1 R_2)\dot{q}_o(t) + q_o(t) = q_i(t) \tag{3.45}$$

This is a second-order differential equation that describes the hydraulic system in Figure 3.10.

Before proceeding, we mention that once a differential equation of a CT system is obtained, the impulse response of the system can be obtained from the differential equation as in the DT case discussed in Section 3.4.2. However, there is a big difference. A difference equation can be solved by direct substitution as we did in Section 3.4.2. Solving differential equations is more complex. It can be solved directly in the time domain or indirectly in the transform domain. The latter method is simpler and will be discussed in Chapter 6.

3.8.2 Op-Amp Circuits

Consider the operational amplifier discussed in Figure 2.16. In order for the op amp to operate as a linear element, we must introduce feedback from the output terminal to the input terminals either directly or through resistors or capacitors. Without feedback, they easily run into saturation region and become nonlinear. Consider the op-amp circuits shown in Figures 3.11(a) and 3.11(b) where all voltages are with respect to ground as shown in Figures 2.16 and 2.17. They are multi-input single-output (MISO) systems. The input signals v_i, for $i = 1, 2, 3$, are connected to the inverting terminal with resistances R/a, R/b, and R/c, where a, b, and c are positive constants. Figure 3.11(a) has a resistor in its feedback path, and Figure 3.11(b) has a capacitor in its feedback path. Both noninverting terminals are grounded, thus we have $e_+ = 0$. We develop in the following their equations. The two circuits can be used to simulate or implement CT LTI and lumped systems as we will discuss in Chapter 7.

Consider the op-amp circuit shown in Figure 3.11(a). Let v_o be its output voltage. If the op amp is modeled as ideal, then we have $e_- = e_+ = 0$ and $i_- = 0$. Consequently the currents, flowing into the inverting terminal, in all resistors are $i_o = (v_o - e_-)/R = v_o/R$, $i_1 = (v_1 - e_-)/(R/a) = av_1/R$, $i_2 = bv_2/R$, and $i_3 = cv_3/R$, respectively. Kirchhoff's current law states that the algebraic sum of all currents entering the node must be zero. Thus we have, using $i_- = 0$,

$$\frac{v_o(t)}{R} + \frac{av_1(t)}{R} + \frac{bv_2(t)}{R} + \frac{cv_3(t)}{R} = 0$$

which implies

$$v_o(t) = -\left(av_1(t) + bv_2(t) + cv_3(t)\right) \tag{3.46}$$

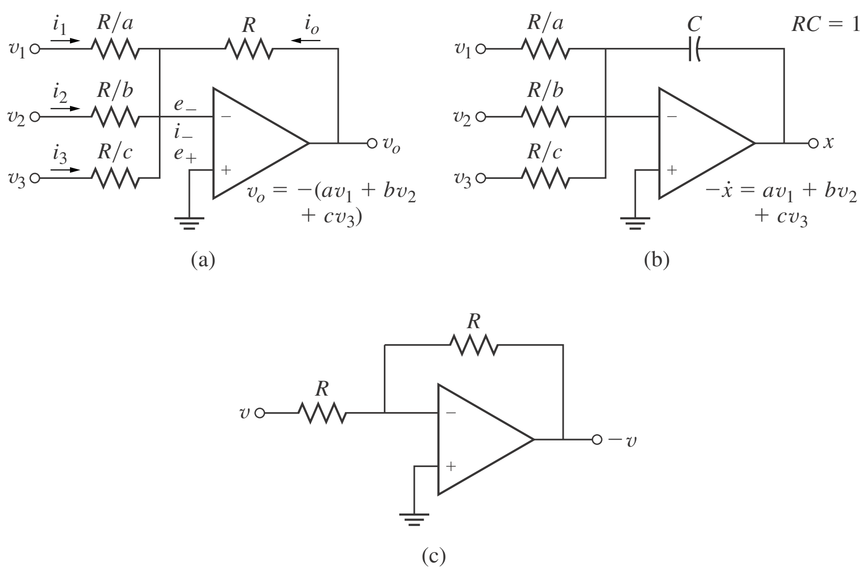


Figure 3.11 (**a**) Adder. (**b**) Integrator. (**c**) Inverter.

This algebraic equation describes the op-amp circuit in Figure 3.11(a). The circuit is memoryless and functions as multipliers and adders. Because all signals are connected to the inverting terminal, there is a minus sign in (3.46). It is important to mention that we have freedom in assigning the output voltage as v_o or $-v_o$. If we assign it as $-v_o$, then we have $v_o = av_1 + bv_2 + cv_3$.

Next we consider the circuit in Figure 3.11(b). Let the output voltage be denoted by $x(t)$. Then the current in the capacitor with capacitance C is $C\dot{x}(t)$. The currents in the three input resistors are av_1/R, bv_2/R, and cv_3/R, respectively. Thus we have

$$C\dot{x}(t) + \frac{av_1(t)}{R} + \frac{bv_2(t)}{R} + \frac{cv_3(t)}{R} = 0$$

or, for $RC = 1$ by setting, for example, $R = 1\ M\Omega = 10^6\ \Omega$ and $C = 1\ \mu F = 10^{-6} F$,

$$\dot{x}(t) = -(av_1(t) + bv_2(t) + cv_3(t)) \tag{3.47}$$

This is a first-order differential equation. This element functions as multipliers, adders, and integrators. Again we have freedom in assigning the output voltage as $x(t)$ or $-x(t)$. If we assign it as $-x(t)$, then we have $\dot{x}(t) = av_1(t) + bv_2(t) + cv_3(t)$.

If we use only the first input terminal in Figure 3.11(a) with $a = 1$, then the circuit becomes the one in Figure 3.11(c). The circuit changes only the sign of the input signal and is called an *inverter*. See Problem 2.12.

PROBLEMS

3.1 Consider the savings account in Example 3.2.1 or Example 3.3.1. If we deposit $u[0] = \$10{,}000.00$, $u[10] = -500.00$, and $u[20] = 1000.00$, what is the total amount of money in the account at the end of one year ($n = 365$)?

3.2 Consider the sequence $h = [1\ \ 2\ \ 3\ \ -2]$ which is located at $n = 0:3$ and $u = [1\ \ 0\ \ -2\ \ 4\ \ 4\ \ 5]$ which is located at $n = 0:5$. Compute their convolution for $n = -2:12$.

3.3 Consider the polynomials

$$h(s) = s^3 + 2s^2 + 3s - 2$$

and

$$u(s) = s^5 - 2s^3 + 4s^2 + 4s + 5$$

Compute $h(s)u(s)$. Verify that its coefficients are the same as those for $n = 0:8$ in Problem 3.2. Thus the multiplication of two polynomials can be computed as the convolution of their coefficients.

3.4 Consider a DT LTI system with impulse response $h[n] = 2e^{0.2n}$, for $n = 0, 1, 2, \ldots$. What is its output $y[n]$ excited by the input $u[n]$? Compute $y[n]$ for $n = 0:3$ if $u[n] = 1/(n+1)$ for $n = 0, 1, 2, \ldots$.

3.5 Consider the system in Problem 3.4. Find a difference equation to describe the system.

3.6 Consider a DT LTI system with impulse response $h[0] = 0$ and $h[n] = 1$ for all $n \geq 1$. Find a difference equation to describe the system.

3.7 Consider a DT system described by the advanced-form difference equation

$$2y[n+3] - y[n+2] + 3y[n+1] = u[n+3] + 2u[n+1]$$

What is its order? Can you simplify the equation and reduce its order?

3.8 Change the simplified equation in Problem 3.7 into a delayed-form difference equation.

3.9 Consider the difference equation

$$y[n] + 2y[n-1] = u[n-1] + 3u[n-2] + 2u[n-3]$$

Does it describe a causal system? What is its order?

3.10 Compute the impulse response of the DT system in Problem 3.9. Is it FIR or IIR? If it is FIR, what is its length? Can you find a nonrecursive difference equation to describe the system?

3.11 Compute the impulse response of the difference equation

$$y[n] + 2y[n-1] = u[n-1] + 3u[n-2] - 2u[n-3]$$

Is it FIR or IIR?

3.12 Design a 20-day moving average filter. Find a nonrecursive difference equation and a recursive difference equation to describe it. Which is preferable?

3.13 Consider the positive-feedback system shown in Figure 3.12(a). Find its impulse response.

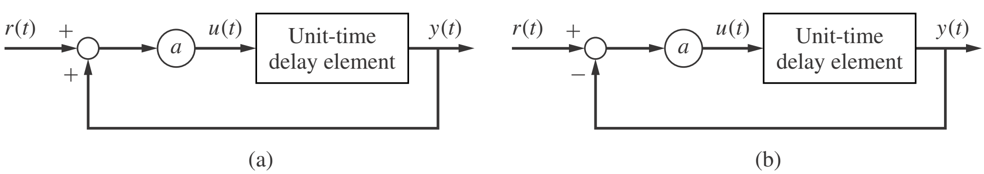


Figure 3.12

3.14 Consider the negative-feedback system shown in Figure 3.12(b). Find its impulse response.

3.15 Compute the integral convolution of $h_i(t)$ and $u_i(t)$, for $i = 1, 2$, shown in Figure 3.13.

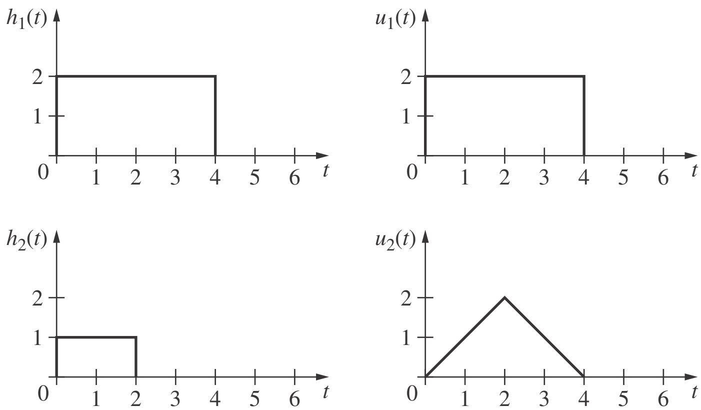


Figure 3.13

3.16 Compute

$$y_1(t) = \int_{-\infty}^{\infty} f_1(t-\tau) f_2(\tau)\, d\tau$$

and

$$y_2 = \int_{-\infty}^{\infty} f_1(t) f_2(t)\, dt$$

for $f_1(t)$ and $f_2(t)$ shown in Figure 3.14. Note that the first integration is a convolution and yields a function of t. The second integration yields just a number.

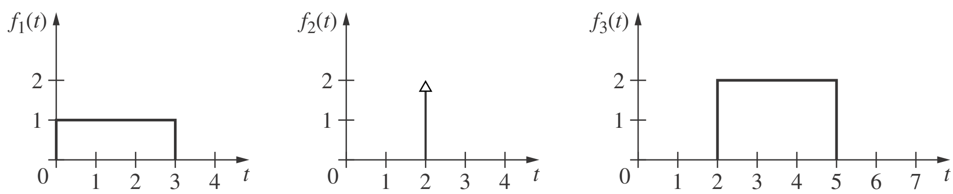


Figure 3.14

3.17 Repeat Problem 3.16 if $f_2(t)$ is replaced by $f_3(t)$ shown in Figure 3.14.

3.18 Find a differential equation to describe the network shown in Figure 3.15.

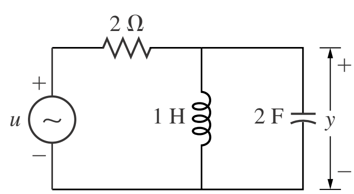


Figure 3.15

3.19 Consider the pendulum system shown in Figure 3.16, where the input is the horizontal force $u(t)$ and the output is the angular displacement θ. It is assumed that there is no friction in the hinge and no air resistance. Find a differential equation to describe the system. Is the equation a linear equation? If θ is small so that $\sin\theta$ and $\cos\theta$ can be approximated as θ and 1, respectively, what is its differential equation?

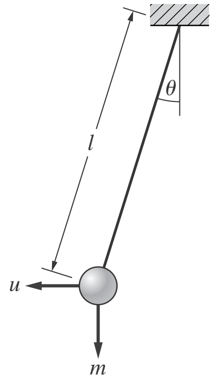


Figure 3.16

3.20 Find an equation to describe the op-amp circuit shown in Figure 3.17(a). It is a *differentiator*.

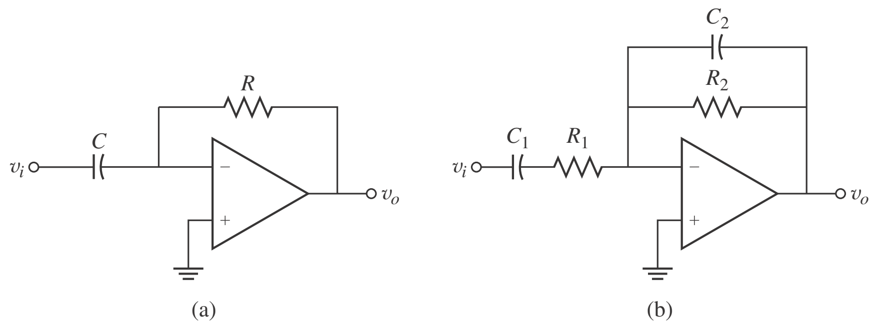


Figure 3.17

3.21 Verify that the circuit in Figure 3.17(b) is described by the second-order differential equation

$$\ddot{v}_o(t) + \left(\frac{1}{R_1C_1} + \frac{1}{R_2C_2}\right)\dot{v}_o(t) + \frac{1}{R_1R_2C_1C_2}v_o(t) = -\frac{1}{R_1C_2}\dot{v}_i(t)$$

CHAPTER 4 Frequency Spectra of CT Signals

4.1 INTRODUCTION

A CT signal can be expressed as $x(t)$, a function of time t. This is called the time-domain description. In this chapter, we shall develop a different expression $X(\omega)$, a function of frequency ω, to describe the signal. This is called the frequency-domain description. The two descriptions are equivalent and contain the same amount of information. However, the latter will reveal *explicitly* the frequency content of the signal. This information, as we will show in the next chapter, is needed in selecting a sampling period if the CT signal is to be processed digitally. It will also be used to specify systems to be designed. Thus the frequency-domain description is indispensable in engineering.

It is possible to introduce the frequency-domain description axiomatically—that is, by defining it directly from $x(t)$. We adopt the approach of introducing an intermediate step by expressing a periodic time signal in the Fourier series, in which the signal is expressed in terms of time and frequency. We then modify it so that the new function depends only on frequency.

Even though all signals encountered in practice are real-valued, because mathematical developments are almost identical for real-valued and complex-valued signals, all equations in this chapter are developed for complex-valued signals. We will then specialize the results to the real-valued case. We also assume signals to be defined for all t in $(-\infty, \infty)$ to simplify the discussion of spectra of periodic signals.

4.1.1 Orthogonality of Complex Exponentials

The frequency of $\sin\omega_0 t$, $\cos\omega_0 t$, and $e^{j\omega_0 t}$ are well defined and equals ω_0 rad/s or $\omega_0/2\pi$ Hz. Consider two sinusoids with frequency ω_1 and ω_2. As discussed in Section 1.7, if ω_1 and ω_2 have no common divisor, then any linear combination of the two sinusoids is not periodic. If ω_1 and ω_2 have common divisors, any linear combination of the two sinusoids is periodic with fundamental frequency equal to the greatest common divisor (gcd) of ω_1 and ω_2. For example, because 2 and π have no common divisor, the signal

$$x(t) = \sin 2t + \sin \pi t$$

is not periodic. On the other hand, the signal

$$x(t) = -2 + 3\sin 0.6t - 4\cos 0.6t - 2.4\cos 2.7t \tag{4.1}$$

is periodic with fundamental frequency 0.3 rad/s, which is the gcd of 0.6 and 2.7. The periodic function in (4.1) can be expressed using sine and cosine functions as shown, exclusively sine functions, or exclusively cosine functions. Such expressions, however, are not as convenient as the expression using complex exponentials, thus we use mostly complex exponentials in signal analysis. As discussed in Sections 1.7.2 and 1.7.3, in using complex exponentials, we encounter both positive and negative frequencies and complex coefficients.

Let us consider the set of complex exponentials

$$\phi_m(t) := e^{jm\omega_0 t} = \cos m\omega_0 t + j \sin m\omega_0 t, \qquad m = 0, \pm 1, \pm 2, \ldots \tag{4.2}$$

where m is an integer, ranging from $-\infty$ to ∞. There are infinitely many $\phi_m(t)$. It is clear that $\phi_m(t)$ is a complex-valued time function with frequency $m\omega_0$. For $m = 0$, the function equals 1 for all t and is called a dc signal. For $m = 1$, the function is $\cos \omega_0 t + j \sin \omega_0 t$ and has fundamental period $P := 2\pi/\omega_0$. For $m = 2$, the function is $\cos 2\omega_0 t + j \sin 2\omega_0 t$ and has fundamental period $2\pi/2\omega_0 = P/2$. Thus the function $e^{j2\omega_0 t}$ repeats itself twice in every time interval P. In general, $e^{jm\omega_0 t}$ is periodic with fundamental period P/m and repeats itself m times in every time interval P. Thus all $\phi_m(t)$ are periodic with (not necessarily fundamental) period P and are said to be *harmonically related*. There are infinitely many harmonically related complex exponentials.

Because every time interval P contains *complete* cycles of $e^{jm\omega_0 t}$, we have

$$\begin{aligned}\int_{<P>} e^{jm\omega_0 t}\,dt &:= \int_{t_0}^{t_0+P} e^{jm\omega_0 t}\,dt = \int_{t_0}^{t_0+P} (\cos m\omega_0 t + j \sin m\omega_0 t)\,dt \\ &= \begin{cases} P & \text{if } m = 0 \\ 0 & \text{if } m \neq 0 \end{cases}\end{aligned} \tag{4.3}$$

for any t_0. The notation $<P>$ is used to denote *any* time interval P. If $m = 0$, the integrand becomes 1 and its integration over P equals P. If $m \neq 0$, the positive part and negative part of $\cos m\omega_0 t$ and $\sin m\omega_0 t$ cancel out completely over P, thus the integration is 0. This establishes (4.3). Of course, (4.3) can also be established by direct integration.

The set of complex exponentials $\phi_m(t)$, for all integers m, has a very important property. Let us use an asterisk to denote complex conjugation. For example, we have

$$\begin{aligned}\phi_m^*(t) &= (e^{jm\omega_0 t})^* = (\cos m\omega_0 t + j \sin m\omega_0 t)^* \\ &= \cos m\omega_0 t - j \sin m\omega_0 t = e^{-jm\omega_0 t}\end{aligned}$$

Then we have

$$\begin{aligned}\int_{<P>} \phi_m(t)\phi_k^*(t)\,dt &= \int_{t_0}^{t_0+P} e^{jm\omega_0 t} e^{-jk\omega_0 t}\,dt = \int_{<P>} e^{j(m-k)\omega_0 t}\,dt \\ &= \begin{cases} P & \text{if } m = k \\ 0 & \text{if } m \neq k \end{cases}\end{aligned} \tag{4.4}$$

for any t_0. This follows directly from (4.3) and is called the *orthogonality* property of the set $\phi_m(t)$.

Consider

$$x(t) = \sum_{m=-\infty}^{\infty} c_m e^{jm\omega_0 t} \tag{4.5}$$

where c_m are real or complex constants. It is a linear combination of the periodic functions in (4.2). Because their frequencies have the greatest common divisor ω_0, the function in (4.5) is periodic with fundamental frequency ω_0. In fact, most periodic signals can be so expressed as we develop next.

4.2 FOURIER SERIES OF PERIODIC SIGNALS—FREQUENCY COMPONENTS

A CT signal $x(t)$ is, as discussed in Section 1.7, periodic with period P if $x(t) = x(t + P)$ for all t. The smallest such P is called the *fundamental period*. Its *fundamental frequency* is defined as $\omega_0 = 2\pi/P$ in rad/s or $f_0 = 1/P$ in Hz. We show that such a signal can be expressed as

$$x(t) = \sum_{m=-\infty}^{\infty} c_m e^{jm\omega_0 t} \qquad \text{(Synthesis equation)} \tag{4.6}$$

and

$$c_m = \frac{1}{P}\int_{<P>} x(t)e^{-jm\omega_0 t}\,dt = \frac{1}{P}\int_{-P/2}^{P/2} x(t)e^{-jm\omega_0 t}\,dt \qquad \text{(Analysis equation)} \tag{4.7}$$

for $m = 0, \pm 1, \pm 2, \ldots$, and m is called the *frequency index* because it is associated with frequency $m\omega_0$. The set of two equations in (4.6) and (4.7) is called the *CT Fourier series* (CTFS). We call c_m the *CTFS coefficients* or the *frequency components*. As we shall discuss shortly, the coefficients c_m reveal the frequency content of the signal. Thus (4.7) is called the *analysis equation*. The function $x(t)$ can be constructed from c_m by using (4.6). Thus (4.6) is called the *synthesis equation*.

To develop (4.7) from (4.6), we multiply (4.6) by $e^{-jk\omega_0 t}$ and then integrate the resulting equation over a time interval P to yield

$$\begin{aligned}\int_{t=<P>} x(t)e^{-jk\omega_0 t}\,dt &= \int_{t=<P>}\left(\sum_{m=-\infty}^{\infty} c_m e^{jm\omega_0 t}e^{-jk\omega_0 t}\right)dt \\ &= \sum_{m=-\infty}^{\infty} c_m\left(\int_{t=<P>} e^{j(m-k)\omega_0 t}\,dt\right)\end{aligned} \tag{4.8}$$

where we have changed the order of integration and summation.[1] Note that in the equation, the integer k is fixed and m ranges over all integers. The orthogonality property in (4.4) implies that the integral in the last parentheses of (4.8) equals 0 if $m \neq k$ and P if $m = k$. Thus (4.8) reduces to

$$\int_{<P>} x(t)e^{-jk\omega_0 t}\,dt = c_k P$$

This becomes (4.7) after renaming the index k to m. This establishes the Fourier series pair. We give some examples.

---

[1] The condition of uniform convergence is needed to permit such a change of order. We will not be concerned with this mathematical condition and will proceed intuitively.

EXAMPLE 4.2.1

Consider the CT periodic signal with period $P = 4$ shown in Figure 4.1(a). It is a train of boxcars with width $2a = 2$ and height 1 and can be expressed as

$$x(t) = \sum_{k=-\infty}^{\infty} x_p(t + kP)$$

where

$$x_p(t) = \begin{cases} 1 & \text{for } |t| \le a \\ 0 & \text{for all } |t| > a \end{cases}$$

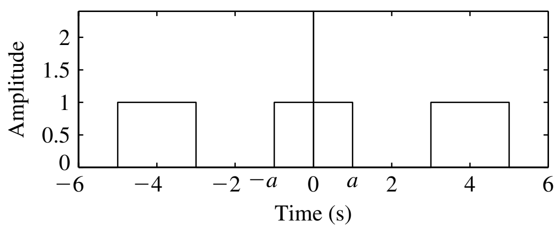
This expression or the plot in Figure 4.1(a) is called the *time-domain* description.

We now develop its Fourier series. Its fundamental frequency is $\omega_0 = 2\pi/P$. First we compute

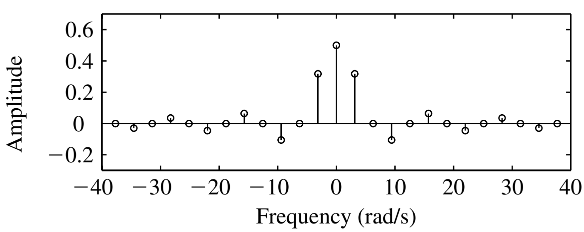
$$c_0 = \frac{1}{P}\int_{-P/2}^{P/2} x(t)\,dt = \frac{1}{P}\int_{-a}^{a} 1\,dt = \frac{2a}{P}$$

and

$$\begin{aligned} c_m &= \frac{1}{P}\int_{-P/2}^{P/2} x(t)e^{-jm\omega_0 t}\,dt = \frac{1}{P}\int_{-a}^{a} e^{-jm\omega_0 t}\,dt \\ &= \frac{1}{-jm\omega_0 P}(e^{-jm\omega_0 a} - e^{jm\omega_0 a}) = \frac{2\sin m\omega_0 a}{m\omega_0 P} \end{aligned} \tag{4.9}$$



(a)



(b)

Figure 4.1 (**a**) Train of boxcars. (**b**) Its CTFS coefficients.

Note that c_0 is the average value of the signal over one period and can also be obtained from (4.9) using l'Hôpital's rule. Thus the train of boxes can also be expressed as

$$x(t) = \sum_{m=-\infty}^{\infty} c_m e^{jm\omega_0 t}$$

with $\omega_0 = 2\pi/P$ and c_m given in (4.9). It is a function of time and frequency. Although it is not exactly the frequency-domain description, it does reveal the frequency content of the signal. It has frequency components at $m\omega_0$ with coefficient c_m. Because c_m are real-valued, they can be plotted as shown in Figure 4.1(b) for $P = 4$ and $a = 1$. Note that its vertical ordinate is amplitude, not magnitude. A magnitude cannot assume a negative number.

EXAMPLE 4.2.2 (Switching Function)

We discuss a special case of Example 4.2.1 in which $a = P/4$. Substituting $\omega_0 = 2\pi/P$ and $a = P/4$ into (4.9) yields

$$c_m = \frac{2\sin(m(2\pi/P)(P/4))}{m(2\pi/P)P} = \frac{\sin(m\pi/2)}{m\pi}$$

which implies $c_m = c_{-m}$, $c_0 = 1/2$, and

$$c_m = \begin{cases} 0 & \text{for } m = \pm 2, \pm 4, \pm 6, \ldots \\ 1/m\pi & \text{for } m = \pm 1, \pm 5, \pm 9, \ldots \\ -1/m\pi & \text{for } m = \pm 3, \pm 7, \pm 11, \ldots \end{cases}$$

Thus the train of boxcars with $a = P/4$, called $p(t)$, can be expressed as

$$\begin{aligned} p(t) &:= \sum_{m=-\infty}^{\infty} c_m e^{jm\omega_0 t} \\ &= \cdots + c_{-3}e^{-j3\omega_0 t} + c_{-1}e^{-j\omega_0 t} + c_0 + c_1 e^{j\omega_0 t} + c_3 e^{j3\omega_0 t} + \cdots \\ &= c_0 + 2c_1 \cos\omega_0 t + 2c_3 \cos 3\omega_0 t + 2c_5 \cos 5\omega_0 t + \cdots \\ &= \frac{1}{2} + \frac{2}{\pi}\cos\omega_0 t - \frac{2}{3\pi}\cos 3\omega_0 t + \frac{2}{5\pi}\cos 5\omega_0 t + \cdots \end{aligned} \tag{4.10}$$

It consists of only cosine functions. The function $p(t)$ can act as a switch as we will discuss in Example 4.2.5 and is therefore called a *switching function*.

EXAMPLE 4.2.3

Consider the signal in (1.24) or

$$x(t) = -2 + 3\sin 0.6t - 4\cos 0.6t - 2.4\cos 2.7t \tag{4.11}$$

It is periodic with fundamental frequency $\omega_0 = 0.3$ rad/s and fundamental period $2\pi/\omega_0 = 20.9$. For this signal, there is no need to use (4.7) to compute c_m. Using

Euler's formula, we can obtain its Fourier series as

$$x(t) = 2e^{j\pi}e^{j0\cdot t} + 2.5e^{-j2.5}e^{j0.6t} + 2.5e^{j2.5}e^{-j0.6t} + 1.2e^{j\pi}e^{j2.7t} + 1.2e^{j\pi}e^{-j2.7t} \quad (4.12)$$

See (1.31). Thus we have

$$c_0 = 2e^{j\pi}, \qquad c_{\pm 2} = 2.5e^{\mp j2.5}, \qquad c_{\pm 9} = 1.2e^{j\pi}$$

and the rest of c_m equal 0. Note that c_m is associated with complex exponential $e^{jm\omega_0}$ that has frequency $m\omega_0$. Because c_m are complex, plotting them against frequency requires three dimensions and is difficult to visualize. Thus we plot their magnitudes and phases against ω as shown in Figure 1.25. Note that we may plot their real and imaginary parts against frequencies, but such plots have no physical meaning.

EXAMPLE 4.2.4 (Sampling Function)

Consider the sequence of impulses shown in Figure 4.2. The impulse at $t = 0$ can be denoted by $\delta(t)$, the ones at $t = \pm T$ can be denoted by $\delta(t \mp T)$, and so forth. Thus the sequence can be expressed as

$$r(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT) \quad (4.13)$$

It is periodic with fundamental period $P = T$ and fundamental frequency $\omega_0 = 2\pi/T$. Its CT Fourier series is

$$r(t) = \sum_{m=-\infty}^{\infty} c_m e^{jm\omega_0 t} \quad (4.14)$$

with, using (4.7) and the sifting property of the impulse in (1.12),

$$c_m = \frac{1}{T}\int_{-T/2}^{T/2} \delta(t)e^{-jm\omega_0 t}\,dt = \frac{1}{T}e^{-jm\omega_0 t}\Big|_{t=0} = \frac{1}{T}$$

for all m. Thus we have

$$r(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT) = \sum_{m=-\infty}^{\infty} \frac{1}{T}e^{jm\omega_0 t} \quad (4.15)$$

This is called the *sampling function* for reasons to be given in the next chapter. The first expression has the time index n ranging from $-\infty$ to ∞, and the second expression has the frequency index m ranging from $-\infty$ to ∞. This is a very useful equation.

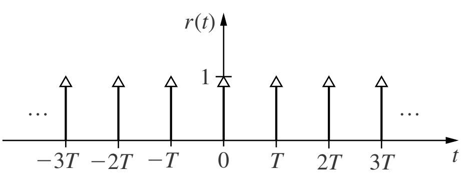


Figure 4.2 Impulse sequence.

EXAMPLE 4.2.5

Consider the half-wave rectifier shown in Figure 2.20. The diode can be considered as a switch. If the applied voltage v_i is positive, the diode acts as a short circuit or the switch is closed. If $v_i < 0$, the diode acts as an open circuit or the switch is open. If $v_i(t) = \cos\omega_0 t$, then the output $v_o(t)$ is as shown in Figure 4.3 and can be expressed as

$$v_o(t) = v_i(t)p(t) = \cos\omega_0 t \cdot p(t)$$

where $p(t)$ is given in (4.10) and plotted in Figure 4.1(a). The switch is closed when $p(t) = 1$ and open when $p(t) = 0$. Using (4.10), we have

$$v_o(t) = \cos\omega_0 t\left[\frac{1}{2} + \frac{2}{\pi}\cos\omega_0 t - \frac{2}{3\pi}\cos 3\omega_0 t + \frac{2}{5\pi}\cos 5\omega_0 t + \cdots\right]$$

which, using $\cos\theta\cos\phi = [\cos(\theta+\phi) + \cos(\theta-\phi)]/2$, can be written as

$$\begin{aligned}
v_o(t) &= \frac{1}{2}\cos\omega_0 t + \frac{1}{\pi}(\cos 2\omega_0 t + 1) - \frac{1}{3\pi}(\cos 4\omega_0 t + \cos 2\omega_0 t) \\
&\quad + \frac{1}{5\pi}(\cos 4\omega_0 t + \cos 6\omega_0 t) + \cdots \\
&= \frac{1}{\pi} + \frac{1}{2}\cos\omega_0 t + \left(\frac{1}{\pi} - \frac{1}{3\pi}\right)\cos 2\omega_0 t - \left(\frac{1}{3\pi} - \frac{1}{5\pi}\right)\cos 4\omega_0 t + \cdots \\
&= \frac{1}{\pi} + \frac{1}{2}\cos\omega_0 t + \frac{2}{3\pi}\cos 2\omega_0 t - \frac{2}{15\pi}\cos 4\omega_0 t + \cdots
\end{aligned}$$

This output contains frequencies $0, \omega_0, 2\omega_0, 4\omega_0, \ldots$, even though the input contains only frequency ω_0. Thus the switching, a nonlinear operation, can generate new frequencies. This fact will be used in modulation and demodulation in Chapter 8.

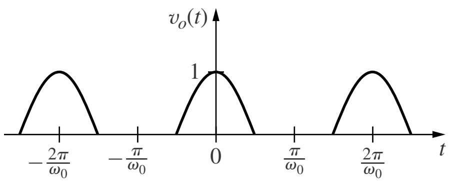


Figure 4.3 Output of a half-wave rectifier.

Not every CT periodic signal can be expressed as a Fourier series. The conditions for the existence of Fourier series and the questions of convergence are complicated. See, for example, Reference 13. We mention only the conditions often cited in engineering texts. A signal is said to be of *bounded variation* if it has a finite number of finite discontinuities and a finite number of maxima and minima in any finite time interval. Figure 4.4(a) shows the function $x(t) = \sin 1/t$, for t in [0, 1]. It has infinitely many minima and maxima. Figure 4.3(b) shows a function in [0, 1]; it equals 1 in [0, 0.5], 0.5 in [0.5, 0.75], 0.25 in $[0.75, 0.75 + (1 - 0.75)/2]$, and so forth. It has infinitely many discontinuities. These functions are mathematically contrived and cannot arise in practice. Thus we assume that *all signals in this text are of bounded variation*.

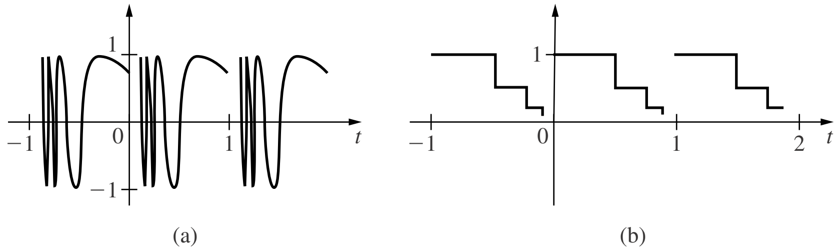


Figure 4.4 (**a**) Signal that has infinitely many maxima and minima. (**b**) Signal that has infinitely many discontinuities.

Next we discuss a sufficient condition. If $x(t)$ is absolutely integrable over one period, that is,

$$\int_0^P |x(t)|\,dt \le M < \infty$$

then its Fourier series exists. Furthermore, all its coefficients are finite. Indeed, because $|e^{-jm\omega_0 t}| = 1$, (4.7) implies

$$\begin{aligned}|c_m| &= \left|\frac{1}{P}\int_0^P x(t)e^{-jm\omega_0 t}\,dt\right| \le \frac{1}{P}\int_0^P |x(t)||e^{-jm\omega_0 t}|\,dt \\ &= \frac{1}{P}\int_0^P |x(t)|\,dt < \infty\end{aligned}$$

Moreover, the infinite series in (4.6) converges to a value at every t. The value equals $x(t)$ if $x(t)$ is continuous at t. If $x(t)$ is not continuous at t—that is, $x(t_+) \neq x(t_-)$—then we have

$$\frac{x(t_+) + x(t_-)}{2} = \sum_{m=-\infty}^{\infty} c_m e^{jm\omega_0 t} \tag{4.16}$$

That is, the infinite series converges to the midpoint at every discontinuity. Note that a simpler sufficient condition is that $x(t)$ is bounded. If $x(t)$ is bounded, then it is automatically absolutely integrable in one period. The two conditions, bounded variations and absolute integrability, are called the *Dirichlet conditions*.

In the literature, the set of c_m is often called the *discrete frequency spectrum*. We will not use the terminology as we will explain in a later section. Instead, we call c_m the CTFS coefficients or the frequency components.

4.2.1 Properties of Fourier Series Coefficients

We discuss a general property of Fourier series coefficients. In general, c_m are complex-valued. However, if $x(t)$ is real-valued, then we have, as shown in Section 1.7.3,

$$|c_m| = |c_{-m}| \quad \text{and} \quad \measuredangle c_m = -\measuredangle c_{-m} \tag{4.17}$$

In other words, their magnitudes are even with respect to frequency $m\omega_0$ or frequency index m and their phases are odd. We establish these properties once again using (4.7). First we write

(4.7) as

$$c_m = \frac{1}{P}\int_{t=-P/2}^{P/2} x(t)(\cos m\omega_0 t - j\sin m\omega_0 t)\,dt$$

If $x(t)$ is real, then we have

$$\text{Re}(c_m) = \frac{1}{P}\int_{-P/2}^{P/2} x(t)\cos m\omega_0 t\,dt \tag{4.18}$$

$$\text{Im}(c_m) = -\frac{1}{P}\int_{-P/2}^{P/2} x(t)\sin m\omega_0 t\,dt \tag{4.19}$$

These imply, because $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$,

$$\text{Re}(c_m) = \text{Re}(c_{-m}), \qquad \text{Im}(c_m) = -\text{Im}(c_{-m}) \tag{4.20}$$

and $c_m = c_{-m}^*$. Such c_m are said to be *conjugate symmetric;* their real parts are even and their imaginary parts are odd. Because

$$|c_m| = \sqrt{(\text{Re}(c_m))^2 + (\text{Im}(c_m))^2}$$

and

$$\measuredangle c_m = \tan^{-1}(\text{Im}(c_m))/\text{Re}(c_m))$$

the properties in (4.20) imply the properties in (4.17). In conclusion, if $x(t)$ is a real-valued periodic signal, its CTFS coefficients are conjugate symmetric, their real parts and magnitudes are even, their imaginary parts and phases are odd. If c_m are complex, plotting them against frequencies requires three dimensions and the resulting plot is difficult to visualize. Thus we plot their magnitudes and phases against frequencies. Recall from Section 1.7.3 that two degrees are considered the same if they differ by 360° or its integer multiple. Thus we have $180^\circ = -180^\circ$ (modulo 360) and may discard 180° in discussing the oddness of phases.

If $x(t)$ is real and even $[x(t) = x(-t)]$, the product $x(t)\sin m\omega_0 t$ is odd and the integration in (4.19) is zero. Thus, if $x(t)$ is real and even, so are its CTFS coefficients. The signal in Figure 4.1(a) is real and even, and so are its CTFS coefficients in (4.9) as shown in Figure 4.1(b).

EXERCISE 4.2.1

Show that if $x(t)$ is real and odd $[x(t) = -x(-t)]$, then its Fourier series coefficients are pure imaginary and odd.

Answer

The integration in (4.18) is zero.

We can develop many other properties of Fourier series coefficients. However, they are the same as those to be developed in Section 4.4 and will not be discussed here.

4.2.2 Distribution of Average Power in Frequencies

This section discusses the physical significance of CTFS coefficients. Suppose $x(t)$ is a voltage signal. If it is applied to a 1-Ω resistor, then the current passing through it is $x(t)/1$ and the power is $x(t)x(t)$. Thus the total energy provided by $x(t)$ from $-\infty$ to ∞ is

$$E = \int_{-\infty}^{\infty} x^2(t)\, dt$$

where we have assumed implicitly that $x(t)$ is real. If $x(t)$ is complex, then its total energy is defined as

$$E := \int_{-\infty}^{\infty} x(t)x^*(t)\, dt = \int_{-\infty}^{\infty} |x(t)|^2\, dt \tag{4.21}$$

It is clear that every periodic signal has $E = \infty$. Thus it is meaningless to discuss its total energy. Instead, we discuss its average power over one period defined as

$$P_{av} = \frac{1}{P}\int_0^P x(t)x^*(t)\, dt = \frac{1}{P}\int_0^P |x(t)|^2\, dt \tag{4.22}$$

It turns out that this average power can also be computed from the CTFS coefficients of $x(t)$.

If $x(t)$ is absolutely integrable in one period, then its Fourier series exists. This condition, however, does not guarantee that the signal has a finite average power. For example, the sequence of impulses in Figure 4.2 is absolutely integrable in one period, but has infinite average power (Problem 4.3). See Problem 4.4 for a different example. These two time signals are not bounded. If a periodic signal is bounded, then it automatically has a finite average power.

EXERCISE 4.2.2

Consider a periodic signal $x(t)$ with period P. Show that if $x(t)$ is bounded by M, then its average power is less than M^2/P.

Now we show that the average power can also be computed from CTFS coefficients. Substituting (4.6) into (4.22) and interchanging the order of integration and summation yield

$$P_{av} = \frac{1}{P}\int_0^P \left[\sum_{-\infty}^{\infty} c_m e^{jm\omega_0 t}\right] x^*(t)\, dt = \sum_{-\infty}^{\infty} c_m \left[\frac{1}{P}\int_0^P x(t)e^{-jm\omega_0 t}\, dt\right]^*$$

The term inside the brackets is the c_m in (4.7). Thus we have

$$P_{av} = \frac{1}{P}\int_0^P |x(t)|^2\, dt = \sum_{m=-\infty}^{\infty} c_m c_m^* = \sum_{m=-\infty}^{\infty} |c_m|^2 \tag{4.23}$$

This is called a *Parseval's formula*. It states that the average power equals the sum of squared magnitudes of c_m. Because $|c_m|^2$ is the power at frequency $m\omega_0$, the set c_m reveals the distribution of power in frequencies. Note that the average power depends only on the magnitudes of c_m and is independent of the phases of c_m.

EXAMPLE 4.2.6

Consider the train of boxcars shown in Figure 4.1(a) with $P = 4$ and $a = 1$. Find the percentage of the average power lying inside the frequency range $[-2, 2]$ in rad/s.

The average power of the signal can be computed directly in the time domain as

$$P_{av} = \frac{1}{P}\int_{-a}^{a} 1^2\,dt = \frac{1}{P}[a - (-a)] = \frac{1+1}{4} = 0.5$$

Computing the average power using c_m is complicated. However, to compute the average power inside the frequency range $[-2, 2]$, we must use the c_m computed in Example 4.2.1. Because $\omega_0 = 2\pi/P = 2\pi/4 = 1.57$, only $m\omega_0$, with $m = -1, 0, 1$, lie inside the range $[-2, 2]$. From (4.9), we compute $c_0 = 0.5$,

$$c_{-1} = \frac{2\sin(-(\pi/2)\cdot 1)}{-(\pi/2)4} = \frac{2(-1)}{-2\pi} = \frac{1}{\pi} = 0.32$$

and $c_1 = 0.32$. Thus the average power in $[-2, 2]$ is

$$(0.32)^2 + (0.5)^2 + (0.32)^2 = 0.4548$$

It is $0.4548/0.5 = 0.91 = 91\%$ of the total average power. Because most power is located at low frequencies, the train is a low-frequency signal.

EXAMPLE 4.2.7

Consider

$$x(t) = A\sin\omega_0 t = A\frac{e^{j\omega_0 t} - e^{-j\omega_0 t}}{2j} = -0.5Aje^{j\omega_0 t} + 0.5Aje^{-j\omega_0 t}$$

Thus $x(t)$ has Fourier series coefficients $c_1 = -0.5Aj$ and $c_{-1} = 0.5Aj$ and average power

$$P_{av} = |c_1|^2 + |c_{-1}|^2 = 0.25A^2 + 0.25A^2 = 0.5A^2 = A^2/2$$

It can also be computed by direct integration in the time domain. Thus a sinusoid with *peak* amplitude A is said to have *effective* amplitude $A/\sqrt{2}$. The household electricity in the United States has 120 V and 60 Hz or, equivalently, the waveform $120\sqrt{2}\sin(60\cdot 2\pi t)$.

4.3 FOURIER TRANSFORM—FREQUENCY SPECTRA

The CT Fourier series is applicable only to CT periodic signals. Now we shall modify it so that it is applicable to aperiodic signals as well. Before proceeding, we use the periodic signal in Figure 4.1 as an example to illustrate the idea. The signal in Figure 4.1(a) can be expressed as

$$x(t) = \sum_{m=-\infty}^{\infty} c_m e^{jm\omega_0 t}$$

with $\omega_0 = 2\pi/P$ and

$$c_m = \frac{2\sin m\omega_0 a}{m\omega_0 P} \tag{4.24}$$

The periodic signal $x(t)$ repeats itself every P seconds. Now if P becomes infinity, the signal never repeats itself and becomes an aperiodic signal. However, as P becomes infinity, c_m in (4.24) becomes zero for all m and will not reveal the frequency content of the aperiodic signal. Now we consider Pc_m or

$$Pc_m = \frac{2\sin m\omega_0 a}{m\omega_0} \tag{4.25}$$

Its right-hand side is independent of P. We plot in Figure 4.5 the time signal for $P = P_1$, $2P_1$, and ∞ and the corresponding Pc_m. We see that the envelopes of Pc_m are the same for all P. As P increases, the fundamental frequency $\omega_0 = 2\pi/P$ becomes smaller and the discrete frequencies $m\omega_0$ at which Pc_m appear become closer or denser. Eventually, Pc_m appear at every frequency as shown in Figure 4.5(c).

With this preliminary, we are ready to modify the Fourier series. First we write (4.7) as

$$X(m\omega_0) := Pc_m = \int_{t=-P/2}^{P/2} x(t)e^{-jm\omega_0 t}\,dt \tag{4.26}$$

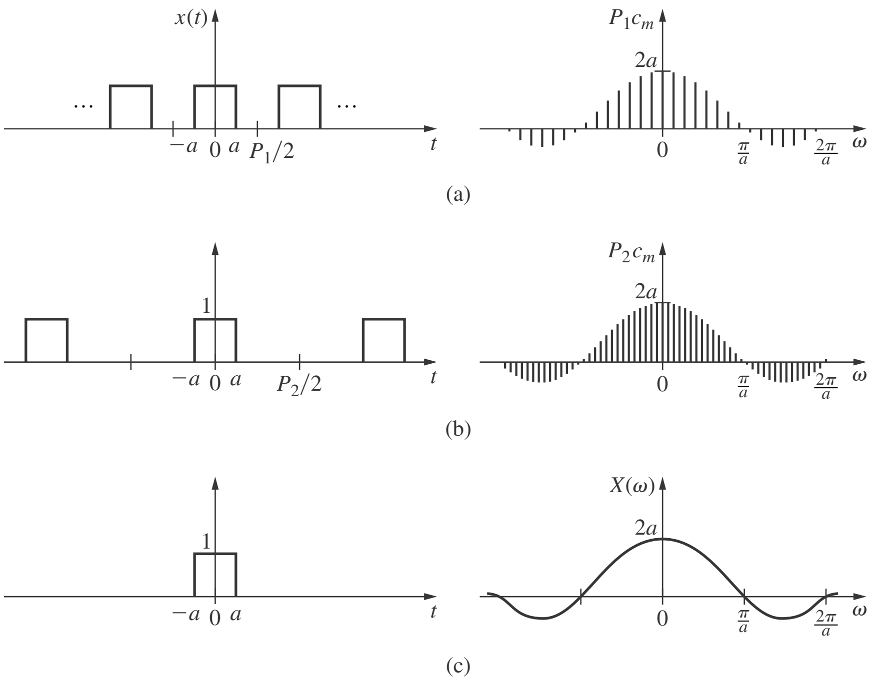


Figure 4.5 Trains of boxcars with period P and their CTFS coefficients multiplied by P: (**a**) $P = P_1$, (**b**) $P = 2P_1$, and (**c**) $P = \infty$.

where we have defined $X(m\omega_0) := Pc_m$. This is justified by the fact that the coefficient c_m is associated with frequency $m\omega_0$. We then use $X(m\omega_0)$ and $\omega_0 = 2\pi/P$ or $1/P = \omega_0/2\pi$ to rewrite (4.6) as

$$x(t) = \frac{1}{P}\sum_{m=-\infty}^{\infty} Pc_m e^{jm\omega_0 t} = \frac{1}{2\pi}\sum_{m=-\infty}^{\infty} X(m\omega_0)e^{jm\omega_0 t}\omega_0 \tag{4.27}$$

A periodic signal with period P becomes aperiodic if P approaches infinity. Define $\omega := m\omega_0$. As $P \to \infty$, we have $\omega_0 = 2\pi/P \to 0$. In this case, ω becomes a continuum and ω_0 can be written as $d\omega$. Furthermore, the summation in (4.27) becomes an integration. Thus the modified Fourier series pair in (4.26) and (4.27) becomes, as $P \to \infty$,

$$X(\omega) = \mathcal{F}[x(t)] := \int_{t=-\infty}^{\infty} x(t)e^{-j\omega t}\,dt \tag{4.28}$$

$$x(t) = \mathcal{F}^{-1}[X(\omega)] := \frac{1}{2\pi}\int_{\omega=-\infty}^{\infty} X(\omega)e^{j\omega t}\,d\omega \tag{4.29}$$

This is the *CT Fourier transform* pair. $X(\omega)$ is called the Fourier transform of $x(t)$ and $x(t)$ the inverse Fourier transform of $X(\omega)$. Equation (4.28) is called the *analysis equation* because it reveals the frequency content of $x(t)$. Equation (4.29) is called the *synthesis equation* because it can be used to construct $x(t)$ from $X(\omega)$. We first give an example.

EXAMPLE 4.3.1

Consider the positive-time function

$$x(t) = \begin{cases} e^{2t} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases}$$

Its Fourier transform is

$$\begin{aligned} X(\omega) &= \int_{t=-\infty}^{\infty} x(t)e^{-j\omega t}\,dt = \int_{t=0}^{\infty} e^{2t}e^{-j\omega t}\,dt \\ &= \frac{1}{2-j\omega}\, e^{(2-j\omega)t}\Big|_{t=0}^{\infty} = \frac{1}{2-j\omega}\left[e^{(2-j\omega)t}\Big|_{t=\infty} - 1\right] \end{aligned}$$

where we have used $e^{(2-j\omega)\times 0} = e^0 = 1$. Because the function, for any ω,

$$e^{(2-j\omega)t} = e^{2t}[\cos\omega t - j\sin\omega t]$$

grows unbounded as $t \to \infty$, the value of $e^{(2-j\omega)t}$ at $t = \infty$ is not defined. Thus the Fourier transform $X(\omega)$ is not defined or does not exist.

As we saw in the preceding example, the Fourier transforms of some functions may not be defined. In general, if a function grows to ∞ or $-\infty$ as $t \to \infty$ or $-\infty$, then it Fourier transform is not defined. We discuss in the following a number of sufficient conditions. If $x(t)$ is *absolutely* *integrable* in $(-\infty, \infty)$, that is,

$$\int_{-\infty}^{\infty} |x(t)|\, dt < \infty$$

then its Fourier transform exists. Furthermore, its Fourier transform is bounded and continuous. Indeed, (4.28) implies, because $|e^{-j\omega t}| = 1$,

$$\begin{aligned} |X(\omega)| &= \left| \int_{t=-\infty}^{\infty} x(t) e^{-j\omega t}\, dt \right| \le \int_{-\infty}^{\infty} |x(t)||e^{-j\omega t}|\, dt \\ &= \int_{-\infty}^{\infty} |x(t)|\, dt \end{aligned}$$

Thus if $x(t)$ is absolutely integrable, then its Fourier transform is bounded. To show that it is a continuous function of ω is more complex, see References 2 (pp. 92–93) and 33.

Another sufficient condition for the existence of Fourier transform is that $x(t)$ is periodic and bounded. Its Fourier transform, as we will derive later, consists of impulses that are not bounded or continuous. Yet another sufficient condition is that $x(t)$ has a finite total energy or, equivalently, is squared absolutely integrable[2] in $(-\infty, \infty)$, that is,

$$\int_{-\infty}^{\infty} |x(t)|^2\, dt < \infty$$

In this case, its Fourier transform $X(\omega)$ is squared absolutely integrable in $(-\infty, \infty)$ and may not be bounded or continuous.

If the Fourier transform of a CT signal exists or is well defined, the transform is called the *frequency spectrum* or, simply, the *spectrum* of the signal. From now on, *the Fourier transform, frequency spectrum, and spectrum will be used interchangeably.*

No practical signal can grow to ∞ or $-\infty$ and last forever. In other words, every practical signal $x(t)$ is bounded and of finite duration. That is, there exists a constant q such that

$$|x(t)| \le q \qquad \text{for all } t \text{ in } (-\infty, \infty)$$

and there exist finite t_1 and t_2 such that

$$x(t) = 0 \qquad \text{for } t < t_1 \text{ and } t > t_2$$

In this case, the signal is automatically absolutely integrable for

$$\int_{-\infty}^{\infty} |x(t)|\, dt \le \int_{t_1}^{t_2} q\, dt = q(t_2 - t_1) < \infty$$

Thus the frequency spectrum of every practical signal is defined and well behaved (no jumps and no infinity).

The frequency spectrum is generally complex-valued. Its magnitude $|X(\omega)|$ is called the *magnitude spectrum,* and its phase $\measuredangle X(\omega)$ is called the *phase spectrum*. As in the Fourier series, if a signal is real-valued, then its magnitude spectrum is even, that is,

$$|X(\omega)| = |X(-\omega)| \qquad \text{(even)}$$

---

[2]If $x(t)$ is real, then $x(t)$ is squared integrable.

and its phase spectrum is odd, that is,

$$\measuredangle X(\omega) = -\measuredangle X(-\omega) \qquad \text{(odd)}$$

If $x(t)$ is real and even, then its spectrum $X(\omega)$ is also real and even. Their proofs are identical to those for Fourier series and will not be repeated. We now give some examples.

EXAMPLE 4.3.2

Consider the positive-time function $x(t) = e^{-at}$ for $t \geq 0$, where a is a positive constant. The function for $a = 0.1$ is plotted in Figure 4.6(a) for $t \geq 0$. It vanishes exponentially as $t \to \infty$. Because

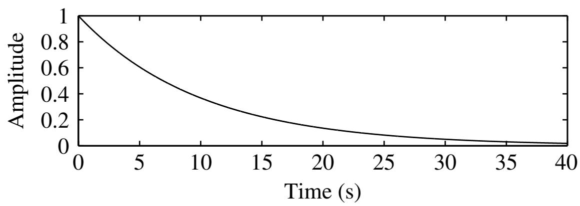
$$\int_{t=-\infty}^{\infty} |x(t)|\,dt = \int_{t=0}^{\infty} e^{-at}\,dt = \frac{1}{-a}\,e^{-at}\Big|_{t=0}^{\infty} = \frac{1}{-a}[0-1] = \frac{1}{a}$$

the function, for any $a > 0$, is absolutely integrable in $(-\infty, \infty)$. Its Fourier transform is

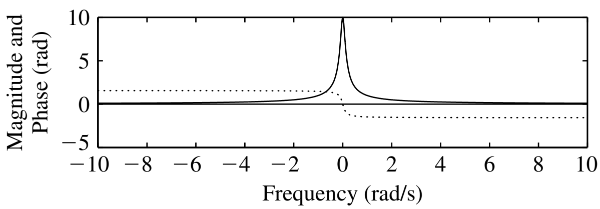
$$\begin{aligned} X(\omega) &= \int_{t=-\infty}^{\infty} x(t)e^{-j\omega t}\,dt = \int_{t=0}^{\infty} e^{-at}e^{-j\omega t}\,dt \\ &= \frac{1}{-a-j\omega}\,e^{(-a-j\omega)t}\Big|_{t=0}^{\infty} = \frac{1}{-a-j\omega}\left[e^{(-a-j\omega)t}\Big|_{t=\infty} - 1\right] \end{aligned}$$

Because $e^{-at}e^{-j\omega t}$, for $a > 0$ and any ω, equals zero as $t \to \infty$, we have

$$X(\omega) = \frac{1}{-a-j\omega}[0-1] = \frac{1}{j\omega + a} \tag{4.30}$$



(a)



(b)

Figure 4.6 (**a**) Time function $e^{-0.1t}$ for $t \geq 0$. (**b**) Its magnitude spectrum (solid line) and phase spectrum (dotted line).

This is the Fourier transform or frequency spectrum of the positive-time function e^{-at}. It is a complex-valued function of ω. Its direct plot against ω requires three dimensions and is difficult to visualize. We may plot its real part and imaginary part. However, these two plots have no physical meaning. Thus we plot its magnitude and phase spectra.

We compute the values of $X(\omega)$, for $a = 0.1$, at $\omega = 0, 0.1, 4,$ and 10 as

$$\omega = 0: \qquad X(0) = \frac{1}{0.1} = 10e^{j0}$$

$$\omega = 0.1: \qquad X(j0.1) = \frac{1}{0.1 + j0.1} = \frac{10}{0.1 \cdot \sqrt{2}e^{j\pi/4}} = 7.07e^{-j\pi/4}$$

$$\omega = 4: \qquad X(j4) = \frac{1}{0.1 + j4} = \frac{1}{4.001e^{j1.55}} = 0.25e^{-j1.55}$$

$$\omega = 10: \qquad X(j10) \approx \frac{1}{j10} = 0.1e^{-j\pi/2}$$

From these, we can plot in Figure 4.6(b) its magnitude spectrum (solid line) and phase spectrum (dotted line). The more ω we compute, the more accurate the plot. Note that the phase approaches $\mp\pi/2 = \mp 1.57$ rad/s as $\omega \to \pm\infty$. Figure 4.6(b) is in fact obtained using MATLAB by typing

```
Program 4.1
w=-10:0.02:10;
X=1.0./(j*w+0.1);
plot(w,abs(X),w,angle(X),':')
```

The first line generates ω from -10 to 10 with increment 0.02. Note that every statement in MATLAB is executed and stored in memory. A semicolon (;) at its end suppresses its display or printout. The second line is (4.30) with $a = -0.1$. Note the use of dot division (./) for element by element division in MATLAB. If we use division without dot (/), then an error message will appear. We see that using MATLAB to plot frequency spectra is very simple. As shown in Figure 4.6, the magnitude spectrum is an even function of ω, and the phase spectrum is an odd function of ω.

The signal in Example 4.3.2 can be represented by e^{-at}, for $t \geq 0$, or by $1/(j\omega + a)$. The former is a function of t and is called the *time-domain description;* the latter is a function of ω and is called the *frequency-domain description*. These descriptions are equivalent in the sense that they contain exactly the same amount of information about the signal and either description can be obtained from the other. However, the frequency-domain description reveals explicitly the frequency content of the signal and is essential, as we will see in this text, in many applications.

The relationship between a time signal $x(t)$ and its spectrum $X(\omega)$ is a global one (defined from $-\infty$ to ∞ in time or in frequency). Thus it is difficult to detect a local property of $x(t)$ from $X(\omega)$ and vice versa. In other words, even though the two descriptions are equivalent, their relationships are not exactly transparent. One thing we can say is that if the rate of change of $x(t)$ is large, then the signal contains high-frequency components. In particular, if $x(t)$ is discontinuous (its rate of change is infinity), then the signal has nonzero, albeit very small, frequency components at $\pm\infty$. Note that the signal in Figure 4.6(a) is discontinuous at $t = 0$. We will discuss other relationships as we proceed.

EXAMPLE 4.3.3 (CT Rectangular Window)

Consider the function shown in Figure 4.7(a) or

$$w_a(t) = \begin{cases} 1 & \text{for } |t| \le a \\ 0 & \text{for } |t| > a \end{cases} \tag{4.31}$$

It is called a CT *rectangular window* with width $L{:=}2a$. Its Fourier transform is

$$\begin{aligned} W_a(\omega) &:= \int_{-\infty}^{\infty} w_a(t)e^{-j\omega t}\,dt = \int_{-a}^{a} e^{-j\omega t}\,dt = \frac{1}{-j\omega}e^{-j\omega t}\Big|_{t=-a}^{a} \\ &= \frac{e^{-j\omega a} - e^{j\omega a}}{-j\omega} = \frac{2(e^{j\omega a} - e^{-j\omega a})}{j2\omega} = \frac{2\sin a\omega}{\omega} \end{aligned} \tag{4.32}$$

It equals (4.25) if we identify $\omega = m\omega_0$. Because the window is real and even, so is its spectrum $W_a(\omega)$.

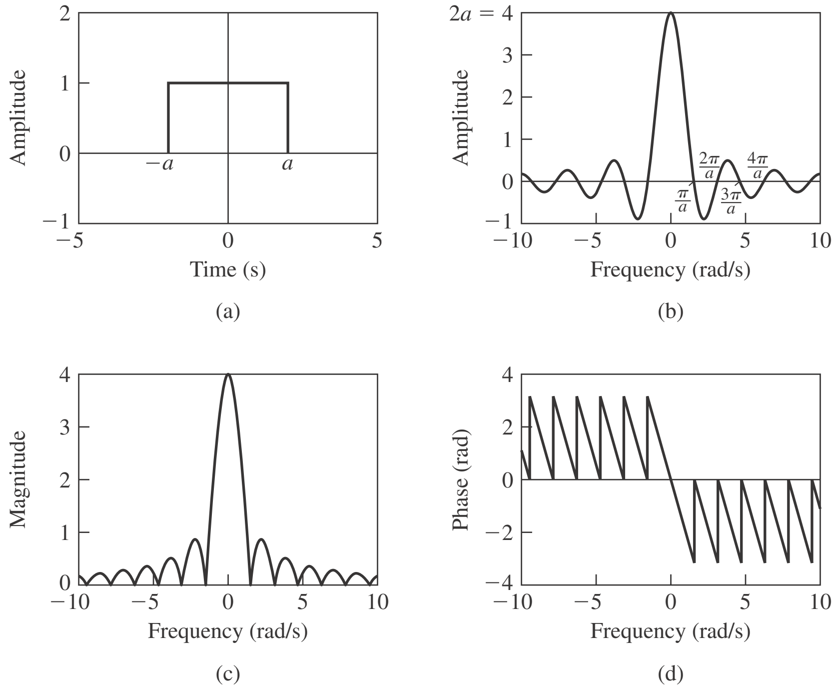


Figure 4.7 (**a**) Rectangular window with $a = 2$. (**b**) Its spectrum. (**c**) Magnitude spectrum of the shifted window in (4.38) with $L = 2a = 4$. (**d**) Phase spectrum of (4.38).

Because $W_a(\omega)$ is real-valued, we can plot it directly as shown in Figure 4.7(b). Of course we can also plot its magnitude and phase spectra (Problem 4.10). Because $\sin a\omega = 0$ at $\omega = k\pi/a$, for all integers k, we have $W_a(k\pi/a) = 0$ for all nonzero integers k. The value of $W_a(0)$, however, is different from zero and can be computed, using l'Hôpital's rule, as

$$W_a(0) = \left.\frac{d(2\sin a\omega)/d\omega}{d\omega/d\omega}\right|_{\omega=0} = \left.\frac{2a\cos a\omega}{1}\right|_{\omega=0} = 2a = L$$

The spectrum consists of (a) one main lobe with base width $2\pi/a$ and height $L = 2a$ and (b) infinitely many side lobes with base width π/a and decreasing magnitudes.

Note that the wider the window, the narrower and higher the main lobe. In fact, as $a \to \infty$, the spectrum approaches an impulse with weight 2π, as we will see in (4.52). Thus the function in (4.32) can also be used to define an impulse.

EXAMPLE 4.3.4 (Analog Ideal Lowpass Filter)

In this example, we will search for a time function that has the spectrum shown in Figure 4.8(a) or

$$H(\omega) = \begin{cases} 1 & \text{for } |\omega| \le \omega_c \\ 0 & \text{for } \omega_c < |\omega| < \infty \end{cases}$$

This function, as we will discuss in Chapter 6, characterizes a CT ideal lowpass filter with cutoff frequency ω_c. We use (4.29) to compute its inverse Fourier transform:

$$\begin{aligned} h(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)e^{j\omega t}\,d\omega = \frac{1}{2\pi}\int_{-\omega_c}^{\omega_c} e^{j\omega t}\,d\omega \\ &= \frac{1}{2\pi jt}\,e^{j\omega t}\Big|_{\omega=-\omega_c}^{\omega_c} = \frac{e^{j\omega_c t} - e^{-j\omega_c t}}{2j\pi t} = \frac{\sin\omega_c t}{\pi t} \end{aligned} \tag{4.33}$$

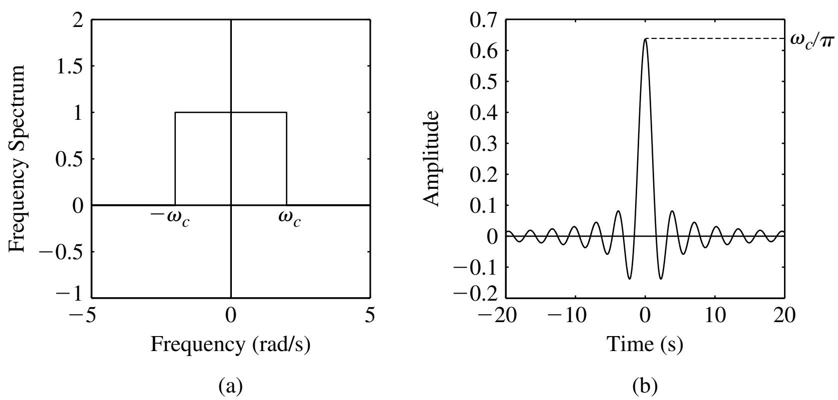


Figure 4.8 (**a**) Ideal lowpass filter with $\omega_c = 2$. (**b**) Its inverse DT Fourier transform.

It is the impulse response of the analog ideal lowpass filter and is plotted in Figure 4.8(b). The value of $h(t)$ at $t = 0$ can be computed, using l'Hôpital's rule, as ω_c/π. Because $h(t) \neq 0$, for $t < 0$, the ideal filter is not causal. We mention that $h(t)$ is not absolutely integrable, thus its spectrum is not continuous as shown in Figure 4.8(a).[3] We also mention that $h(t)$ can also be obtained from Example 4.3.3 by using the duality property of the CT Fourier transform and its inverse. See Problem 4.12.

EXERCISE 4.3.1

Find the frequency spectrum of $x(t) = \delta(t)$.

Answer

$\Delta(\omega) = 1$

Sinc Function To conclude this section, we introduce the function

$$\text{sinc}\,\theta := \frac{\sin\theta}{\theta} \tag{4.34}$$

It is called the *sinc function* and is plotted in Figure 4.9. It has one main lobe with base width 2π and peak magnitude 1, as well as infinitely many side lobes with base width π and decreasing peak magnitudes. Using this function, we can write (4.32) as $W_a(\omega) = 2a\,\text{sinc}\,a\omega$, and (4.33) as $h(t) = (\omega_c/\pi)\text{sinc}\,\omega_c t$. The sinc function appears in the impulse response of an ideal lowpass filter and, as we will show in the next chapter, in an interpolating formula. Thus it is also called the *filtering or interpolating* function.

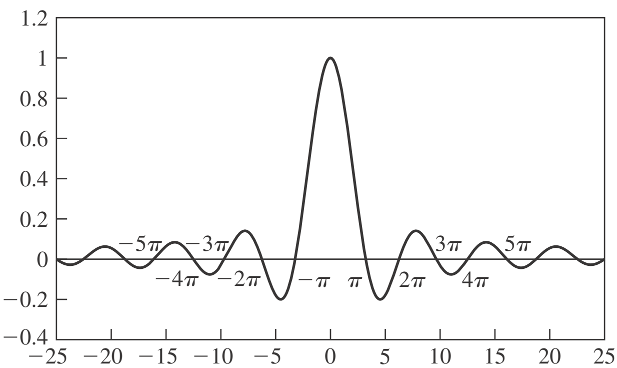


Figure 4.9 Sinc function.

[3]The function is integrable ($\int_{-\infty}^{\infty} h(t)\,dt = 1$) but not absolutely integrable ($\int_{-\infty}^{\infty} |h(t)|\,dt = \infty$). The function, however, is squared integrable ($\int_{-\infty}^{\infty} h^2(t)\,dt = \omega_c/\pi$), thus its Fourier transform is still defined. See Problem 4.24.

4.4 PROPERTIES OF FREQUENCY SPECTRA

In this section, we discuss some properties of frequency spectra. Before proceeding, we mention that the CT Fourier transform is a linear operator in the sense that

$$\mathcal{F}[a_1x_1(t) + a_2x_2(t)] = a_1\mathcal{F}[x_1(t)] + a_2\mathcal{F}[x_2(t)]$$

for any constants a_1 and a_2. This can be directly verified from the definition.

Time Shifting Consider a CT signal $x(t)$ with frequency spectrum $X(\omega)$. Let t_0 be a constant. Then the signal $x(t - t_0)$, as discussed in Section 1.4.1, is the shifting of $x(t)$ to t_0. Let $X_0(\omega)$ be the spectrum of $x(t - t_0)$. Then we have

$$X_0(\omega) := \mathcal{F}[x(t - t_0)] = \int_{-\infty}^{\infty} x(t - t_0)e^{-j\omega t}\,dt$$
$$= e^{-j\omega t_0}\int_{t=-\infty}^{\infty} x(t - t_0)e^{-j\omega(t-t_0)}\,dt$$

which becomes, after introducing the new variable $\tau := t - t_0$,

$$X_0(\omega) = \mathcal{F}[x(t - t_0)] = e^{-j\omega t_0}\int_{\tau=-\infty}^{\infty} x(\tau)e^{-j\omega\tau}\,d\tau = e^{-j\omega t_0}X(\omega) \tag{4.35}$$

Because $|e^{-j\omega t_0}| = 1$ and $\measuredangle e^{-j\omega t_0} = -\omega t_0$, we have

$$|X_0(\omega)| = |e^{-j\omega t_0}||X(\omega)| = |X(\omega)| \tag{4.36}$$

and

$$\measuredangle X_0(\omega) = \measuredangle e^{-j\omega t_0} + \measuredangle X(\omega) = \measuredangle X(\omega) - \omega t_0 \tag{4.37}$$

Thus time shifting will not affect the magnitude spectrum but will introduce a linear phase into the phase spectrum.

EXAMPLE 4.4.1

Consider

$$w_L(t) = \begin{cases} 1 & \text{for } 0 \le t \le L \\ 0 & \text{for } t < 0 \text{ and } t > L \end{cases} \tag{4.38}$$

It is the shifting to the right of the window in (4.31) by $a = L/2$ and is called a CT *shifted rectangular window* with length L. Its Fourier transform can be computed directly (Problem 4.13) or obtained from (4.32) as, using (4.35),

$$W_L(\omega) = e^{-ja\omega}\cdot\frac{2\sin a\omega}{\omega} \tag{4.39}$$

where $a = L/2$. It is complex-valued. Its magnitude and phase spectra are plotted in Figures 4.7(c) and 4.7(d). Its magnitude spectrum equals the one in Figure 4.7(b) if its negative part is flipped to positive. The phase in Figure 4.7(b) is either 0 or π depending on whether $W_a(\omega)$ is positive or negative. The phase of (4.39) differs from the one of (4.32) by $-a\omega$, a linear function of ω with slope $-a$. Because phases are plotted within $\pm 180°$ or $\pm\pi$, the phase spectrum in Figure 4.7(d) shows sections of straight lines bounded by $\pm\pi$, instead of a single straight line.

Frequency Shifting Let $x(t)$ be a CT signal with spectrum $X(\omega)$. Then we have

$$\begin{aligned}\mathcal{F}[e^{j\omega_0 t}x(t)] &= \int_{-\infty}^{\infty} e^{j\omega_0 t}x(t)e^{-j\omega t}\,dt = \int_{-\infty}^{\infty} x(t)e^{-j(\omega-\omega_0)t}\,dt \\ &= X(\omega-\omega_0) \end{aligned} \tag{4.40}$$

This is called frequency shifting because the frequency spectrum of $x(t)$ is shifted to ω_0 if $x(t)$ is multiplied by $e^{j\omega_0 t}$. Using (4.40) and the linearity of the Fourier transform, we have

$$\begin{aligned}\mathcal{F}[x(t)\cos\omega_0 t] &= \mathcal{F}\left[x(t)\frac{e^{j\omega_0 t}+e^{-j\omega_0 t}}{2}\right] \\ &= 0.5[X(\omega-\omega_0)+X(\omega+\omega_0)] \end{aligned} \tag{4.41}$$

If $x(t)$ has a spectrum as shown in Figure 4.10(a), then the spectrum of $x(t)\cos\omega_0 t$ is as shown in Figure 4.10(b). It consists of the spectra of $x(t)$ shifted to $\pm\omega_0$, and the magnitude is cut in half. The multiplication of $x(t)$ by $\cos\omega_0 t$ is called *modulation* and will be discussed further in a later chapter.

Time Compression and Expansion Consider the signal $x(t)$ shown in Figure 4.11(a). We plot in Figures 4.11(b) and 4.11(c) $x(2t)$ and $x(0.5t)$. We see that $x(2t)$ lasts only half of the duration of $x(t)$, whereas $x(0.5t)$ lasts twice the time duration of $x(t)$. In general, we have

- $x(at)$ with $a > 1$: time compression or speed up
- $x(at)$ with $0 < a < 1$: time expansion or slow down

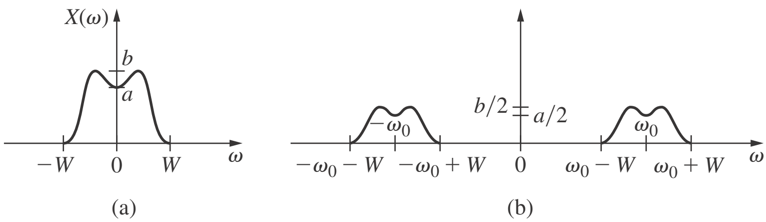


Figure 4.10 (**a**) Spectrum of $x(t)$. (**b**) Spectrum of $x(t)\cos\omega_0 t$.

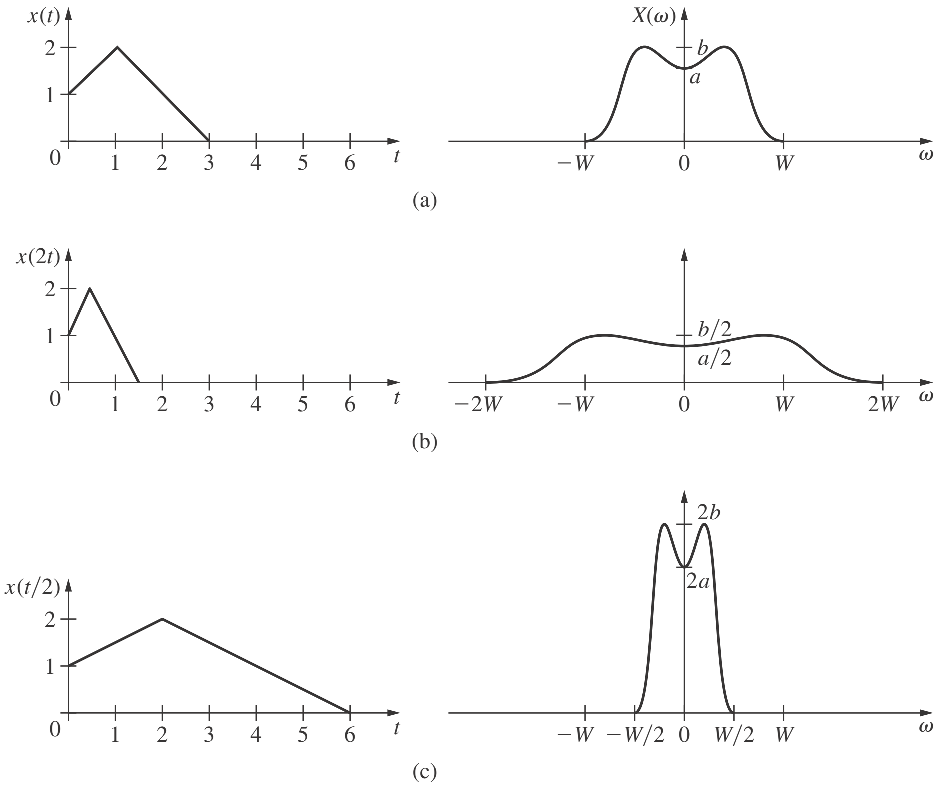


Figure 4.11 (**a**) Time signal and its frequency spectrum. (**b**) Time compression and frequency expansion. (**c**) Time expansion and frequency compression.

Now we study their effects on frequency spectra. By definition, we have

$$\mathcal{F}[x(at)] = \int_{-\infty}^{\infty} x(at)e^{j\omega t}\, dt = \frac{1}{a}\int_{-\infty}^{\infty} x(at)e^{j(\omega/a)at}\, d(at)$$

$$= \frac{1}{a}\int_{-\infty}^{\infty} x(\tau)e^{j(\omega/a)\tau}\, d\tau$$

where we have used $\tau = at$. Thus we conclude

$$\mathcal{F}[x(at)] = \frac{1}{a}X\left(\frac{\omega}{a}\right)$$

Note that this holds only for a positive. If the spectrum of $x(t)$ is as shown on the right-hand side of Figure 4.11(a), then the spectra of $x(2t)$ and $x(0.5t)$ are as shown in Figures 4.11(b) and 4.11(c). Thus we have

- Time compression ⇔ frequency expansion
- Time expansion ⇔ frequency compression

This is intuitively apparent. In time compression, the rate of change of the time signal increases, thus the spectrum contains higher-frequency components. In time expansion, the rate of change of the time signal decreases, thus the spectrum contains lower-frequency components. This fact can be used to explain the change of pitch of an audio tape when its speed is changed. When we increase the speed of the tape (time compression), its frequency spectrum is expanded and, consequently, contains higher-frequency components. Thus the pitch is higher. When we decrease the speed (time expansion), its spectrum is compressed and contains lower-frequency components. Thus its pitch is lower.

Time Duration and Frequency Bandwidth Consider a time signal $x(t)$. If $x(t)$ starts to appear at t_1 and ends at t_2 or it lasts from t_1 to t_2 and is identically zero before t_1 and after t_2, then we may define the *time duration* of $x(t)$ as $L := t_2 - t_1$. If we use this definition, then a step function has infinite time duration, and so have the time signals $e^{-0.1t}$ and e^{-2t}, for $t \geq 0$, shown in Figure 1.9. Clearly, this is not a good definition. Instead we may define the time duration to be the width of the time interval in which

$$|x(t)| > ax_{max}$$

where x_{max} is the peak magnitude of $x(t)$, and a is some small constant. For example, if $a = 0.01$, then e^{-2t} and $e^{-0.1t}$ have time durations 2.5 and 50, respectively, as discussed in Section 1.3. If we select a different a, then we will obtain different time durations. Another way is to define the time duration as the width of the time interval that contains, say, 90% of the total energy of the signal. Thus there are many ways of defining the time duration of a signal.

Likewise, there are many ways of defining the bandwidth of the spectrum of a signal. We may define the frequency bandwidth of a signal to be the width of frequency interval in which

$$|X(\omega)| > bX_{max}$$

where X_{max} is the peak magnitude of $X(\omega)$ and b is a constant such as 0.01. We can also define the bandwidth as the width of frequency interval which contains 90% of the total energy. In any case, no matter how the time duration and frequency bandwidth are defined, generally, we have

$$\text{Time duration} \sim \frac{1}{\text{Frequency bandwidth}} \tag{4.42}$$

It means that the larger the frequency bandwidth, the smaller the time duration and vice versa. For example, the time function $\delta(t)$ has zero time duration and its frequency spectrum $\Delta(\omega) = \mathcal{F}[\delta(t) = 1$, for all ω, has infinite frequency bandwidth. The time function $x(t) = 1$, for all t, has infinite time duration; its spectrum (as we will show later) is $\mathcal{F}[1] = 2\pi\delta(\omega)$, which has zero frequency bandwidth. It is also consistent with our earlier discussion of time expansion and frequency compression.

A mathematical proof of (4.42) is difficult if we use any of the aforementioned definitions. However, if we define the time duration as

$$L = \frac{\left(\int_{-\infty}^{\infty} |x(t)|\, dt\right)^2}{\int_{-\infty}^{\infty} |x(t)|^2\, dt} \tag{4.43}$$

and frequency bandwidth as

$$B = \frac{\int_{-\infty}^{\infty} |X(\omega)|^2 \, d\omega}{2|X(0)|^2} \tag{4.44}$$

then we can show $BL \geq \pi$. See Problem 4.16. However, the physical meanings of (4.43) and (4.44) are not transparent. In any case, the inverse relationship in (4.42) is widely accepted in engineering.

Convolution in Time and Multiplication in Frequency[4] Consider the convolution in (3.33) or

$$y(t) = \int_{\tau=-\infty}^{\infty} h(t-\tau)u(\tau)\, d\tau \tag{4.45}$$

It relates the input $u(t)$ and output $y(t)$ of a system with impulse response $h(t)$. Applying the Fourier transform to $y(t)$ yields

$$\begin{aligned}
Y(\omega) &= \int_{t=-\infty}^{\infty} y(t)e^{-j\omega t}\, dt \\
&= \int_{t=-\infty}^{\infty} \left(\int_{\tau=-\infty}^{\infty} h(t-\tau)u(\tau)\, d\tau \right) e^{-j\omega(t-\tau)}e^{-j\omega\tau}\, dt \\
&= \int_{\tau=-\infty}^{\infty} \left(\int_{t=\infty}^{\infty} h(t-\tau)e^{-j\omega(t-\tau)}\, dt \right) u(\tau)e^{-j\omega\tau}\, d\tau \\
&= \left(\int_{\bar{t}=-\infty}^{\infty} h(\bar{t})e^{-j\omega\bar{t}}\, d\bar{t} \right) \left(\int_{\tau=-\infty}^{\infty} u(\tau)e^{-j\omega\tau}\, d\tau \right)
\end{aligned}$$

where we have changed the order of integrations and introduced the new variable $\bar{t} = t - \tau$. Let $H(\omega)$ and $U(\omega)$ be the Fourier transforms of $h(t)$ and $u(t)$, respectively. Then we have

$$Y(\omega) = H(\omega)U(\omega) \tag{4.46}$$

We see that the Fourier transform transforms the convolution in (4.45) into the algebraic equation in (4.46). Thus we can use (4.46) to carry out system analysis. However, its use is not convenient, as we will discuss in Section 6.9.1.

Multiplication in Time and Convolution in Frequency[5] Consider two time functions $x_1(t)$ and $x_2(t)$ with their Fourier transforms $X_1(\omega)$ and $X_2(\omega)$. Let us compute the Fourier transform of $x_1(t)x_2(t)$:

$$\begin{aligned}
\mathcal{F}[x_1(t)x_2(t)] &= \int_{t=-\infty}^{\infty} x_1(t)x_2(t)e^{-j\omega t}\, dt \\
&= \int_{t=-\infty}^{\infty} x_1(t)\left(\frac{1}{2\pi} \int_{\bar{\omega}=-\infty}^{\infty} X_2(\bar{\omega})e^{j\bar{\omega}t}\, d\bar{\omega} \right) e^{-j\omega t}\, dt
\end{aligned}$$

[4] This property may be skipped without loss of continuity.

[5] This property may be skipped without loss of continuity.

where we have used $\bar{\omega}$ in the inverse Fourier transform of $X_2(\bar{\omega})$ to differentiate it from the ω in the Fourier transform of $x_1(t)x_2(t)$. Interchanging the order of integrations yields

$$\mathcal{F}[x_1(t)x_2(t)] = \frac{1}{2\pi}\int_{\bar{\omega}=-\infty}^{\infty}\left(\int_{t=-\infty}^{\infty} x_1(t)e^{-j(\omega-\bar{\omega})t}\,dt\right)X_2(\bar{\omega})\,d\bar{\omega}$$

which becomes, because the term inside the parentheses is $X_1(\omega-\bar{\omega})$,

$$\begin{aligned}\mathcal{F}[x_1(t)x_2(t)] &= \frac{1}{2\pi}\int_{\bar{\omega}=-\infty}^{\infty} X_1(\omega-\bar{\omega})X_2(\bar{\omega})\,d\bar{\omega}\\ &= \frac{1}{2\pi}\int_{\bar{\omega}=-\infty}^{\infty} X_1(\bar{\omega})X_2(\omega-\bar{\omega})\,d\bar{\omega}\end{aligned}\tag{4.47}$$

The last equality can be obtained from the first by changing variables. The convolution is called *complex convolution*. It has the commutative property as shown.

We discuss two special cases of (4.47). Let $x_2(t) = x_1^*(t)$. Then we have

$$X_2(\omega) = \mathcal{F}[x_1^*(t)] = \int_{-\infty}^{\infty} x_1^*(t)e^{-j\omega t}\,dt = \left(\int_{-\infty}^{\infty} x_1(t)e^{j\omega t}\,dt\right)^* = X_1^*(-\omega)$$

and $X_2(-\omega) = X_1^*(\omega)$. Thus (4.47) becomes, after setting $\omega = 0$,

$$\int_{-\infty}^{\infty}|x_1(t)|^2\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} X_1(\bar{\omega})X_2(-\bar{\omega})\,d\bar{\omega} = \frac{1}{2\pi}\int_{-\infty}^{\infty} X_1(\bar{\omega})X_1^*(\bar{\omega})\,d\bar{\omega}$$

This is called a *Parseval's formula.* We will derive it again in the next subsection without using (4.47).

The modulation formula in (4.41) can also be obtained from (4.47). If $x_1(t) = x(t)$ and $x_2(t) = \cos\omega_o t$, then we have $X_1(\omega) = X(\omega)$ and, as we will derive later,

$$X_2(\omega) = \pi\delta(\omega-\omega_o) + \pi\delta(\omega+\omega_o)$$

Substituting these into (4.47) and using the sifting property of impulses, we obtain

$$\begin{aligned}\mathcal{F}[x(t)\cos\omega_c t] &= \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega-\bar{\omega})[\pi\delta(\bar{\omega}-\omega_o)+\pi\delta(\bar{\omega}+\omega_o)]\,d\bar{\omega}\\ &= 0.5[X_1(\omega-\bar{\omega})|_{\bar{\omega}=\omega_o} + X_1(\omega-\bar{\omega})|_{\bar{\omega}=-\omega_o}]\\ &= 0.5[X(\omega-\omega_o)+X(\omega+\omega_o)]\end{aligned}$$

This is the equation in (4.41).

4.4.1 Distribution of Energy in Frequencies

We now discuss the physical meaning of frequency spectra. The total energy of a signal is, as discussed in (4.21),

$$E := \int_{t=-\infty}^{\infty} x(t)x^*(t)\,dt = \int_{-\infty}^{\infty}|x(t)|^2\,dt \tag{4.48}$$

where $x^*(t)$ is the complex conjugate of $x(t)$. An absolutely integrable CT signal may not have a finite total energy. See Problem 4.4. However, if it is also bounded ($|x(t)| < q$), then we have

$$E = \int_{-\infty}^{\infty} |x(t)||x(t)|\,dt < \int_{-\infty}^{\infty} q|x(t)|\,dt = q\int_{-\infty}^{\infty} |x(t)|\,dt < \infty$$

and the signal has a finite total energy.

The total energy of a CT signal can be computed directly in the time domain or indirectly from its frequency spectrum. Let $X(\omega)$ be the CT Fourier transform of $x(t)$. Then we have

$$x(t) = \frac{1}{2\pi}\int_{\omega=-\infty}^{\infty} X(\omega)e^{j\omega t}\,d\omega$$

Substituting this into (4.48) yields

$$\begin{aligned} E &= \int_{t=-\infty}^{\infty} x(t)\left[\frac{1}{2\pi}\int_{\omega=-\infty}^{\infty} X(\omega)e^{j\omega t}\,d\omega\right]^* dt \\ &= \frac{1}{2\pi}\int_{\omega=-\infty}^{\infty} X^*(\omega)\left[\int_{t=-\infty}^{\infty} x(t)e^{-j\omega t}\,dt\right] d\omega \end{aligned}$$

where we have changed the order of integrations. The term inside the brackets equals $X(\omega)$, thus we have

$$\int_{-\infty}^{\infty} |x(t)|^2\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} X^*(\omega)X(\omega)\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} |X(\omega)|^2\,d\omega \tag{4.49}$$

This is called a *Parseval's formula*.

The total energy of $x(t)$ is independent of its phase spectrum. It can be computed, using (4.49), from its magnitude spectrum. More importantly, the magnitude spectrum reveals the distribution of energy in frequencies. For example, the energy contained in the frequency range $[\omega_1, \omega_2]$, with $\omega_1 < \omega_2$, is given by

$$\frac{1}{2\pi}\int_{\omega_1}^{\omega_2} |X(\omega)|^2\,d\omega$$

We mention that if $x(t)$ is absolutely integrable, then its spectrum contains no impulses and

$$\int_{\omega_{0-}}^{\omega_{0+}} |X(\omega)|^2\,d\omega = 0$$

Thus it is meaningless to talk about energy at a discrete or isolated frequency. For this reason, $X(\omega)$ is also called the *spectral density* of $x(t)$.

The integration in (4.49) is carried out over positive and negative frequencies. If $x(t)$ is real, then $|X(-\omega)| = |X(\omega)|$ and (4.49) can also be written as

$$E = \frac{1}{2\pi}\int_{-\infty}^{\infty} |X(\omega)|^2\,d\omega = \frac{1}{\pi}\int_{0}^{\infty} |X(\omega)|^2\,d\omega \tag{4.50}$$

Thus the total energy of a real-valued signal can be computed over positive frequencies. The quantity $|X(\omega)|^2/\pi$ may be called the *energy spectral density*.

EXAMPLE 4.4.2

Find the frequency range $[-\omega_0, \omega_0]$ in which $x(t) = e^{-t}$, for $t \geq 0$, contains half of its total energy.

The total energy of $x(t)$ is

$$E = \int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{0}^{\infty} e^{-2t} \, dt = \frac{1}{-2} e^{-2t} \Big|_{t=0}^{\infty}$$
$$= -0.5[0 - 1] = 0.5$$

The spectrum of $x(t)$ was computed in (4.30) with $a = 1$. Thus we have

$$X(\omega) = \frac{1}{j\omega + 1} \quad \text{and} \quad |X(\omega)|^2 = \frac{1}{\omega^2 + 1}$$

We compute, using an integration table,

$$\frac{1}{2\pi} \int_{-\omega_0}^{\omega_0} \frac{1}{\omega^2 + 1} \, d\omega = \frac{1}{\pi} \int_{0}^{\omega_0} \frac{1}{\omega^2 + 1} \, d\omega = \frac{1}{\pi} \tan^{-1} \omega_0$$

Equating

$$\frac{1}{\pi} \tan^{-1} \omega_0 = \frac{0.5}{2}$$

yields $\omega_0 = 1$. In other words, although the energy of e^{-t} is spread over the frequency range $(-\infty, \infty)$, half of its total energy lies inside the frequency range $[-1, 1]$ in rad/s. Thus the signal is a low-frequency signal.

EXERCISE 4.4.1

Compute the total energy of $x(t) = \delta(t)$. Compute it directly in the time domain and then verify it in the frequency domain.

Answer

∞

4.5 FREQUENCY SPECTRA OF CT PERIODIC SIGNALS

This section develops the frequency spectrum of periodic signals. Instead of computing directly the frequency spectrum of periodic signals, we compute first the inverse Fourier transform of a spectrum that consists of an impulse at ω_0 or $\delta(\omega - \omega_0)$. Using the sifting property in (1.12), we compute

$$\mathcal{F}^{-1}[\delta(\omega - \omega_0)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \delta(\omega - \omega_0) e^{j\omega t} \, d\omega = \frac{1}{2\pi} e^{j\omega t} \Big|_{\omega=\omega_0} = \frac{1}{2\pi} e^{j\omega_0 t}$$

which implies

$$\mathcal{F}[e^{j\omega_0 t}] = 2\pi \delta(\omega - \omega_0) \tag{4.51}$$

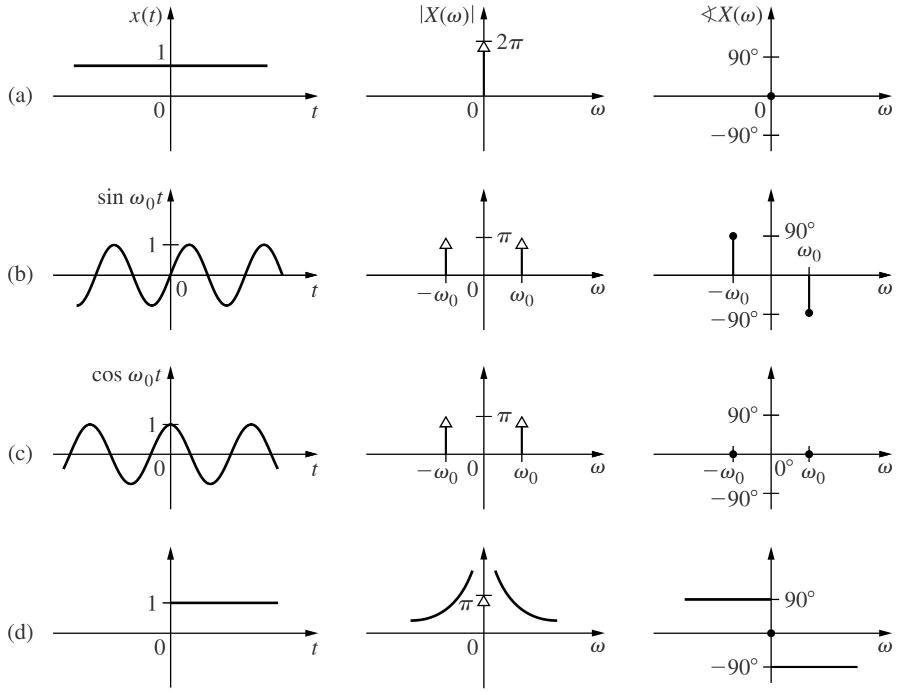


Figure 4.12 Frequency spectra of (**a**) $x(t) = 1$, for all t, (**b**) $\sin \omega_0 t$, (**c**) $\cos \omega_0 t$, and (**d**) $q(t)$.

Thus the frequency spectrum of the complex exponential $e^{j\omega_0 t}$ is an impulse at ω_0 with weight 2π. If $\omega_0 = 0$, (4.51) becomes

$$\mathcal{F}[1] = 2\pi\delta(\omega) \tag{4.52}$$

where 1 is the time function $x(t) = 1$ for all t in $(-\infty, \infty)$. Its frequency spectrum is zero everywhere except at $\omega = 0$ as shown in Figure 4.12(a). Because $x(t) = 1$ contains no nonzero frequency component, it is called a dc signal.

Using the linearity property of the Fourier transform and (4.51), we have

$$\begin{aligned}\mathcal{F}[\sin \omega_0 t] &= \mathcal{F}\left[\frac{e^{j\omega_0 t} - e^{-j\omega_0 t}}{2j}\right] = \frac{\pi}{j}[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)] \\ &= -j\pi\delta(\omega - \omega_0) + j\pi\delta(\omega + \omega_0)\end{aligned} \tag{4.53}$$

and

$$\mathcal{F}[\cos \omega_0 t] = \mathcal{F}\left[\frac{e^{j\omega_0 t} + e^{-j\omega_0 t}}{2}\right] = \pi\delta(\omega - \omega_0) + \pi\delta(\omega + \omega_0) \tag{4.54}$$

The magnitude and phase spectra of (4.53) are shown in Figure 4.12(b). It consists of two impules with weight π at $\omega = \pm\omega_0$ and phase -90° at $\omega = \omega_0$ and 90° at $\omega = -\omega_0$ as shown. The magnitude and phase spectra of (4.54) are plotted in Figure 4.12(c). The frequency spectra of $\sin \omega_0 t$ and $\cos \omega_0 t$ are zero everywhere except at $\pm\omega_0$. Thus their frequency spectra as defined are consistent with our perception of their frequency.

EXERCISE 4.5.1

Find the spectrum of $x(t) = \cos 50t + 2\cos 70t$.

Answer

$X(\omega) = \pi\delta(\omega - 50) + \pi\delta(\omega + 50) + 2\pi\delta(\omega - 70) + 2\pi\delta(\omega + 70)$

EXERCISE 4.5.2

Find the spectrum of $x(t) = \sin 50t + 2\sin 70t$.

Answer

$X(\omega) = -j\pi\delta(\omega - 50) + j\pi\delta(\omega + 50) - 2\pi j\delta(\omega - 70) + 2\pi j\delta(\omega + 70)$

The sine and cosine functions are not absolutely integrable in $(-\infty, \infty)$ and their frequency spectra are still defined. Their spectra consist of impulses that are neither bounded nor continuous. In fact, this is the case for every periodic signal with a well-defined Fourier series. Consider a periodic signal $x(t)$ with Fourier series

$$x(t) = \sum_{-\infty}^{\infty} c_m e^{jm\omega_0 t}$$

where c_m are its CTFS coefficients. The equation contains both t and ω and is not exactly a frequency-domain description. Applying the Fourier transform and using (4.51), we obtain

$$X(\omega) = \mathcal{F}[x(t)] = \sum_{-\infty}^{\infty} c_m \mathcal{F}[e^{jm\omega_0 t}] = \sum_{-\infty}^{\infty} 2\pi c_m \delta(\omega - m\omega_0) \tag{4.55}$$

It is independent of t and is the frequency-domain description of the periodic signal. It consists of a sequence of impulses with weight $2\pi c_m$. Clearly, it is neither bounded nor continuous. For example, the periodic signal in (1.24) can be expressed in Fourier series as in (1.31). Thus its frequency spectrum is as shown in (1.32). Because the spectrum consists of nonzero values only at discrete frequencies, it may be called *discrete frequency spectrum*. We mentioned earlier that we do not call the set c_m discrete frequency spectrum. If we do so, then it will be confused with the discrete frequency spectrum defined in (4.55). Thus we call c_m frequency components or CTFS coefficients.

We encounter in practice only positive-time signals. However, in discussing spectra of periodic signals, it is simpler to consider them to be defined for all t. For example, the frequency spectrum of the step function $q(t)$ can be computed as

$$Q(\omega) = \mathcal{F}[q(t)] = \pi\delta(\omega) + \frac{1}{j\omega} \tag{4.56}$$

See Reference 4. It consists of an impulse at $\omega = 0$ with weight π and $1/j\omega$. It has nonzero magnitude spectrum for all ω as shown in Figure 4.12(d). Even so, we often call a step function a dc signal. Strictly speaking, a dc signal must be a constant defined for all t in $(-\infty, \infty)$. Likewise, the spectrum of $\cos \omega_0 t$, for t in $(-\infty, \infty)$, consists of only two impulses at $\pm\omega_0$, and the spectrum of $\cos \omega_0 t$ for t in $[0, \infty)$ or $[\cos \omega_0 t]q(t)$ is

$$\mathcal{F}[\cos \omega_0 t \cdot q(t)] = 0.5\pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] + \frac{j\omega}{\omega_0^2 - \omega^2} \tag{4.57}$$

It is nonzero for all ω. Thus it is much simpler to use the former to introduce the concept of spectra for sinusoids.

4.6 EFFECTS OF TRUNCATION

If we use a computer to compute the spectrum of an infinitely long signal, the signal must be truncated to have a finite length. This section studies its effects on the spectrum.[6]

Consider a CT signal $x(t)$ with spectrum $X(\omega)$. Let us truncate the signal before $t = -a$ and after a. This truncation is the same as multiplying $x(t)$ by the CT rectangular window $w_a(t)$ defined in (4.31) and plotted in Figure 4.7(a). The spectrum of $w_a(t)$ is computed in (4.32) as

$$W_a(\omega) = \frac{2 \sin a\omega}{\omega} \tag{4.58}$$

Thus the spectrum of the truncated $x(t)$ or the signal $x(t)w_a(t)$ is, using (4.47),

$$\mathcal{F}[x(t)w_a(t)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\bar{\omega}) W_a(\omega - \bar{\omega})\, d\bar{\omega} \tag{4.59}$$

which implies that the spectrum of a truncated signal equals the convolution of $W_a(\omega)$ and the spectrum of the untruncated signal. Thus the waveform of $W_a(\omega)$ plays a crucial role in determining the effects of truncation. As shown in Figure 4.7(b), the spectrum $W_a(\omega)$ has one main lobe with base width $2\pi/a$ and height $2a$. All side lobes have base width π/a, half of the base width of the main lobe. The largest magnitude of the side lobe on either side is roughly 20% of the height of the main lobe.

We now use examples to study the effects of truncation. Consider $\cos \omega_0 t$, for all t. Its spectrum is

$$X(\omega) = \pi\delta(\omega - \omega_0) + \pi\delta(\omega + \omega_0)$$

and is plotted in Figure 4.13(a) for $\omega_0 = 0.5$ and in Figure 4.13(b) for $\omega_0 = 3$. If we truncate $\cos \omega_0 t$ before $t = -a$ and after a, then its spectrum can be computed as, using the sifting property of impulses,

$$\begin{aligned}\mathcal{F}[\cos \omega_0 t \cdot w_a(t)] &= \frac{1}{2\pi} \int_{-\infty}^{\infty} [\pi\delta(\bar{\omega} - \omega_0) + \pi\delta(\bar{\omega} + \omega_0)] W_a(\omega - \bar{\omega})\, d\bar{\omega} \\ &= 0.5[W_a(\omega - \omega_0) + W_a(\omega + \omega_0)]\end{aligned}$$

where $W_a(\omega)$ is given in (4.58). Thus the spectrum of the truncated $\cos \omega_0 t$ is the sum of $0.5W_a(\omega)$ shifted to $\pm\omega_0$. Note that the height of $0.5W_a(\omega)$ is a at $\omega = 0$. Figures 4.13(aa)

[6] This section and the following section may be skipped without loss of continuity. However, the reader is recommended to read through them.

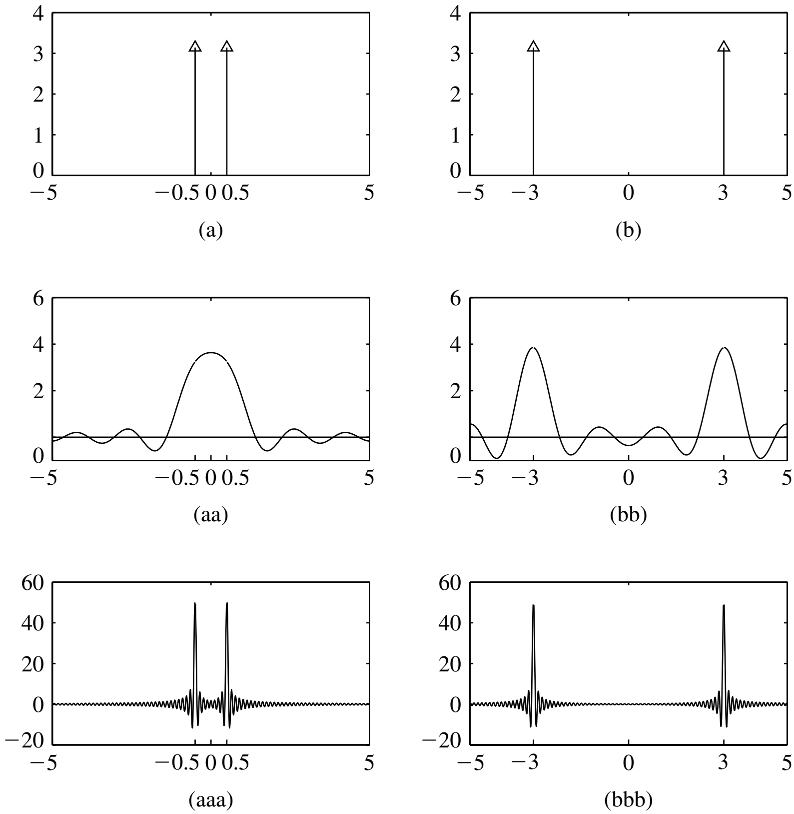


Figure 4.13 Effects of truncation: (**a**) and (**b**) Spectra of $\cos 0.5t$ and $\cos 3t$ for all t. (**aa**) and (**bb**) Spectra of truncated $\cos \omega_0 t$ with $a = 4$. (**aaa**) and (**bbb**) Spectra of truncated $\cos \omega_0 t$ with $a = 51.2$. All horizontal coordinates are frequency in rad/s.

and 4.13(bb) show the frequency spectra of truncated $\cos 0.5t$ and $\cos 3t$ with $a = 4$. For $a = 4$, the width of the main lobe is $2\pi/a = 1.57$ and the two main lobes shifted to ± 0.5 overlap with each other, and the resulting plot shows only one lobe in Figure 4.13(aa). The two main lobes shifted to ± 3 do not overlap with each other as shown in Figure 4.13(bb). In conclusion, because of the main lobe, truncation will introduce *leakage* or *smearing* into a spectrum. The wider the main lobe, the larger the leakage.

The spectrum of the window consists of infinitely many side lobes, all with the same base width but decreasing magnitudes. When these side lobes convolve with a spectrum, they will introduce ripples into the spectrum as shown in Figures 4.13(aa) and 4.13(bb). Thus truncation will also introduce *ripples* into a spectrum.

Figures 4.13(aaa) and 4.13(bbb) show the spectra of truncated $\cos 0.5t$ and $\cos 3t$ for $a = 51.2$. The effects of truncation (leakage and ripples) are clearly visible but not as large as the ones for $a = 4$. As a increases, the two spikes become higher and narrower. Eventually, the two spikes become the two impulses shown in Figures 4.13(a) and 4.13(b). In other words, leakage will disappear.

Will ripples disappear as a increases? As shown in Figure 4.13, as a increases, the ripples become narrow and move closer to $\pm\omega_0$. However, there is an unusual phenomenon: The largest magnitude of the ripples, relative to the peak magnitude of the spectrum, will not decrease as a increases. It remains roughly constant as shown. This phenomenon is called the *Gibbs phenomenon*.

We demonstrate once again the Gibbs phenomenon for the analog ideal lowpass filter studied in Example 4.3.4 or

$$h(t) = \frac{\sin \omega_c t}{\pi t}$$

Its spectrum is

$$H(\omega) = \begin{cases} 1 & \text{for } |\omega| \le \omega_c \\ 0 & \text{for } |\omega| > \omega_c \end{cases}$$

If we truncate $h(t)$ using the rectangular window $w_a(t)$, then the spectrum of the truncated $h(t)$ is

$$\mathcal{F}[h(t)w_a(t)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\bar{\omega})W_a(\omega - \bar{\omega})\,d\bar{\omega} = \int_{-\omega_c}^{\omega_c} \frac{\sin a(\omega - \bar{\omega})}{\pi(\omega - \bar{\omega})}\,d\bar{\omega}$$

If $\omega_c = 2$, then the spectra of the truncated CT ideal lowpass filter for $a = 5$ and $a = 50$ are as shown in Figures 4.14(a) and 4.14(b). See Reference 2 for its plotting. We see that the sharp edges at $\omega_c = \pm 2$ are smeared, and ripples are introduced into the spectra. However, the magnitude of the largest ripple remains roughly the same, about 9% of the amount of the discontinuity. This is the Gibbs phenomenon.

To conclude this section, we mention that the Gibbs phenomenon occurs only if the original spectrum has discontinuities. If the original spectrum is continuous, then as a increases, the spectrum of the truncated signal will approach to the original spectrum without the Gibbs phenomenon. We also mention that the preceding discussion still holds if the rectangular window $w_a(t)$ is replaced by the shifted rectangular window $w_L(t)$ defined in (4.38).

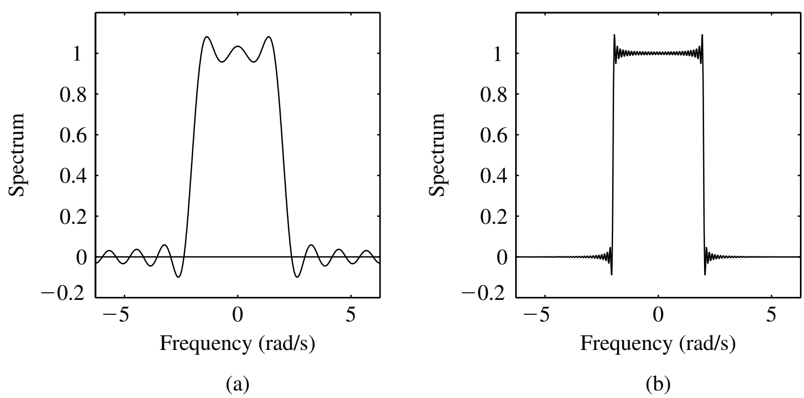


Figure 4.14 (**a**) Spectrum of $\sin 2t/\pi t$ with $|t| \le 5$. (**b**) Spectrum of $\sin 2t/\pi t$ with $|t| \le 50$.

4.7 TIME-LIMITED BAND-LIMITED THEOREM

To conclude this chapter, we discuss a fundamental relationship between a time signal and its frequency spectrum. A CT signal $x(t)$ is band-limited to W if its frequency spectrum $X(\omega)$ is 0 for $|\omega| > W$. It is time-limited to b if

$$x(t) = 0 \qquad \text{for } |t| > b$$

It turns out that if a CT signal is time-limited, then it cannot be band-limited and vice versa. The only exception is the trivial case $x(t) = 0$ for all t. To establish this assertion, we show that if $x(t)$ is band-limited to W and time-limited to b, then it must be identically zero. Indeed, if $x(t)$ is band-limited to W, then (4.29) implies

$$x(t) = \frac{1}{2\pi}\int_{-W}^{W} X(\omega)e^{j\omega t}\,d\omega \tag{4.60}$$

Its differentiation repeatedly with respect to t yields

$$x^{(k)}(t) = \frac{1}{2\pi}\int_{-W}^{W} X(\omega)(j\omega)^k e^{j\omega t}\,d\omega \tag{4.61}$$

for $k = 0, 1, 2, \ldots,$ where $x^{(k)}(t) := d^k x(t)/dt^k$. Because $x(t)$ is time-limited to b, its derivatives are identically zero for all $|t| > b$. Thus (4.61) implies

$$\int_{-W}^{W} X(\omega)(\omega)^k e^{j\omega a}\,d\omega = 0 \tag{4.62}$$

for any a with $a > b$. Next we use

$$e^c = 1 + \frac{c}{1!} + \frac{c^2}{2!} + \cdots = \sum_{k=0}^{\infty}\frac{c^k}{k!}$$

to rewrite (4.60) as

$$\begin{aligned}
x(t) &= \frac{1}{2\pi}\int_{-W}^{W} X(\omega)e^{j\omega(t-a)}e^{j\omega a}\,d\omega \\
&= \frac{1}{2\pi}\int_{-W}^{W} X(\omega)\left[\sum_{k=0}^{\infty}\frac{(j\omega(t-a))^k}{k!}\right]e^{j\omega a}\,d\omega \\
&= \sum_{k=0}^{\infty}\frac{(j(t-a))^k}{2\pi k!}\int_{-W}^{W} X(\omega)(\omega)^k e^{j\omega a}\,d\omega
\end{aligned}$$

which, following (4.62), is zero for all t. Thus a band-limited and time-limited CT signal must be identically zero. In conclusion, no nontrivial CT signal can be both time-limited and band-limited.

This fact imposes a mathematical limitation in computer computation of the spectrum of a CT signal $x(t)$ from its sampled sequence $x(nT)$. If $x(t)$ is frequency band-limited, then its time duration is infinite. Thus $x(t)$ or $x(nT)$ must be truncated in its computer computation, and the error due to truncation will occur. If $x(t)$ is time-limited, then its spectrum is not band-limited. In this case, frequency aliasing, as we will discuss in the next chapter, will occur in using $x(nT)$ to compute the spectrum of $x(t)$. Thus computing the spectrum of $x(t)$ from $x(nT)$ will introduce errors either from truncation or from frequency aliasing. In conclusion, it is mathematically impossible to compute the *exact* spectrum of a CT signal from its sampled sequence.

Fortunately, most engineering applications and designs involve some approximations. Although no CT signal can be both time-limited and band-limited in theory, most CT signals can be so considered in practice. For example, consider the exponential signal in Figure 4.6(a). It is not time-limited. However, the signal decays to zero rapidly; its magnitude is less than 1% of its peak magnitude for $t \geq 50$ and less than 0.01% for $t \geq 100$. Therefore, the exponential signal can be considered to be time-limited in practice. Likewise, its frequency spectrum shown in Figure 4.6(b) approaches 0 as $|\omega|$ approaches infinity. Thus its frequency spectrum can also be considered to be band-limited. In conclusion, most practical signals can be considered to be both time-limited and band-limited, and their frequency spectra are widely computed from their time samples. This is the topic of the next chapter.

PROBLEMS

4.1 What is the fundamental frequency and fundamental period of the signal

$$x(t) = 3 + \sin 6t - 2\cos 6t + \pi \sin 9t - \cos 12t$$

Express it in complex Fourier series and plot the magnitudes and phases of its frequency components.

4.2 Consider the full-wave rectifier in Figure 2.21(a). What is its output $y(t)$ if the input is $u(t) = \sin 2t$? What are the fundamental period and fundamental frequency of $y(t)$? Express the output in Fourier series. Does the output contains frequency other than 2 rad/s?

4.3 Use (1.7) and (1.8) to show that the periodic signal in Figure 4.2 has infinite average power.

4.4 Consider a periodic signal with period 1 and $x(t) = 1/\sqrt{t}$ for $0 \leq t < 1$. Show that the signal is absolutely integrable in one period but has infinite average power.

4.5 What is the total energy of the signal in Problem 4.1? What is its average powers in one period? How many percentage of the average power lies inside the frequency range $[-7, 7]$?

4.6 What are the average powers of the input and output in Example 4.2.5? How many percentage of the input power is transmitted to the output?

4.7 What are the average powers of the input and output in Problem 4.2? What percentage of the input power is transmitted to the output?

4.8 The signal-to-noise ratio (SNR) is defined in Example 3.8.2 as the ratio of their peak magnitudes. Now if we define it as the ratio of their average power, what are the SNRs at the input and output of Example 3.8.2?

4.9 A periodic signal $x(t)$ with fundamental period P is also periodic with period $2P$. What will happen in computing its Fourier series if we use $\omega_1 = 2\pi/2P$ instead of the fundamental frequency $\omega_0 = 2\pi/P$?

4.10 Plot the magnitude and phase spectra of the rectangular window in Example 4.3.3.

4.11 Find the frequency spectrum, if it exists, for each of the following signals:

(a) e^{-2t}, for all t in $(-\infty, \infty)$

(b) e^{-2t}, for $t \geq 0$ and zero for $t < 0$

(c) $e^{-2|t|}$, for all t in $(-\infty, \infty)$

(d) e^{-3t}, for $t \geq 0$, and $2e^{4t}$ for $t < 0$

Plot also their magnitude and phase spectra. Is there any frequency spectrum that is real-valued? Is every magnitude spectrum even? Is every phase spectrum odd?

4.12 Show that if $X(\omega) = \mathcal{F}[x(t)]$, then

$$x(-\omega) = \frac{1}{2\pi}\mathcal{F}[X(t)]$$

This is called the *duality property* of the CT Fourier transform. Use this property to derive (4.33) from (4.32).

4.13 Compute directly the Fourier transform of the shifted rectangular window in (4.38). Does it equal (4.39)?

4.14 Consider the signal

$$x(t) = Aw_a(t)\cos\omega_c t$$

where w_a is defined in (4.31). It is called a radio-frequency (RF) pulse when ω_c lies inside the radio-frequency band. What is its frequency spectrum?

4.15 Find the spectrum of

$$x(t) = e^{-0.1t}\sin(10t + \pi/4)$$

for $t \geq 0$. [*Hint:* Expand the sine function and then use (4.40) and (4.41). It can also be computed using the method in Section 6.9.]

4.16 Use the definitions in (4.43) and (4.44) to show $LB \geq \pi$. Thus the time duration is inversely related to the frequency bandwidth.

4.17 Compute the total energy of the signal whose magnitude spectrum is shown in Figure 4.15.

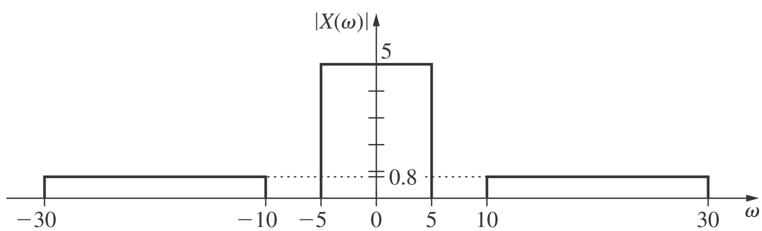


Figure 4.15

4.18 Compute the total energy of the signal whose squared magnitude spectrum is shown in Figure 4.16.

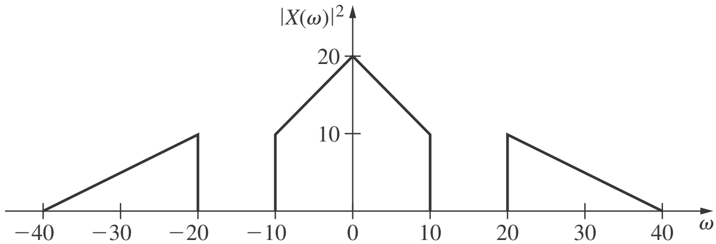


Figure 4.16

4.19 Consider $x(t)$ and its modulated signal $x_m(t) = x(t)\cos\omega_c t$. It is assumed that $x(t)$ is bandlimited to W; that is, $X(\omega) = 0$, for $|\omega| > W$. Verify that, if $\omega_c > 2W$, the total energy of $x_m(t)$ is half of the total energy of $x(t)$. Is the assertion correct without the condition $\omega_c > 2W$? Give your reasons.

4.20 What is the frequency spectrum of the signal in Problem 4.1?

4.21 Find the frequency spectrum of

$$x(t) = \begin{cases} 3\sin 2t + e^{-t} & \text{for } t \geq 0 \\ 3\sin 2t & \text{for } t < 0 \end{cases}$$

4.22 A signal is called a *power signal* if its average power defined by

$$P_{av} = \lim_{T\to\infty} \frac{1}{2T}\int_{t=-T}^{T} |x(t)|^2\,dt$$

is *nonzero* and finite. A signal is called an *energy signal* if its total energy defined by

$$E = \int_{-\infty}^{\infty} |x(t)|^2\,dt$$

is finite. Show that if a signal is a power signal, then it cannot be an energy signal. Show that if a signal is an energy signal, then it cannot be a power signal. Thus power and energy signals are mutually exclusive.

4.23 Show that the signal e^{-t}, for all t in $(-\infty, \infty)$, is neither an energy signal nor a power signal. Thus energy and power signals do not cover all signals.

4.24 Use (4.47) to show $\int_{-\infty}^{\infty} h(t)\,dt = 1$ and $\int_{-\infty}^{\infty} h^2(t)\,dt = \omega_c/\pi$, where $h(t)$ is given in (4.33). The proof of $\int_{-\infty}^{\infty} |h(t)|\,dt = \infty$ is complicated. It can be shown that the areas of the lobes are proportional to $\pm 1/n$, for $n = 1, 2, 3, \ldots$. The sum of their absolute values diverges as shown on page 361. This argument was provided to the author by Professors A. Zemanian and D. Kammler, the authors of References 13 and 23.

CHAPTER 5

Sampling Theorem and FFT Spectral Computation

5.1 INTRODUCTION

Frequency spectra of most CT signals encountered in practice cannot be expressed in closed form and cannot be computed analytically. The only way to compute them is numerically using their time samples. This chapter introduces the numerical method. In order to do so, we develop first frequency spectra of DT signals. We then develop the Nyquist sampling theorem, which relates the spectra of $x(t)$ and $x(nT)$. Finally we introduce the fast Fourier transform (FFT) to compute spectra of DT and CT signals.

There are two ways to develop frequency spectra of DT signals. One way is to develop the DT Fourier series, DT Fourier transform, and their properties. This development will be parallel to Sections 4.1 through 4.5 but is more complex. See, for example, Reference 2. Another way is to modify DT signals so that the CT Fourier transform can be directly applied. This approach is more efficient and is adopted in this chapter.

Before proceeding, we discuss two formulas. First we have

$$\sum_{n=0}^{N} r^n = 1 + r + r^2 + \cdots + r^N = \frac{1 - r^{N+1}}{1 - r} \tag{5.1}$$

where r is a real or complex constant. The formula can be directly verified by multiplying its both sides by $1 - r$. If $|r| < 1$, then $r^N \to 0$ as $N \to \infty$ and we have

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1 - r} \qquad \text{for } |r| < 1 \tag{5.2}$$

It diverges if $|r| > 1$ and $r = 1$. If $r = -1$, the summation in (5.1) equals alternatively 1 and -1 as N increases, thus (5.2) does not converge.[1] In conclusion the condition $|r| < 1$ is essential in (5.2).

[1]To see a difficulty involving ∞, consider the value of (5.2) with $r = -1$. It can be considered as 0 if we group it as $(1 - 1) + (1 - 1) + \cdots$, or 0.5 if we use $1/(1 + x) = 1 - x + x^2 - x^3 + \cdots$, with $x = 1$. To be precise, its value depends on how the summation in (5.2) is defined. See Reference 14 (p. 5).

5.2 FREQUENCY SPECTRA OF DT SIGNALS—DT FOURIER TRANSFORM

Consider a DT signal $x(nT)$, where n is the time index and $T > 0$, the sampling period. As discussed in Section 1.6, in processing or manipulation of $x(nT)$, the sampling period T does not play any role and can be suppressed or assumed to be 1. The frequency content of $x(nT)$, however, depends on T as is evident from the fact that the frequency of sinusoidal sequences is, as discussed in Section 1.8, limited to $(-\pi/T, \pi/T]$. Thus the sampling period T must appear explicitly in our discussion

We first modify the DT impulse sequence defined in (1.17) as

$$\delta_d([n-k]T) = \begin{cases} 1 & \text{for } n = k \\ 0 & \text{for } n \neq k \end{cases} \tag{5.3}$$

It differs from (1.17) only in containing the sampling period T. Note that the notation used is consistent with our convention that variables inside brackets are integers, and those inside parentheses are real numbers. Using (5.3), we can express any DT sequence $x(nT)$ as

$$x(nT) = \sum_{k=-\infty}^{\infty} x(kT)\delta_d([n-k]T) \tag{5.4}$$

This is essentially the equation in (1.21) except that the sequence in (1.21) is positive time, and the sequence in (5.4) is two-sided or is defined for all integers in $(-\infty, \infty)$. If we apply the CT Fourier transform to (5.4), the result will be identically zero, because (5.4) consists of a sequence of numbers with zero time duration. See Section 1.5. Now we modify (5.4) as, replacing nT by t,

$$x_d(t) := \sum_{k=-\infty}^{\infty} x(kT)\delta(t-kT) = \sum_{n=-\infty}^{\infty} x(nT)\delta(t-nT) \tag{5.5}$$

where $\delta(t)$ is the CT impulse defined in (1.9) and (1.10). Because $x_d(t)$ contains no n, we have changed the running index k to n after the last equality of (5.5). The signal in (5.5) is defined for all t, but is zero everywhere except at time instants nT where it is an impulse with weight $x(nT)$. Thus we may consider $x_d(t)$ as a CT representation of the DT sequence $x(nT)$.

The frequency spectrum of a CT signal $x(t)$ is defined in Section 4.3 as the CT Fourier transform of $x(t)$. Thus it is natural to define the frequency spectrum of $x(nT)$ as the CT Fourier transform of $x_d(t)$, that is,

$$X_d(\omega) := \text{Spectrum of } x(nT) = \mathcal{F}[x_d(t)] \tag{5.6}$$

Using the linearity of the CT Fourier transform and

$$\mathcal{F}[\delta(t-nT)] = \int_{t=-\infty}^{\infty} \delta(t-nT)e^{-j\omega t}\,dt = e^{-j\omega t}\big|_{t=nT} = e^{-j\omega nT}$$

we can write (5.6) as

$$X_d(\omega) = \sum_{n=-\infty}^{\infty} x(nT)\mathcal{F}[\delta(t-nT)] = \sum_{n=-\infty}^{\infty} x(nT)e^{-jn\omega T} =: \mathcal{F}_d[x(nT)] \tag{5.7}$$

This is, if it exists, the *frequency spectrum* or, simply, *spectrum* of the DT signal $x(nT)$. The function $X_d(\omega)$ in (5.7) is in fact the *DT Fourier transform* of $x(nT)$, denoted as $\mathcal{F}_d[x(nT)]$.

Just as the CT case, not every DT sequence has a frequency spectrum. If $x(nT)$ grows unbounded as $n \to \infty$ or $-\infty$, then its frequency spectrum is not defined. A sufficient condition is that the sequence is absolutely summable in $(-\infty, \infty)$, that is,

$$\sum_{n=-\infty}^{\infty} |x(nT)| < \infty$$

Another sufficient condition is that $x(nT)$ is periodic in the sense that there exists an integer N such that $x([n+N]T) = x(nT)$, for all integer n. Yet another sufficient condition is that $x(nT)$ is squared absolutely summable in $(-\infty, \infty)$.[2]

If a CT signal is absolutely integrable, then its frequency spectrum is bounded and continuous. Likewise, if a DT signal is absolutely summable, then its spectrum is bounded and continuous. Every practical DT signal will not grow to $\pm\infty$ and will not last forever, thus it is absolutely summable and its spectrum is bounded and continuous. In other words, frequency spectra we encounter in practice are all well behaved (no jumps and no infinity). We now give examples.

EXAMPLE 5.2.1

Consider the sequence

$$x(nT) = \begin{cases} e^{-0.1nT} & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases}$$

as shown in Figure 5.1(a) with $T = 0.5$. The sequence decreases exponentially to zero. Because $e^{-0.1T} = e^{-0.05} = 0.95$, (5.2) implies

$$\sum_{n=-\infty}^{\infty} |x(nT)| = \sum_{n=0}^{\infty} (e^{-0.1T})^n = \frac{1}{1-0.95} = 20 < \infty$$

Thus the sequence is absolutely summable. Its frequency spectrum is, using (5.2),

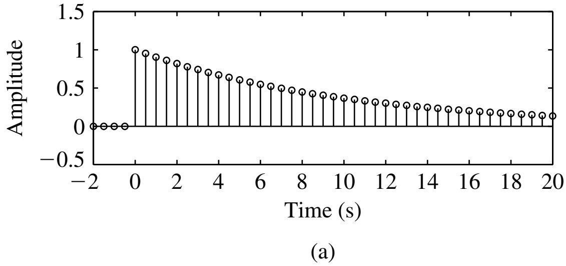
$$\begin{aligned} X_d(\omega) &= \sum_{n=0}^{\infty} e^{-0.1nT} e^{-jn\omega T} = \sum_{n=0}^{\infty} (e^{-0.1T} e^{-j\omega T})^n \\ &= \frac{1}{1-e^{-0.1T}e^{-j\omega T}} = \frac{1}{1-0.95e^{-j0.5\omega}} \end{aligned} \tag{5.8}$$

---

[2]Note that $\sum |x(nT)| < \infty$ implies $\sum |x(nT)|^2 < \infty$, but not conversely. If $\sum |x(nT)| < \infty$, then $|x(nT)| \leq q$, for all n and some finite q. Thus $\sum |x(nT)||x(nT)| \leq \sum q|x(nT)| = q \sum |x(nT)|$ and the first assertion follows. The sequence $1/n$ for $n \geq 1$ is squared absolutely summable but not absolutely summable. Thus we have the second assertion. Note that in the CT case, the condition that $h(t)$ is absolutely integrable does not imply that $h(t)$ is bounded or that $h(t)$ is squared absolutely integrable. See Problem 4.4. Thus the CT case is more complex than the DT case.

where we have used the fact that for any $T > 0$, $|e^{j\omega T}| = 1$, for all ω, and

$$|e^{-0.1T}e^{-j\omega T}| = e^{-0.1T} < 1$$



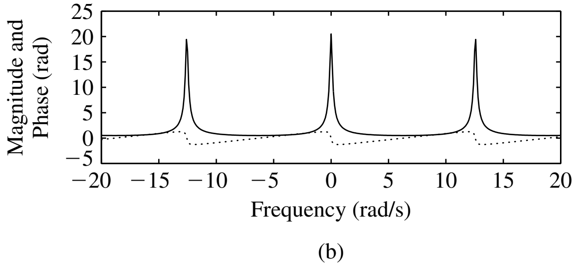


Figure 5.1 (**a**) Time sequence. (**b**) Its magnitude and phase spectra.

In general, the frequency spectrum is complex-valued as shown in (5.8). Its magnitude is called the *magnitude spectrum,* and its phase is called the *phase spectrum*. We compute the magnitudes and phases of (5.8) with $T = 0.5$ at $\omega = 0$ and 1 as

$$\omega = 0: \qquad X_d(0) = \frac{1}{1 - 0.95 \times 1} = 20.5$$

$$\omega = 1: \qquad X_d(1) = \frac{1}{1 - 0.95 \times (0.88 - j0.48)} = \frac{1}{0.16 + j0.46}$$

$$= \frac{1}{0.49e^{j1.24}} = 2.04e^{-j1.24}$$

We see that its computation by hand is complicated. Thus we use MATLAB. Typing

```
w=-20:0.01:20;T=0.5;
X=1./(1-exp(-0.1*T).*exp(-j*w*T));
plot(w,abs(X),w,angle(X),':')
```

yields the magnitude (solid line) and phase (dotted line) spectra of (5.8) in Figure 5.1(b).

As in CT signals, if a DT signal is real-valued, then we have

$$|X_d(\omega)| = |X_d(-\omega)| \quad \text{and} \quad \measuredangle X_d(\omega) = -\measuredangle X_d(-\omega) \tag{5.9}$$

that is, its magnitude spectrum is even and its phase spectrum is odd. This is indeed the case as shown in Figure 5.1(b). If $x(nT)$ is real and even, so is its frequency spectrum. The situation is identical to the CT case.

A common misconception is that the spectrum of a DT sequence is defined only at discrete frequencies. This is not the case as shown in Figure 5.1. We give one more example to dispel this misconception.

EXAMPLE 5.2.2

Consider a sequence that consists of only four nonzero entries $x[0] = 2$, $x[1] = 3$, $x[2] = 2$, and $x[3] = -3$, with $T = 0.3$. Its frequency spectrum is

$$X_d(\omega) = \sum_{n=-\infty}^{\infty} x(nT)e^{-jn\omega T} = 2 + 3e^{-j\omega T} + 2e^{-j2\omega T} - 3e^{-j3\omega T} \tag{5.10}$$

Typing

```
w=-20:0.01:20;T=0.3;
Xd=2+3*exp(-j*T.*w)+2*exp(-j*2*T.*w)-3*exp(-j*3*T.*w);
subplot(1,2,1)
plot(w,abs(Xd))
subplot(1,2,2)
plot(w,angle(Xd))
```

in MATLAB yields the magnitude and phase spectra in Figures 5.2(a) and 5.2(b). We see that the spectrum is defined for all ω, bounded, and continuous. Note that the jumps in the phase plot are due to the way it is plotted. The phase is plotted within $\pm\pi$. Thus when a phase crosses $\pm\pi$, it jumps to $\mp\pi$.

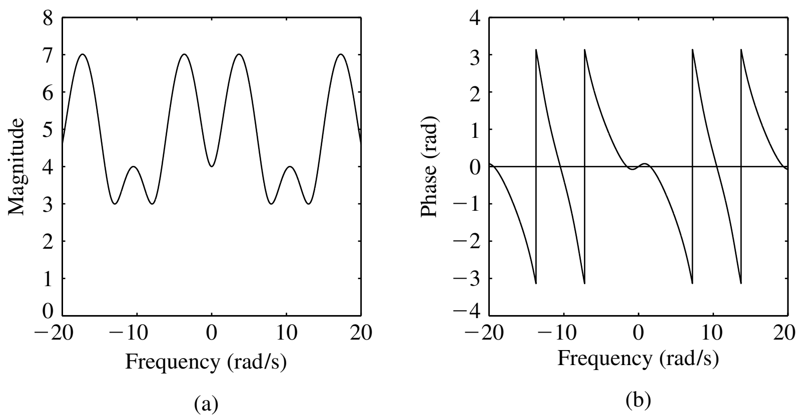


Figure 5.2 Frequency spectrum.

EXERCISE 5.2.1

Find the spectrum of $x[0] = -2$, $x[n] = 0$ for all nonzero integer n, and $T = 1$. Plot its magnitude and phase spectra.

Answer

$X_d(\omega) = -2$. Its magnitude spectrum is 2 for all ω and is even. Its phase spectrum can be selected as π for all ω, or π for $\omega \geq 0$ and $-\pi$ for $\omega < 0$. In either case, the phase spectrum is mathematically odd. See Section 1.7.3.

EXERCISE 5.2.2

Find the spectrum of $x[1] = 2$, $x[n] = 0$ for all $n \neq 1$, and $T = 1$. Plot its magnitude and phase spectra.

Answer

$X_d(\omega) = 2e^{-j\omega}$, $|X_d(\omega)| = 2$, and $\measuredangle X_d(\omega) = -\omega$, for all ω.

5.2.1 Nyquist Frequency Range

Because frequency spectra of DT signals are defined through the CT Fourier transform, all properties discussed for CT signals are directly applicable to DT signals. There is, however, an important difference as we discuss next.

The spectrum of $x(nT)$ is, as defined in (5.7),

$$X_d(\omega) = \sum_{n=-\infty}^{\infty} x(nT)e^{-jn\omega T}$$

We show that the spectrum $X_d(\omega)$ is periodic with period $2\pi/T$. Indeed, because

$$e^{jn(2\pi)} = 1$$

for all integers n (negative, 0, and positive), we have

$$e^{jn(\omega+2\pi/T)T} = e^{jn\omega T}e^{jn(2\pi/T)T} = e^{jn\omega T}e^{jn(2\pi)} = e^{jn\omega T}$$

Thus we have

$$X_d(\omega) = X_d(\omega + 2\pi/T) \tag{5.11}$$

for all ω. This is indeed the case as shown in Figures 5.1(b) and 5.2. The sequence in Example 5.1 has sampling period $T = 0.5$, and its spectrum in Figure 5.1(b) is periodic with period $2\pi/T = 12.6$. The sequence in Example 5.2 has sampling period $T = 0.3$, and its spectrum in Figure 5.2 is periodic with period $2\pi/0.3 = 20.9$. Because of the periodicity of $X_d(\omega)$, we need to plot it only over one period. As in Section 1.8, we select the period as $(-\pi/T, \pi/T]$, called the *Nyquist frequency range*. As mentioned there, the Nyquist frequency range can also be selected as $[-\pi/T, \pi/T)$, but not $[-\pi/T, \pi/T]$. If $X_d(\omega)$ is bounded and continuous as in Figures 5.1(b) and 5.2, there is no difference in using $[-\pi/T, \pi/T]$, $[-\pi/T, \pi/T)$, or $(-\pi/T, \pi/T]$.

5.2.2 Inverse DT Fourier Transform

Consider a DT sequence $x(nT)$. Its DT Fourier transform is

$$X_d(\omega) = \mathcal{F}_d[x(nT)] = \sum_{n=-\infty}^{\infty} x(nT)e^{-jn\omega T} \tag{5.12}$$

We now discuss its inverse DT Fourier transform—that is, computing $x(nT)$ from $X_d(\omega)$.[3]

The DT Fourier transform $X_d(\omega)$, as shown in (5.11), is periodic with period $2\pi/T$. Comparing (5.12) with (4.6) and identifying $P = 2\pi/T$, $m = n$, $t = \omega$, and $\omega_0 = T$, we see that (5.12) is actually the CT Fourier series of the periodic function $X_d(\omega)$. Thus the DT Fourier transform is intimately related to the CT Fourier series, and the procedure for computing the CTFS coefficient c_m in (4.7) can be used to compute $x(nT)$.

As in (4.4), the set $e^{-jn\omega T}$, for $n = 0, \pm 1, \pm 2, \ldots$, is *orthogonal* in the sense

$$\int_{-\pi/T}^{\pi/T} e^{-j(n-k)\omega T} d\omega = \begin{cases} 2\pi/T & \text{if } n = k \\ 0 & \text{if } n \neq k \end{cases}$$

Indeed, if $n = k$, then the integrand is 1, for all ω, and its integration over the interval $2\pi/T$ is $2\pi/T$. If $n \neq k$, then, using $\sin l\pi = 0$ for all integers l, we have

$$\begin{aligned}\int_{-\pi/T}^{\pi/T} e^{-j(n-k)\omega T} d\omega &= \frac{1}{-j(n-k)T} e^{-j(n-k)\omega T}\Big|_{\omega=-\pi/T}^{\pi/T} \\ &= \frac{e^{-j(n-k)\pi} - e^{j(n-k)\pi}}{-j(n-k)T} = \frac{2\sin(n-k)\pi}{(n-k)T} = 0\end{aligned}$$

This shows the orthogonality of the set $e^{-jn\omega T}$.

Let us multiply (5.12) by $e^{jk\omega T}$ and then integrate the resulting equation from $-\pi/T$ to π/T to yield

$$\int_{-\pi/T}^{\pi/T} X_d(\omega)e^{jk\omega T}\, d\omega = \sum_{n=-\infty}^{\infty} x(nT)\left(\int_{-\pi/T}^{\pi/T} e^{-j(n-k)\omega T}\, d\omega\right)$$

The right-hand side, using the orthogonality property, can be reduced to $x(kT)(2\pi/T)$. Thus we have

$$\int_{-\pi/T}^{\pi/T} X_d(\omega)e^{jk\omega T}\, d\omega = \frac{2\pi}{T}x(kT)$$

---

[3]This subsection and the next one may be skipped without loss of continuity if we accept the spectra of sinusoidal sequences in (5.16) and (5.17).

or, renaming k to n,

$$x(nT) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X_d(\omega) e^{jn\omega T} \, d\omega \tag{5.13}$$

This is the *inverse DT Fourier transform,* denoted as $x(nT) = \mathcal{F}_d^{-1}[X_d(\omega)]$.

Physically, a DT signal has no energy because it consists of only a sequence of numbers. Mathematically, its total energy, however, can be defined as, similar to (4.48),

$$E := \sum_{n=-\infty}^{\infty} x(nT)x^*(nT) = \sum_{-\infty}^{\infty} |x(nT)|^2$$

where $x^*(nT)$ is the complex conjugate of $x(nT)$. Using (5.13), we obtain

$$\begin{aligned} E &= \sum_{n=-\infty}^{\infty} x(nT) \left(\frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X_d^*(\omega) e^{-jn\omega T} d\omega \right) \\ &= \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X_d^*(\omega) \left(\sum_{n=-\infty}^{\infty} x(nT) e^{-jn\omega T} \right) d\omega \end{aligned}$$

where we have changed the order of summation and integration. Using (5.12), we finally obtain

$$\sum_{n=-\infty}^{\infty} |x(nT)|^2 = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X_d^*(\omega) X_d(\omega) \, d\omega = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} |X_d(\omega)|^2 \, d\omega$$

We see that the spectrum reveals the distribution of energy in frequencies in the Nyquist frequency range. The equation is a *Parseval's formula* and is the DT counterpart of (4.49). Using the formula, we can compute the total energy of $x(nT)$ either directly in the time domain or indirectly in the frequency domain.

5.2.3 Frequency Spectra of DT Sinusoidal Sequences

Even though a DT sinusoid is not absolutely summable from $n = -\infty$ to ∞, its frequency spectrum is still defined. Instead of computing it directly from (5.12), we compute the inverse DT Fourier transform of $X_d(\omega) = \delta(\omega - \omega_0)$, an impulse at ω_0, with $|\omega_0| < \pi/T$. Using the sifting property in (1.12), we compute

$$\mathcal{F}_d^{-1}[\delta(\omega - \omega_0)] = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \delta(\omega - \omega_0) e^{jn\omega T} d\omega = \frac{T}{2\pi} e^{j\omega_0 nT}$$

which implies

$$\mathcal{F}_d[e^{j\omega_0 nT}] = \frac{2\pi}{T} \delta(\omega - \omega_0) \tag{5.14}$$

Thus the frequency spectrum of the complex exponential sequence $e^{j\omega_0 nT}$ is an impulse at ω_0 with weight $2\pi/T$. If $\omega_0 = 0$, (5.14) becomes

$$\mathcal{F}_d[1] = \frac{2\pi}{T} \delta(\omega) \tag{5.15}$$

where 1 is the time sequence $x(nT) = 1$ for all n in $(-\infty, \infty)$. Its frequency spectrum is zero everywhere except at $\omega = 0$.

Using Euler's formula, we have

$$\mathcal{F}_d[\sin \omega_0 nT] = \mathcal{F}_d\left[\frac{e^{j\omega_0 nT} - e^{-j\omega_0 nT}}{2j}\right] = \frac{\pi}{jT}[\delta(\omega-\omega_0) - \delta(\omega+\omega_0)]$$
$$= -\frac{j\pi}{T}\delta(\omega-\omega_0) + \frac{j\pi}{T}\delta(\omega+\omega_0) \tag{5.16}$$

and

$$\mathcal{F}_d[\cos \omega_0 nT] = \mathcal{F}_d\left[\frac{e^{j\omega_0 nT} + e^{-j\omega_0 nT}}{2}\right] = \frac{\pi}{T}[\delta(\omega-\omega_0) + \delta(\omega+\omega_0)] \tag{5.17}$$

We see that they are the same as those in (4.53) and (4.54) except the factor $1/T$.

EXERCISE 5.2.3

Find the spectrum of $x(nT) = \cos 50nT + 2\cos 70nT$, with $70 < \pi/T$.

Answer

$X_d(\omega) = (\pi/T)\delta(\omega-50) + (\pi/T)\delta(\omega+50) + (2\pi/T)\delta(\omega-70) + (2\pi/T)\delta(\omega+70)$

EXERCISE 5.2.4

Find the spectrum of $x(nT) = \sin 50nT + 2\sin 70nT$, with $70 < \pi/T$.

Answer

$X_d(\omega) = -(j\pi/T)\delta(\omega-50) + (j\pi/T)\delta(\omega+50) - (2\pi j/T)\delta(\omega-70) + (2\pi j/T)\delta(\omega+70)$

5.3 NYQUIST SAMPLING THEOREM

Let $x(t)$ be a CT signal with frequency spectrum $X(\omega)$, that is,

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\, dt$$

Let $x(nT)$ be its sampled signal with sampling period T. The spectrum $X_d(\omega)$ of $x(nT)$ is defined as the CT Fourier transform of

$$x_d(t) = \sum_{n=-\infty}^{\infty} x(nT)\delta(t-nT) \tag{5.18}$$

that is,

$$X_d(\omega) = \int_{-\infty}^{\infty} x_d(t)e^{-j\omega t} dt = \sum_{n=-\infty}^{\infty} x(nT)e^{-j\omega nT} \tag{5.19}$$

Now we discuss the relationship between $X(\omega)$ and $X_d(\omega)$.

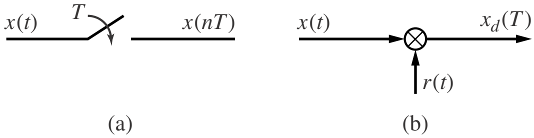


Figure 5.3 (**a**) Sampling. (**b**) Modulation.

Because $x(nT)\delta(t-nT) = x(t)\delta(t-nT)$, as discussed in (1.11), we write (5.18) as

$$x_d(t) = \sum_{n=-\infty}^{\infty} x(t)\delta(t-nT) = x(t)\sum_{n=-\infty}^{\infty}\delta(t-nT) =: x(t)r(t) \tag{5.20}$$

where $r(t)$ is the sampling function defined in (4.13). Thus the process of sampling shown in Figure 5.3(a) can also be considered as the modulation shown in Figure 5.3(b).

The sampling function $r(t)$, as shown in Figure 4.2, consists of a sequence of impulses, is periodic with period T, and can be expressed in Fourier series as in (4.15). Let us define $\omega_s = 2\pi/T$. It is the *sampling frequency* in rad/s. We also define $f_s = 1/T = \omega_s/2\pi$. It is the sampling frequency in Hz. Substituting (4.15), with ω_0 replaced by ω_s, into (5.20) yields

$$x_d(t) = x(t)\sum_{m=-\infty}^{\infty}\frac{1}{T}e^{jm\omega_s t} = \frac{1}{T}\sum_{m=-\infty}^{\infty} x(t)e^{jm\omega_s t} \tag{5.21}$$

Applying the CT Fourier transform to (5.21) and using the frequency shifting property in (4.40), we obtain

$$X_d(\omega) = \frac{1}{T}\sum_{m=-\infty}^{\infty} X(\omega - m\omega_s) = \frac{1}{T}\sum_{m=-\infty}^{\infty} X\left(\omega - m\frac{2\pi}{T}\right) \tag{5.22}$$

This is a fundamental equation relating the frequency spectrum of a CT signal and the frequency spectrum of its sampled sequence.

A CT signal $x(t)$ is said to be *band-limited* to ω_{max} if $[-\omega_{max}, \omega_{max}]$ is the *smallest frequency band* that contains all nonzero spectrum $X(\omega)$ of $x(t)$ or, equivalently,

$$X(\omega) = 0 \qquad \text{for } |\omega| > \omega_{max}$$

and $X(\omega_{max}) \neq 0$ or $X(\omega_{max-}) \neq 0$, where ω_{max-} is a frequency infinitesimally smaller than ω_{max}. Now we discuss the implication of (5.22). We assume $X(\omega)$ to be real-valued and of the form shown in Figure 5.4(a). It is band-limited to ω_{max} (rad/s). The function $X(\omega - m\omega_s)$ is the shifting of $X(\omega)$ to $\omega = m\omega_s$. Thus $X_d(\omega)$ is the sum of all repetitive shiftings of $X(\omega)/T$ to $m\omega_s$, for $m = 0, \pm 1, \pm 2, \ldots$. Because $X_d(\omega)$ is periodic, we need to plot it only in the Nyquist frequency range bounded by the two vertical dashed lines shown in Figure 5.4. Furthermore, *if* $X(\omega)$ *is even,* the two shiftings immediately outside the range can be obtained by folding $X(\omega)$ with respect to $\pm\omega_s/2$. Using either shifting or folding, we plot in Figure 5.4 the frequency spectra $TX_d(\omega)$ for three different T. Note that the vertical ordinates of Figures 5.4(b) through (d) are $TX_d(\omega)$, not $X_d(\omega)$. We see that if $\omega_s/2 > \omega_{max}$ or $\omega_s > 2\omega_{max}$, the repetitions will not overlap as in Figure 5.4(b) and the resulting $TX_d(\omega)$ is identical to $X(\omega)$ inside the Nyquist frequency range $[-\omega_s/2, \omega_s/2)$. Note that they are different outside the range because $X_d(\omega)$ can be extended periodically but $X(\omega)$ is zero. In conclusion, if $x(t)$ is band-limited to ω_{max} rad/s and if the sampling period T is chosen to be less than π/ω_{max} or the sampling frequency ω_s is

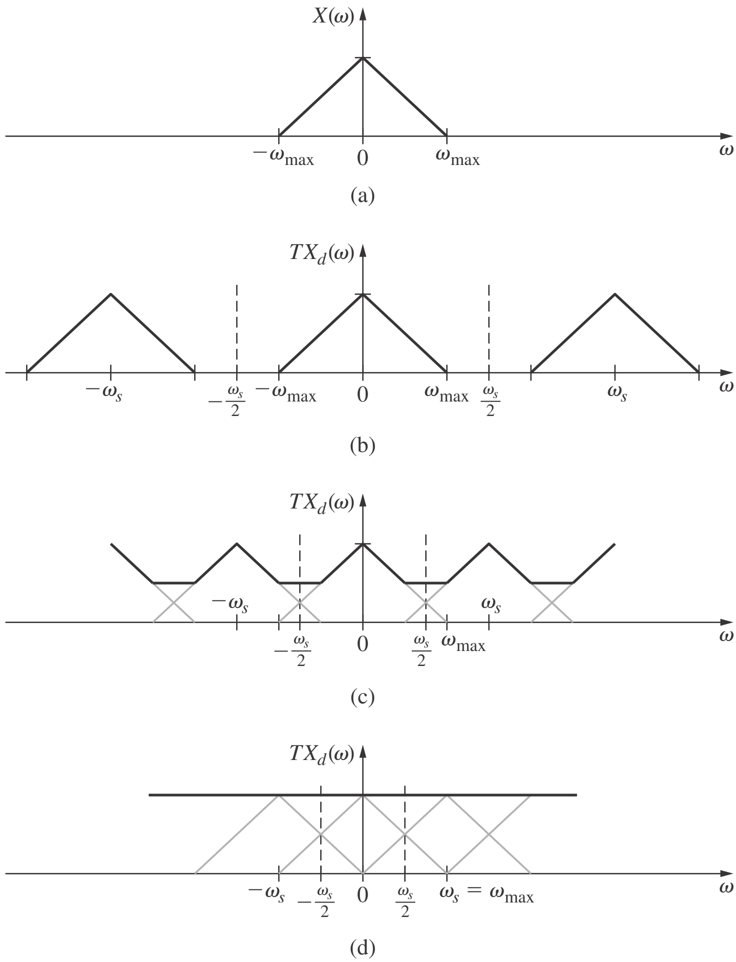


Figure 5.4 (**a**) Frequency spectrum of a CT signal $x(t)$ band-limited to ω_{max}. (**b**) Frequency spectrum of $Tx(nT)$ with sampling period $T < \pi/\omega_{max}$ or $\omega_s > 2\omega_{max}$. (**c**) With $\omega_{max} < \omega_s < 2\omega_{max}$. (**d**) With $\omega_s = \omega_{max} < 2\omega_{max}$.

chosen to be larger than $2\omega_{max}$, then

$$\text{Spec. of } x(t) = \begin{cases} T \times (\text{Spec. of } x(nT)) & \text{for } |\omega| \le \pi/T = \omega_s/2 \\ 0 & \text{for } |\omega| > \pi/T = \omega_s/2 \end{cases} \tag{5.23}$$

where Spec. stands for frequency spectrum.

Consider a CT signal $x(t)$. Suppose only its values at sampling instants are known, can we determine all values of $x(t)$? Or can $x(t)$ be recovered from $x(nT)$? The answer is affirmative if $x(t)$ is band-limited. This follows directly from (5.23). We use $x(nT)$ to compute the spectrum of $x(t)$. Its inverse Fourier transform yields $x(t)$. This is essentially the theorem that follows.

Nyquist Sampling Theorem Let $x(t)$ be a CT signal band-limited to ω_{max} (in rad/s) or $f_{max} = \omega_{max}/2\pi$ (in Hz), that is,

$$X(\omega) = 0 \qquad \text{for } |\omega| > \omega_{max}$$

where $X(\omega)$ is the frequency spectrum of $x(t)$. Then $x(t)$ can be recovered from its sampled sequence $x(nT)$ if the sampling period T is less than π/ω_{max}, or the sampling frequency $f_s = 1/T$ is larger than $2f_{max}$.

If $x(t)$ is band-limited to ω_{max}, then it can be computed from its spectrum as

$$\begin{aligned} x(t) &= \frac{1}{2\pi}\int_{\omega=-\infty}^{\infty} X(\omega)e^{j\omega t}\,d\omega \\ &= \frac{1}{2\pi}\int_{\omega=-\pi/T}^{\pi/T} X(\omega)e^{j\omega t}\,d\omega \end{aligned} \tag{5.24}$$

where $\pi/T > \omega_{max}$ or $T < \pi/\omega_{max}$. If $T < \pi/\omega_{max}$, then the spectrum of $x(t)$ equals the spectrum of $Tx(nT)$ inside the integration range. Thus, using (5.23) and (5.19), we can write (5.24) as

$$\begin{aligned} x(t) &= \frac{1}{2\pi}\int_{\omega=-\pi/T}^{\pi/T} \left[T\sum_{n=-\infty}^{\infty} x(nT)e^{-j\omega nT}\right] e^{j\omega t}\,d\omega \\ &= \frac{T}{2\pi}\sum_{n=-\infty}^{\infty} x(nT)\int_{\omega=-\pi/T}^{\pi/T} e^{j\omega(t-nT)}\,d\omega \\ &= \frac{T}{2\pi}\sum_{n=-\infty}^{\infty} x(nT)\frac{e^{j(\pi/T)(t-nT)} - e^{-j(\pi/T)(t-nT)}}{j(t-nT)} \\ &= \sum_{n=-\infty}^{\infty} x(nT)\frac{\sin[\pi(t-nT)/T]}{\pi(t-nT)/T} \\ &= \sum_{n=-\infty}^{\infty} x(nT)\text{sinc}\,[\pi(t-nT)/T] \end{aligned} \tag{5.25}$$

where sinc θ is the sinc function defined in (4.34). Thus $x(t)$ is a weighted sum of sinc functions. We verify (5.25) at $t = lT$, where l is an integer. Because sinc $0 = 1$ and sinc $k\pi = 0$ for every nonzero integer k, as shown in Figure 4.9, if $t = lT$, we have sinc $[\pi(lT - nT)/T] =$ sinc $(l-n)\pi = 0$ for all n except at $n = l$. Thus the summation in (5.25) reduces to $x(lT)$ at $t = lT$.

If $x(nT)$, for all n, are known, the CT signal $x(t)$, for all t, can be computed using (5.25). Thus we conclude that if $x(t)$ is band-limited to ω_{max} and if $T < \pi/\omega_{max}$, then $x(t)$ can be recovered from $x(nT)$. We call (5.25) the *ideal interpolation formula*. This establishes the Nyquist sampling theorem. The discussion in Section 1.9 is a special case of this theorem.

To conclude this section, we mention that in the literature, the *sampling frequency* is also called the *sampling rate*. But the Nyquist frequency ($\omega_s/2$) is different from the Nyquist rate ($2\omega_{max}$). Thus frequency and rate are not always interchangeable.

5.3.1 Frequency Aliasing Due to Time Sampling

Consider again (5.22) and the CT signal with the frequency spectrum shown in Figure 5.4(a). If $\omega_s < 2\omega_{max}$, then the repetitions of $X(\omega - m\omega_s)$ will overlap as shown in Figures 5.4(c) and 5.4(d). This can be obtained by shifting or folding. This type of overlapping is called *frequency aliasing*. In this case, the resulting $TX_d(\omega)$ in the frequency range $[-\pi/T, \pi/T)$ will be different from $X(\omega)$ and we cannot compute the spectrum of $x(t)$ from its sampled sequence. We give two more examples.

EXAMPLE 5.3.1

Consider the CT signal studied in Example 1.9.2 or $x_1(t) = \cos 50t + 2\cos 70t$. Its spectrum was computed in Exercise 4.5.1 and is plotted in Figure 5.5(a). The signal is

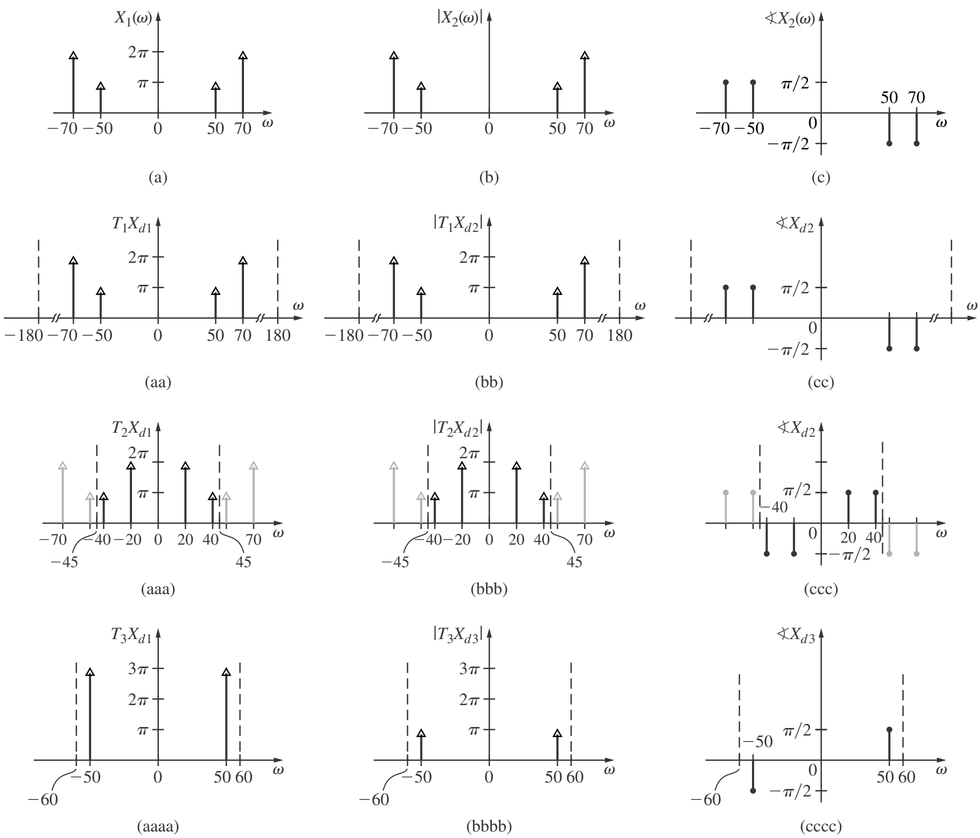


Figure 5.5 (**a**) Spectrum of $x_1(t) = \cos 50t + 2\cos 70t$. (**aa**) Spectrum of $Tx_1(nT)$ with $T = \pi/180$. (**aaa**) With $T = \pi/45$. (**aaaa**) With $T = \pi/60$. (**b**) Magnitude spectrum of $x_2(t) = \sin 50t + 2\sin 70t$. (**bb**) Magnitude spectrum of $Tx_2(nT)$ with $T = \pi/180$. (**bbb**) With $T = \pi/45$. (**bbbb**) With $T = \pi/60$. (**c**) Phase spectrum of $x_2(t) = \sin 50t + 2\sin 70t$. (**cc**) Phase spectrum of $Tx_2(nT)$ with $T = \pi/180$. (**ccc**) With $T = \pi/45$. (**cccc**) With $T = \pi/60$.

band-limited to 70 rad/s. If we sample it with sampling period $T_1 = \pi/180$, then the spectrum of its sampled sequence $x_1(nT_1) = \cos 50nT_1 + 2\cos 70nT_1$ multiplied by T_1 is shown in Figure 5.5(aa). This is consistent with the result in Exercise 5.2.3. We see that the spectra in Figures 5.5(a) and 5.5(aa) are the same inside the Nyquist frequency range (NFR) $(-180, 180]$ bounded by the two vertical dashed lines. Thus there is no frequency aliasing, and $x_1(t)$ can be recovered from $x_1(nT_1)$.

Next we select $T_2 = \pi/45$. Its NFR is $(-45, 45]$ and is bounded by the two vertical dashed lines in Figure 5.5(aaa). Because ± 50 and ± 70 are outside the range, we must fold them into ± 40 and ± 20 as shown in Figure 5.5(aaa). The resulting sampled sequence consists of frequencies ± 20 and ± 40 rad/s. They are the aliased frequencies. This is consistent with (1.43).

Finally we select $T_3 = \pi/60$. Its NFR is $(-60, 60]$ and is bounded by the two vertical dashed lines in Figure 5.5(aaaa). Because ± 70 are outside the range, we must fold them into ± 50, which coincide with the existing ones. Thus the resulting spectrum is as shown in Figure 5.5(aaaa). It is the spectrum of $3T_3 \cos 50nT_3$ and is consistent with (1.42).

EXAMPLE 5.3.2

Consider the CT signal studied in Problem 1.34 or $x_2(t) = \sin 50t + 2\sin 70t$. Its spectrum was computed in Exercise 4.5.2. It is complex-valued. We plot its magnitude and phase spectra, respectively, in Figures 5.5(b) and 5.5(c). The signal is band-limited to 70 rad/s. If we sample it with sampling period $T_1 = \pi/180$, then the magnitude and phase spectra of its sampled sequence $x_2(nT_1) = \sin 50nT_1 + 2\sin 70nT_1$ multiplied by T_1 are as shown in Figures 5.5(bb) and 5.5(cc). This is consistent with the result in Exercise 5.2.4. Because the frequency spectra of $x(t)$ and $T_1 x(nT_1)$ are the same inside the Nyquist frequency range (NFR) $(-180, 180]$ bounded by the two vertical dashed lines, there is no frequency aliasing. Thus $x_2(t)$ can be recovered from $x_2(nT_1)$.

Next we select $T_2 = \pi/45$. Its NFR is $(-45, 45]$ and is bounded by the two vertical dashed lines in Figures 5.5(bbb) and 5.5(ccc). Because the magnitude spectrum is even, the magnitude spectrum of the sampled sequence can be obtained by shifting or folding. Because the phase spectrum is not even, the phase spectrum of the sampled sequence can be obtained *only* by shifting (no folding). We see that frequency aliasing occurs and it is not possible to recover $x_2(t)$ from $x_2(nT_2)$.

Finally we select $T_3 = \pi/60$. Its NFR is $(-60, 60]$ and is bounded by the two vertical dashed lines in Figures 5.5(bbbb) and 5.5(cccc). Because ± 70 are outside the range, we must shift them into ∓ 50, which coincide with the existing ones. Because the spectrum is complex-valued, the effects of frequency aliasing involve additions of complex numbers. The final results are as shown in Figures 5.5(bbbb) and 5.5(cccc). See Problem 1.34. Again, it is not possible to recover $x_2(t)$ from $x_2(nT_3)$.

If the spectrum of a CT signal is real-valued, then the effect of frequency aliasing can be easily obtained by shifting or folding as shown in Figures 5.4 and 5.5(a) through 5.5(aaaa). If the spectrum of a CT signal is complex-valued as is always the case in practice, then the effect of frequency aliasing must be obtained by shifting (no folding) and is complicated because it involves additions of complex numbers. Fortunately in practice, there is no need to compute the effect of frequency aliasing so long as it is small or negligible.

If a CT signal is not bandlimited, frequency aliasing always occurs no matter how small the sampling period is chosen. However, frequency spectra of most practical CT signals decrease to zero as $|\omega| \to \infty$. Thus in practice, even though $x(t)$ is not band-limited, there exists a sufficiently small T such that

$$X(\omega) \approx 0 \qquad \text{for } |\omega| > \pi/T = \omega_s/2 \tag{5.26}$$

which implies

$$X(\omega - k\omega_s) \approx 0 \qquad \text{for } \omega \text{ in } (-\omega_s/2, \omega_s/2]$$

and for all $k \neq 0$. Thus (5.23) becomes, under the condition in (5.26),

$$X(\omega) \approx \begin{cases} TX_d(\omega) & \text{for } |\omega| \le \omega_s/2 = \pi/T \\ 0 & \text{for } |\omega| > \omega_s/2 = \pi/T \end{cases} \tag{5.27}$$

for T sufficiently small. This equation implies

$$|X(\omega)| \approx \begin{cases} T|X_d(\omega)| & \text{for } |\omega| \le \omega_s/2 = \pi/T \\ 0 & \text{for } |\omega| > \omega_s/2 = \pi/T \end{cases} \tag{5.28}$$

The situation for the phase spectrum is different.[4] The phase of a complex number can be very large even if the complex number is very small. For example, the phase of $-0.0001 + j0.0001$ is $3\pi/4$ rad or 135°. Thus in general, we have

$$\measuredangle X(\omega) \neq \measuredangle(TX_d(\omega)) = \measuredangle X_d(\omega) \qquad \text{for } |\omega| \le \omega_s/2 \tag{5.29}$$

This is demonstrated in the next example.

EXAMPLE 5.3.3

Consider the CT signal $x(t) = e^{-0.1t}$, for $t \ge 0$, discussed in Example 4.3.2 with $a = 0.1$. Its spectrum was computed in (4.30) as $X(\omega) = 1/(j\omega + 0.1)$ and plotted in Figure 4.6(b). Its magnitude approaches zero and its phase approaches $\mp\pi/2$ as $\omega \to \pm\infty$. It is not band-limited.

The sampled sequence of $x(t)$ with sampling period T is $x(nT) = e^{-0.1nT}$, for $n \ge 0$. Its spectrum was computed in Example 5.2.1 as

$$X_d(\omega) = \frac{1}{1 - e^{-0.1T}e^{-j\omega T}}$$

[4]Equation (5.27) does imply $\mathrm{Re}[X(\omega)] \approx \mathrm{Re}[TX_d(\omega)]$ and $\mathrm{Im}[X(\omega)] \approx \mathrm{Im}[TX_d(\omega)]$. But the plots of the real and imaginary parts of a spectrum against frequency have no physical meaning and are rarely plotted.

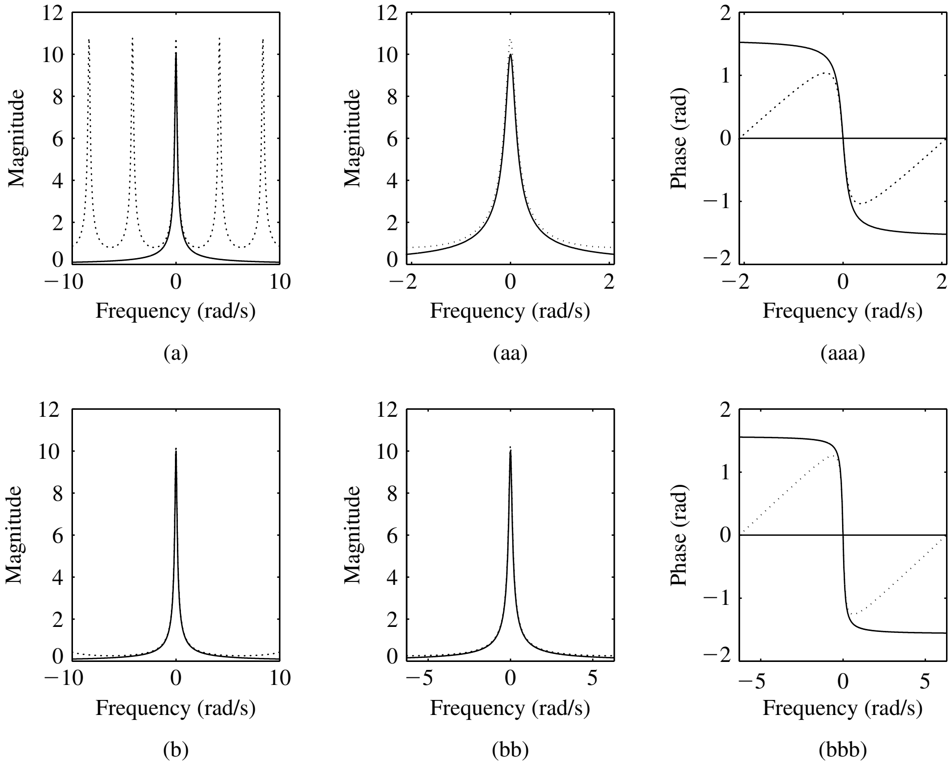


Figure 5.6 (**a**) Magnitude spectra of $x(t)$ (solid line) and $Tx(nT)$ with $T = 1.5$ (dotted line) for ω in $[-10, 10]$. (**aa**) Plot of (a) for ω in the Nyquist frequency range $[-2.1, 2.1)$. (**aaa**) Corresponding phase spectra. (**b**), (**bb**), and (**bbb**) Repetitions of (a), (aa), and (aaa) with $T = 0.5$.

In Figure 5.6(a) we plot the magnitude spectrum of $TX_d(\omega)$ for $T = 1.5$ with a dotted line and the magnitude spectrum of $X(\omega)$ with a solid line for ω in $[-10, 10]$. For $T = 1.5$, the spectrum $X_d(\omega)$ is periodic with period $2\pi/T = 4.2$ as shown. Thus we plot in Figure 5.6(aa) only the part of Figure 5.6(a) for ω in the Nyquist frequency range $(-\pi/T, \pi/T] = (-2.1, 2.1]$. We plot the corresponding phase spectra in Figure 5.6(aaa). We repeat the plots in Figures 5.6(b), 5.6(bb), and 5.6(bbb) using $T = 0.5$. Note that the Nyquist frequency range for $T = 0.5$ is $(-6.28, 6.28]$.

We now compare the spectra of $x(t)$ and $Tx(nT)$. Their spectra are always different outside the Nyquist frequency range as shown in Figure 5.6(a), thus we compare them only inside the range. For $T = 1.5$, their magnitude spectra (solid and dotted lines) shown in Figure 5.6(aa) differ appreciably. Thus $T = 1.5$ is not small enough. For $T = 0.5$, they are very close as shown in Figure 5.6(bb). If we select a smaller T, then the result will be even better. This verifies (5.28). For the phase spectra in Figure 5.6(aaa) for $T = 1.5$, the solid and dotted lines differ greatly except in the immediate neighborhood of $\omega = 0$. The situation does not improve by selecting a smaller T as shown in Figure 5.6(bbb). This verifies (5.29). Thus we cannot compute the phase spectrum of $x(t)$ from its sampled sequence.

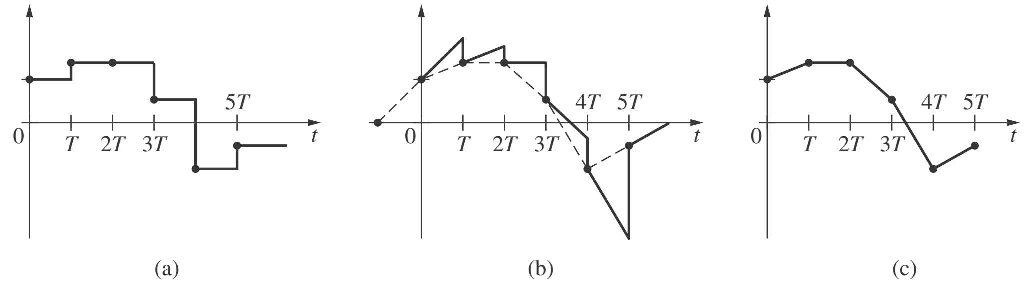


Figure 5.7 (**a**) Zero-order hold. (**b**) First-order hold. (**c**) Linear interpolation.

In conclusion, if frequency aliasing due to time sampling is negligible, the magnitude spectrum of $x(t)$ can be computed from its sampled sequence even if $x(t)$ is not band-limited. One way to check frequency aliasing is to compare the spectra of $x(t)$ and $Tx(nT)$ for all ω in $(-\pi/T, \pi/T] = (-\omega_s/2, \omega_s/2]$. If the spectrum of $x(t)$ is not known, the way to check frequency aliasing is to check the spectrum of $Tx(nT)$ in the neighborhood of $\pm\omega_s/2$. If the magnitude spectrum of $Tx(nT)$ in the neighborhood of $\pm\omega_s/2$ is significantly different from 0, frequency aliasing is large and we must select a smaller T. If the magnitude spectrum in the neighborhood of $\pm\omega_s/2$ is practically zero, then frequency aliasing is negligible and (5.28) holds. In general, the phase spectrum of $x(t)$ cannot be computed from $x(nT)$ no matter how small T is chosen.

5.3.2 Construction of CT Signals from DT Signals

If $x(t)$ is band-limited to $\omega_{\max}$ and if $T < \pi/\omega_{\max}$, then $x(t)$ can be recovered exactly, using (5.25), from its sampled sequence $x(nT)$. However, (5.25) is not used in practice for two reasons. First it requires infinitely many additions and multiplications. This is not feasible in practice. Second, we can start to compute $x(t)$ only after $x(nT)$, for all n, are received. Thus the computation cannot be carried out in real time. Thus (5.25) is more of theoretical interest. In practice, we use simpler schemes to recover $x(t)$ from $x(nT)$, as we discuss next.

Let $n_0T \le t < (n_0+1)T$. In recovering $x(t)$ from $x(nT)$, if we use only its current and past data ($x(nT)$ with $n \le n_0$) but not future data ($x(nT)$ with $n \ge n_0 + 1$), then $x(t)$ can appear as soon as we receive $x(n_0T)$ or immediately thereafter. We call this *real-time processing*. This real-time recovering is essential in digital transmission of telephone conversations and in music CDs.

The simplest way of constructing $x(t)$, for $nT < t \le (n+1)T$, from $x(nT)$ is to hold the value of $x(nT)$ until the arrival of the next value $x((n+1)T)$ as shown in Figure 5.7(a). It is called the *zero-order hold*. If we connect $x((n-1)T)$ and $x(nT)$ by a straight line and then extend it to yield the value of $x(t)$ for $nT < t \le (n+1)T$ as shown in Figure 5.7(b), then it is called the *first-order hold*.[5] Because of its simplicity, the zero-order hold is used as outputs of most DACs (digital-to-analog converters).

[5]The straight line can be described by a polynomial of t of degree 1.

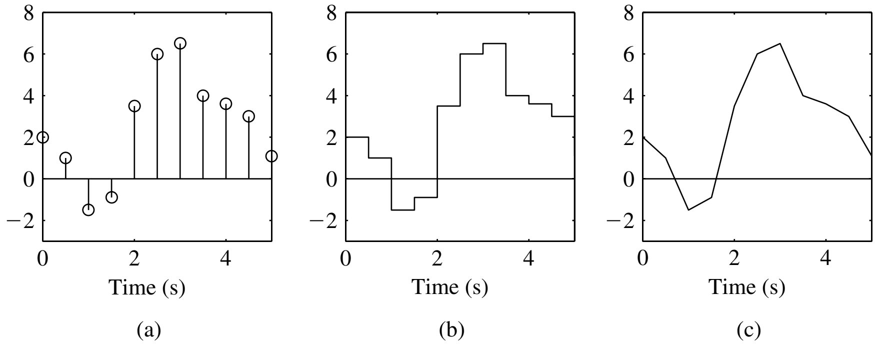


Figure 5.8 (**a**) DT signal. (**b**) CT signal using zero-order hold. (**c**) CT signal using linear interpolation.

The zero- and first-order holds use past data to carry out extrapolation and can be used in real time. We now discuss the MATLAB function plot. It connects two adjacent data by a straight line to carry out interpolation as shown in Figure 5.7(c). It is called the *linear interpolation*. Because it uses the future sample value, it cannot be used in real time. However, if the delay of one sample is permitted, then it can be used. The MATLAB program

```
t=0:0.5:5;
x=[2 1 -1.5 -0.9 3.5 6 6.5 4 3.6 3 1.1];
subplot(1,3,1)
stem(t,x),title('(a)')
subplot(1,3,2)
stairs(t,x),title('(b)')
subplot(1,3,3)
plot(t,x),title('(c)')
```

generates the plots in Figure 5.8. Figure 5.8(a) shows the original DT signal and is generated using the function stem. Figure 5.8(b) shows the constructed CT signal using a zero-order hold and is generated using stairs. Figure 5.8(c) shows the constructed CT signal using the linear interpolation and is generated using plot.

When we use a zero-order hold to reconstruct $x(t)$, errors always occur. Clearly, the smaller the sampling period, the smaller the errors, as shown in Figure 1.20. However, a smaller T requires more computation and more memory. Cost and performance specification also play a role. Thus a sampling period is selected by compromising among many conflicting factors. It is not determined by the sampling theorem alone. Furthermore, what is ω_{max} is not so clear cut for most practical signals. See Section 4.7.

To conclude this section, we mention that, on telephone lines, the sampling frequency has been selected as 8 kHz. Broadcasting companies are using digital recording with sampling frequency 32 kHz. Music compact discs that have the capability of error correcting (to remove scratches) use sampling frequency 44.1 kHz.

5.4 COMPUTING FREQUENCY SPECTRA OF DT SIGNALS

This section discusses the computation of frequency spectra of DT signals.[6] To simplify the discussion, we study only positive-time signals. Let $x(nT)$ be a positive-time sequence. Then its frequency spectrum is

$$X_d(\omega) = \sum_{n=0}^{\infty} x(nT)e^{-jn\omega T} \tag{5.30}$$

If $x(nT)$ lasts forever, the equation cannot be computed because it requires infinitely many operations. However, if $x(nT)$ is absolutely summable, then it must approach zero as $n \to \infty$. Thus there exists an N such that $x(nT) \approx 0$, for $n \geq N$, and we can approximate (5.30) by

$$X_d(\omega) = \sum_{n=0}^{N-1} x(nT)e^{-jn\omega T} \tag{5.31}$$

In other words, we use the first N data of $x(nT)$, for $n = 0 : N - 1$, to compute the spectrum of $x(nT)$. Clearly, the larger the value of N, the more accurate the result.

The spectrum in (5.31) is defined for all ω in $(-\pi/T, \pi/T]$. There are infinitely many ω in the range, and it is not possible to compute them all. We can compute $X_d(\omega)$ only at a finite number of discrete frequencies. If the number of frequencies equals the number of data, then we can develop a very efficient computational method. Thus we will select N equally spaced frequencies in $(-\pi/T, \pi/T]$ or $[-\pi/T, \pi/T)$. Because $X_d(\omega)$ is periodic with period $2\pi/T$, it is much simpler for indexing by selecting N equally spaced frequencies in $[0, 2\pi/T)$. Let us define

$$D := \frac{2\pi}{NT} \tag{5.32}$$

and

$$\omega_m := mD = \frac{2m\pi}{NT} \tag{5.33}$$

for $m = 0 : N - 1$. They are N equally spaced frequencies in $[0, 2\pi/T)$ with frequency increment D. We call D the *frequency resolution*. If we compute $X_d(\omega)$ at these ω_m, then (5.31) becomes

$$\begin{aligned} X_d(\omega_m) = X_d(mD) &= \sum_{n=0}^{N-1} x(nT)e^{-jn\omega_m T} \\ &= \sum_{n=0}^{N-1} x(nT)e^{-jn(2m\pi/NT)T} = \sum_{n=0}^{N-1} x(nT)e^{-j2nm\pi/N} \end{aligned} \tag{5.34}$$

for $m = 0 : N - 1$. This is the equation we will use to compute $X_d(\omega)$ at N discrete frequencies. Once $X_d(mD)$ are computed, we must interpolate their values to all ω in $[0, 2\pi/T)$.

---

[6]The remainder of this chapter may be skipped without loss of continuity.

5.4.1 Fast Fourier Transform (FFT)

This section discusses computer computation of (5.34). We first define $X_d[m] := X_d(mD)$, $x[n] = x(nT)$, and $W := e^{-j2\pi/N}$. Then (5.34) can be written as

$$X_d[m] = \sum_{n=0}^{N-1} x[n]W^{nm} \tag{5.35}$$

for $m = 0 : N-1$. In other words, given the time sequence $x[n] = x(nT)$, for $n = 0 : N-1$, the equation computes its frequency spectrum $X_d(\omega)$ at $\omega = mD = m(2\pi)/NT$, for $m = 0 : N-1$ in the frequency range $[0, 2\pi/T)$. Note that the sampling period T does not appear in (5.35). The computed data $X_d[m]$, however, are to be located at frequency $m(2\pi)/NT$.

Equation (5.35) has two integer indices n and m both ranging from 0 to $N-1$. The former is the time index, and the latter is the frequency index. The equation is called the *discrete Fourier transform* (DFT). It is actually the discrete-time Fourier series if we extend $x[n]$ periodically, with period N, to all n. It is possible to give a complete treatment of the DFT as in the CT Fourier series. See Reference 2. We will not do so. Instead, we discuss only one property that is needed in its computer computation.

Because $X_d[m]$ computes N samples of $X_d(\omega)$ in $[0, 2\pi/T)$ and because $X_d(\omega)$ is periodic with period $2\pi/T$, $X_d[m]$ is periodic with period N. By this, we mean

$$X_d[m] = X_d[m+N] \tag{5.36}$$

for all integer m. Indeed, because $W = e^{-j2\pi/N}$, we have

$$W^{nN} = e^{-j(2\pi/N)nN} = e^{-j2\pi n} = 1$$

and

$$W^{n(m+N)} = W^{nm}W^{nN} = W^{nm}$$

for any integers n and m. Thus we have

$$X_d[m+N] = \sum_{n=0}^{N-1} x[n]W^{n(m+N)} = \sum_{n=0}^{N-1} x[n]W^{nm} = X_d[m]$$

This establishes (5.36).

We now discuss the number of operations in computing (5.35). The summation in (5.35) has N terms, and each term is the product of $x[n]$ and W^{nm}. Thus for each m, (5.35) requires N multiplications and $N-1$ additions. To compute $X[m]$, for $m = 0 : N-1$, (5.35) requires a total of N^2 multiplications and $N(N-1)$ additions. These numbers increase rapidly as N increases. Thus the equation, for N large, is not computed directly on a computer.

Suppose N in (5.35) is divisible by 2, then we can decompose $x[n]$ into two subsequences $x_1[n]$ and $x_2[n]$ as

$$x_1[n] := x[2n], \quad x_2[n] := [2n+1] \qquad \text{for } n = 0, 1, \ldots, (N/2)-1 \tag{5.37}$$

where $x_1[n]$ and $x_2[n]$ consist, respectively, of the even and odd terms of $x[n]$. Clearly, both have length $N_1 := N/2$. Their DFT are, respectively,

$$X_{d1}[m] = \sum_{n=0}^{N_1-1} x_1[n]W_1^{mn} \tag{5.38}$$

and

$$X_{d2}[m] = \sum_{n=0}^{N_1-1} x_2[n]W_1^{mn} \tag{5.39}$$

with

$$W_1 := e^{-j2\pi/N_1} = e^{-j4\pi/N} = W^2 \tag{5.40}$$

for $m = 0, 1, \ldots, N_1 - 1$. Note that $X_{d1}[m]$ and $X_{d2}[m]$ are periodic with period $N_1 = N/2$ or

$$X_{di}[m] = X_{di}[m + N_1] \tag{5.41}$$

for any integer m and for $i = 1, 2$. Let us express $X_d[m]$ in terms of $X_{d1}[m]$ and $X_{d2}[m]$. We use $W_1 = W^2$ to write (5.35) explicitly as

$$\begin{aligned} X_d[m] &= x[0]W^0 + x[2]W^{2m} + \cdots + x[N-2]W^{(N-2)m} \\ &\quad + x[1]W^m + x[3]W^{3m} + \cdots + x[N-1]W^{(N-1)m} \\ &= x_1[0]W_1^0 + x_1[1]W_1^m + \cdots + x_1[N_1-1]W_1^{(N_1-1)m} \\ &\quad + W^m\left[x_2[0]W_1^0 + x_2[1]W_1^m + \cdots + x_2[N_1-1]W_1^{(N_1-1)m}\right] \end{aligned}$$

which becomes, after substituting (5.38) and (5.39),

$$X_d[m] = X_{d1}[m] + W^m X_{d2}[m] \tag{5.42}$$

for $m = 0, 1, \ldots, N-1$. Thus it is possible to compute $X_d[m]$ using (5.38), (5.39), and (5.42).

Let us compute the number of multiplications needed in (5.38), (5.39), and (5.42). Each $X_{di}[m]$ in (5.38) and (5.39) requires $N_1^2 = (N/2)^2$ multiplications. In addition, we need N multiplications in (5.42) to generate $X_d[m]$. Thus the total number of multiplications in computing $X_d[m]$ from (5.38), (5.39), and (5.42) is

$$\left(\frac{N}{2}\right)^2 + \left(\frac{N}{2}\right)^2 + N = N + \frac{N^2}{2} \approx \frac{N^2}{2}$$

for N large. Likewise, the total number of additions needed in (5.38), (5.39), and (5.42) is

$$\frac{N}{2}\left(\frac{N}{2} - 1\right) + \frac{N}{2}\left(\frac{N}{2} - 1\right) + N = \frac{N^2}{2}$$

Computing $X_d[m]$ directly requires N^2 multiplications and $N(N-1)$ additions. If we decompose $x[n]$ into two subsequences of length $N/2$ and compute $X_d[m]$ using (5.42), then the amount of computation, for N large, is reduced by almost a factor of 2.

Now if N is not only even but also a power of 2, say, $N = 2^k$, then each subsequence $x_i[n]$ has an even length and can be divided into two subsequences, and the amount of computation can again be cut in half. We can repeat the process for each subsequence. If $N = 2^k$, then the process can be repeated $\log_2 N = k$ times and the total numbers of multiplications and additions in computing (5.35) can both be reduced to roughly $N \log_2 N$. This process of computing (5.35) is called the *fast Fourier transform* (FFT). For large N, the saving is very significant. For example, for $N = 2^{10} = 1024$, direct computation requires roughly 10^7 multiplications and additions, but the FFT requires only roughly 5120 multiplications and 10240 additions, a saving of over 100 times. It means that a program that takes 2 minutes in direct computation will take only one second by using FFT.

5.5 FFT SPECTRAL COMPUTATION OF DT SIGNALS

This section uses the FFT function in MATLAB to compute frequency spectra of DT signals. In MATLAB, if x is an $\bar{N}$-point sequence, the function fft(x) or fft(x,$\bar{N}$) generates $\bar{N}$ data of (5.35), for $m = 0 : \bar{N} - 1$. If $N \leq \bar{N}$, the function fft(x,N) uses the first N data of x to compute N data of (5.35) for $m = 0 : N - 1$. If $\bar{N} < N$, the function fft(x,N) automatically pads trailing zeros to x and then computes (5.35) for $m = 0 : N - 1$. In using fft(x,N), N can be selected as any positive integer such as 3 or 1111. However, if N is a power of 2, then the computation will be most efficient. We use the sequence in Example 5.2.2 to discuss the issues involved in using FFT.

Consider the sequence $x[0] = 2$, $x[1] = 3$, $x[2] = 2$, $x[3] = -3$ with sampling period $T = 0.3$. Its frequency spectrum was computed in Example 5.2.2, and its magnitude and phase spectra were plotted in Figure 5.2. We now discuss the use of FFT. We express the time sequence as a row vector: x=[2 3 2 -3]. Typing fft(x) in MATLAB yields

$$4 \quad -6i \quad 4 \quad 6i$$

They are four values of $X_d[m]$, for $m = 0 : 3$. Note that i and j both denote $\sqrt{-1}$ in MATLAB. Typing X=fft(x) yields

$$X = [4 \quad -6i \quad 4 \quad 6i]$$

The output has been named X. Indices of all variables in MATLAB start from 1 and are enclosed by parentheses. Thus typing X(1) yields 4. Typing X(4) yields $6i$. Typing X(0) yields an error message. Thus we have

$$X_d[m] = X(m + 1) \tag{5.43}$$

for $m = 0 : 3$. Note that if the time sequence is expressed as a column vector, then the FFT will yield a column vector.

In using FFT, the sampling period T does not appear in the equation and the result holds for any T. However, if the time sequence has sampling period T, then the output $X_d[m] = X(m+1)$ are to be located at

$$mD := m\frac{2\pi}{NT}$$

for $m = 0 : N - 1$. For our example, we have $N = 4$ and $T = 0.3$. Thus the four data of X are to be located at $mD = 2m\pi/(4 \times 0.3) = 5.24m$, $m = 0 : 3$, or at

$$\omega = 0, 5.24, 10.48, 15.72$$

Typing

```
Program 5.1
x=[2 3 2 -3]; N=4; T=0.3;
X=fft(x,N);
m=0:N-1;D=2*pi/(N*T);
subplot(1,2,1)
stem(m*D,abs(X),'fill'),title('(a)')
subplot(1,2,2)
stem(m*D,phase(X),'fill'),title('(aa)')
```

in MATLAB yields the solid dots in Figures 5.9(a) (magnitudes) and 5.9(aa) (phases). We plot there with dotted lines also the complete magnitude and phase spectra of the 4-point sequence. The dots are indeed the samples of the spectrum $X_d(\omega)$ as shown. They are equally spaced in the frequency range $[0, 2\pi/T) = [0, 20.9)$. Note that the program shows only key functions. Functions such as sizing the plot and drawing the horizontal coordinate are not shown. The subprogram that generates the dotted lines (see Example 5.2.2) is neither shown.

If a DT signal is real valued as is always the case in practice, then its magnitude spectrum is even and its phase spectrum is odd. Thus we often plot them only in the positive frequency range $[0, \pi/T]$. For N even, if $m_p = 0 : N/2$, then $m_p D$ ranges from 0 to $(N/2)(2\pi/NT) = \pi/T$. The set m_p ranges from 0 to $N/2$, and thus we must use the corresponding entries of X. As mentioned earlier, the index of X starts from 1. Thus the entries corresponding to m_p are X(mp+1). Let us modify Program 5.1 as

```
Program 5.2
x=[2 3 2 -3];N=4;T=0.3;
X=fft(x,N);
mp=0:N/2;D=2*pi/(N*T);
subplot(1,2,1)
stem(mp*D,abs(X(mp+1)),'fill'),title('(a)')
subplot(1,2,2)
stem(m*D,phase(X(mp+1)),'fill'),title('(aa)')
```

Then the program yields the solid dots in Figures 5.9(b) (magnitudes) and (bb) (phases). They are the samples of the spectrum $X_d(\omega)$ in $[0, \pi/T] = [0, 10.45]$ as shown. Note that if we type X(mp), instead of X(mp+1), in Program 5.2, then an error message will appear.

To conclude this section, we mention that if N is odd, then mp=0:N/2 in Program 5.2 must be changed to mp=0:(N-1)/2 or mp=0:floor(N/2), where floor rounds a number downward to an integer. No other changes are needed.

5.5.1 Interpolation and Frequency Resolution

As shown in Figure 5.2, the spectrum of the 4-point sequence $x = [2\ \ 3\ \ 2\ \ -3]$ is a continuous function of ω. When we use 4-point FFT to compute its 4-point samples, we must carry out

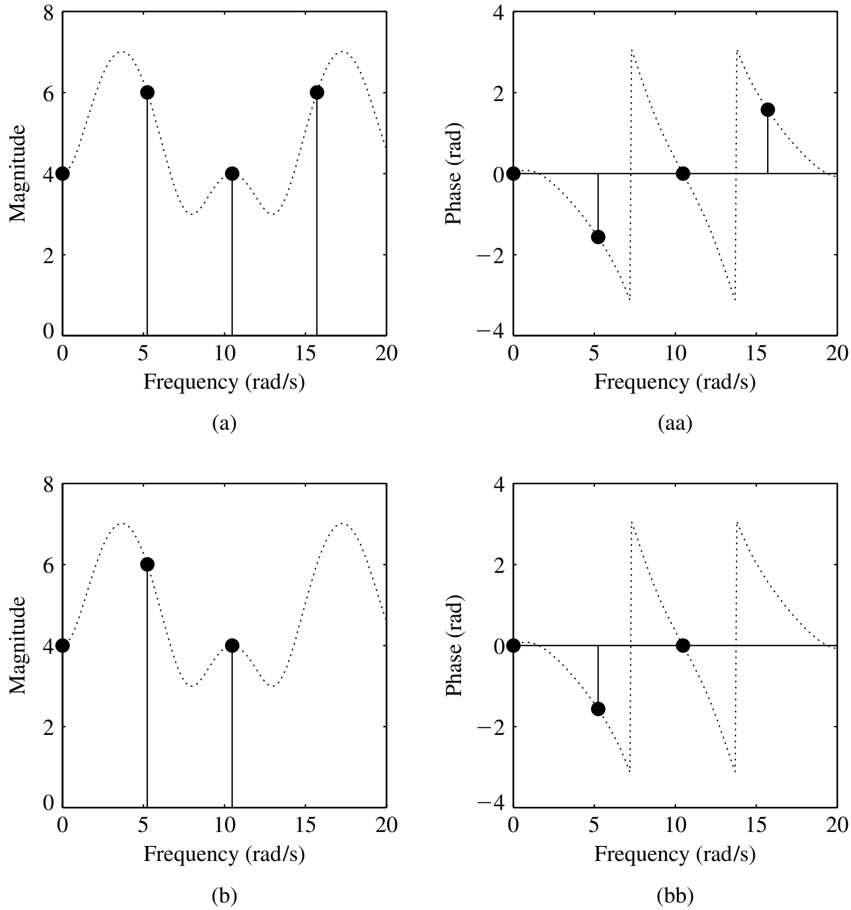


Figure 5.9 (**a**) Magnitude spectrum of $x = [2\ \ 3\ \ 2\ \ -3]$ (dotted line) and its four samples in [0, 20.9) (solid dots). (**aa**) Its phase spectrum (dotted line) and its four samples (solid dots). (**b**) Magnitude spectrum of x (dotted line) and its three samples in [0, 10.45]. (**bb**) Its phase spectrum and its four samples in [0, 10.45].

an interpolation to all ω. A simple interpolation scheme is to connect the computed samples by straight lines. This can be achieved by replacing the function stem in Programs 5.1 and 5.2 by plot. The MATLAB function plot connects two adjacent points by a straight line. This is the linear interpolation discussed in Section 5.3.2. The program that follows carries out the interpolation.

```
Program 5.3
x=[2 3 2 -3];N=4;T=0.3;
X=fft(x,N);
m=0:N-1;D=2*pi/(N*T);
plot(m*D,abs(X),m*D,angle(X),':')
```

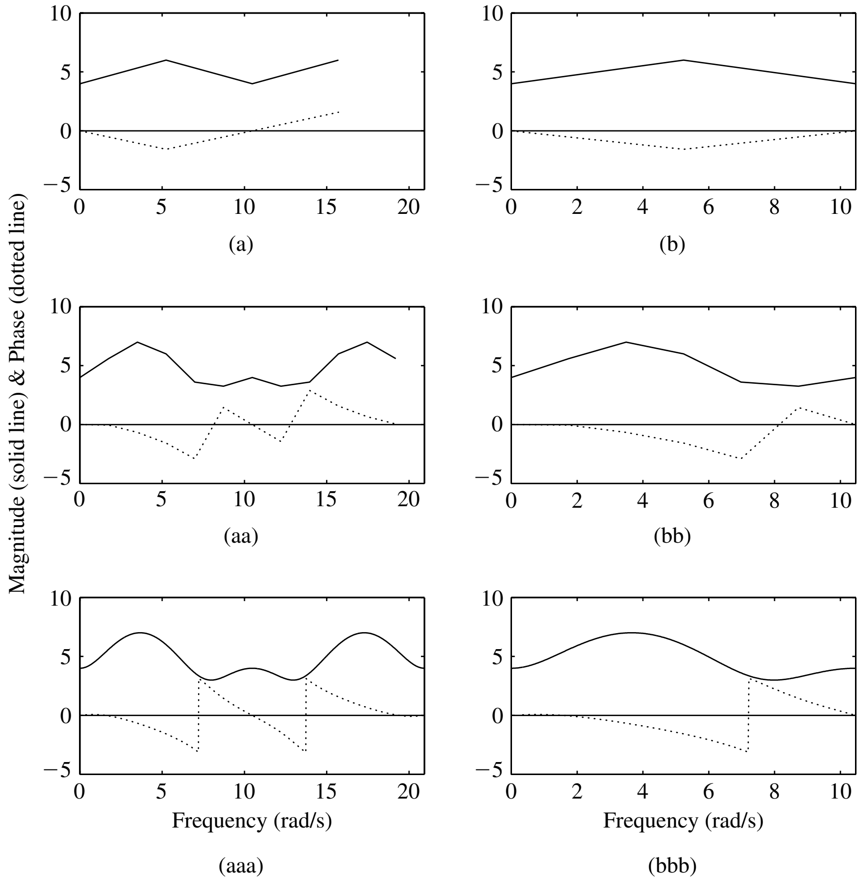


Figure 5.10 (**a**) Spectrum of $x = [2\ \ 3\ \ 2\ \ -3]$ in $[0, 2\pi/T) = [0, 20.94)$ computed using 4-point FFT and linear interpolation. (**aa**) Using 12-point FFT and linear interpolation. (**aaa**) Using 1024-point FFT and linear interpolation. (**b**), (**bb**), and (**bbb**) Corresponding plots in $[0, \pi/T] = [0, 10.47]$.

The interpolated magnitude and phase spectra are plotted in Figure 5.10(a) with solid and dotted lines, respectively. The interpolated spectra are quite different from the actual spectra.

The frequency resolution in Figure 5.10(a) is $2\pi/NT = 2\pi/(0.3 \times 4) = 5.24$. It is very large or very poor. For the given x=[2 3 2 -3] with sampling period $T = 0.3$, the only way to improve the resolution is to increase N. It is clear that the spectrum of x=[2 3 2 -3] is identical to the spectrum of xe=[2 3 2 -3 0 0 ...]. Thus by padding trailing zeros to x, we can improve the resolution of the computed spectrum. Recall that the function fft(x,N) automatically pads trailing zeros to x if the length of x is less than N. Now if we change N in Program 5.3 from 4 to 12, then the program yields the plot in Figure 5.10(aa). The generated spectrum is closer to the actual spectrum. If we select $N = 1024$, then the program yields the plot in Figure 5.10(aaa).

It is indistinguishable, in the Nyquist frequency range $[0, 2\pi/T)$, from the plot using the exact formula in (5.10). Note that for N small, the frequency closest to $2\pi/T$ is still quite far from $2\pi/T$ as shown in Figure 5.10(a). Thus the use of $[0, 2\pi/T)$ is essential. For N large, we cannot tell visually the difference in using $[0, 2\pi/T]$ and $[0, 2\pi/T)$.

If we modify Program 5.3 as

```
Program 5.4
x=[2 3 2 -3];N=4;T=0.3;
X=fft(x,N);
mp=0:N/2;D=2*pi/(N*T);
plot(mp*D,abs(X(mp+1)),mp*D,angle(X(mp+1)),':')
```

then it generates the spectrum in the positive Nyquist frequency range $[0, \pi/T]$. Figures 5.10(b), 5.10(bb), and 5.10(bbb) show the results for $N = 4$, 12, and 1024. We see that the one using $N = 1024$ is indistinguishable from the exact one in $[0, \pi/T]$. From now on, we plot mostly spectra in the positive Nyquist frequency range.

As a final example, we discuss FFT computation of the spectrum of the positive-time sequence $x(nT) = e^{-0.1nT}$ for $T = 0.5$ and all $n \geq 0$. This is an infinite sequence and we must truncate it before carry out its computation. Let us use N data, that is, use $x(nT)$ for $n = 0 : N - 1$. The integer N should be selected so that the magnitude of $x(nT)$ is practically zero for all $n \geq N$. Arbitrarily, we select $N = 1024$. Then the program

```
Program 5.5
N=1024;T=0.5;D=2*pi/(N*T);
n=0:N-1;x=exp(-0.1*n*T);
X=fft(x);
mp=0:N/2;
plot(mp*D,abs(X(mp+1)),mp*D,angle(X(mp+1)))
```

generates the magnitude (upper solid line) and phase (lower solid line) spectra in Figure 5.11(a) in the positive Nyquist frequency range $[0, \pi/T = 6.28]$. We superpose on them with dotted lines the spectra obtained from its exact spectrum $X_d(\omega) = 1/(1 - e^{-0.1T}e^{-j\omega T})$. In order to see better, we plot in Figure 5.11(b) only the frequency range in [0, 1]. This is obtained by adding axis([0 1 -5 25]) at the end of Program 5.5. Indeed the computed spectrum is indistinguishable from the exact one.

5.5.2 Plotting Spectra in $[-\pi/T, \pi/T)$

FFT generates the spectrum of a DT signal in $[0, 2\pi/T)$. We discuss how to transform it into the frequency range $[-\pi/T, \pi/T)$. As mentioned earlier, the Nyquist frequency range can be selected as $(-\pi/T, \pi/T]$ or $[-\pi/T, \pi/T)$. We select the latter in order to use the MATLAB function fftshift. We first discuss the range of m so that mD lie inside the Nyquist frequency range. For the example in Figure 5.9 with $T = 0.3$ and $N = 4$, the Nyquist frequency range is $[-\pi/T, \pi/T) = [-10.47, 10.47)$ and the frequency resolution is $D = 2\pi/NT = 5.235$. If

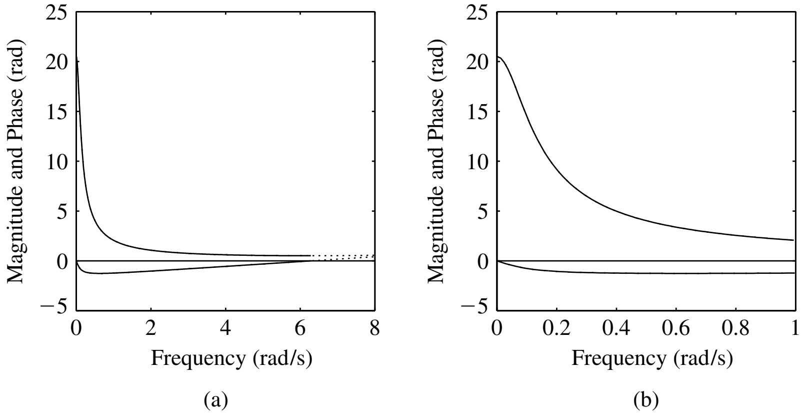


Figure 5.11 (**a**) Exact and computed spectra using 1024-point FFT and linear interpolation. (**b**) Zoom of (a).

we select m as $-2 : 1$, then the four frequencies mD are -10.47, -5.235, 0, 5.235. They lie inside the Nyquist frequency range $[-10.47, 10.47)$ and are equally spaced. In general, if N is even and if

$$m = -N/2 : N/2 - 1 \tag{5.44}$$

then mD lie inside $[-\pi/T, \pi/T)$. Note that there are N different m in (5.44).

The function fft(x) yields $X_d[m]$, for $m = 0 : N - 1$. Because $X_d[m]$ is, as shown in (5.36), periodic with period N, the values of $X_d[m]$ for m in (5.44) can be obtained by periodic extension. This can be achieved by shifting the values of $X_d[m]$ for $m = N/2 : N - 1$ to $m = -N/2 : -1$ or by shifting the second half of $X_d[m]$ to the first half. The MATLAB function fftshift carries out this shifting. For example, for $N = 4$, typing

```
X=[4 -6i 4 6i]
Xs=fftshift(X)
```

yields

$$Xs = [4 \;\; 6i \;\; 4 \;\; -6i]$$

which are the values of $X_d[m]$ to be located at $\omega = mD$ for $m = -2 : 1$. Thus if we modify Program 5.3 as

```
x=[2 3 2 -3];T=0.3;N=4;
Xs=fftshift(fft(x,N));
ms=-N/2:N/2-1;D=2*pi/(N*T);
plot(ms*D,abs(Xs),ms*D,angle(Xs),':')
```

then it generates the magnitude (solid line) and phase (dotted line) spectra in the Nyquist frequency range $[-\pi/T, \pi/T)$ in Figure 5.12(a). The result is poor because the program uses 4-point FFT. If we use $N = 1024$, then the result is as shown in Figure 5.12(b), which is

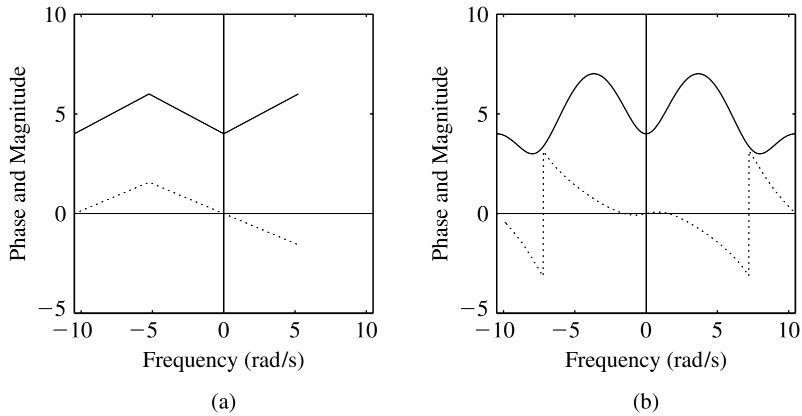


Figure 5.12 (**a**) Magnitude and phase spectra in $[-\pi/T, \pi/T)$, using 4-point FFT and linear interpolation. (**b**) Using 1024-point FFT and linear interpolation.

indistinguishable from the one obtained using (5.10). Because the resolution is so small, we cannot tell the difference between $[-\pi/T, \pi/T)$ and $[-\pi/T, \pi/T]$.

5.6 FFT SPECTRAL COMPUTATION OF CT SIGNALS

This section discusses FFT computation of frequency spectra of CT positive-time signals. As discussed in (5.29) and demonstrated in Example 5.3.3, we cannot computed the phase spectrum of $x(t)$ from its sample sequence $x(nT)$. Thus we compute only the magnitude spectrum. Furthermore, we plot only the positive frequency part of (5.28), that is,

$$|X(\omega)| \approx \begin{cases} T|X_d(\omega)| & \text{for } 0 \leq \omega \leq \pi/T = \omega_{s/2} \\ 0 & \text{for } \omega > \pi/T \end{cases} \tag{5.45}$$

where $X(\omega)$ is the spectrum of $x(t)$ and $X_d(\omega)$ is the spectrum of $x(nT)$. We use examples to illustrate the procedure and the issues involved.

EXAMPLE 5.6.1

Consider the CT signal

$$x(t) = 0.5t^2 + (\cos \omega_0 t - 1)/\omega_0{}^2 \qquad \text{for } 0 \leq t \leq 2\pi/\omega_0 \tag{5.46}$$

with $\omega_0 = 0.1$, and $x(t) = 0$ outside the range. This signal will be used in Section 8.5 to model the distance traveled by an object. The signal has a finite time duration $L = 2\pi/\omega_0 = 62.8$; thus it is, as discussed in Section 4.7, not frequency band-limited. Consequently, frequency aliasing always occurs when we use its sampled sequence to compute its spectrum.

The signal $x(t)$ is bounded and of finite duration. Thus it is absolutely integrable. It also has a finite total energy. Consequently, its spectrum is bounded and continuous, and it approaches zero as $|\omega| \to \infty$. Thus if the sampling period T is selected to be sufficiently small, we can compute its magnitude spectrum from its sampled sequence. Clearly, the smaller the value of T, the lesser the frequency aliasing, and the more accurate the computed magnitude spectrum.

We select arbitrarily $N = 2^{13} = 8192$ equally spaced samples of $x(t)$ in $[0, 62.8)$. Then the sampling period T is

$$T = \frac{62.8}{8192} = 0.00076$$

The program

```
Program 5.6
wo=0.1;L=2*pi/wo;N=8192;T=L/N;D=2*pi/(N*T);
n=0:N-1;t=n*T;
x=0.5*t.*t+(cos(wo*t)-1)./(wo.*wo);
X=T*fft(x);
mp=0:N/2;
plot(mp*D,abs(X(mp+1)))
```

generates the magnitude spectrum in Figure 5.13(a). For this T, the positive Nyquist frequency range is $[0, \pi/T] = [0, 409.6]$. The generated magnitude spectrum is practically zero in the neighborhood of π/T; thus frequency aliasing is negligible, and the computed magnitude spectrum should be closed to the actual spectrum of $x(t)$. Figure 5.13(b) shows the part of Figure 5.13(a) in the frequency range [0, 10]. It is obtained by adding the function axis([0 10 0 4*10ˆ4]) at the end of Program 5.6.

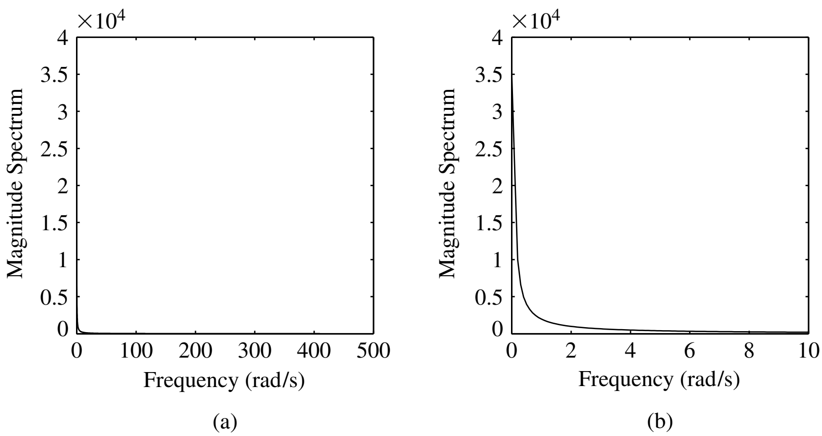


Figure 5.13 (**a**) Magnitude spectrum in $[0, \pi/T]$. (**b**) Zoom of (a) in [0, 10].

EXAMPLE 5.6.2

Consider again the signal in (5.46) with $\omega_0 = 5$. If we change `wo=0.1` to `wo=5`, then Program 5.6 yields the solid line in Figure 5.14(a) for the positive Nyquist frequency range $[0, \pi/T = 20490]$. Figure 5.14(aa) shows the part of Figure 5.14(a) in the frequency range [0, 10]. Clearly, frequency aliasing is very small and the computed result should be close to the actual spectrum. Recall that FFT computes only the spectrum at discrete frequencies, and the plots in Figures 5.14(a) and 5.14(aa) are obtained by linear interpolation. Is the result good? To answer this, we compute its frequency resolution. The time duration of the signal in (5.46) with $\omega_0 = 5$ is $L = 2\pi/\omega_0$. If we take N time samples, then the sampling period is $T = L/N = 2\pi/\omega_0 N$. Thus

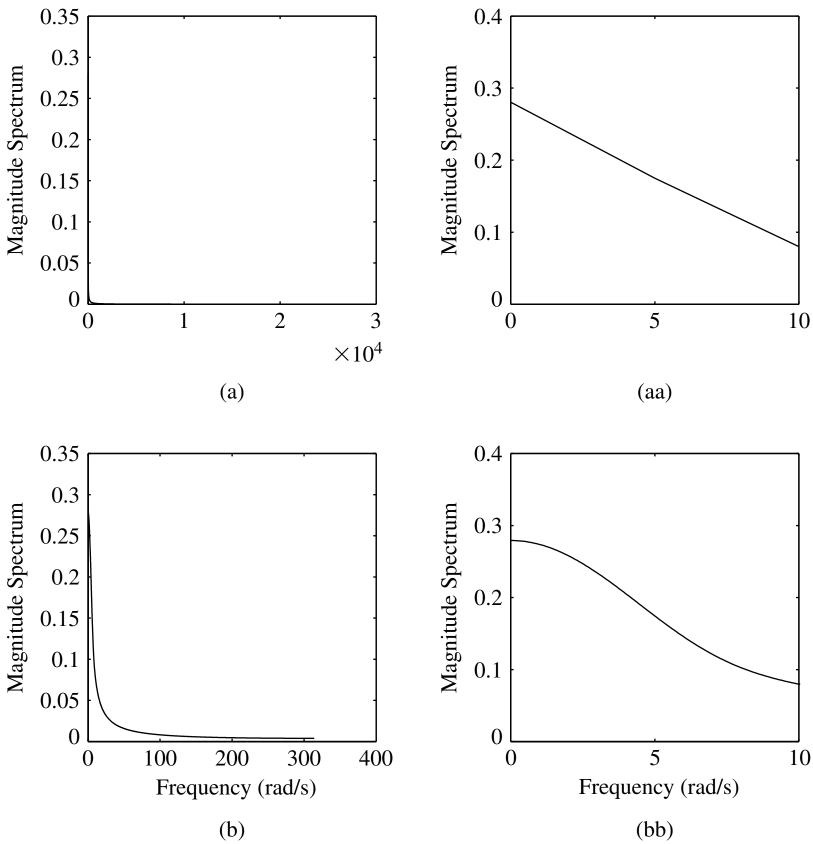


Figure 5.14 (**a**) Magnitude spectrum in $[0, \pi/T]$ using $N = 8192$ and $T = 2\pi/5N = 0.00015$. (**aa**) Zoom of (a) in [0, 10]. (**b**) Magnitude spectrum in $[0, \pi/T]$ using $T = 0.01$ and $N = 8192$ by padding trailing zeros. (**bb**) Zoom of (b) in [0, 10].

the frequency resolution is

$$D = \frac{2\pi}{NT} = \frac{2\pi}{N \cdot (2\pi/\omega_o N)} = \omega_o = 5$$

It means that in the frequency range [0, 10] in Figure 5.14(aa), the program computes only 3 points at $\omega = 0$, 5, and 10, and then it connects them by straight lines. Thus the result is poor. Note that the frequency resolution in Example 5.6.1 is $D = \omega_0 = 0.1$ and Program 5.6 computes 101 points in Figure 5.13(b), and the result is good.

The frequency resolution is $D = 2\pi/(NT)$; thus it can be improved by increasing T or N. For example, if we use the original $T = 2\pi/(0.5 \times 8192) = 0.0015$ and increase N to, for example, 8192×124 by padding trailing zeros to $x(t)$, then the frequency resolution will be small and the result will be good.

However, MATLAB 5.3 Student Version is limited to 16284, thus we cannot use $N = 8192 \times 124$. If we use $N = 8192$, the only way to improve D is to use a larger T. From Figure 5.14(a), we see that the signal is practically band-limited to roughly $W = 100$. Thus we can select $T < \pi/W = 0.0314$. Arbitrarily we select $T = 0.01$. The time duration of $x(t)$ in (5.46) with $\omega_0 = 5$ is $L = 2\pi/5 = 1.256$. If $T = 0.01$, the number of time samples in [0, 1.256] is 126. If we use $N = 126$, then the frequency resolution will be poor. Instead, we use $N = 2^{13} = 8192$. Recall that fft(x,N) automatically pads trailing zeros to x if the length of x is less than N. We type

```
Program 5.7
wo=5;L=2*pi/wo;N=8192;T=0.01;
t=0:0.01:L;
x=0.5*t.*t+(cos(wo*t)-1)./(wo*wo);
X=T*fft(x,N);
mp=0:N/2;D=2*pi/(N*T);
plot(mp*D,abs(X(mp+1)))
```

The program generates the magnitude spectrum of (5.46) with $\omega_0 = 5$ in the positive Nyquist frequency range $[0, \pi/T] = [0, 314]$ in Figure 5.14(b). The spectrum in the neighborhood of π/T is small, thus the frequency aliasing is negligible. Therefore the result should be close to the exact magnitude spectrum. We plot in Figure 5.14(bb) the part of Figure 5.14(b) in the frequency range [0, 10]. It is obtained by adding axis([0 10 0 0.4]) at the end of Program 5.7.

EXAMPLE 5.6.3

Compute the spectra of $x_1(t) = 2e^{-0.3t} \sin t$ and $x_2(t) = -x_1(t) \cos 20t$, for $t \geq 0$. The signals will be used in Section 8.5 to model earthquakes.

We compute first the spectrum of $x_1(t)$. We select arbitrarily $T = 0.1$ and $N = 8192$. Then we have $L = NT = 819.2$. Because $|x(t)| \leq 2e^{-0.3t} \leq 3.7 \times 10^{-107}$, for $t > L$,

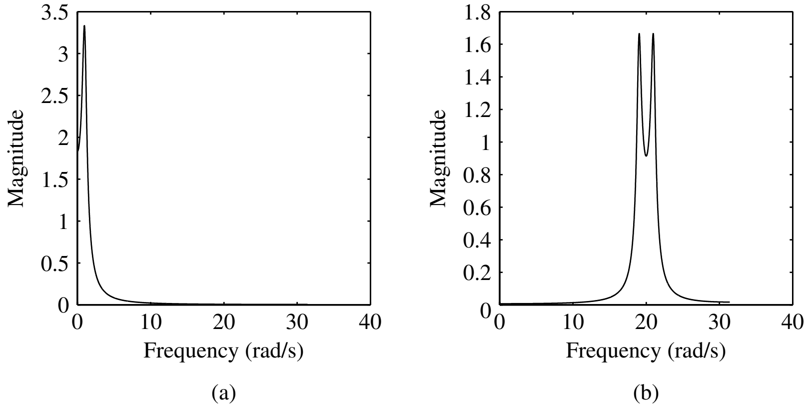


Figure 5.15 FFT computed magnitude spectra of (**a**) $x_1(t)$ and (**b**) $x_2(t)$.

the record length is sufficiently large. The program

```
Program 5.8
T=0.1;N=8192;D=2*pi/(N*T);
n=0:N-1;t=n*T;
x=2*exp(-0.3*t).*sin(t);
X=T*fft(x);
mp=0:N/2;w=mp*D;
plot(w,abs(X(mp+1)))
```

generates the magnitude spectrum in Figure 5.15(a). The spectrum is practically band-limited to, roughly, 10 rad/s. The sampling period $T = 0.1$ is smaller than $\pi/10 = 0.31$, thus the computed spectrum has negligible frequency aliasing and should be close to the actual magnitude spectrum.

The signal $x_2(t)$ is the modulation of $-x_1(t)$ with carrier frequency 20 rad/s. Thus the spectrum of $x_2(t)$ should be, as derived in (4.41), the shifting of the spectrum of $-x_1(t)$ to ± 20 and with magnitude cut in half. Because $x_1(t)$ is roughly band-limited to 10 rad/s, $x_2(t)$ is roughly band-limited to 30 rad/s. Thus the sampling period for computing the spectrum of $x_2(t)$ should be less than $\pi/30 = 0.105$. Clearly, $T = 0.1$ in Program 5.8 can still be used in this computation. If we replace x in Program 5.8 by

```
x=-2*exp(-0.3*t).*sin(t).*cos(20*t)
```

then the program yields the plot in Figure 5.15(b). Its magnitude spectrum is indeed the shifting of the one in Figure 5.15(a) to 20 rad/s and with magnitude cut in half.

From the preceding examples, we see that FFT computation of frequency spectra of CT signals may involve three issues as we discuss next.

1. *Frequency aliasing:* If $x(t)$ is not band-limited as is often the case in practice, the sampling period T must be selected so that the effect of frequency aliasing is negligible. A

necessary condition to have negligible frequency aliasing is that the computed spectrum in the neighborhood of π/T is practically zero.

2. *Frequency resolution:* The frequency spectrum of a CT signal is defined for all ω, whereas FFT computes the spectrum at only a finite number of frequencies. Thus the result must be interpolated. The smaller the frequency resolution $D = 2\pi/(NT) = 2\pi/L$, the better the interpolated result. If the data length L is small, we may pad trailing zeros to improve the resolution.
3. *Effect of truncation:* If a positive-time signal is of infinite duration, then it must be truncated, as in Example 5.6.3, in computer computation. Truncation will introduce in the spectrum, as discussed in Section 4.6, leakage, ripples, and Gibbs phenomenon. However, if $x(t)$ is absolutely integrable as is often the case, its spectrum is bounded and continuous. Thus Gibbs phenomenon will not appear. As the data length L increases, leakage and ripples will disappear and the computed magnitude spectrum will approach the exact one.

In spite of the preceding three issues, if the number of data N used in a computer and in a software can be arbitrarily large, computing the magnitude spectrum of $x(t)$ from its sampled sequence should be simple.

5.6.1 Selecting *T* and *N*

As shown in the preceding section, computing the magnitude spectrum of a CT positive-time signal from its time samples is simple and straightforward. We select a very small T and a very large N and then carry out computation. Clearly, the smaller the value of T and the larger the value of N, the more accurate the computed magnitude spectrum. However, it also requires more computation and more memory, and thus it is more costly. In practice, we often can obtain a good result without using unnecessarily small T and large N, as we demonstrate in the next example.

EXAMPLE 5.6.4

We compute the magnitude spectrum of $x(t) = e^{-0.1t} \sin 10t$, for $t \geq 0$. The signal has infinite length and must be truncated in computer computation. Arbitrarily, we select $T = 0.01$ and $N = 4096$. Then the program

```
Program 5.9
T=0.01;N=4096;D=2*pi/(N*T);
n=0:N-1;t=n*T;
x=exp(-0.1*t).*sin(10*t);
X=T*fft(x);
mp=0:N/2;
subplot(1,2,1)
plot(mp*D,abs(X(mp+1))),title('(a)')
subplot(1,2,2)
plot(mp*D,abs(X(mp+1))),title('(b)')
axis([0 30 0 5])
```

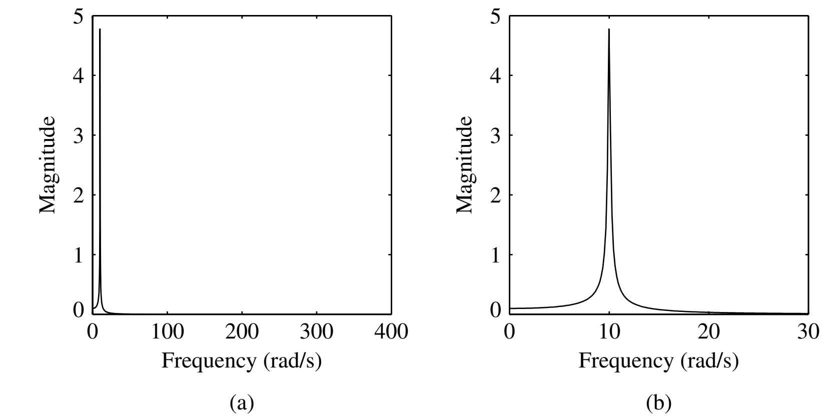


Figure 5.16 (**a**) Magnitude spectrum of $x(t)$ computed using $T = 0.01$ and $N = 4096$ for ω in the positive Nyquist frequency range [0, 314]. (**b**) Part of (a) for ω in [0, 30].

generates the magnitude spectrum in Figure 5.16(a) for ω in the positive Nyquist frequency range $[0, \pi/T] = [0, 314]$, and in Figure 5.16(b) for ω in the frequency range [0, 30]. We see that the nonzero frequency spectrum is centered around 10 rad/s and the signal is practically band-limited to 30 rad/s. Thus the computed magnitude spectrum should be very close to the exact one.

If the signal $x(t) = e^{-0.1t} \sin 10t$ is practically band-limited to 30, the sampling period can be selected as $T \leq \pi/30 = 0.105$. Thus $T = 0.01$ used in Program 5.9 is unnecessarily small. Let us compare its result with one using a larger T and a smaller N. We repeat Program 5.9 using $T = 0.1$ and $N = 512$. The result is plotted in Figure 5.17(a) with a solid line for the frequency range [0, 31.4] and in Figure 5.17(b) for the frequency range [9, 11]. We superpose there with dotted lines the result using $T = 0.01$ and $N = 4096$. Because the two plots are quite different in the neighborhood of $\omega = 10$, we conclude that the selected $T = 0.1$ and $N = 512$ may not be appropriate.

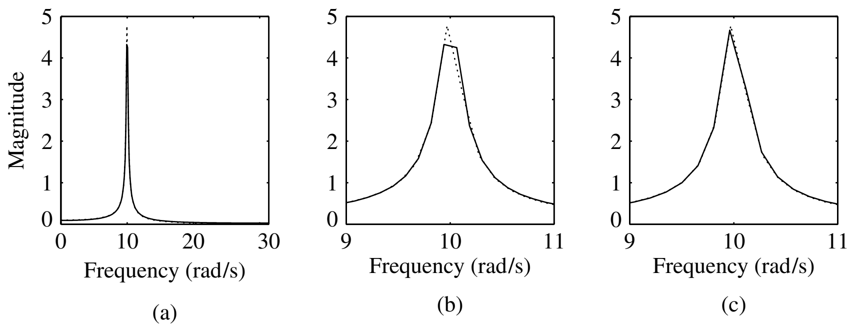


Figure 5.17 (**a**) Magnitude spectra of $x(t)$ using $T = 0.1$ and $N = 512$ (solid line) and using $T = 0.01$ and $N = 4096$ (dotted line) for ω in [0, 30]. (**b**) Part of (a) for ω in [9, 11]. (**c**) Magnitude spectra of $x(t)$ using $T = 0.1$ and $N = 410$ (solid line) and using $T = 0.01$ and 4096 for ω in [9, 11].

Before selecting a smaller T and a larger N, let us discuss the differences of the plots in Figure 5.17(b). For $T = 0.01$ and $N = 4096$, the frequency resolution is

$$D_1 = \frac{2\pi}{NT} = \frac{2\pi}{40.96} = 0.1534$$

Let us compute $10/D_1 = 65.12$. Thus in the neighborhood of $\omega = 10$, we compute the spectrum at frequencies $65D_1 = 9.97$, $66D_1 = 10.12$, and $67D_1 = 10.28$. For $T = 0.1$ and $N = 512$, we have

$$D_2 = \frac{2\pi}{0.1 \times 512} = 0.1227$$

and $10/D_2 = 81.5$. Thus in the neighborhood of $\omega = 10$, we compute the spectrum at frequencies $81D_2 = 9.94$, $82D_2 = 10.06$, and $83D_2 = 10.18$. Therefore we compare in Figure 5.17(b) the magnitude spectra at different frequencies; consequently, the comparison is not exactly fair.

Now we use the same $T = 0.1$ but select a different N so that its frequency resolution is close to D_1. Equating $2\pi/0.1N = D_1$ yields $N = 409.59$. We select $N = 410$.[7] Program 5.9 using $T = 0.1$ and $N = 410$ yields the solid line in Figure 5.17(c) in the frequency range [9, 11]. It is very close to the dotted line using $T = 0.01$ and $N = 4096$. In other words, the result using $T = 0.1$ and $N = 410$ is comparable to the one using $T = 0.01$ and $N = 4096$, and yet it requires only about one-tenth of computation.

We showed in Example 5.6.4 that it is possible to compute the spectrum of a CT signal without using an unnecessarily small T and an unnecessarily large N. This subsection discusses a procedure of selecting such T and N. Of course, the selection is subjective and is not unique.

Consider a CT positive-time signal $x(t)$. We first select an L so that $|x(t)|$ is practically zero or less than $0.01x_{max}$ for $t \geq L$, where x_{max} is the peak magnitude of $x(t)$. We then select a very large N that is a power of 2. We compute the spectrum of $x(t)$ using its sampled sequence $x(nT)$ with $T = L/N$. From the computed spectrum, we can find a W such that the signal is roughly band-limited to W. If no such W exists, we must select a larger N. Once W is found, we then select T_a as π/W. This is a largest acceptable sampling period.

Alternatively, instead of starting with a very large N, we select a fairly large T_1 and a smallest possible N_1 that is a power of 2 and $T_1N_1 \geq L$. We then compute the spectrum X_1 using T_1 and N_1 and the spectrum X_2 using $T_2 = T_1/2$ and $N_2 = 2N_1$. Because $T_1N_1 = T_2N_2$, the two spectra use the same data length, thus they have the same frequency resolution and the same effects of truncation. Therefore, any difference between X_1 and X_2 is due entirely to frequency aliasing. The Nyquist frequency range using T_1 is $(-\pi/T_1, \pi/T_1]$ and the Nyquist frequency range using $T_2 = T_1/2$ is $(-2\pi/T_1, 2\pi/T_1]$. If $|X_1|$ and $|X_2|$ differ appreciably inside $[0, \pi/T_1]$ or X_2 is not practically zero inside $[\pi/T_1, 2\pi/T_1]$, frequency aliasing is large and we must repeat the process. That is, we compare the spectrum X_2 computed using T_2 and N_2 with the spectrum X_3 computed using $T_3 = T_2/2$ and $N_3 = 2N_2$. If X_2 and X_3 are close inside $[0, \pi/T_2]$ and X_3 is practically zero inside $[\pi/T_2, 2\pi/T_2]$, we stop the repetition and the last T, called T_a, can be used in subsequent computation. This T_a is selected based entirely on the effect of frequency aliasing.

[7] In using FFT, N can be any positive integer. However, to use Program 5.9, N must be even.

Once T_a is selected, we select $\bar{N}_1$ that is a power of 2 such that $T_a\bar{N}_1 \geq L$. We then compute the spectrum $\bar{X}_1$ using T_a and $\bar{N}_1$ and the spectrum $\bar{X}_2$ using T_a and $\bar{N}_2 = 2\bar{N}_1$. Note that they have the same Nyquist frequency range but $\bar{X}_2$ uses twice the data length of $\bar{X}_1$ and has half of the frequency resolution of $\bar{X}_1$. Thus any difference between $\bar{X}_1$ and $\bar{X}_2$ is due to frequency resolution and the effects of truncation. If $\bar{X}_1$ and $\bar{X}_2$ differ appreciably, we repeat the process. That is, we compare the spectrum $\bar{X}_2$ computed using T_a and $\bar{N}_2$ with the spectrum $\bar{X}_3$ computed using T_a and $\bar{N}_3 = 2\bar{N}_2$. If they are close, the frequency resolution is good and the effect of truncation is negligible. We stop the computation and the last $\bar{N}_3$ will be acceptable.

5.6.2 FFT Spectral Computation of CT Sinusoids

This subsection computes the magnitude spectrum of a CT sine function. Consider the sinusoid $\sin 20t$ for $t \geq 0$. It is not absolutely integrable and its spectrum consists of an impulse at $\omega = 20$. The spectrum of the truncated $\sin 20t$ of length L will consist of, as discussed in Section 4.6, a main lobe centered at $\omega = 20$ with height $L/2$ and base width π/L and infinitely many side lobes. Thus it will show leakage and ripples and, consequently, Gibbs phenomena. We will see whether this is the case.

First we must select a sampling period. The sinusoid $\sin 20t$ is band-limited to 20, thus the sampling period can be selected as any number smaller than $\pi/20 = 0.15$. Arbitrarily, we select $T = 0.1$. Then the program

```
T=0.1;N=512;D=2*pi/(N*T);
n=0:N-1;t=n*T;
x=sin(20*t);
X=T*fft(x);
mp=0:N/2:
plot(mp*D,abs(X(mp+1)))
axis square,axis([18 22 0 30])
```

generates the magnitude spectrum in Figure 5.18(a). It consists of a triangular pulse centered at $\omega = 20$ and with height roughly 25. Note that the data length is $L = TN = 51.5$. The height is

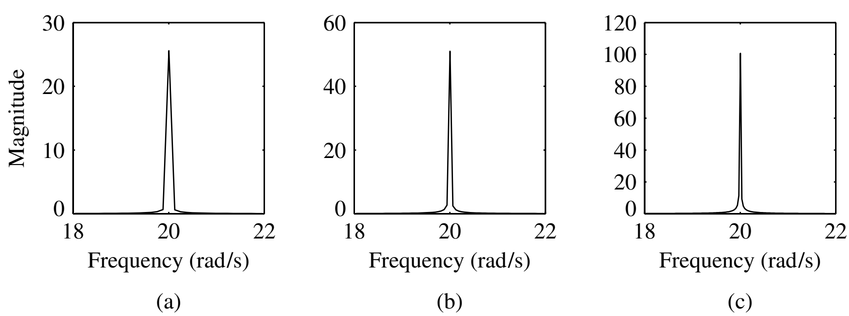


Figure 5.18 FFT computed magnitude spectrum of $\sin 20t$ using $T = 0.1$ and (**a**) $N = 512$, (**b**) $N = 1024$, and (**c**) $N = 2048$.

roughly half of L. Now if we double N from $N = 512$ to 1024 or, equivalently, double L, then the program generates the triangular pulse in Figure 5.18(b). Its base width reduces by half, and its height doubles. If we double N to $N = 2048$, then the program generates the pulse in Figure 5.18(c). Its base width reduces by half, and its height doubles. This is consistent with the theoretical result. As N increases, the pulse approaches an impulse at $\omega = 20$.

The results in Figure 5.18 show leakage, one of the effects of truncation. There are no ripples and, consequently, no Gibbs phenomenon. It turns out that the reason is poor frequency resolution. If we pad trailing zeros to the truncated $\sin 20t$, then ripples will appear. See Reference 2. However, when we have more data, there is no reason to pad trailing zeros. Thus the FFT computed spectrum suits us well.

PROBLEMS

5.1 Compute, if it exists, the frequency spectrum of the positive-time sequence $x[n] = 2^n$, for $n \geq 0$ and $T = 0.5$. Is the sequence absolutely summable?

5.2 Compute the frequency spectrum of the DT sequence $x[-1] = 2$, $x[0] = -1$, $x[1] = 2$, with sampling period $T = 1$. Is the spectrum real-valued? Is the spectrum periodic? What is its period? Plot its spectrum in its Nyquist frequency range. Also plot its magnitude and phase spectra.

5.3 Repeat Problem 5.2 for the positive-time sequence $x[0] = 2$, $x[1] = -1$, $x[2] = 2$, with sampling period $T = 1$. Is its magnitude spectrum the same as the one in Problem 5.2? What is the relationship between their phase spectra?

5.4 Compute the frequency spectrum of the sequence $x[-1] = 2$, $x[0] = 0$, and $x[1] = -2$, and the rest of $x[n]$ are zero. Its sampling period is $T = 0.5$. Plot its magnitude and phase spectra inside the NFR.

5.5 Compute the frequency spectrum of $x(nT) = 1$, for $n = 0 : 4$ and $T = 0.2$. Plot its magnitude and phase spectra inside the NFR.

5.6 Find the frequency spectrum of $x[n] = x(nT) = 0.8^n$, for $n \geq 0$ and $T = 0.2$. Compute its values at $\omega = m\pi$, for $m = 0 : 5$. Use the values to sketch roughly its magnitude and phase spectra. Is it a low- or high-frequency signal?

5.7 Repeat Problem 5.6 for $x[n] = x(nT) = (-0.8)^n$.

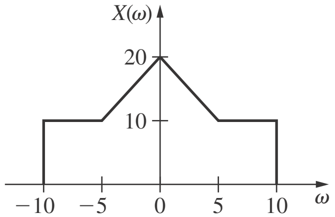


Figure 5.19

5.8 Let the frequency spectrum of a CT signal $x(t)$ be as shown in Figure 5.19. It is bandlimited to $\omega_{\max} = 10$. Find the spectra of its sampled sequences with $T_1 = 0.2$, $T_2 = 0.4$, and $T_3 = 0.6$. Under what condition on T can the spectrum of $x(t)$ be obtained from the spectrum of $x(nT)$?

5.9 Consider the CT signal $x(t) = \cos 40t - \cos 60t$. What is its frequency spectrum? Is it real-valued? Plot its spectrum against frequency. What is the spectrum of its sampled sequence $x(nT)$ if the sampling period is $T_1 = \pi/100$? Are the the spectra of $x(t)$ and $T_1 x(nT_1)$ the same inside the NFR?

5.10 Repeat Problem 5.9 for $T_2 = \pi/50$ and $T_3 = \pi/30$. Verify your results by writing explicitly $x(nT)$.

5.11 Consider the CT signal $x(t) = \sin 40t + \cos 60t$. What is its spectrum? Plot its magnitude and phase spectra. What is the spectrum of its sampled sequence with $T_1 = \pi/100$? Is there frequency aliasing? Can $x(t)$ be recovered from $x(nT_1)$? Repeat the questions for $T_2 = \pi/50$ and $T_3 = \pi/30$.

5.12 What is the frequency spectrum of $x(t) = e^{-0.05t}$, for $t \geq 0$? Is $x(t)$ band-limited? What is the frequency spectrum of its sampled sequence $x(nT)$ with $T = 2$? Is frequency aliasing appreciable?

5.13 Repeat Problem 5.12 with $T = 0.2$.

5.14 Use 3-point FFT to compute the frequency spectrum of the sequence in Problem 5.3. Are they the samples of the magnitude and phase spectra computed in Problem 5.3?

5.15 Use 8-point FFT to compute the frequency spectrum of the sequence in Problem 5.3 and to plot, using the linear interpolation, its magnitude and phase spectra in $[0, \pi/T]$. Do the results close to the exact ones?

5.16 Repeat Problem 5.15 using 1024-point FFT.

5.17 Use 5-point FFT to compute the frequency spectrum of the sequence in Problem 5.5. Where are the 5-point FFT output located? Are they the samples of the magnitude and phase spectra computed in Problem 5.5?

5.18 Use 8-point FFT to compute the spectrum of the sequence in Problem 5.5 and to plot, using the linear interpolation, its magnitude and phase spectra in $[0, \pi/T]$. Do the results close to the exact ones?

5.19 Repeat Problem 5.16 using 1024-point FFT.

5.20 Use FFT to compute the frequency spectrum of $x(t) = e^{-0.01t} \sin t$, for $t \geq 0$.

5.21 Use FFT to compute the frequency spectrum of $e^{-0.01t} \sin t \cos 20t$, for $t \geq 0$.

CHAPTER 6

CT Transfer Functions—Laplace Transform

6.1 INTRODUCTION

We showed in Chapter 3 that every CT LTI system can be described by an integral convolution. If the system is lumped as well, then it can also be described by a differential equation such as

$$a_1 y^{(N)}(t) + a_2 y^{(N-1)}(t) + \cdots + a_{N+1} y(t) = b_1 u^{(M)}(t) + b_2 u^{(M-1)}(t) + \cdots + b_{M+1} u(t) \qquad (6.1)$$

where $y^{(k)}(t) = d^k y(t)/dt^k$, and a_i and b_i are real constants. Generally we require $N \geq M$ to avoid amplification of high-frequency noise. For CT LTIL systems, we prefer differential equations to convolutions just as in the DT case discussed in Section 3.5.

Once the mathematical description of a system is available, the next step is to carry out analyses. There are two types of analyses: quantitative and qualitative. In quantitative analysis, we are interested in specific input–output responses. In qualitative analysis, we are interested in general properties of the system. It turns out that the differential equation is not suitable in either analysis. For quantitative analysis, we will transform the high-order differential equation in (6.1) into a set of first-order differential equations. This will be discussed in the next chapter. This chapter studies only qualitative analysis. We first introduce the Laplace transform to transform integral convolutions and differential equations into *algebraic equations*. We then introduce transfer functions and use them to develop general properties of systems. We also introduce frequency responses under the stability condition and establish the relationship between frequency responses of systems and frequency spectra of signals. We conclude the chapter by showing the phenomenon of resonance.

6.2 LAPLACE TRANSFORM

Consider a CT signal $x(t)$. Its Laplace transform pair is defined as

$$X(s) := \mathcal{L}[x(t)] := \int_0^\infty x(t) e^{-st}\, dt \qquad (6.2)$$

and

$$x(t) = \mathcal{L}^{-1}[X(s)] = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} X(s) e^{st}\, ds \qquad (6.3)$$

where s is a complex variable, called the Laplace-transform variable, and c is a constant to be discussed shortly. The function $X(s)$ is called the Laplace transform of $x(t)$, and $x(t)$ is called the inverse Laplace transform of $X(s)$. The Laplace transform as defined is often called the one-sided Laplace transform. If the lower integration limit 0 in (6.2) is replaced by $-\infty$, then (6.2) is called the two-sided Laplace transform. The two-sided Laplace transform is rarely used in practice,[1] thus its discussion is omitted.

EXAMPLE 6.2.1

Consider the signal $x(t) = e^{2t}$, defined for all t. This function grows exponentially to ∞ as $t \to \infty$.[2] Its Laplace transform is

$$X(s) = \int_0^\infty e^{2t} e^{-st}\, dt = \int_0^\infty e^{-(s-2)t}\, dt = \frac{-1}{s-2} e^{-(s-2)t}\Big|_{t=0}^{\infty}$$

$$= \frac{-1}{s-2}\left[e^{-(s-2)t}|_{t=\infty} - e^{-(s-2)t}|_{t=0}\right] \tag{6.4}$$

The value of $e^{-(s-2)t}$ at $t = 0$ is 1 for all s. However, the value of $e^{-(s-2)t}$ at $t = \infty$ can be, depending on s, zero, infinity, or undefined. To see this, we express the complex variable as $s = \sigma + j\omega$, where $\sigma = \text{Re}\, s$ and $\omega = \text{Im}\, s$ are, respectively, the real and imaginary part of s. Then we have

$$e^{-(s-2)t} = e^{-(\sigma-2)t} e^{-j\omega t} = e^{-(\sigma-2)t}(\cos\omega t - j\sin\omega t)$$

If $\sigma = 2$, the function reduces to $\cos\omega t - j\sin\omega t$ whose value is not defined for $t \to \infty$. If $\sigma < 2$, then $e^{-(\sigma-2)t}$ approaches infinity as $t \to \infty$. If $\sigma > 2$, then the function approaches 0 as $t \to \infty$. In conclusion, we have

$$e^{-(s-2)t}\Big|_{t=\infty} = \begin{cases} \infty \text{ or } -\infty & \text{for Re } s < 2 \\ \text{undefined} & \text{for Re } s = 2 \\ 0 & \text{for Re } s > 2 \end{cases}$$

Thus (6.4) is not defined for $\text{Re}\, s \le 2$. However, if $\text{Re}\, s > 2$, then (6.4) reduces to

$$X(s) = \mathcal{L}[e^{2t}] = \frac{-1}{s-2}[0-1] = \frac{1}{s-2} \tag{6.5}$$

This is the Laplace transform of e^{2t}.

The Laplace transform in (6.5) is, strictly speaking, defined only for $\text{Re}\, s > 2$. The region $\text{Re}\, s > 2$, shown in Figure 6.1, is called the *region of convergence*. The region of convergence is needed in using (6.3) to compute the inverse Laplace transform of $1/(s-2)$. For example,

[1] To the best of this author's knowledge, the two-sided Laplace transform was used *only* in S. S. L. Chang, *Synthesis of Optimal Control Systems,* New York: McGraw-Hill, 1961, to derive optimal systems.

[2] Its Fourier transform, as discussed in Example 4.3.1, is not defined.

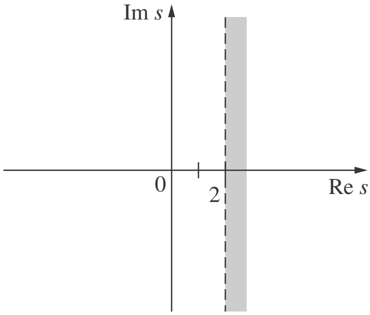


Figure 6.1 Region of convergence.

if c is selected as 2.5 which is inside the region of convergence, then

$$\frac{1}{2\pi j}\int_{2.5-j\infty}^{2.5+j\infty}\frac{1}{s-2}e^{-st}\,ds=\begin{cases}e^{2t} & \text{for } t\geq 0\\ 0 & \text{for } t<0\end{cases}$$

It yields a positive-time signal. Thus we may say that the Laplace transform is defined for positive-time signals. Because we encounter only positive-time signals in practice, the Laplace transform as introduced suits us well.

For the preceding example, if c is selected outside the region of convergence such as $c = 1, 0$ or -2, then its inverse Laplace transform will yield a negative-time signal. More generally, let $X(s)$ be the Laplace transform of a positive-time signal. Then the inverse Laplace transform of $X(s)$ may yield a positive-time, negative-time, or two-sided signal depending on the c selected. In other words, without specifying the region of convergence, the inverse Laplace transform in (6.3) may not yield the original $x(t)$. Fortunately, in our application, we do not use (6.3) to compute the inverse Laplace transform. Instead we use table lookup as we will discuss later. Furthermore, even though (6.5) is developed under the assumption Re $s > 2$, the Laplace transform $1/(s-2)$ can be considered to be defined for all s except at $s = 2$. Thus we will not be concerned further with the region of convergence.

We give two more examples to conclude this section.

EXAMPLE 6.2.2

Find the Laplace transform of the impulse $\delta(t)$. By definition, we have

$$\Delta(s)=\int_0^\infty \delta(t)e^{-st}\,dt$$

As discussed in Section 1.5, when an integration touches an impulse, we agreed to include the whole impulse inside the integration interval. Using the sifting property, we have

$$\Delta(s)=e^{-st}\Big|_{t=0}=e^0=1$$

Thus we have $\mathcal{L}[\delta(t)] = 1$.

EXAMPLE 6.2.3

Find the Laplace transform of $x(t) = 2e^{at}$ where a is a real or complex constant. By definition, we have

$$X(s) = \int_0^\infty 2e^{at}e^{-st}\,dt = \frac{2}{a-s}e^{(a-s)t}\Big|_{t=0}^{\infty}$$
$$= \frac{2}{a-s}\left[e^{(a-s)t}\Big|_{t=\infty} - e^{(a-s)\cdot 0}\right] = \frac{2}{a-s}[0-1] = \frac{2}{s-a}$$

where we have used $e^{(a-s)t} = 0$ at $t = \infty$.

EXERCISE 6.2.1

What are the Laplace transforms of $0.2e^{-2t}$ and $-e^t$?

Answers

$0.2/(s+2)$, $-1/(s-1)$

We discuss some special cases of $\mathcal{L}[e^{at}] = 1/(s-a)$. If $a = 0$, then $x(t) = q(t)$, the step function defined in (1.1). Thus we have

$$\mathcal{L}[q(t)] = \frac{1}{s}$$

If $a = \pm j\omega_0$, where ω_0 is a positive real constant, then

$$\mathcal{L}[e^{\pm j\omega_0 t}] = \frac{1}{s \mp j\omega_0}$$

The Laplace transforms of the signals in the preceding examples are rational functions of s. This is, however, not always the case. In fact, given a signal, the following situations may occur:

1. Its Laplace transform does not exist. In other words, there is no region of convergence. The functions e^{t^2} and e^{e^t} are such examples. These two functions are mathematically contrived and do not arise in practice. Thus it is fair to say that all signals encountered in practice are Laplace transformable.
2. Its Laplace transform exists and is an irrational function of s. For example, the Laplace transforms of $1/\sqrt{\pi t}$ and $(\sinh at)/t$ are $1/\sqrt{s}$ and $0.5\ln[(s+a)/(s-a)]$, respectively. See Reference 33.
3. Its Laplace transform exists but cannot be expressed in closed form. Most signals encountered in practice belong to this type.
4. Its Laplace transform exists and is a rational function of s.

We encounter in this chapter mostly Laplace transforms that belong to the last class.

6.3 TRANSFER FUNCTIONS

Every LTI system that is initially relaxed at $t = 0$ can be described by the convolution

$$y(t) = \int_{\tau=0}^{\infty} h(t-\tau)u(\tau)\,d\tau \tag{6.6}$$

Here we have assumed that the input $u(t)$ is applied from $t = 0$ onward. If the system is causal, then $h(t) = 0$ for $t < 0$ and the output will start to appear from $t = 0$ onward. Thus all signals in (6.6) are positive-time signals. Let us apply the Laplace transform to $y(t)$:

$$Y(s) = \mathcal{L}[y(t)] = \int_{t=0}^{\infty} y(t)e^{-st}\,dt = \int_{t=0}^{\infty}\left[\int_{\tau=0}^{\infty} h(t-\tau)u(\tau)\,d\tau\right] e^{-s(t-\tau+\tau)}\,dt$$

Interchanging the order of integrations yields

$$Y(s) = \int_{\tau=0}^{\infty} u(\tau)e^{-s\tau}\left[\int_{t=0}^{\infty} h(t-\tau)e^{-s(t-\tau)}\,dt\right] d\tau \tag{6.7}$$

Let us introduce a new variable $\bar{t}$ as $\bar{t} := t - \tau$. Then the term inside the brackets becomes

$$\int_{t=0}^{\infty} h(t-\tau)e^{-s(t-\tau)}\,dt = \int_{\bar{t}=-\tau}^{\infty} h(\bar{t})e^{-s\bar{t}}\,d\bar{t} = \int_{\bar{t}=0}^{\infty} h(\bar{t})e^{-s\bar{t}}\,d\bar{t} =: H(s) \tag{6.8}$$

where we have used the fact that $h(\bar{t}) = 0$ for $\bar{t} < 0$. The last integration is, by definition, the Laplace transform of $h(t)$. Because it is independent of τ, $H(s)$ can be moved outside the integration. Thus (6.7) becomes

$$Y(s) = H(s)\int_{\tau=0}^{\infty} u(\tau)e^{-s\tau}\,d\tau$$

or

$$Y(s) = H(s)U(s) \tag{6.9}$$

where $Y(s)$ and $U(s)$ are the Laplace transforms of the output and input, respectively. The function $H(s)$ is the Laplace transform of the impulse response and is called the *transfer function*. Because the convolution in (6.6) describes only zero-state responses, so does the transfer function. In other words, whenever we use (6.9), the system is implicitly assumed to be initially relaxed.

EXERCISE 6.3.1

Consider an amplifier with gain 10. What is its impulse response? What is its transfer function?

Answers

$10\delta(t)$. 10.

Using (6.9), we can also define the transfer function as

$$\text{Transfer function} = \left.\frac{\mathcal{L}[\text{output}]}{\mathcal{L}[\text{input}]}\right|_{\text{all initial conditions zero}} \tag{6.10}$$

Using this equation, we can obtain the transfer function using any input–output pair. This establishes formally the assertion in Chapter 2 that the characteristics of an LTI system can be obtained from any single pair of input and output. The definition in (6.10) includes as a special case the definition that the transfer function is the Laplace transform of the impulse response. Indeed, if the input is an impulse, then its Laplace transform is 1 and the transfer function simply equals the Laplace transform of the corresponding output, which by definition is the impulse response. We mention that transfer functions are also called *system functions* or *network functions*.

EXAMPLE 6.3.1

Consider the system studied in Example 3.7.1. We assumed there that its step response can be measured as

$$y_q(t) = 1 - e^{-0.5t}$$

Its impulse response can then be obtained as $dy_q(t)/dt$. But differentiation is not desirable in practice because of ever presence of high-frequency noise.

Now if we use (6.10), then there is no need to carry out differentiation. The Laplace transform of $y_q(t)$ is

$$Y_q(s) = \frac{1}{s} - \frac{1}{s+0.5} = \frac{s+0.5-s}{s(s+0.5)} = \frac{0.5}{s(s+0.5)}$$

The input is a step function $q(t)$ and its Laplace transform is $Q(s) = 1/s$. Thus the transfer function of the network in Figure 3.3(a) is

$$H(s) = \frac{Y_q(s)}{Q(s)} = \frac{\dfrac{0.5}{s(s+0.5)}}{\dfrac{1}{s}} = \frac{0.5}{s+0.5}$$

Its inverse Laplace transform is $h(t) = 0.5e^{-0.5t}$. This is the impulse response obtained in Example 3.7.1.

Every CT LTI system that is initially relaxed can be described by the algebraic equation in (6.9). If the system is distributed, then its transfer function is either an irrational function of s or cannot be expressed in closed form. However, if the system is lumped, then its transfer function is a rational function of s as we derive in the next subsection.

6.3.1 From Differential Equations to Rational Transfer Functions

Every CT LTI lumped system can be described by, in addition to a convolution, a differential equation such as the one in (6.1). In order to apply the Laplace transform to (6.1), we first develop some formulas. Let $X(s)$ be the Laplace transform of $x(t)$. Then we have

$$\mathcal{L}\left[\frac{dx(t)}{dt}\right] := \mathcal{L}[\dot{x}(t)] = s\,[\mathcal{L}[x(t)]] - x(0) = sX(s) - x(0) \tag{6.11}$$

This can be established using integration by parts. By definition, we have

$$\begin{aligned}\mathcal{L}\left[\frac{dx(t)}{dt}\right] &= \int_0^\infty \frac{dx(t)}{dt}e^{-st}\,dt \\ &= x(t)e^{-st}\Big|_{t=0}^{\infty} - \int_0^\infty x(t)\frac{de^{-st}}{dt}\,dt \\ &= 0 - x(0)e^{-s\cdot 0} - (-s)\int_0^\infty x(t)e^{-st}\,dt \\ &= sX(s) - x(0)\end{aligned}$$

where we have assumed $e^{-st} = 0$ at $t = \infty$. This holds only in the region of convergence. However, in application, this restriction is completely disregarded. Using (6.11), we have

$$\begin{aligned}\mathcal{L}[\ddot{x}(t)] = \mathcal{L}\left[\frac{d\dot{x}(t)}{dt}\right] &= s[\mathcal{L}[\dot{x}(t)]] - \dot{x}(0) \\ &= s[sX(s) - x(0)] - \dot{x}(0) = s^2X(s) - sx(0) - \dot{x}(0)\end{aligned}$$

Proceeding forward, we can show

$$\mathcal{L}[x^{(k)}(t)] = s^kX(s) - s^{k-1}x(0) - s^{k-2}\dot{x}(0) - \cdots - x^{(k-1)}(0) \tag{6.12}$$

where $x^{(k)}(t)$ is the kth derivative of $x(t)$. We use examples to show the use of the formulas.

EXAMPLE 6.3.2

Consider a CT system described by the second-order differential equation

$$2\ddot{y}(t) + 3\dot{y}(t) + 5y(t) = \dot{u}(t) - 2u(t) \tag{6.13}$$

Applying the Laplace transform yields

$$2[s^2Y(s) - sy(0) - \dot{y}(0)] + 3[sY(s) - y(0)] + 5Y(s) = sU(s) - u(0) - 2U(s)$$

which can be grouped as

$$(2s^2 + 3s + 5)Y(s) = (s-2)U(s) + (2s+3)y(0) + 2\dot{y}(0) - u(0)$$

Thus we have

$$Y(s) = \frac{s-2}{2s^2+3s+5}U(s) + \frac{(2s+3)y(0) + 2\dot{y}(0) - u(0)}{2s^2+3s+5} \tag{6.14}$$

We see that the output in the transform domain consists of two parts: One part is excited by the input $U(s)$, and the other part is excited by the initial conditions $y(0)$, $\dot{y}(0)$, and $u(0)$. The former is the zero-state or forced response, and the latter is the zero-input or natural response. Their sum is the total response. See (2.18).

If all initial conditions are zero, then (6.14) reduces to

$$Y(s) = \frac{s-2}{2s^2+3s+5}U(s) =: H(s)U(s)$$

Thus the transfer function of the system described by (6.13) is $(s-2)/(2s^2+3s+5)$.

The advantage of using the Laplace transform is evident by comparing (6.13) and (6.14). The former is a differential equation, and the latter is an algebraic equation. Algebraic manipulations (addition and multiplication) are clearly simpler than calculus manipulations (integration and differentiation). Equation (6.14) also reveals the fact that the response of a linear system can be decomposed as the zero-input and zero-state responses. A great deal more can be said about the system from the algebraic equation, as we will do in this chapter.

To conclude this subsection, we compute the transfer function of the differential equation in (6.1). If all initial conditions are zero, then (6.12) reduces to

$$\mathcal{L}[x^{(k)}(t)] = s^k X(s)$$

for $k = 0, 1, 2, \ldots$. Thus one differentiation in the time domain is equivalent to multiplication by one s in the Laplace-transform domain, denoted as

$$\frac{d}{dt} \leftrightarrow s$$

Applying the Laplace transform to (6.1) and assuming zero initial conditions yield

$$a_1s^NY(s) + a_2s^{N-1}Y(s) + \cdots + a_{N+1}Y(s) = b_1s^MU(s) + b_2s^{M-1}U(s) + \cdots + b_{M+1}U(s)$$

Thus its transfer function is

$$H(s) := \frac{Y(s)}{U(s)} = \frac{b_1s^M + b_2s^{M-1} + \cdots + b_{M+1}}{a_1s^N + a_2s^{N-1} + \cdots + a_{N+1}} \tag{6.15}$$

This is a rational function of s (a ratio of two polynomials of s). We see that transforming a differential equation into a transfer function is simple and straightforward. Likewise, so is transforming a transfer function into a differential equation.

EXERCISE 6.3.2

Write an equation to relate the input $u(t)$ and output $y(t)$ of an amplifier with gain 10. Take its Laplace transform and find its transfer function. Is the result the same as the one obtained in Exercise 6.3.1?

Answers

$y(t) = 10u(t)$. $Y(s) = 10U(s)$. 10. Same.

EXERCISE 6.3.3

Find the transfer function of the mechanical system studied in Example 3.8.4.

Answer

$1/(ms^2 + fs + k)$

EXERCISE 6.3.4

Transform the transfer function

$$H(s) = \frac{Y(s)}{U(s)} = \frac{2s+3}{2s^3 - 2s^2 + 5}$$

into a differential equation.

Answer

$2y^{(3)}(t) - 2\ddot{y}(t) + 5y(t) = 2\dot{u}(t) + 3u(t)$

6.3.2 Transform Impedances

This subsection discusses how to compute transfer functions for RLC networks. One way to find the transfer function of a network is to compute its differential equation and then take its Laplace transform. However, it is generally simpler to compute its transfer function directly as we discuss next.

The branch characteristics of an inductor with inductance L, a capacitor with capacitance C, and a resistor with resistance R are

$$v(t) = L\frac{di(t)}{dt}, \qquad i(t) = C\frac{dv(t)}{dt}, \qquad v(t) = Ri(t)$$

respectively, where $v(t)$ is the voltage across each branch and $i(t)$ is its current flowing from high to low potentials. Taking their Laplace transforms and assuming zero initial conditions yield

$$V(s) = LsI(s), \qquad V(s) = \frac{1}{Cs}I(s), \qquad V(s) = RI(s)$$

If we consider the current as the input and the excited voltage as the output, then the transfer functions of L, C, and R are Ls, $1/Cs$, and R, respectively. They are called *transform impedances, Laplace impedances,* or, simply, *impedances*.[3] Using impedances, the voltage and current of every branch can be written as $V(s) = Z(s)I(s)$ with $Z(s) = Ls$ for inductors, $Z(s) = 1/Cs$ for capacitors, and $Z(s) = R$ for resistors. In basic network analysis, impedances for R, L, and C are also defined as R, $j\omega L$, and $1/jC\omega$. This will be discussed later in Section 6.9.2.

[3] If we consider the voltage as the input and the current as the output, then their transfer functions are called admittances.

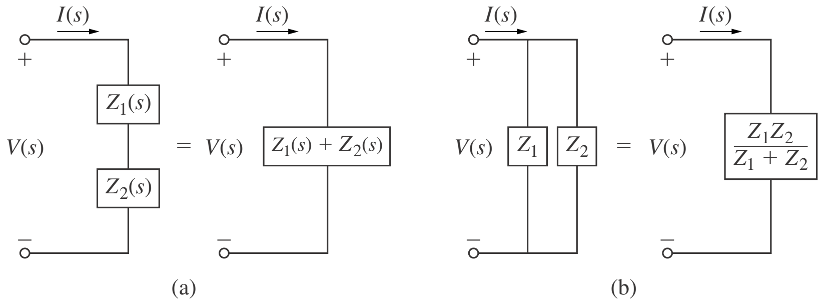


Figure 6.2 (**a**) Series connection of two impedances. (**b**) Parallel connection of two impedances.

The manipulation involving impedances is purely algebraic, identical to the manipulation of resistances. For example, the resistance of the series connection of R_1 and R_2 is $R_1 + R_2$. The resistance of the parallel connection of R_1 and R_2 is $R_1 R_2/(R_1 + R_2)$. Similarly, the impedance of the series connection of two impedances $Z_1(s)$ and $Z_2(s)$ is $Z_1(s) + Z_2(s)$; the impedance of the parallel connection of $Z_1(s)$ and $Z_2(s)$ is

$$Z(s) = \frac{Z_1(s)Z_2(s)}{Z_1(s) + Z_2(s)}$$

as shown in Figure 6.2. The only difference is that we now deal with rational functions rather than real numbers as in the resistive case.

EXAMPLE 6.3.3

Compute the transfer function of the network shown in Figure 3.8. The network is redrawn in Figure 6.3(a) using impedances.

The impedance of the series connection of 3 and s is $Z_1(s) = 3 + s$; the impedance of the parallel connection of 4 and $1/2s$ is

$$Z_2(s) = \frac{4 \times (1/2s)}{4 + (1/2s)} = \frac{4}{8s + 1}$$

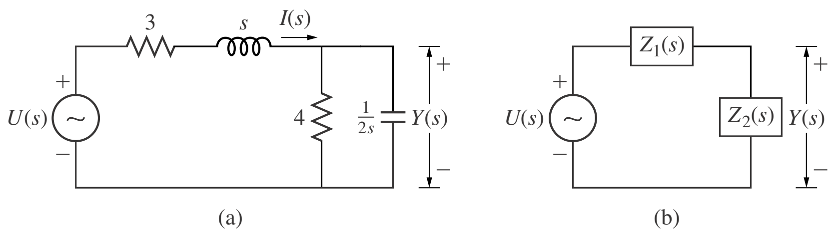


Figure 6.3 (**a**) Network. (**b**) Its equivalent network.

as shown in Figure 6.3(b). Clearly, it is a voltage divider. Thus we have

$$Y(s) = \frac{Z_2(s)}{Z_1(s) + Z_2(s)} U(s) = \frac{\dfrac{4}{8s+1}}{3 + s + \dfrac{4}{8s+1}} U(s)$$

$$= \frac{4}{(s+3)(8s+1)+4} U(s) = \frac{4}{8s^2 + 25s + 7} U(s)$$

and the transfer function from u to y is $4/(8s^2 + 25s + 7)$.

EXERCISE 6.3.5

Consider the network in Figure 6.3(a). Let us consider the current $i(t)$ passing through the inductor as the output. What is the transfer function from $u(t)$ to $i(t)$?

Answer

$(8s + 1)/(8s^2 + 25s + 7)$

EXERCISE 6.3.6

Show that the transfer function of the differential equation in (3.36) equals the one computed in Example 6.3.3.

EXERCISE 6.3.7

Use impedances to compute the transfer function of the network in Figure 3.3(a). Is the result the same as the one in Example 6.3.1?

Answers

$1/(2s + 1)$. Same.

6.3.3 Proper Rational Transfer Functions

Consider the rational function

$$H(s) = \frac{N(s)}{D(s)} \tag{6.16}$$

where $N(s)$ and $D(s)$ are polynomials of s. We use "deg" to denote the degree of a polynomial. Depending on the relative degrees of $N(s)$ and $D(s)$, we have the following definitions:

$$\begin{aligned}
H(s) \text{ improper} \quad &\leftrightarrow \quad \deg N(s) > \deg D(s) \quad \leftrightarrow \quad H(\infty) = \pm\infty \\
H(s) \text{ proper} \quad &\leftrightarrow \quad \deg N(s) \le \deg D(s) \quad \leftrightarrow \quad H(\infty) = \text{constant} \\
H(s) \text{ biproper} \quad &\leftrightarrow \quad \deg N(s) = \deg D(s) \quad \leftrightarrow \quad H(\infty) = \text{nonzero constant} \\
H(s) \text{ strictly proper} \quad &\leftrightarrow \quad \deg N(s) < \deg D(s) \quad \leftrightarrow \quad H(\infty) = 0
\end{aligned}$$

For example, the rational functions

$$\frac{s^2+1}{10s-1}, \qquad s+1, \qquad \frac{s^{10}}{s^3+1}$$

are improper. The rational functions

$$-1, \qquad \frac{2s+1}{s-1}, \qquad \frac{100}{(s^2+3s)}, \qquad \frac{3s^2+s+5}{2s^3}$$

are proper. The first two are also biproper, and last two are strictly proper. Thus proper rational functions include both biproper and strictly proper rational functions. Note that if $H(s) = N(s)/D(s)$ is biproper, so is its inverse $H^{-1}(s) = D(s)/N(s)$.

Properness of $H(s)$ can also be determined from the value of $H(s)$ at $s = \infty$. The rational function $H(s)$ is improper if $|H(\infty)| = \infty$, proper if $H(\infty)$ is a zero or nonzero constant, biproper if $H(\infty)$ is a nonzero constant, and strictly proper if $H(\infty) = 0$.

The transfer functions encountered in practice are mostly proper rational functions. The only exception is transfer functions of tachometers. A *tachometer* is a device whose output is proportional to the derivative of its input. A mechanical tachometer can be a spindle driven by a rotational shaft. The spread of the spindle due to centrifugal force can be used to indicate the shaft's speed.[4] An electromechanical tachometer can be a permanent magnetic dc generator driven by a shaft. If the input θ is the angular position of the shaft, then the generated voltage is $kd\theta/dt$ for some constant k. Thus its transfer function is ks, an improper transfer function. As shown in Figure 3.4 and discussed in Example 3.8.2, if a signal contains high-frequency noise, then the noise will be greatly amplified by differentiation. However, the input of the tachometer is a mechanical signal that does not contain high-frequency noise, thus electromechanical tachometers are widely used in practice. For electronic differentiators with electrical signals as inputs, we design proper transfer functions to approximate ks. This will be discussed in Chapter 8—in particular, (8.10). From now on, we study only proper rational transfer functions or transfer functions of the form shown in (6.15) with $N \geq M$.

6.3.4 Poles and Zeros

Consider a proper rational transfer function $H(s)$. A finite real or complex number λ is called a zero of $H(s)$ if $H(\lambda) = 0$. It is called a pole if $|H(\lambda)| = \infty$. For example, consider

$$H(s) = \frac{2s^2+2s}{s^3+5s^2+9s+5} \tag{6.17}$$

We compute its value at $s = 0$:

$$H(0) = \frac{0+0}{0+0+0+5} = \frac{0}{5} = 0$$

Thus 0 is a zero of $H(s)$. Next we compute

$$H(-1) = \frac{2(-1)^2+(-1)}{(-1)^3+5(-1)^2+9(-1)+5} = \frac{0}{0}$$

[4]Because the centrifugal force is proportional to the square of the speed, the calibration of speed is not linear.

It is not defined. Thus we use l'Hôpital's rule to find its value:

$$H(-1) = \left.\frac{4s+2}{3s^2+10s+9}\right|_{s=-1} = \frac{4(-1)+2}{3(-1)^2+10(-1)+9} = \frac{-2}{2} = -1$$

It is neither 0 nor ∞. Thus -1 is neither a zero nor a pole. Because $H(\infty) = 0$, the number $\lambda = \infty$ could be a zero. However, we require λ to be finite, thus ∞ is not a zero. In other words, we consider only finite poles and finite zeros.

To find the poles and zeros of $H(s)$ in (6.17), we factor it as[5]

$$H(s) = \frac{2s(s+1)}{(s+1)(s^2+4s+5)} = \frac{2s}{s^2+4s+5} = \frac{2s}{(s+2+j)(s+2-j)} \tag{6.18}$$

Its numerator and denominator have the common factor $s+1$. After canceling the common factor, we can readily show that the transfer function $H(s)$ has one zero at 0 and a pair of conjugate poles at $-2 \pm j$. If $H(s)$ has only real coefficients and if a complex number is a pole (zero), then its complex conjugate must also be a pole (zero). In conclusion, given a proper rational transfer function $H(s) = N(s)/D(s)$, if $N(s)$ and $D(s)$ are *coprime* in the sense that they have no common factor of degree 1 or higher,[6] then all roots of $N(s)$ are zeros of $H(s)$ and all roots of $D(s)$ are poles of $H(s)$. This can be argued as follows: If $N(s)$ and $D(s)$ are coprime and if a is a root of $D(s)$ or $D(a) = 0$, then $N(a) \neq 0$ and $H(a) = N(a)/D(a) = \infty$ or $-\infty$. If $N(a) = 0$, then $D(a) \neq 0$ and $H(a) = N(a)/D(a) = 0$. This establishes the assertion. Note that if $N(s)$ and $D(s)$ have a common root at a, then $H(a) = N(a)/D(a) = 0/0$, an indeterminate, and a may or may not be a zero of $H(s)$. From now on, all transfer functions will be assumed to be coprime. Indeed, if $H(s)$ is not coprime, then it can be expressed in many ways such as

$$H(s) = \frac{2(s-1)}{s^2+3s-4} = \frac{2(s^2-1)}{s^3+4s^2-s-4} = \frac{2(s-1)R(s)}{(s^2+3s-4)R(s)}$$

for any polynomial $R(s)$. Such non-coprime expressions serve no purpose.

Let us define the degree of a rational function $H(s) = N(s)/D(s)$, where $N(s)$ and $D(s)$ are coprime. The *degree* of $H(s)$ is defined as the larger of the degrees of $N(s)$ and $D(s)$. For proper $H(s)$, its degree is simply the degree of its denominator. From the preceding discussion, we see that the number of poles of $H(s)$ equals the degree of $H(s)$. In defining the degree of $H(s)$, the coprimeness condition between its numerator and denominator is essential. Without the condition, the degree of $H(s)$ is not uniquely defined.

The transfer function after the last equality of (6.18) is said to be in the *zero/pole/gain form*. It has zero at 0, poles at $-2 \pm j$, and gain 2. A transfer function expressed as a ratio of two polynomials can be transformed into the form in MATLAB by calling the function `tf2zp`, an

[5]To factor the quadratic term $s^2 + as + b$, we first use the first two terms to complete a square such as $s^2 + as + (a/2)^2 - (a/2)^2 + b = (s + a/2)^2 + [b - (a/2)^2]$. If $c := b - (a/2)^2 \leq 0$, then the quadratic term can be factored as $(s + a/2 + \sqrt{-c})(s + a/2 - \sqrt{-c})$ and has two real roots $-a/2 - \sqrt{-c}$ and $-a/2 + \sqrt{-c}$. If $c > 0$, the term can be factored as $(s + a/2 + j\sqrt{c})(s + a/2 - j\sqrt{c})$ and has a pair of complex conjugate roots $-a/2 \pm j\sqrt{c}$.

[6]$N(s)$ and $D(s)$ have any nonzero constant, a polynomial of degree 0, as a common factor. Such a common factor is called a trivial common factor and is excluded in defining coprimeness.

acronym for transfer function to zero/pole. For example, consider

$$H(s) = \frac{8s^3 - 24s - 16}{2s^5 + 20s^4 + 98s^3 + 268s^2 + 376s + 208} \tag{6.19}$$

It is a ratio of two polynomials. A polynomial is expressed in MATLAB as a row vector such as `n=[8 0 -24 -16]` for the numerator in (6.19). Note the missing term s^2. To transform the transfer function in (6.19) into the zero/pole/gain form, we type

```
n=[8 0 -24 -16];d=[2 20 98 268 376 208];
[z,p,k]=tf2zp(n,d)
```

in MATLAB, which yields `z=[-1 -1 2]; p=[-2 -2 -2 -2-3j -2+3j]; k=4`. In other words, the transfer function has zeros at $-1, -1, 2$; poles at $-2, -2, -2, -2 - 3j, -2 + 3j$; and gain 4. Thus the zero/pole/gain form of (6.19) is

$$H(s) = \frac{4(s-2)(s+1)^2}{(s+2)^3(s+2+3j)(s+2-3j)} \tag{6.20}$$

A pole or zero is called *simple* if it appears only once, and it is called *repeated* if it appears twice or more. For example, the transfer function in (6.19) or (6.20) has a simple zero at 2 and two simple poles at $-2 \pm 3j$. It has a repeated zero at -1 with multiplicity 2 and a repeated pole at -2 with multiplicity 3.

6.4 PROPERTIES OF LAPLACE TRANSFORM

We discuss some properties of the Laplace transform and develop some Laplace transform pairs. To be brief, we call the Laplace-transform domain s-domain.

Linearity The Laplace transform is a linear operator. That is, if $X_1(s) = \mathcal{L}[x_1(t)]$ and $X_2(s) = \mathcal{L}[x_2(t)]$, then for any constants α_1 and α_2 we have

$$\mathcal{L}[\alpha_1 x_1(t) + \alpha_2 x_2(t)] = \alpha_1 \mathcal{L}[x_1(t)] + \alpha_2 \mathcal{L}[x_2(t)] = \alpha_1 X_1(s) + \alpha_2 X_2(s)$$

This can directly be verified from the definition of the Laplace transform. This property was used to compute $Y_q(s)$ in Example 6.3.1.

EXAMPLE 6.4.1

We use $\mathcal{L}[e^{at}] = 1/(s-a)$ to compute the Laplace transform of $\sin \omega_0 t$. Using Euler's identity

$$\sin \omega_0 t = \frac{e^{j\omega_0 t} - e^{-j\omega_0 t}}{2j}$$

we have

$$\mathcal{L}[\sin \omega_0 t] = \frac{1}{2j}\left[\mathcal{L}[e^{j\omega_0 t}] - \mathcal{L}[e^{-j\omega_0 t}]\right] = \frac{1}{2j}\left[\frac{1}{s - j\omega_0} - \frac{1}{s + j\omega_0}\right] = \frac{\omega_0}{s^2 + \omega_0{}^2}$$

EXERCISE 6.4.1

Verify

$$\mathcal{L}[\cos \omega_0 t] = \frac{s}{s^2 + \omega_0{}^2}$$

EXERCISE 6.4.2

Compute the Laplace transforms of

(a) $e^{2t} - 4e^{-3t}$

(b) $2 - 2e^t + 0.5 \sin 3t$

Answers

(a) $(-3s + 11)/(s^2 + s - 6)$

(b) $(-0.5s^2 - 1.5s - 18)/[s(s-1)(s^2+9)$

Shifting in the s-Domain (Multiplication by e^{-at} in the Time Domain) If $X(s) = \mathcal{L}[x(t)]$, then

$$\mathcal{L}[e^{-at}x(t)] = X(s+a)$$

By definition, we have

$$\mathcal{L}[e^{-at}x(t)] = \int_0^\infty e^{-at}x(t)e^{-st}\,dt = \int_0^\infty x(t)e^{-(s+a)t}\,dt = X(s+a)$$

This establishes the formula. Using the formula, we have

$$\mathcal{L}[e^{-at}\sin \omega_0 t] = \frac{\omega_0}{(s+a)^2 + \omega_0{}^2}$$

and

$$\mathcal{L}[e^{at}\cos \omega_0 t] = \frac{s-a}{(s-a)^2 + \omega_0{}^2}$$

EXERCISE 6.4.3

Find the Laplace transforms of

(a) $e^{-2t} \sin 4t$

(b) $e^{2t} \sin(t - \pi/4)$

Answers

(a) $4/[(s+2)^2 + 16]$

(b) $0.707(-s+3)/[(s-2)^2 + 1]$

Differentiation in the s-Domain (Multiplication by t in the Time Domain) If $X(s) = \mathcal{L}[x(t)]$, then

$$\mathcal{L}[tx(t)] = -\frac{dX(s)}{ds}$$

The differentiation of (6.2) with respect to s yields

$$\frac{d}{ds}X(s) = \frac{d}{ds}\int_0^\infty x(t)e^{-st}\,dt = \int_0^\infty x(t)\left[\frac{d}{ds}e^{-st}\right]dt = \int_0^\infty (-t)x(t)e^{-st}\,dt$$

which is, by definition, the Laplace transform of $-tx(t)$. This establishes the formula.

Using the formula and $\mathcal{L}[e^{-at}] = 1/(s+a)$, we can establish

$$\mathcal{L}[te^{-at}] = -\frac{d}{ds}\left[\frac{1}{s+a}\right] = \frac{1}{(s+a)^2}$$

$$\mathcal{L}[t^2e^{-at}] = -\frac{d}{ds}\left[\frac{1}{(s+a)^2}\right] = \frac{2!}{(s+a)^3}$$

$$\vdots$$

$$\mathcal{L}[t^ke^{-at}] = \frac{k!}{(s+a)^{k+1}}$$

where $k! = 1 \cdot 2 \cdot 3 \cdots k$.

Integration in the Time Domain (Division by s in the s-Domain) If $X(s) = \mathcal{L}[x(t)]$, then

$$\mathcal{L}\left[\int_0^t x(\tau)\,d\tau\right] = \frac{1}{s}X(s)$$

To establish the formula, we define

$$g(t) = \int_0^t x(\tau)\,d\tau$$

and $G(s) = \mathcal{L}[g(t)]$. Clearly we have $g(0) = 0$ and $x(t) = dg(t)/dt = \dot{g}(t)$. The Laplace transform of $x(t)$ is, using (6.11),

$$X(s) = \mathcal{L}[\dot{g}(t)] = sG(s) - g(0) = sG(s)$$

Thus we have

$$G(s) = \mathcal{L}\left[\int_0^t x(\tau)\,d\tau\right] = \frac{1}{s}X(s)$$

This establishes the formula. This formula states that integration in the time domain is equivalent to multiplication by $1/s$ in the s-domain, denoted as

$$\int_0^t \leftrightarrow \frac{1}{s}$$

This is in contrast to $d/dt \leftrightarrow s$ for differentiation. Thus an integrator can be denoted by s^{-1}.

Time Delay by $t_0 > 0$ (Multiplication by e^{-st_0} in the s-Domain) Let $x(t)$ be a positive-time signal and $X(s) = \mathcal{L}[x(t)]$. As discussed in Section 1.4.1, $x(t - t_0)$ is the shifting of $x(t)$ to t_0. If $t_0 < 0$, then $x(t - t_0)$ becomes a two-sided signal. If $t_0 > 0$, then $x(t - t_0)$ remains positive time and equals 0 for $t < t_0$. We compute its Laplace transform:

$$\mathcal{L}[x(t - t_0)] = \int_{t=t_0}^{\infty} x(t - t_0)e^{-st}\,dt = \int_{t=t_0}^{\infty} x(t - t_0)e^{-s(t-t_0)}e^{-st_0}\,dt = e^{-st_0}\int_{\bar{t}=0}^{\infty} x(\bar{t})e^{-s\bar{t}}\,d\bar{t}$$

where we have introduced the new variable $\bar{t} = t - t_0$. Thus we conclude

$$\mathcal{L}[x(t - t_0)] = e^{-st_0}X(s) \tag{6.21}$$

This formula is valid only if $x(t)$ is positive time and $t_0 \geq 0$. We use an example to illustrate this fact.

EXAMPLE 6.4.2

Consider $x(t) = e^{-2t}$, for all t. Its Laplace transform is $X(s) = 1/(s + 2)$. Consider

$$x(t - 3) = e^{-2(t-3)} = e^{-2t}e^{6} = 403.4e^{-2t}$$

where $t_0 = 3$. The Laplace transform of $x(t - 3)$ is $403.4/(s + 2)$, which differs from $e^{-3s}X(s)$. Thus (6.21) does not hold. The reason is that $x(t)$ is not positive time.

To stress that (6.21) holds only for positive-time signals, we often state the formula as follows: If $X(s) = \mathcal{L}[x(t)q(t)]$, then

$$\mathcal{L}[x(t - t_0)q(t - t_0)] = e^{-st_0}X(s)$$

for any $t_0 > 0$, where $q(t)$ is the step function.

EXAMPLE 6.4.3

Consider the shifted rectangular window of length L shown in Figure 1.15(b). The window can be expressed as

$$w_L(t) = q(t) - q(t - L)$$

Thus its Laplace transform is

$$W_L(s) = \frac{1}{s} - \frac{e^{-Ls}}{s} = \frac{1 - e^{-Ls}}{s} \tag{6.22}$$

It is an irrational function of s.

To conclude this section, we list in Table 6.1 some Laplace transform pairs. We see that all except one Laplace transforms in the table are strictly proper. The only exception is the Laplace transform of the impulse $\delta(t)$, which is biproper. One may wonder what type of signals will

TABLE 6.1 Laplace Transform Pairs

$x(t),\ t \geq 0$	$X(s)$
$\delta(t)$	1
1 or $q(t)$	$\dfrac{1}{s}$
t	$\dfrac{1}{s^2}$
t^k (k: positive integer)	$\dfrac{k!}{s^{k+1}}$
e^{-at} (a: real or complex)	$\dfrac{1}{s+a}$
$t^k e^{-at}$	$\dfrac{k!}{(s+a)^{k+1}}$
$\sin \omega_0 t$	$\dfrac{\omega_0}{s^2+\omega_0{}^2}$
$\cos \omega_0 t$	$\dfrac{s}{s^2+\omega_0{}^2}$
$t \sin \omega_0 t$	$\dfrac{2\omega_0 s}{(s^2+\omega_0{}^2)^2}$
$t \cos \omega_0 t$	$\dfrac{s^2-\omega_0{}^2}{(s^2+\omega_0{}^2)^2}$
$e^{-at} \sin \omega_0 t$	$\dfrac{\omega_0}{(s+a)^2+\omega_0{}^2}$
$e^{-at} \cos \omega_0 t$	$\dfrac{s+a}{(s+a)^2+\omega_0{}^2}$

yield improper Laplace transforms. Using (6.11), we have

$$\mathcal{L}\left[\frac{d\delta(t)}{dt}\right] = s\Delta(s) - \delta(0) = s$$

where we have assumed $\delta(0) = 0$. In other words, the Laplace transform of the derivative of $\delta(t)$ is s which is improper. In practice, it is impossible to generate $\delta(t)$, not to mention its derivative. Thus we rarely encounter improper Laplace transforms. This justifies once again our study of only proper rational functions of s.

6.5 INVERSE LAPLACE TRANSFORM

This section discusses the computation of the time function $x(t)$ of a given Laplace transform $X(s)$. If $X(s)$ cannot be expressed in a closed form or is an irrational function of s, then its inverse must be computed using (6.3) which can be computed using FFT. See Reference 2. We discuss in this section only the case where $X(s)$ is a proper rational function of s. The basic procedure is to express $X(s)$ as a summation of terms whose inverse Laplace transforms are available in a table such as Table 6.1. This step is called *partial fraction expansion.* We then use the table to find the inverse Laplace transform of $X(s)$.

6.5.1 Real Simple Poles

We consider first the inverse Laplace transform of $X(s)$ that has only real simple poles. We use an example to illustrate the procedure. Consider

$$X(s) = \frac{2s^3 + 2s - 10}{s^3 + 3s^2 - s - 3} = \frac{2s^3 + 2s - 10}{(s+1)(s-1)(s+3)} \tag{6.23}$$

It is biproper. The first step is to compute the roots of the denominator. If the degree of the denominator is three or higher, computing its roots by hand is difficult. However, it can easily be carried out using a computer. For example, typing `d=[1 3 -1 -3];roots(d)` in MATLAB will yield the three roots `-1, 1, -3`. Once the denominator is factored, we expand it as

$$X(s) = \frac{2s^3 + 2s - 10}{(s+1)(s-1)(s+3)} = k_0 + k_1 \frac{1}{s+1} + k_2 \frac{1}{s-1} + k_3 \frac{1}{s+3} \tag{6.24}$$

We call k_0 the *direct term,* and we call the rest of k_i *residues*. In other words, every residue is associated with a pole and the direct term is not. Once all k_i are computed, then the inverse Laplace transform of $X(s)$ is, using Table 6.1,

$$x(t) = k_0\delta(t) + k_1 e^{-t} + k_2 e^{t} + k_3 e^{-3t}$$

for $t \geq 0$. Thus the remaining problem is to compute k_i.

If $X(s)$ is a proper rational function of degree N, then we must have exactly $N+1$ parameters in (6.24) in order to solve them uniquely. If the number of parameters is N or less, then the equality may not hold. See Exercise 6.5.1 at the end of this section. If the number is $N+2$ or larger, then solutions exist but are not unique. The $X(s)$ in (6.24) has $N = 3$, and its expansion has four parameters k_i, for $i = 0 : 3$. In this case, the k_i can be solved uniquely from (6.24). There are many ways of solving k_i. We discuss in the following only the simplest method.

Equation (6.24) is an identity and holds for any selected s. If we select $s = \infty$, then the equation reduces to

$$X(\infty) = 2 = k_0 + k_1 \times 0 + k_2 \times 0 + k_3 \times 0$$

which implies

$$k_0 = X(\infty) = 2$$

Thus the direct term is simply the value of $H(\infty)$. Note that if $X(s)$ is strictly proper, then the direct term is zero.

Next we select $s = -1$, then we have

$$X(-1) = \infty = k_0 + k_1 \times \infty + k_2 \times \frac{1}{-1-1} + k_3 \times \frac{1}{-1+3}$$

This equation contains ∞ and cannot be used to solve any k_i. However, if we multiply (6.24) by $s+1$ to yield

$$X(s)(s+1) = \frac{2s^3 + 2s - 10}{(s-1)(s+3)} = k_0(s+1) + k_1 + k_2 \frac{s+1}{s-1} + k_3 \frac{s+1}{s+3}$$

and then substitute s by -1, we obtain

$$k_1 = X(s)(s+1)|_{s+1=0} = \left.\frac{2s^3+2s-10}{(s-1)(s+3)}\right|_{s=-1} = \frac{2(-1)^3+2(-1)-10}{(-2)2} = \frac{-14}{-4} = 3.5$$

Using the same procedure, we can obtain

$$k_2 = X(s)(s-1)|_{s-1=0} = \left.\frac{2s^3+2s-10}{(s+1)(s+3)}\right|_{s=1} = \frac{2+2-10}{2\cdot 4} = \frac{-6}{8} = -0.75$$

and

$$k_3 = X(s)(s+3)|_{s+3=0} = \left.\frac{2s^3+2s-10}{(s+1)(s-1)}\right|_{s=-3} = \frac{2(-3)^3+2(-3)-10}{(-2)(-4)}$$
$$= \frac{-54-6-10}{8} = \frac{-70}{8} = -8.75$$

This completes solving k_i. Thus (6.24) becomes

$$X(s) = \frac{2s^3+2s-10}{(s+1)(s-1)(s+3)} = 2 + 3.5\frac{1}{s+1} - 0.75\frac{1}{s-1} - 8.75\frac{1}{s+3} \tag{6.25}$$

and its inverse Laplace transform is

$$x(t) = 2\delta(t) + 3.5e^{-t} - 0.75e^{t} - 8.75e^{-3t}$$

for $t \geq 0$.

In conclusion, the direct term of a proper $H(s)$ can be computed as

$$k_0 = X(\infty)$$

and the residue k_i associated with simple pole p_i can be computed as

$$k_i = X(s)(s-p_i)|_{s=p_i}$$

These simple formulas can be easily employed.

EXAMPLE 6.5.1

Consider a system with transfer function $H(s) = (3s^2-2)/(s^2+3s+2)$. Compute its impulse response and step response. The input $u(t)$ and the output $y(t)$ of the system are related by $Y(s) = H(s)U(s)$. If the input is an impulse $\delta(t)$, then its Laplace transform is, using Table 6.1, $U(s) = 1$. Thus we have $Y(s) = H(s)$ and *the impulse response is simply the inverse Laplace transform of* $H(s)$. We expand $H(s)$ by partial fraction expansion as

$$H(s) = \frac{3s^2 - 2}{(s+2)(s+1)} = k_0 + k_1 \frac{1}{s+2} + k_2 \frac{1}{s+1}$$

with

$$k_0 = H(\infty) = 3$$

$$k_1 = \left. \frac{3s^2 - 2}{s+1} \right|_{s=-2} = \frac{12 - 2}{-2+1} = -10$$

$$k_2 = \left. \frac{3s^2 - 2}{s+2} \right|_{s=-1} = \frac{3 - 2}{-1+2} = 1$$

Thus we have

$$H(s) = 3 - 10 \frac{1}{s+2} + \frac{1}{s+1}$$

and the impulse response of the system is, using Table 6.1,

$$h(t) = 3\delta(t) - 10e^{-2t} + e^{-t}$$

for $t \geq 0$.

Next we compute the step response. If the input is a step function or $u(t) = q(t)$, then $U(s) = 1/s$ and the step response in the Laplace-transform domain is $Y_q(s) = H(s)(1/s)$. Its inverse Laplace transform yields the step response in the time domain. Let us carry out partial fraction expansion of $Y_q(s)$ as

$$Y_q(s) = \frac{3s^2 - 2}{(s+2)(s+1)s} = k_0 + k_1 \frac{1}{s+2} + k_2 \frac{1}{s+1} + k_3 \frac{1}{s}$$

with

$$k_0 = Y_q(\infty) = 0$$

$$k_1 = \left. \frac{3s^2 - 2}{(s+1)s} \right|_{s=-2} = \frac{12 - 2}{(-1)(-2)} = 5$$

$$k_2 = \left. \frac{3s^2 - 2}{(s+2)s} \right|_{s=-1} = \frac{3 - 2}{1(-1)} = -1$$

$$k_3 = \left. \frac{3s^2 - 2}{(s+2)(s+1)} \right|_{s=0} = \frac{-2}{2 \cdot 1} = -1$$

Thus the step response of the system is

$$y_q(t) = 5e^{-2t} - e^{-t} - 1$$

for $t \geq 0$.

From these computations, we see that the parameter or the residue associated with each real simple pole can be computed using a simple formula.

EXERCISE 6.5.1

In Example 6.5.1, if we expand

$$\frac{3s^2-2}{(s+2)(s+1)} = k_1\frac{1}{s+2} + k_2\frac{1}{s+1}$$

can you find constants k_1 and k_2 so that the equality holds? Does the expansion have enough parameters?

Answers

Using the formula, we can compute $k_1 = -10$ and $k_2 = 1$, but the equality does not hold. No.

EXERCISE 6.5.2

Find the impulse and step responses of a system with transfer function $H(s) = 3s/(2s^2+10s+12)$.

Answers

$h(t) = -3e^{-2t} + 4.5e^{-3t}$, $y_q(t) = 1.5e^{-2t} - 1.5e^{-3t}$, for $t \geq 0$.

To conclude this section, we mention that partial fraction expansions can be carried out in MATLAB by calling the function `residue`. For example, for the Laplace transform in (6.23), typing

```
n=[2 0 2 -10];d=[1 3 -1 -3];
[ki,p,k0]=residue(n,d)
```

yields `ki=-8.75 -0.75 3.5`; `p=-3 1 -1`; `k0=2`. The result is the same as the one computed in (6.25).

6.5.2 Repeated Real Poles

This subsection considers the case where $X(s)$ has repeated real poles. Again we use an example to illustrate the procedure. Consider

$$X(s) = \frac{2s^2+3s-1}{(s+1)(s+2)^3} \tag{6.26}$$

It has one simple pole at -1 and a repeated pole at -2 with multiplicity 3. We can expand it as

$$X(s) = k_0 + k_1\frac{1}{s+1} + \frac{\bar{k}_2 s^2 + \bar{k}_3 s + \bar{k}_4}{(s+2)^3} \tag{6.27}$$

The Laplace transform $X(s)$ has degree four and the expansion in (6.27) has five parameters. Thus the parameters can uniquely be determined. However, (6.27) is not a good expansion because it requires additional manipulation in order to use Table 6.1. However, if we expand it as

$$X(s) = k_0 + k_1 \frac{1}{s+1} + r_1 \frac{1}{s+2} + r_2 \frac{1}{(s+2)^2} + r_3 \frac{1}{(s+2)^3} \tag{6.28}$$

then the inverse Laplace transform of every term can be read from Table 6.1. Note that (6.28) still has five parameters. Because $X(s)$ is strictly proper, we have $k_0 = X(\infty) = 0$. Using the procedure in the preceding section, we can compute

$$k_1 = \left. \frac{2s^2 + 3s - 1}{(s+2)^3} \right|_{s=-1} = \frac{2-3-1}{1^3} = -2$$

To compute the parameters associated with the repeated pole, we multiply (6.28) by $(s+2)^3$ to yield

$$X(s)(s+2)^3 = \frac{2s^2 + 3s - 1}{s+1} = k_0(s+2)^3 + k_1 \frac{(s+2)^3}{s+1} + r_1(s+2)^2 + r_2(s+2) + r_3 \tag{6.29}$$

From the equation, we can readily obtain

$$r_3 = X(s)(s+1)^3 \Big|_{s=-2} = \left. \frac{2s^2 + 3s - 1}{s+1} \right|_{s=-2} = \frac{8-6-1}{-2+1} = -1$$

Thus the parameter associated with the highest power of the repeated pole can still be computed using a simple formula. We discuss next two methods of computing r_1 and r_2.

Method I: Solving Equations We rewrite in the following (6.28) with computed k_0, k_1, and r_3 (they have been computed using simple formulas):

$$\frac{2s^2 + 3s - 1}{(s+1)(s+2)^3} = -2 \frac{1}{s+1} + r_1 \frac{1}{s+2} + r_2 \frac{1}{(s+2)^2} - \frac{1}{(s+2)^3} \tag{6.30}$$

This is an identity and holds for every s. If we select two s to yield two equations, then the two unknowns r_1 and r_2 can be solved. Other than the poles of $X(s)$, the two s can be arbitrarily selected. We select $s = 0$ and $s = 1$, then (6.30) yields

$$\frac{-1}{8} = -2 + r_1 \frac{1}{2} + r_2 \frac{1}{4} - \frac{1}{8}$$

$$\frac{2+3-1}{2 \cdot 27} = -2 \frac{1}{2} + r_1 \frac{1}{3} + r_2 \frac{1}{9} - \frac{1}{27}$$

which can be simplified as

$$2r_1 + r_2 = 8$$

$$3r_1 + r_2 = 10$$

From these equations we can easily solve $r_1 = 2$ and $r_2 = 4$. With these r_i, (6.30) becomes

$$X(s) = -2 \frac{1}{s+1} + 2 \frac{1}{s+2} + 4 \frac{1}{(s+2)^2} - \frac{1}{(s+2)^3}$$

Thus its inverse Laplace transform is, using Table 6.1,

$$x(t) = -2e^{-t} + 2e^{-2t} + 4te^{-2t} - 0.5t^2 e^{-2t}$$

for $t \geq 0$.

Method II: Using Formula[7] If we differentiate (6.29) and then substitute $s = -2$, we can obtain

$$\begin{aligned} r_2 &= \frac{d}{ds}[X(s)(s+2)^3]\bigg|_{s=-2} = \frac{d}{ds}\left[\frac{2s^2+3s-1}{s+1}\right]\bigg|_{s=-2} \\ &= \frac{(s+1)(4s+13)-(2s^2+3s-1)\cdot 1}{(s+1)^2}\bigg|_{s=-2} \\ &= \frac{2s^2+4s+4}{(s+1)^2}\bigg|_{s=-2} = \frac{2(-2)^2+4(-2)+4}{1} = 4 \end{aligned} \tag{6.31}$$

If we differentiate (6.29) twice or, equivalently, differentiate (6.31) once and then substitute $s = -2$, we can obtain

$$\begin{aligned} r_1 &= \frac{1}{2}\frac{d^2}{ds^2}[X(s)(s+2)^3]\bigg|_{s=-2} = \frac{1}{2}\frac{d^2}{ds^2}\left[\frac{2s^2+3s-1}{s+1}\right]\bigg|_{s=-2} \\ &= \frac{1}{2}\frac{d}{ds}\left[\frac{2s^2+4s+4}{(s+1)^2}\right]\bigg|_{s=-2} = \frac{(s+1)(4s+4)-2(2s^2+4s+4)}{2(s+1)^3}\bigg|_{s=-2} \\ &= \frac{4-8}{2(-1)} = \frac{-4}{-2} = 2 \end{aligned} \tag{6.32}$$

They are the same as the ones computed in Method I. Method I requires no differentiations and may be preferable.

EXERCISE 6.5.3

Find the impulse and step responses of a system with transfer function $H(s) = (2s^2-1)/[s(s+1)]$.

Answers

$h(t) = 2\delta(t) - 1 - e^{-t}$, $y_q(t) = 1 - t + e^{-t}$, for $t \geq 0$.

6.5.3 Simple Complex Poles—Quadratic Terms

Consider

$$X(s) = \frac{2s^3+3s-1}{(s+1)(s^2+6s+25)} = \frac{2s^3+2s-1}{(s+1)(s+3+4j)(s+3-4j)} \tag{6.33}$$

Its denominator has roots -1 and $-3 \pm j4$. They are all simple, thus the procedure in Section 6.5.1 can be directly applied. We expand it as

$$X(s) = k_0 + k_1\frac{1}{s+1} + \bar{k}_2\frac{1}{s+3+4j} + \bar{k}_3\frac{1}{s+3-4j}$$

---

[7] This method may be skipped without loss of continuity.

with

$$k_0 = X(\infty) = 2$$

$$k_1 = \left.\frac{2s^3 + 3s - 1}{s^2 + 6s + 25}\right|_{s=-1} = \frac{-6}{20} = -0.3$$

$$\bar{k}_2 = \left.\frac{2s^3 + 2s - 1}{(s+1)(s+3-4j)}\right|_{s=-3-4j} = \frac{246.5e^{-j0.4}}{35.8e^{j2.68}} = 6.9e^{-j3.08}$$

$$\bar{k}_3 = \bar{k}_2^* = 6.9e^{j3.08}$$

Note that $X(s)$ has only real coefficients, thus $\bar{k}_3$ must be complex conjugate of $\bar{k}_2$. Thus the inverse Laplace transform of $X(s)$ is

$$\begin{aligned} x(t) &= 2\delta(t) - 0.3e^{-t} + 6.9e^{-j3.08}e^{-(3+j4)t} + 6.9e^{j3.08}e^{-(3-j4)t} \\ &= 2\delta(t) - 0.3e^{-t} + 6.9e^{-3t}\left[e^{-j(4t+3.08)} + e^{j(4t+3.08)}\right] \\ &= 2\delta(t) - 0.3e^{-t} + 13.8e^{-3t}\cos(4t + 3.08) \qquad (6.34) \\ &= 2\delta(t) - 0.3e^{-t} + 13.8e^{-3t}\sin(4t + 1.47) \qquad (6.35) \end{aligned}$$

where we have used $\cos(\theta - \pi/2) = \sin(\theta)$. Thus we conclude that the pair of complex-conjugate poles generates the response $k_3e^{-3t}\cos(4t + k_4)$ or $k_3e^{-3t}\sin(4t + \hat{k}_4)$ for some constants k_3, k_4, and $\hat{k}_4$. The real part of the complex-conjugate poles governs the envelope of oscillation whose frequency is determined by the imaginary part.

To conclude this section, we mention that if we factor $s^2 + 6s + 25$ as $(s+3)^2 + 4^2$, then we can also expand (6.33) as

$$X(s) = \frac{2s^3 + 3s - 1}{(s+1)(s^2 + 6s + 25)} = k_0 + k_1\frac{1}{s+1} + k_2\frac{4}{(s+3)^2 + 4^2} + k_3\frac{s+3}{(s+3)^2 + 4^2} \qquad (6.36)$$

Once k_i are computed, then the inverse Laplace transform of (6.36) is

$$x(t) = k_0\delta(t) + k_1e^{-t} + k_2e^{-3t}\sin 4t + k_3e^{-3t}\cos 4t \qquad (6.37)$$

In this expansion, k_i can be computed without using complex numbers.

6.5.4 Why Transfer Functions Are Not Used in Computer Computation

We give in this section the reasons that transfer functions are not used in computer computation of system responses. In the first place, if the Laplace transform of an input cannot be expressed as a rational function of s as is often the case in practice, then the method discussed so far cannot be used. Even if the Laplace transform of an input is a rational function of s, the method is still not desirable as we explain next.

Every computer computation involves two issues: complexity of the algorithm and its numerical sensitivity. It turns out that the Laplace transform method is poor on these two accounts. To use the Laplace transform to compute a system response requires computing the Laplace transform of an input, computing the roots of a polynomial, carrying out a partial fraction expansion, and a table lookup. This programming is complicated. More seriously, computing the roots of a polynomial is very sensitive to parameter variations. By this we mean that if some coefficients of a polynomial change slightly, the roots may change greatly. For example, the roots of

$$D(s) = s^4 + 7s^3 + 18s^2 + 20s + 8$$

are -1, -2, -2, and -2. However, the roots of

$$D'(s) = s^4 + 7.001s^3 + 17.999s^2 + 20s + 8$$

are -0.998, -2.2357, and $-1.8837 \pm 0.1931j$. We see that the two coefficients change less than 0.1%, but all its roots change greatly. In conclusion, even for those inputs whose Laplace transforms are rational functions of s, transfer functions are not used in computing their responses.

Consider the transfer function

$$H(s) = \frac{2s^2 + 3s - 1}{s^3 + 7s^2 + 18s^2 + 20s + 8}$$

To compute its step response in MATLAB, all we do is type

```
n=[2 3 -1];d=[1 7 18 20 8];
step(n,d)
```

Then its step response will appear on the monitor. This does not mean that the response is computed using the transfer function. Internally, the program first transforms the transfer function into a state-space (ss) equation, and then it uses the ss equation to compute its response. This will be discussed in the next chapter.

6.5.5 A Study of Automobile Suspension Systems

Computer simulation provides a powerful way of study and design of systems. In this subsection, we use MATLAB to study the automobile suspension system discussed in Figure 3.7 and described by the differential equation, as derived in (3.35),

$$m\ddot{y}(t) + f\dot{y}(t) + ky(t) = u(t)$$

where m, f, and k are the mass, viscous friction or *damping* coefficient, and spring constant, respectively. Applying the Laplace transform and assuming zero initial conditions yield

$$ms^2Y(s) + fsY(s) + kY(s) = U(s)$$

Thus the transfer function of the system is

$$H(s) = \frac{Y(s)}{U(s)} = \frac{1}{ms^2 + fs + k} \tag{6.38}$$

a strictly proper rational function of degree 2.

When an automobile hits a pothole or curb, the impact is very short in time duration but has large energy. Thus we may approximate the impact as an impulse with some finite weight. We now study the effect of the damping coefficient on the impulse response of the system. Let us

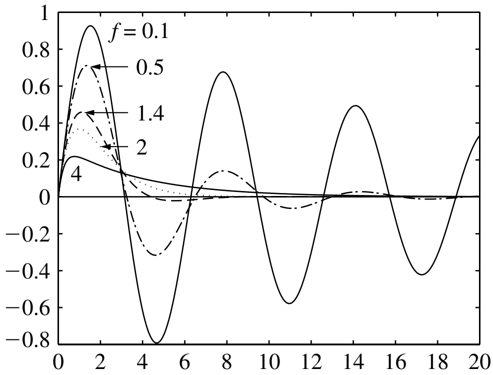


Figure 6.4 Impulse responses of (6.38) for $k = m = 1$ and $f = 4, 2, 1.4, 0.5, 0.1$.

assume $m = 1$ and $k = 1$. Figure 6.4 shows the impulse responses of (6.38) for $f = 4, 2, 1.4, 0.5$ and 0.1. It is generated in MATLAB by typing

```
t=0:0.01:20;n=1;
d1=[1 4 1];d2=[1 2 1];d3=[1 1.4 1];d4=[1 0.5 1];d5=[1 0.1 1];
y1=impulse(n,d1,t);y2=impulse(n,d2,t);y3=impulse(n,d3,t);
y4=impulse(n,d4,t);y5=impulse(n,d5,t);
plot(t,y1,t,y2,':',t,y3,'--',t,y4,'-.',t,y5,[0 20],[0 0])
```

The first line specifies the time interval to be computed from 0 to 20 with time increment selected as 0.01, and the numerator coefficient 1. The second lines are the coefficients of the denominator for $m = 1$, $k = 1$, and $f = 4, 2, 1.4, 0.5, 0.1$. The next two lines use the MATLAB function `impulse` to generate the impulse responses. The last line generates the graph. Note that the last pair `[0 20],[0 0]` in the function `plot` generates the horizontal line through 0. We see that all responses approach zero eventually.[8] Some responses oscillate, some do not. The largest negative value in an impulse response is called an *undershoot*. If an impulse response does not go to negative values, its undershoot is zero.

An acceptable automobile suspension system can be specified as follows: (1) Its response time (the time for the impulse response to return to less than 1% of its peak magnitude) should be less than, say, 8 seconds. (2) Its undershoot should be less than, say, 5% of its peak magnitude. These are *time-domain* specifications. This type of time-domain specifications is widely used in the design of control systems. See Reference 5.

Now we consider the impulse responses in Figure 6.4. The impulse responses for $f = 0.1$ and 0.5 have large response times and large undershoots. Thus they are not acceptable. The impulse response for $f = 4$ has no undershoot but a large response time. Thus it is neither acceptable. The impulse responses for $f = 1.4$ and 2 meet both specifications and are acceptable. Therefore, there is a wide range of f in which the system will be acceptable.

---

[8]This is true only if $f > 0$. If $f = 0$, the response is a sustained oscillation and will not approach zero. If $f < 0$, the response will grow unbounded.

A new shock absorber is designed to have, relative to m and k, a large damping coefficient. It should not be unnecessarily large because it will cause a large response time. The ride will also be stiff. As the shock absorber wears away with time, the damping coefficient becomes smaller. When it reaches the stage that the automobile will vibrate or oscillate greatly when it hits a pothole, it is time to have the shock absorber replaced. We see that using the simple model in (6.38) and computer simulation, we can study the effect of the damping coefficient.

Different automobile companies may have different specifications. Different drivers may also have different preferences. Thus in practice, the set of specifications is not absolute and there is a large leeway in designing a suspension system. This is often the case in designing most practical systems.

To conclude this subsection, we stress once again that the impulse responses in Figure 6.4 are not obtained by taking the inverse Laplace transforms of (6.38). They are obtained using state-space (ss) equations and without generating impulses, as we will discuss at the end of Section 7.4.1.

6.6 SIGNIFICANCE OF POLES AND ZEROS

Although transfer functions are not used in computer computation, they are most convenient in developing general properties of systems. We discuss first the significance of poles and zeros. Consider a transfer function $H(s)$ expressed in the zero/pole/gain form such as

$$H_1(s) = \frac{4(s-3)^2(s+1+j2)(s+1-j2)}{(s+1)(s+2)^2(s+0.5+j4)(s+0.5-j4)} \tag{6.39}$$

Poles and zeros are often plotted on the complex s-plane with small circles and crosses, respectively, as shown in Figure 6.5. The transfer function in (6.39) has three simple poles at -1, $-0.5 \pm j4$ and has a repeated pole at -2 with multiplicity 2, denoted by [2] on the graph. It has simple zeros at $-1 \pm j2$ and a repeated zero at 3 with multiplicity 2, denoted by [2]. Note that the gain 4 does not appear on the graph.

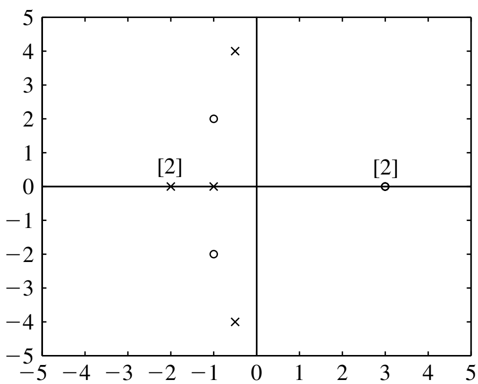


Figure 6.5 Poles and zeros.

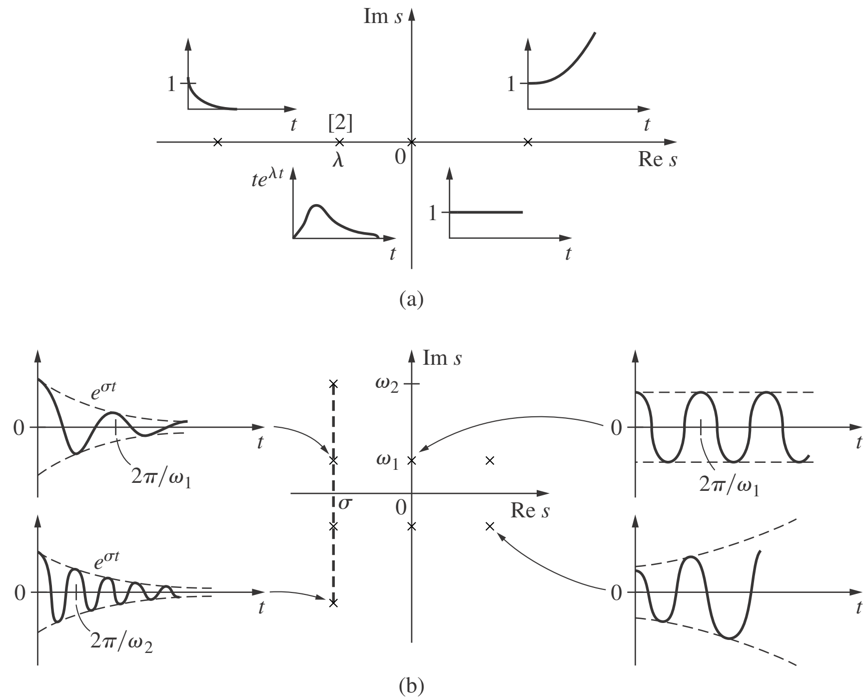


Figure 6.6 Responses of poles.

The complex s-plane can be divided into three parts: left half-plane (LHP), right half-plane (RHP), and the imaginary axis or $j\omega$-axis. We call a half plane closed if it includes the $j\omega$-axis; open if it excludes the $j\omega$-axis. All poles of $H_1(s)$ lie inside the open LHP. The simple zeros $-1 \pm j2$ lie inside the open LHP, whereas the repeated zero 3 lies inside the open RHP.

We plot in Figures 6.6(a) and 6.6(b) time responses of some real poles and some pairs of complex-conjugate poles. We see that the time response of a real simple pole in the open LHP approaches monotonically zero as $t \to \infty$. If it is repeated such as $1/(s+2)^2$, then its time response is te^{-2t}. As t increases, its time response grows and then decreases to zero as shown in the left bottom corner of Figure 6.6(a). This can be verified using l'Hôpital's rule as

$$\lim_{t\to\infty} te^{-2t} = \lim_{t\to\infty} \frac{t}{e^{2t}} = \lim_{t\to\infty} \frac{1}{2e^{2t}} = 0$$

The time response of a pair of complex-conjugate poles in the open LHP approaches oscillatorily zero as $t \to \infty$. In conclusion, time responses of poles, simple or repeated, in the open LHP approach zero as $t \to \infty$. On the other hand, time responses of poles, simple or repeated, in the open RHP grow unbounded as $t \to \infty$.

The situation for poles on the $j\omega$-axis is slightly more complex. The time response of a simple pole at $s = 0$ is constant for all $t \geq 0$. If it is repeated such as $1/s^2$ or $1/s^3$, then its time response is t or $0.5t^2$, for $t \geq 0$. It grows unbounded. The time response of a pair of pure imaginary poles is a sustained oscillation. If the pair is repeated, then its response grows oscillatorily to infinity.

The preceding discussion is summarized in the following:

Pole Location	Response as $t \to \infty$
LHP, simple or repeated	0
RHP, simple or repeated	∞ or $-\infty$
$j\omega$-axis, simple	constant or sustained oscillation
$j\omega$-axis, repeated	∞ or $-\infty$

We see that the response of a pole approaches 0 as $t \to \infty$ *if and only if* the pole lies inside the open LHP.

Let us consider the following transfer functions:

$$H_2(s) = \frac{-12(s-3)(s^2+2s+5)}{(s+1)(s+2)^2(s^2+s+16.25)} \tag{6.40}$$

$$H_3(s) = \frac{-20(s^2-9)}{(s+1)(s+2)^2(s^2+s+16.25)} \tag{6.41}$$

$$H_4(s) = \frac{60(s+3)}{(s+1)(s+2)^2(s^2+s+16.25)} \tag{6.42}$$

$$H_5(s) = \frac{180}{(s+1)(s+2)^2(s^2+s+16.25)} \tag{6.43}$$

They all have the same denominator as (6.39). In addition, we have $H_i(0) = 180/65 = 2.77$, for $i = 1:5$. Even though they have different zeros, their step responses, in the s-domain, are all of the form

$$Y_i(s) = H_i(s)\frac{1}{s} = \frac{k_1}{s+1} + \frac{k_2}{s+2} + \frac{k_3}{(s+2)^2} + \frac{\bar{k}_4 s + \bar{k}_5}{(s+0.5)^2+4^2} + \frac{k_6}{s}$$

with

$$k_6 = H_i(s)\frac{1}{s}\cdot s\Big|_{s=0} = H_i(0) = 2.77$$

Thus their step responses in the time domain are all of the form

$$y_i(t) = k_1 e^{-t} + k_2 e^{-2t} + k_3 t e^{-2t} + k_4 e^{-0.5t}\sin(4t + k_5) + 2.77 \tag{6.44}$$

This form is determined entirely by the poles of $H_i(s)$ and $U(s)$. The pole -1 generates the term $k_1 e^{-t}$. The repeated pole -2 generates the terms $k_2 e^{-2t}$ and $k_3 t e^{-2t}$. The pair of complex poles generates the term $k_4 e^{-0.5t}\sin(4t + k_5)$ as shown in (6.35). The input is a step function with Laplace transform $U(s) = 1/s$. Its pole at $s = 0$ generates a step function with amplitude $H_i(0) = 2.77$. We see that the zeros of $H_i(s)$ do not play any role in determining the form in (6.44). The zeros, however, do affect k_i. Different set of zeros yields different set of k_i. Because all poles of $H_i(s)$ lie inside the open LHP, their time responses all approach zero as $t \to \infty$. Thus $y_i(t)$ approaches 2.77 as $t \to \infty$.

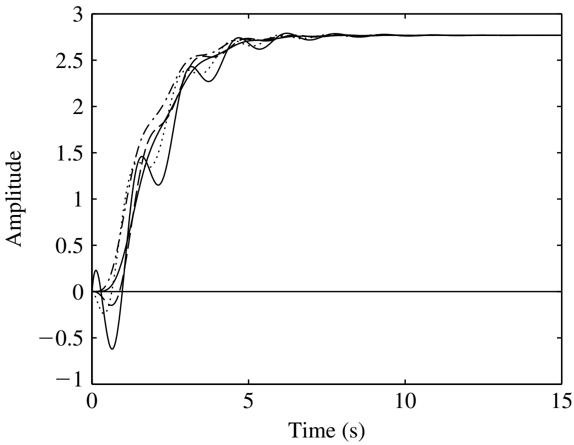


Figure 6.7 Step responses of the transfer functions in (6.39) through (6.43).

We plot the step responses of $H_i(s)$, for $i = 1 : 5$ in Figure 6.7. They are obtained in MATLAB by typing

```
d=[1 6 29.25 93.25 134 65];
n1=[4 -16 8 -48 180];n2=[-12 12 12 180];
n3=[-20 0 180];n4=[60 180];n5=180;
t=0:0.01:15;
y1=step(n1,d,t);y2=step(n2,d,t);y3=step(n3,d,t);
y4=step(n4,d,t);y5=step(n5,d,t);
plot(t,y1,t,y2,':',t,y3,'--',t,y4,'-.',t,y5)
```

All responses approach 2.77 as $t \to \infty$, but the responses right after the application of the input are all different even though their step responses are all of the form in (6.44). This is due to different set of k_i. In conclusion, poles dictate the general form of responses; zeros affect only the parameters k_i. Thus we conclude that zeros play a lesser role than poles in determining responses of systems.

6.7 STABILITY

This section introduces the concept of stability for systems. In general, if a system is not stable, it may burn out or saturate (for an electrical system), disintegrate (for a mechanical system), or overflow (for a computer program). Thus every system designed to process signals are required to be stable. Let us give a formal definition.

DEFINITION 6.1 A system is BIBO (bounded-input bounded-output) stable or, simply, stable if *every* bounded input excites a bounded output. Otherwise, the system is said to be unstable.

A signal is bounded if it does not grow to ∞ or $-\infty$. In other words, a signal $u(t)$ is bounded if there exists a constant M such that $|u(t)| \leq M < \infty$ for all t. We first give an example to illustrate the concept.[9]

EXAMPLE 6.7.1

Consider the network shown in Figure 6.8(a). The input $u(t)$ is a current source; the output $y(t)$ is the voltage across the capacitor. The impedances of the inductor and capacitor are respectively s and $1/s$. The impedance of their parallel connection is $s(1/s)/(s + 1/s)$. Thus the input and output of the network are related by

$$Y(s) = \frac{s(1/s)}{s + (1/s)} U(s) = \frac{s}{s^2 + 1} U(s)$$

If we apply a step input ($u(t) = 1$, for $t \geq 0$), then its output is

$$Y(s) = \frac{s}{s^2 + 1} \frac{1}{s} = \frac{1}{s^2 + 1}$$

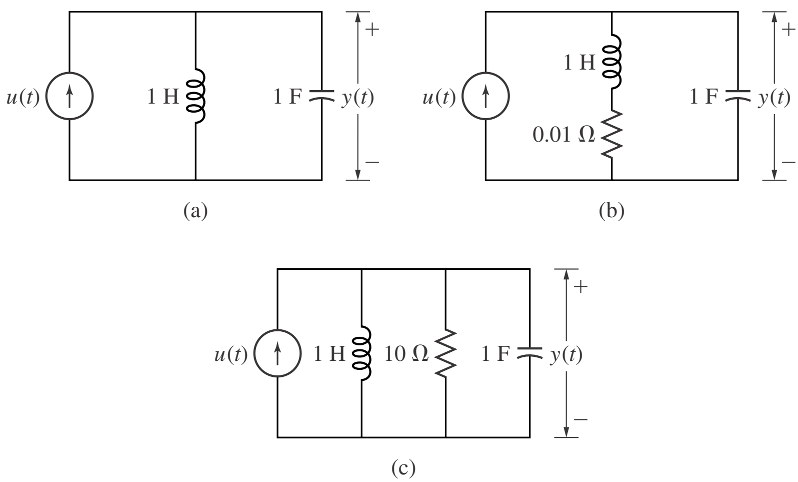


Figure 6.8 (**a**) Unstable network. (**b**) and (**c**) Stable networks.

---

[9]The BIBO stability is defined for zero-state or forced responses. Some texts introduce the concept of stability using a pendulum and an inverted pendulum or a bowl and an inverted bowl. Strictly speaking, such stability is defined for zero-input or natural responses and should be defined for a so-called equilibrium state. This type of stability is called *asymptotic stability* or *marginal stability*. See Reference 3. A pendulum is marginally stable if there is no friction, and it is asymptotically stable if there is friction. An inverted pendulum is neither marginally nor asymptotically stable.

which implies $y(t) = \sin t$. This output is bounded. If we apply $u(t) = \sin \omega_0 t$, with $\omega_0 \neq 1$, then the output is

$$Y(s) = \frac{s}{s^2+1}\frac{\omega_0}{s^2+\omega_0{}^2} = \frac{\omega_0 s\left[\left(s^2+\omega_0{}^2\right)-(s^2+1)\right]}{\left(\omega_0{}^2-1\right)(s^2+1)\left(s^2+\omega_0{}^2\right)}$$

$$= \frac{\omega_0}{\omega_0{}^2-1}\left[\frac{s}{s^2+1} - \frac{s}{s^2+\omega_0{}^2}\right]$$

which implies, using Table 6.1,

$$y(t) = \frac{\omega_0}{\omega_0{}^2-1}[\cos t - \cos \omega_0 t]$$

This output is bounded for any $\omega_0 \neq 1$. Thus the outputs excited by the bounded inputs $u(t) = 1$ and $u(t) = \sin \omega_0 t$ with $\omega_0 \neq 1$ are all bounded. Even so, we still cannot conclude that the network is BIBO stable because we have not yet checked all possible bounded inputs.

Now let us apply the bounded input $u(t) = 2\sin t$. Then the output is

$$Y(s) = \frac{s}{s^2+1}\frac{2}{s^2+1} = \frac{2s}{(s^2+1)^2}$$

or, using Table 6.1,

$$y(t) = t \sin t$$

which grows unbounded. In other words, the bounded input $u(t) = 2\sin t$ excites an unbounded output. Thus the network is not BIBO stable.

EXAMPLE 6.7.2

Consider the op-amp circuits in Figures 2.17(a) and 2.17(b). Because $v_o(t) = v_i(t)$, if an input $v_i(t)$ is bounded, so is the output. Thus the two circuits are BIBO stable. Likewise, any amplifier with a finite gain is BIBO stable.

Except for memoryless systems, we cannot use Definition 6.1 to conclude the stability of a system because there are infinitely many bounded inputs to be checked. In fact, stability of a system can be determined from its mathematical descriptions without applying any input. In other words, stability is a property of a system and is independent of applied inputs. The output of a stable system excited by any bounded input must be bounded; its output is generally unbounded if the applied input is unbounded.

THEOREM 6.1 An LTI system with impulse response $h(t)$ is BIBO stable if and only if $h(t)$ is absolutely integrable in $[0, \infty)$; that is,

$$\int_0^\infty |h(t)|\, dt \leq M < \infty$$

for some constant M.

Proof We first show that the system is BIBO stable under the condition. Indeed, the input $u(t)$ and output $y(t)$ of the system are related by

$$y(t) = \int_0^t h(\tau)u(t-\tau)\,d\tau$$

If $u(t)$ is bounded or $|u(t)| \le M_1$ for all $t \ge 0$, then we have

$$\begin{aligned} |y(t)| &= \left| \int_0^t h(\tau)u(t-\tau)\,d\tau \right| \le \int_0^t |h(\tau)||u(t-\tau)|\,d\tau \\ &\le M_1 \int_0^t |h(\tau)|\,d\tau \le M_1 M \end{aligned}$$

for all $t \ge 0$. This shows that if $h(t)$ is absolutely integrable, then the system is BIBO stable.

Next we show that if $h(t)$ is not absolutely integrable, then there exists a bounded input that will excite an unbounded output. If $h(t)$ is not absolutely integrable, then for any arbitrarily large M_2, there exists a t_1 such that

$$\int_0^{t_1} |h(\tau)|\,d\tau \ge M_2$$

Let us select an input as, for $t_1 \ge \tau \ge 0$,

$$u(t_1 - \tau) = \begin{cases} 1 & \text{for } h(\tau) \ge 0 \\ -1 & \text{for } h(\tau) < 0 \end{cases}$$

For this bounded input, the output $y(t)$ at $t = t_1$ is

$$y(t_1) = \int_0^{t_1} h(\tau)u(t_1-\tau)\,d\tau = \int_0^{t_1} |h(\tau)|\,d\tau \ge M_2$$

This shows that if $h(t)$ is not absolutely integrable, then there exists a bounded input that will excite an output with an arbitrarily large magnitude. Thus the system is not stable. This establishes the theorem.

Let us apply the theorem to the system in Figure 6.8(a). Its transfer function is computed in Example 6.7.1 as $H(s) = s/(s^2+1)$. Thus its impulse response is, using Table 6.1, $h(t) = \cos t$, for $t \ge 0$, and is plotted in Figure 6.9(a). The cosine function has period 2π and we have

$$\int_0^{\infty} \cos t\,dt \le \int_0^{\pi/4} \cos t\,dt = \sin t \Big|_{t=0}^{\pi/4} = 0.707$$

This is so because the positive and negative parts cancel out completely and the largest possible value is as shown. Likewise we have

$$-0.707 \le \int_0^{\infty} \cos t\,dt \quad \text{and} \quad \left| \int_0^{\infty} \cos t\,dt \right| \le 0.707 < \infty$$

Thus the impulse response $h(t) = \cos t$ is said to be *integrable* in $[0, \infty)$. However, this is not the condition in Theorem 6.1. To check stability, we first take the absolute value of $h(t) = \cos t$

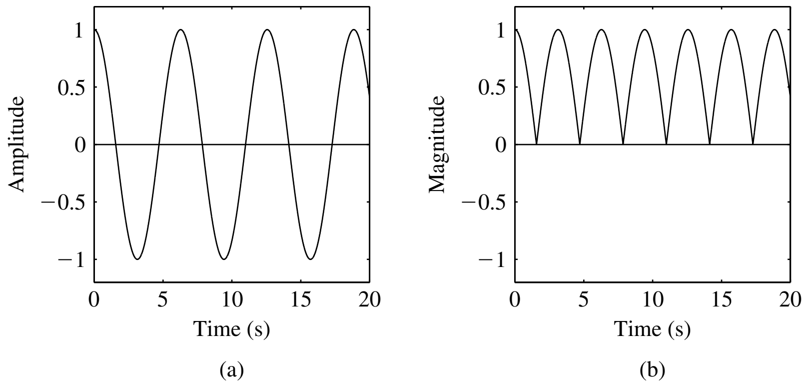


Figure 6.9 (**a**) $\cos t$. (**b**) $|\cos t|$.

as shown in Figure 6.9(b) and then take its integration. Clearly, we have

$$\int_0^\infty |\cos t|\, dt = \infty$$

because there is no more cancellation. Thus $\cos t$ is *integrable* but not *absolutely integrable,* and the system in Figure 6.8(a) is not stable as we demonstrated in Example 6.7.1.

We mention that the condition $h(t) \to 0$ as $t \to \infty$ is not sufficient for $h(t)$ to be absolutely integrable.[10] For example, the function

$$h(t) = \frac{1}{t+1}$$

for $t \geq 0$, approaches 0 as $t \to \infty$ but is not absolutely integrable:

$$\int_0^\infty |h(t)|\, dt = \int_0^\infty \frac{1}{t+1}\, dt = \log(t+1)\bigg|_0^\infty = \infty$$

This is so because $h(t)$ does not approach zero fast enough. If a function approaches 0 sufficiently fast, such as e^{-at} or te^{-at} for $a > 0$, then it is absolutely integrable. Indeed, we have

$$\int_0^\infty e^{-at}\, dt = \frac{1}{-a} e^{-at}\bigg|_0^\infty = \frac{1}{-a}(0-1) = \frac{1}{a}$$

and, using an integration table,

$$\int_0^\infty te^{-at}\, dt = \frac{1}{(-a)^2}[(-at-1)e^{-at}]\bigg|_0^\infty = \frac{1}{a^2}[0+1] = \frac{1}{a^2}$$

Proceeding forward, we can show that $t^k e^{-at}$, for any positive integer k and any $a > 0$, is absolutely integrable. It approaches zero *exponentially.*

---

[10]Neither is it necessary. See the function in Figure 5.1 of Reference 3.

THEOREM 6.2 An LTI lumped system with proper rational transfer function $H(s)$ is stable if and only if every pole of $H(s)$ has a negative real part or, equivalently, all poles of $H(s)$ lie inside the open left half s-plane.

If $H(s)$ has one or more poles in the open RHP, then its impulse response grows unbounded and is thus not absolutely integrable. If it has poles on the $j\omega$-axis, then its impulse response contains a constant or a sustained oscillation and is thus not absolutely integrable. In conclusion, if $H(s)$ has one or more poles in the closed RHP, then the system is not stable. On the other hand, if every pole of $H(s)$ has a negative real part, then its time response approaches zero exponentially and is absolutely integrable. Thus we conclude that the system is stable.

COROLLARY 6.2 An LTI lumped system with impulse response $h(t)$ is BIBO stable if and only if $h(t) \to 0$ as $t \to \infty$.

The impulse response of a system is the inverse Laplace transform of its transfer function. If the system is lumped, its transfer function $H(s)$ is a rational function. For simplicity, we assume $H(s)$ to have only simple poles. Then it can be expanded as

$$H(s) = \sum_i \frac{k_1}{s - p_i} + k_0$$

and

$$h(t) = \sum_i k_i e^{p_i t} + k_0 \delta(t)$$

If $H(s)$ is stable, all poles lie inside the open LHP and all $e^{p_i t}$ approach zero as $t \to \infty$. Note that $\delta(t) = 0$, for $t > 0$. Thus we have $h(t) \to 0$ as $t \to \infty$. If $H(s)$ is not stable, it has at least one pole with zero or negative real part and its time response will not approach 0 as $t \to \infty$. Consequently, so is $h(t)$. This establishes Corollary 6.2.

The difference between Theorem 6.1 and Corollary 6.2 is that the former is applicable to distributed and lumped systems whereas the latter is applicable only to lumped systems. For example, the system with impulse response $h(t) = 1/(1 + t)$ has an irrational transfer function. Thus it is a distributed system and Theorem 6.2 and Corollary 6.2 are not applicable. We encounter mostly lumped LTI systems, thus we use mostly Theorem 6.2 and its corollary to check their stabilities. For example, consider the network studied in Example 6.7.1. Its transfer function $s/(s^2 + 1)$ has poles $\pm j1$. The real parts of the poles are 0, thus the system is not BIBO stable. It also follows from the fact that the poles are on the $j\omega$-axis, not inside the open LHP. It also follows from Corollary 6.2 because its impulse response $\cos t$ does not approach zero as $t \to \infty$. We mention that stability is independent of the zeros of $H(s)$. A system can be stable with zeros anywhere (inside the RHP or LHP or on the $j\omega$-axis).

If a system is known to be LTI and lumped, then its stability can easily be determined by measurement. We apply an arbitrary input to excite the system and then remove the input (for example, $u(t) = 1$ for $0 \le t \le 1$), then the system is stable if and only if the output approaches zero as $t \to \infty$. Indeed, if $H(s)$ has only simple poles p_i, then the response of the system after the input is removed is of the form

$$y(t) = \sum_i k_i e^{p_i t}$$

with all k_i nonzero. The chance for some k_i to be zero is very small. See Problem 6.17. In other words, all poles of the system will be excited by the input. Now the output $y(t)$ approaches zero if and only if all p_i have negative real parts or the system is stable. This establishes the assertion.

Resistors with positive resistances, inductors with positive inductances, and capacitors with positive capacitances are called *passive elements*. Resistors dissipate energy. Although inductors and capacitors can store energy, they cannot generate energy. Thus when an input is removed from any RLC network, the energy stored in the inductors and capacitors will eventually dissipate in the resistors and consequently the response eventually approaches zero. Thus all RLC networks are stable. Note that the LC network in Figure 6.8(a) is a model. In fact, every practical inductor has a small resistance in series as shown in Figure 6.8(b), and every practical capacitor has a large resistance in parallel as shown in Figure 6.8(c). Thus all practical RLC networks are stable, and the stability is not an issue. Consequently, before the advent of active elements such as op amps, stability is not studied in passive RLC networks. It is studied only in designing feedback systems. This practice remains to this date. See, for example, References 8, 10–12, and 23.

6.7.1 Routh Test

Consider a system with proper rational transfer function $H(s) = N(s)/D(s)$. We assume that $N(s)$ and $D(s)$ are coprime or have no common factors. Then the poles of $H(s)$ are the roots of $D(s)$. If $D(s)$ has degree three or higher, computing its roots by hand is not simple. This, however, can be carried out using a computer. For example, consider

$$D(s) = 2s^5 + s^4 + 7s^3 + 3s^2 + 4s + 2 \tag{6.45}$$

Typing

```
d=[2 1 7 3 4 2];
roots(d)
```

in MATLAB yields $-0.0581 \pm j1.6547$, $0.0489 \pm j0.8706$, -0.4799. It has a real root and two pairs of complex conjugate roots. One pair has a positive real part 0.0489. Thus any transfer function with (6.45) as its denominator is not BIBO stable. Of course, if we use the MATLAB function tf2zp to transform $H(s)$ into the zero/pole/gain form, we can also check its stability from its pole locations.

We next discuss a method of checking stability of $H(s) = N(s)/D(s)$ without computing its poles or, equivalently, the roots of $D(s)$.[11] We call $D(s)$ a *stable polynomial* if all its roots have negative real parts. The Routh test to be introduced checks whether a polynomial is stable or not.

[11] The remainder of this subsection may be skipped without loss of continuity.

TABLE 6.2 The Routh Table

	s^6	a_0	a_2	a_4	a_6	
$k_1 = a_0/a_1$	s^5	a_1	a_3	a_5		$[b_0\ \ b_1\ \ b_2\ \ b_3] =$ (1st row)-k_1(2nd row)
$k_2 = a_1/b_1$	s^4	b_1	b_2	b_3		$[c_0\ \ c_1\ \ c_2] =$ (2nd row)-k_2(3rd row)
$k_3 = b_1/c_1$	s^3	c_1	c_2			$[d_0\ \ d_1\ \ d_2] =$ (3rd row)-k_3(4th row)
$k_4 = c_1/d_1$	s^2	d_1	d_2			$[e_0\ \ e_1] =$ (4th row)-k_4(5th row)
$k_5 = d_1/e_1$	s	e_1				$[f_0\ \ f_1] =$ (5th row)-k_5(6th row)
	s^0	f_1				

We use the following $D(s)$ to illustrate the procedure:

$$D(s) = a_0 s^6 + a_1 s^5 + a_2 s^4 + a_3 s^3 + a_4 s^2 + a_5 s + a_6 \qquad \text{with } a_0 > 0 \tag{6.46}$$

We call a_0 the leading coefficient. If the leading coefficient is negative, we apply the procedure to $-D(s)$. Because $D(s)$ and $-D(s)$ have the same set of roots, if $-D(s)$ is stable, so is $D(s)$. The polynomial in (6.46) has degree 6 and seven coefficients a_i, $i = 0 : 6$. We use the coefficients to form the first two rows of the table, called the Routh table, in Table 6.2. They are placed, starting from a_0, alternatively in the first and second rows. Next we compute $k_1 = a_0/a_1$, the ratio of the first entries of the first two rows. We then subtract from the first row the product of the second row and k_1. The result $[b_0\ \ b_1\ \ b_2\ \ b_3]$ is placed at the right-hand side of the second row, where

$$b_0 = a_0 - k_1 a_1 = 0, \qquad b_1 = a_2 - k_1 a_3, \qquad b_2 = a_4 - k_1 a_5, \qquad b_3 = a_6 - k_1 \cdot 0 = a_6$$

Note that $b_0 = 0$ because of the way k_1 is defined. We discard b_0 and place $[b_1\ \ b_2\ \ b_3]$ in the third row as shown in the table. The fourth row is obtained using the same procedure from its previous two rows. That is, we compute $k_2 = a_1/b_1$, the ratio of the first entries of the second and third rows. We subtract from the second row the product of the third row and k_2. The result $[c_0\ \ c_1\ \ c_2]$ is placed at the right-hand side of the third row. We discard $c_0 = 0$ and place $[c_1\ \ c_2]$ in the fourth row as shown. We repeat the process until the row corresponding to $s^0 = 1$ is computed. If the degree of $D(s)$ is N, the table contains $N + 1$ rows.

We discuss the size of the Routh table. If $N = \deg D(s)$ is even, the first row has one more entry than the second row. If N is odd, the first two rows have the same number of entries. In both cases, the number of entries decreases by one at odd powers of s. For example, the numbers of entries in the rows of s^5, s^3, s decrease by one from their previous rows. The last entries of all rows corresponding to even powers of s are all the same. For example, we have $a_6 = b_3 = d_2 = f_1$ in Table 5.2.

THEOREM 6.3 A polynomial with a positive leading coefficient is a stable polynomial if and only if every entry in the Routh table is positive. If a zero or a negative number appears in the table, then $D(s)$ is not a stable polynomial.

The proof of this theorem is beyond the scope of this text. We discuss only its employment. A necessary condition for $D(s)$ to be stable is that all its coefficients are positive. If $D(s)$ has missing terms (coefficients are zero) or negative coefficients, then it is not a stable polynomial. For example, the polynomials

$$D(s) = s^4 + 2s^2 + 3s + 10$$

and

$$D(s) = s^4 + 3s^3 - s^2 + 2s + 1$$

are not stable polynomials. On the other hand, a polynomial with all positive coefficients is not necessarily stable. The polynomial in (6.45) has all positive coefficients but is not stable as discussed earlier. We can also verify this by applying the Routh test. We use the coefficients of

$$D(s) = 2s^5 + s^4 + 7s^3 + 3s^2 + 4s + 2$$

to form

$$\begin{array}{rr|ccc}
 & s^5 & 2 & 7 & 4 \\
k_1 = 2/1 & s^4 & 1 & 3 & 2 \qquad [0 \quad 1 \quad 0] \\
 & s^3 & 1 & 0 &
\end{array}$$

A zero appears in the table. Thus we conclude that the polynomial is not stable. There is no need to complete the table.

EXAMPLE 6.7.3

Consider the polynomial

$$D(s) = 4s^5 + 2s^4 + 14s^3 + 6s^2 + 8s + 3$$

We form

$$\begin{array}{rr|cccl}
 & s^5 & 4 & 14 & 8 & \\
k_1 = 2/1 & s^4 & 2 & 6 & 3 & [0 \quad 2 \quad 2] \\
k_2 = 2/2 & s^3 & 2 & 2 & & [0 \quad 4 \quad 3] \\
k_3 = 2/4 & s^2 & 4 & 3 & & [0 \quad 0.5] \\
k_4 = 4/0.5 & s^1 & 0.5 & & & [0 \quad 3] \\
 & s^0 & 3 & & &
\end{array}$$

Every entry in the table is positive, thus the polynomial is a stable polynomial.

EXERCISE 6.7.1

Which of the following are stable polynomials?

(a) $5s^4 + 3s^3 + s + 10$

(b) $s^4 + s^3 + s^2 + s + 1$

(c) $s^4 + 2.5s^3 + 2.5s^2 + s + 0.5$

(d) $-s^5 - 3s^4 - 10s^3 - 12s^2 - 7s - 3$

Answers

No. No. Yes. Yes.

EXERCISE 6.7.2

Find the transfer functions of the networks in Figures 6.8(b) and 6.8(c) and then verify that the two networks are stable.

Answers

$(s + 0.01)/(s^2 + 0.01s + 1)$, $10s/(10s^2 + s + 10)$

6.8 FREQUENCY RESPONSES

Before discussing the implication of stability, we introduce the concept of frequency responses. Consider a transfer function $H(s)$. Its values along the $j\omega$-axis—that is, $H(j\omega)$—is called the *frequency response*. In general, $H(j\omega)$ is complex-valued and can be expressed in polar form as

$$H(j\omega) = A(\omega)e^{j\theta(\omega)} \tag{6.47}$$

where $A(\omega)$ and $\theta(\omega)$ are real-valued functions of ω and $A(\omega) \geq 0$. We call $A(\omega)$ the *magnitude response* and call $\theta(\omega)$ the *phase response*. If $H(s)$ contains only real coefficients, then

$$A(\omega) = A(-\omega) \qquad \text{(even)} \tag{6.48}$$

and

$$\theta(\omega) = -\theta(-\omega) \qquad \text{(odd)} \tag{6.49}$$

This follows from $H(j\omega) = H^*(-j\omega)$, which imples

$$A(\omega)e^{j\theta(\omega)} = \left[A(-\omega)e^{j\theta(-\omega)}\right]^* = A(-\omega)e^{-j\theta(-\omega)}$$

Thus if $H(s)$ has only real coefficients, then its magnitude response is even and its phase response is odd. Because of these properties, we often plot magnitude and phase responses only for $\omega \geq 0$ instead of for all ω in $(-\infty, \infty)$. We give an example.

EXAMPLE 6.8.1

Consider $H(s) = 2/(s + 2)$. We first write

$$H(j\omega) = \frac{2}{j\omega + 2}$$

and then compute

$$\begin{aligned}
\omega = 0: &\qquad H(0) = \frac{2}{2} = 1 = 1 \cdot e^{j0^\circ} \\
\omega = 1: &\qquad H(j1) = \frac{2}{j1 + 2} = \frac{2e^{j0^\circ}}{2.24e^{j26.6^\circ}} = 0.9e^{-j26.6^\circ} \\
\omega = 2: &\qquad H(j2) = \frac{2}{j2 + 2} = \frac{1}{1.414e^{j45^\circ}} = 0.707e^{-j45^\circ} \\
\omega \to \infty: &\qquad H(j\omega) \approx \frac{2}{j\omega} = re^{-j90^\circ}
\end{aligned}$$

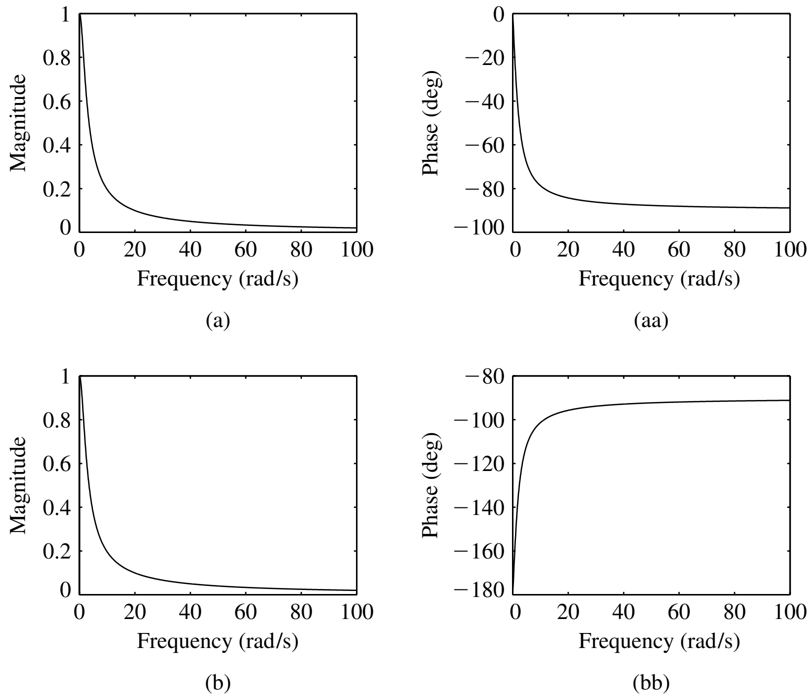


Figure 6.10 (**a**) Magnitude response of $2/(s+2)$. (**aa**) Its phase response. (**b**) Magnitude response of $2/(s-2)$. (**bb**) Its phase response.

Note that as $\omega \to \infty$, the frequency response can be approximated by $2/j\omega$. Thus its magnitude $r = 2/\omega$ approaches zero but its phase approaches $-90°$. From the preceding computation, we can plot the magnitude and phase responses as shown in Figures 6.10(a) and 6.10(aa). Clearly, the more ω we compute, the more accurate the plot.

The plot in Figure 6.10 is actually generated using the MATLAB function freqs(n,d,w), where n and d are the numerator's and denominator's coefficients of the transfer function and w specifies the frequencies at which the frequency responses are to be computed. Note that the last character "s" in freqs denotes the Laplace transform. For the transfer function in Example 6.8.1, typing

```
n=2;d=[1  2];w=0:0.1:100;
H=freqs(n,d,w);
subplot(1,2,1)
plot(w,abs(H)),title('(a)')
subplot(1,2,2)
plot(w,angle(H)*180/pi),title('(b)')
```

yields the magnitude and phase responses in Figures 6.10(a) and 6.10(aa). Note that in using freqs, if frequencies are not specified such as in

```
[H,w1]=freqs(n,d)
```

then the function freqs selects automatically 200 frequencies in $[0, \infty)$, denoted by w1, and computes the frequency spectrum at those frequencies.

EXAMPLE 6.8.2

We repeat Example 6.8.1 for $H_1(s) = 2/(s-2)$. The result is shown in in Figures 6.10(b) and 6.10(bb). We see that its magnitude response is identical to the one in Figure 6.10(a).

EXERCISE 6.8.1

Verify analytically that $H(s) = 2/(s+2)$ and $H_1(s) = 2/(s-2)$ have the same magnitude response.

We now discuss the implication of stability and physical meaning of frequency responses. Consider a system with stable transfer function $H(s)$ with real coefficients. Let us apply the input $u(t) = ae^{j\omega_0 t}$, where a is a real constant. Its Laplace transform is $a/(s - j\omega_0)$. Thus the output of the system is given by

$$Y(s) = H(s)U(s) = H(s)\frac{a}{s - j\omega_0} = k_1\frac{1}{s - j\omega_0} + \text{Terms due to poles of } H(s)$$

with

$$k_1 = Y(s)(s - j\omega_0)|_{s - j\omega_0 = 0} = aH(s)|_{s = j\omega_0} = aH(j\omega_0)$$

If $H(s)$ is stable, then all its poles have negative real parts and their time responses all approach 0 as $t \to \infty$. Thus we conclude that if $H(s)$ is stable and if $u(t) = ae^{j\omega_0 t}$, then we have

$$y_{ss}(t) := \lim_{t\to\infty} y(t) = aH(j\omega_0)e^{j\omega_0 t} \tag{6.50}$$

or, using (6.47),

$$\begin{aligned} y_{ss}(t) &= aA(\omega_0)e^{j\theta(\omega_o)}e^{j\omega_o t} = aA(\omega_0)e^{j(\omega_o t + \theta(\omega_o))} \\ &= aA(\omega_0)\left[\cos(\omega_0 t + \theta(\omega_0)) + j\sin(\omega_0 t + \theta(\omega_0))\right] \end{aligned} \tag{6.51}$$

We call $y_{ss}(t)$ the *steady-state response*. We list some special cases of (6.51) as a theorem.

THEOREM 6.4 Consider a system with proper rational transfer function $H(s)$. If the system is BIBO stable, then

$$\begin{aligned}
u(t) &= a \quad && \text{for } t \geq 0 \rightarrow\ y_{ss}(t) = aH(0) \\
u(t) &= a\sin\omega_0 t \quad && \text{for } t \geq 0 \rightarrow\ y_{ss}(t) = a|H(j\omega_0)|\sin(\omega_0 t + \measuredangle H(j\omega_0)) \\
u(t) &= a\cos\omega_0 t \quad && \text{for } t \geq 0 \rightarrow\ y_{ss}(t) = a|H(j\omega_0)|\cos(\omega_0 t + \measuredangle H(j\omega_0))
\end{aligned}$$

The steady-state response in (6.51) is excited by the input $u(t) = ae^{j\omega_0 t}$. If $\omega_0 = 0$, then the input is a step function with amplitude a, and the output approaches a step function with amplitude $aH(0)$. We call $H(0)$ the dc gain. If $u(t) = a\sin\omega_0 t = \text{Im}\ ae^{j\omega_0 t}$, where Im stands for the imaginary part, then the output approaches the imaginary part of (6.51) or $aA(j\omega)\sin(\omega_0 t + \theta(\omega_0))$. Using the real part of $ae^{j\omega_0 t}$, we will obtain the next equation. In conclusion, if we apply a sinusoidal input to a system, then the output approaches a sinusoidal signal with the same frequency, but its amplitude will be modified by $A(\omega_0) = |H(j\omega)|$ and its phase by $\measuredangle H(j\omega_0) = \theta(\omega_0)$. We stress that the system must be stable. If it is not, then the output generally grows unbounded and $H(j\omega_0)$ has no physical meaning as we will demonstrate later. We first use an example to discuss the implication of Theorem 6.4.

EXAMPLE 6.8.3

Consider a system with transfer function $H(s) = 2/(s+2)$. Find the steady-state response of the system excited by

$$u(t) = 1 + \sin 0.1t + 0.2\cos 20t \tag{6.52}$$

where $1 + \sin 0.1t$ will be considered as the desired signal and $0.2\cos 20t$ noise.

In order to apply Theorem 6.4, we read from Figures 6.10(a) and 6.10(aa) $H(0) = 1$, $H(j0.1) = 0.999e^{-j3^\circ} = 0.999e^{-j0.05}$, and $H(j20) = 0.1e^{-j84^\circ} = 0.1e^{-j1.47}$. Clearly, the reading cannot be very accurate. Note that the angles in degrees have been changed into radians. Then Theorem 6.4 implies

$$\begin{aligned}
y_{ss}(t) &= 1\times 1 + 0.999\sin(0.1t - 0.05) + 0.2\times 0.1\cos(20t - 1.47) \\
&= 1 + 0.999\sin(0.1t - 0.05) + 0.02\cos(20t - 1.47)
\end{aligned} \tag{6.53}$$

We plot in Figure 6.11(a) the input $u(t)$ of the system and in Figure 6.11(b) with a solid line its output. We also plot in Figure 6.11(b) with a dotted line the desired signal $1 + \sin 0.1t$. The output starts from 0 and approaches rapidly the desired signal. We see that the noise is essentially eliminated by the system. The system passes low-frequency signals and stops high-frequency signals and is therefore called a lowpass filter.

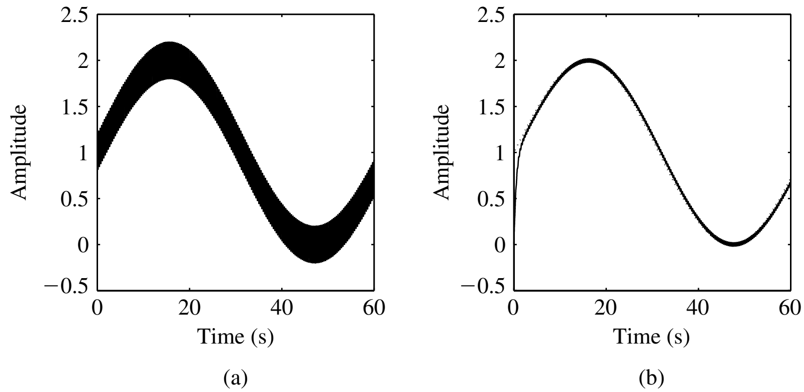


Figure 6.11 (**a**) Input signal (desired signal corrupted by noise). (**b**) Output of the system (solid line) and desired signal (dotted line).

In view of Theorem 6.4, if we can design a stable system with the magnitude response (solid line) and the phase response (dotted line) shown in Figure 6.12(a), then the system will pass sinusoids with frequency $|\omega| < \omega_c$ and stop sinusoids with frequency $|\omega| > \omega_c$. We require the phase response to be linear to avoid distortion as we will discuss in a later section. We call such a system an *ideal lowpass filter* with cutoff frequency ω_c. The frequency range $[0, \omega_c]$ is called the *passband,* and $[\omega_c, \infty)$ is called the *stopband*. Figures 6.12(b) and 6.12(c) show the characteristics of ideal bandpass and highpass filters. The ideal bandpass filter will pass sinusoids with frequencies lying inside the range $[\omega_l, \omega_u]$, where ω_l and ω_u are the lower and upper cutoff frequencies, respectively. The ideal highpass filter will pass sinusoids with frequencies larger than the cutoff frequency ω_c. They are called *frequency-selective filters*. Note that filters are special types of systems that are designed to pass some frequency bands of signals.

The impulse response $h(t)$ of the ideal lowpass filter in Figure 6.12(a) with $\omega_c = 2$ and $t_0 = 0$ is shown in Figure 4.8(b). Because $h(t) \neq 0$, for $t < 0$, the ideal filter is not causal. In fact, no causal systems can have the frequency responses shown in Figure 6.12. Thus, *ideal filters cannot be built in the real world*. In practice, we modify the magnitude responses in Figure 6.12 to the ones shown in Figure 6.13. We insert a *transition band* between the passband and stopband. Furthermore, we specify a passband tolerance and stopband tolerance as shown with shaded

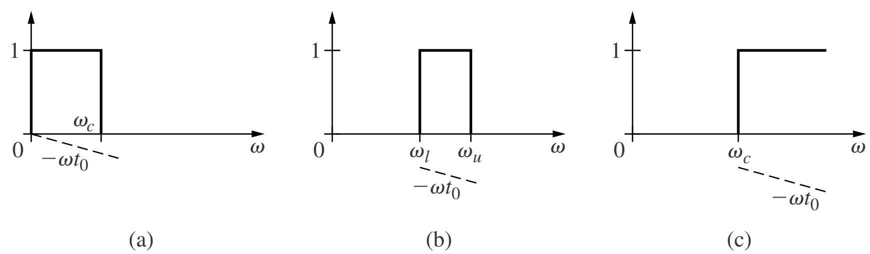


Figure 6.12 (**a**) Ideal lowpass filter with cutoff frequency ω_c. (**b**) Ideal bandpass filter with upper and lower cutoff frequencies ω_u and ω_l. (**c**) Ideal highpass filter with cutoff frequency ω_c.

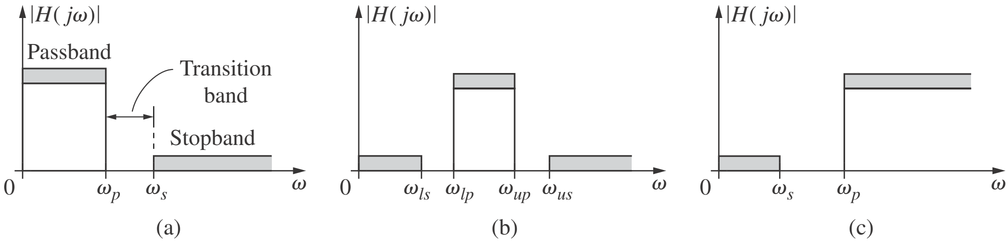


Figure 6.13 Specifications of practical (**a**) lowpass filter, (**b**) bandpass filter, and (**c**) highpass filter.

areas. The transition band is generally not specified and is the "don't care" region. We also introduce the *group delay* defined as

$$\text{Group delay} = \tau(\omega) = -\frac{d\theta(\omega)}{d\omega} \tag{6.54}$$

For an ideal filter with linear phase such as $\theta(\omega) = -t_0\omega$, its group delay is t_0, a constant. Thus instead of specifying the phase response to be a linear function of ω, we specify the group delay in the passband to be roughly constant or $t_0 - \epsilon < \tau(\omega) < t_0 + \epsilon$, for some constants t_0 and ϵ. Even with these more relaxed specifications on the magnitude and phase responses, if we specify both of them, it is still difficult to find causal filters to meet both specifications. Thus in practice, we often specify only magnitude responses as shown in Figure 6.13. The design problem is then to find a proper stable rational function of a degree as small as possible to have its magnitude response lying inside the specified region. See, for example, Reference 2.

To conclude this section, we mention that the stability condition is essential in Theorem 6.4. This is demonstrated in the next example.

EXAMPLE 6.8.4

Consider a system with transfer function $H(s) = s/(s-1)$. We compute its response excited by the input $u(t) = \cos 2t$. The Laplace transform of $u(t)$ is $U(s) = s/(s^2+4)$. Thus the output is

$$\begin{aligned} Y(s) = H(s)U(s) &= \frac{s}{s-1}\frac{s}{s^2+4} \\ &= \frac{1}{5(s-1)} + \frac{4s}{5(s^2+4)} + \frac{4}{5(s^2+4)} \end{aligned}$$

and, in the time domain,

$$\begin{aligned} y(t) &= 0.2e^t + 0.8\cos 2t + 0.4\sin 2t \\ &= 0.2e^t + 0.8944\cos(2t - 0.4636) \\ &= 0.2e^t + |H(j2)|\cos(2t + \measuredangle H(j2)) \end{aligned}$$

where we have used $|H(j2)| = 0.8944$ and $\measuredangle H(j2) = -0.4636$ (Problem 6.30).

Even though the output contains the sinusoid $|H(j2)|\cos(2t + \measuredangle H(j2))$, it also contains the exponentially increasing function $0.2e^t$. As $t \to \infty$, the former is buried or overwhelmed by the latter. Thus the output is $0.2e^t$ as $t \to \infty$ and Theorem 6.4 does not hold.

The preceding example shows that if $H(s)$ has RHP poles, then the output excited by a sinusoid will grow unbounded. If $H(s)$ has a pair of complex-conjugate poles at ω_1 and has the rest of the poles lying inside the open LHP, then the output excited by $\sin \omega_0 t$ will grow unbounded if $\omega_0 = \omega_1$; if $\omega_0 \neq \omega_1$, it approaches

$$y_{ss}(t) = |H(j\omega_0)| \sin(\omega_0 t + \measuredangle H(j\omega_0)) + k_1 \sin(\omega_1 t + k_2)$$

for some constants k_1 and k_2. See Problems 6.15 and 6.16. In conclusion, if $H(s)$ is not stable, then Theorem 6.4 is not applicable. Furthermore, its frequency response $H(j\omega)$ has no physical meaning.

6.8.1 Speed of Response—Time Constant

Consider a stable system with transfer function

$$H(s) = \frac{24}{s^3 + 3s^2 + 8s + 12} = \frac{24}{(s+2)(s+0.5+j2.4)(s+0.5-j2.4)} \tag{6.55}$$

Its step response approaches the steady state $y_{ss}(t) = H(0) = 2$ as shown in Figure 6.14(a). Mathematically speaking, it takes an infinite amount of time for the output to reach steady state. In practice, we often consider the output to have reached steady state if the output remains within $\pm 1\%$ of its steady state as indicated by the two horizontal dotted lines shown in Figure 6.14(a). One may then wonder: How fast will the response reach steady state? This question is important in designing practical systems.

The step response of (6.55) is of the form

$$y(t) = H(0) + k_1 e^{-2t} + k_2 e^{-0.5t} \sin(2.4t + k_3)$$

Its steady-state response is $y_{ss}(t) = H(0) = 2$. We call $y(t) - y_{ss}(t)$ or

$$y_{tr}(t) := y(t) - y_{ss}(t) = k_1 e^{-2t} + k_2 e^{-0.5t} \sin(2.4t + k_3)$$

the *transient response* and is shown in Figure 6.14(b). Clearly, the faster the transient response approaches zero, the faster the (total) response reaches steady state.

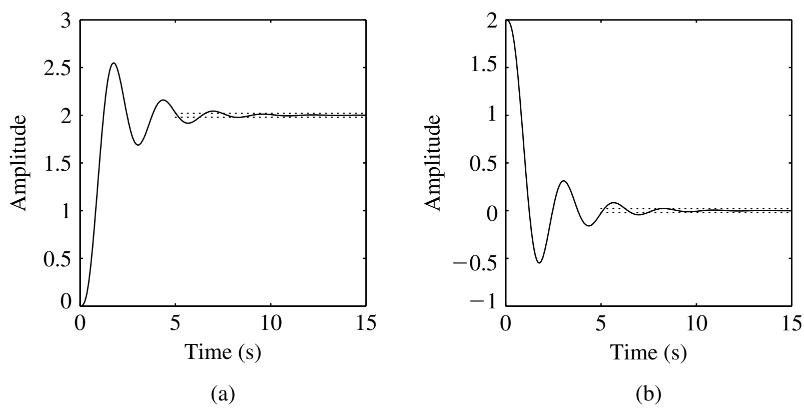


Figure 6.14 (**a**) Step response. (**b**) Its transient response.

We introduced in Section 1.3 the concept of time constant for $e^{-at} = \mathcal{L}^{-1}[1/(s+a)]$ with $a > 0$. We showed there that the function e^{-at} decreases to less than 1% of its peak magnitude in five time constants. We now extend the concept to stable proper rational transfer functions. The *time constant* of $H(s)$ is defined as

$$
\begin{aligned}
t_c &= \frac{1}{\text{Smallest real part in magnitude of all poles}} \\
&= \frac{1}{\text{Smallest distance from all poles to the } j\omega\text{-axis}}
\end{aligned}
\tag{6.56}
$$

The transfer function in (6.55) has three poles -2 and $-0.5 \pm j2.4$. The smallest real part in magnitude is 0.5. Thus the time constant of (6.55) is $1/0.5 = 2$. Note that its transient response consists of e^{-2t} and $e^{-0.5t}\sin 2.4t$. The larger the real part in magnitude of a pole, the faster its time response approaches zero. Thus the pole with the smallest real part in magnitude dictates the time for the transient response to approach zero.

It turns out that the transient response of a stable $H(s)$ will approach practically zero (less than 1% of its peak magnitude) in five time constants. Thus in practice, we may consider the response to have reached steady state in five time constants. For example, the transfer function in (6.55) has time constant 2 and its response reaches steady state in roughly 10 seconds as shown in Figure 6.14. All transfer functions in (6.39) through (6.43) have the same set of poles as plotted in Figure 6.5. The smallest distance from all poles to the imaginary axis is 0.5. Thus they have the same time constant 2, and their responses reach steady state in roughly $5 \times 2 = 10$ seconds as shown in Figure 6.7. The system in Example 6.8.3 has only one pole -2 and its time constant is 0.5. Its response reaches steady state roughly in $5 \times 0.5 = 2.5$ seconds as shown in Figure 6.11(b). Thus the rule of five time constants appears to be widely applicable. Note that time constants are defined only for stable transfer functions and are defined using only poles; they are independent of zeros.

It is important to mention that the rule of five time constants should be used only as a guide. It is possible to construct examples whose transient responses do not decrease to less than 1% of their peak magnitudes in five time constants. See Problem 6.31. However, it is generally true that the smaller the time constant, the faster a system responds.

EXERCISE 6.8.2

What are the time constants of the following transfer functions:

(a) $\dfrac{1}{s-2}$

(b) $\dfrac{s-5}{(s+1)(s+2)(s+5)}$

(c) $\dfrac{3(s^2-4)}{(s+1)(s^2+0.1s+1)}$

Answers

Not defined. 1. 20.

6.8.2 Bandwidth of Frequency-Selective Filters

This subsection introduces the concept of bandwidth for frequency-selective filters. The *bandwidth* of a filter is defined as the width of its passband in the positive frequency range. The bandwidth is ω_c for the ideal lowpass filter in Figure 6.12(a), $\omega_u - \omega_l$ for the ideal bandpass filter in Figure 6.12(b), and infinity for the ideal highpass filter in Figure 6.12(c). Clearly, for highpass filters, the passband cutoff frequency is more meaningful than bandwidth.

For practical filters, different specification of passband tolerance will yield different bandwidth. We discuss one specification in the following. The bandwidth of a filter can be defined as the width of the positive frequency range in which the magnitude response meets the condition

$$|H(j\omega)| \geq 0.707 H_{max} \tag{6.57}$$

where H_{max} is the peak magnitude of $H(j\omega)$. For example, the peak magnitude of $H(s) = 2/(s+2)$ is 1 as shown in Figure 6.10(a). From the plot we see that the magnitude is 0.707 or larger for ω in [0, 2]. Thus the bandwidth of $2/(s+2)$ is 2 rad/s. We call $\omega_p = 2$ the passband edge frequency. See Problem 6.33.

We give two more examples. Consider the transfer functions

$$H_1(s) = \frac{1000s^2}{s^4 + 28.28s^3 + 5200s^2 + 67882.25s + 5760000} \tag{6.58}$$

and

$$H_2(s) = \frac{921.27s^2}{s^4 + 38.54s^3 + 5742.73s^2 + 92500.15s + 5760000} \tag{6.59}$$

Their magnitude responses are plotted respectively in Figures 6.15(a) and 6.15(b). They are obtained using the MATLAB function freqs. For the transfer function in (6.58), typing

```
n=[1000 0 0];d=[1 28.28 5200 67882.25 5760000];w=0:0.1:200;
H=freqs(n,d,w);plot(w,abs(H))
```

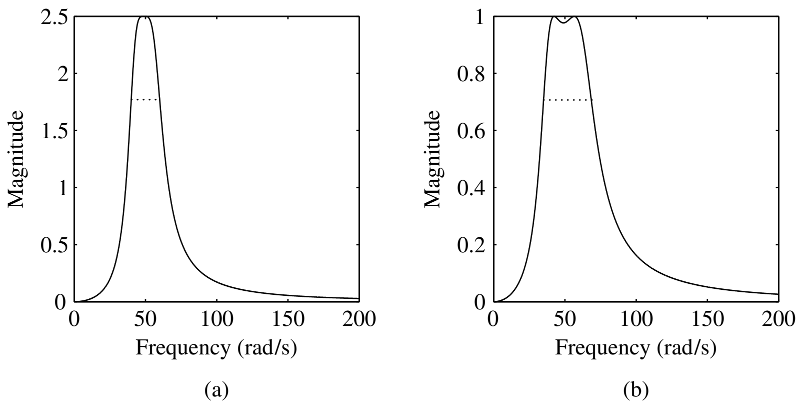


Figure 6.15 (**a**) Magnitude response and bandwidth of (6.58). (**b**) Magnitude response of bandwidth of (6.59).

yields the plot in Figure 6.15(a). The magnitude plot in Figure 6.15(a) has peak magnitude $H_{1max} = 2.5$. From the plot, we see that for ω in [40, 60], its magnitude is $0.707 \times 2.5 = 1.77$ or larger as shown with a dotted line. We call 40 the lower passband edge frequency, and we call 60 the upper passband edge frequency. Thus the bandwidth of the filter is $60 - 40 = 20$ rad/s. For the transfer function in (6.59), we have $H_{2max} = 1$ and the magnitude response is 0.707 or larger in the frequency range [35, 70]. Thus its bandwidth is $70 - 35 = 35$ rad/s.

The passband tolerance is often defined using the unit of decibel (dB). The dB is the unit of $20 \log_{10} |H(j\omega)|$. Because the bandwidth is defined with respect to the peak magnitude, we define

$$a = 20 \log_{10} \left(\frac{|H(j\omega)|}{H_{max}} \right)$$

where H_{max} is the peak magnitude. If $|H(j\omega)| = H_{max}$, then a is 0 dB. If $|H(j\omega)| = 0.707 H_{max}$, then

$$a = 20 \log_{10} 0.707 = -3 \text{ dB}$$

Thus the passband is the positive frequency range in which its magnitude is -3 dB or larger, and the width of the range is also called the *3-dB bandwidth* (not -3-dB bandwidth). Note that we can also define 2-dB or 1-dB bandwidth. We mention that the 3-dB bandwidth is also called the *half-power* bandwidth. The power of $y(t)$ is $y^2(t)$ and, consequently, proportional to $|H(\omega)|^2$. If the power at H_{max} is A, then the power at $0.707 H_{max}$ is $(0.707)^2 A = 0.5A$. Thus the 3-dB bandwidth is also called the *half-power bandwidth*.

6.8.3 An Alternative Derivation of Frequency Responses

This subsection discusses a different way of developing frequency responses. Consider an LTI system described by the convolution

$$y(t) = \int_0^\infty h(t - \tau) u(\tau)\, d\tau \tag{6.60}$$

In this equation, the system is assumed to be initially relaxed at $t = 0$ and the input is applied from $t = 0$ onward. Now if the system is assumed to be initially relaxed at $t = -\infty$ and the input is applied from $t = -\infty$, then the equation must be modified as

$$y(t) = \int_{-\infty}^\infty h(t - \tau) u(\tau)\, d\tau = \int_{-\infty}^\infty h(\tau) u(t - \tau)\, d\tau \tag{6.61}$$

Note that (6.61) has the commutative property but (6.60) does not. Now if we apply $u(t) = e^{j\omega_0 t}$ from $t = -\infty$, then (6.61) becomes

$$\begin{aligned} y(t) &= \int_{-\infty}^\infty h(\tau) e^{j\omega_0 (t-\tau)}\, d\tau = \left(\int_{-\infty}^\infty h(\tau) e^{-j\omega_0 \tau}\, d\tau \right) e^{j\omega_0 t} \\ &= \left(\int_0^\infty h(\tau) e^{-j\omega_0 \tau}\, d\tau \right) e^{j\omega_0 t} = H(j\omega_0) e^{j\omega_0 t} \end{aligned} \tag{6.62}$$

where we have used $h(t) = 0$ for $t < 0$ and (6.2) with s replaced by $j\omega_0$. The equation is similar to (6.50) with $a = 1$. This is another way of developing frequency responses.

The derivation of (6.62) appears to be straightforward, but it is not yet complete without discussing its validity conditions.[12] The function $H(j\omega_0)$ in (6.62) is in fact the Fourier transform of $h(t)$. If $h(t) = e^{2t}$, for $t \geq 0$, the integration diverges and (6.62) is meaningless. If $h(t) = \sin 2t$, $H(j\omega)$ will contain impulses and the meaning of (6.62) is not clear. See Problems 6.15 and 6.16. However, if $h(t)$ is absolutely integrable, then $H(j\omega)$ is bounded and continuous and (6.62) holds. Thus the derivation of (6.62) requires the same stability condition as (6.50).

They, however, have a big difference. Equation (6.62) holds for all t in $(-\infty, \infty)$, whereas (6.50) holds only for $t \to \infty$. In other words, there is no transient response in (6.62), and (6.62) describes only the steady-state response. In reality, it is not possible to apply an input from time $-\infty$. An input can be applied only from some finite time where we may call it time zero. Recall that time zero is relative and is defined by us. Thus there is always a transient response. In conclusion, the derivation of (6.50), which uses explicitly the stability condition and is valid only for $t \to \infty$, is more revealing than the derivation of (6.62).

6.9 FROM LAPLACE TRANSFORM TO FOURIER TRANSFORM

This section shows that the Fourier transform of a two-sided signal can be obtained from the (one-sided) Laplace transform. However, this is possible only if the signal is absolutely integrable.

Let $x(t)$ be a two-sided signal and let $\bar{X}(\omega)$ be its Fourier transform, that is,

$$\begin{aligned}\bar{X}(\omega) &= \mathcal{F}[x(t)] = \int_{t=-\infty}^{\infty} x(t)e^{-j\omega t}\,dt \\ &= \int_{t=-\infty}^{0} x(t)e^{-j\omega t}\,dt + \int_{t=0}^{\infty} x(t)e^{-j\omega t}\,dt \\ &= \int_{\tau=0}^{\infty} x(-\tau)e^{j\omega\tau}\,d\tau + \int_{t=0}^{\infty} x(t)e^{-j\omega t}\,dt \end{aligned} \tag{6.63}$$

where we have introduced the new variable $\tau = -t$.

Let $X(s)$ be the (one-sided) Laplace transform of $x(t)$. That is,

$$X(s) = \mathcal{L}[x(t)] = \int_{t=0}^{\infty} x(t)e^{-st}\,dt \tag{6.64}$$

Note that this transform uses only the positive-time part of $x(t)$. Let us write $x(t) = x_+(t) + x_-(t)$, where $x_+(t) = x(t)q(t)$ is the positive-time part of $x(t)$ and $x_-(t) = x(t)q(-t)$ is the negative-time part of $x(t)$. See (1.6). Let $X_+(s)$ be the Laplace transform of $x_+(t)$. Then we have

$$X_+(s) = \mathcal{L}[x_+(t)] = \int_{t=0}^{\infty} x_+(t)e^{-st}\,dt = \int_{t=0}^{\infty} x(t)e^{-st}\,dt = \mathcal{L}[x(t)] \tag{6.65}$$

where we have used the fact that $x_+(t) = x(t)$ for t in $[0, \infty)$ and (6.64). Because $x_-(t)$ is negative time, we have $\mathcal{L}[x_-(t)] = 0$. Let us flip $x_-(t)$ to become a positive-time signal

[12] Most texts use (6.62) to develop the concept of frequency responses without discussing its validity conditions.

$x_-(-t) = x(-t)q(t)$. Then its Laplace transform is

$$X_-(s) := \mathcal{L}[x_-(-t)] = \int_{t=0}^{\infty} x(-t)q(t)e^{-st}\,dt = \int_{t=0}^{\infty} x(-t)e^{-st}\,dt = \mathcal{L}[x(-t)] \tag{6.66}$$

If $x(t)$ is absolutely integrable, so are $x_+(t)$ and $x_-(-t)$. Thus their Laplace transform variable s can be replaced by $j\omega$. Comparing (6.63) with (6.65) and (6.66), we have

$$\mathcal{F}[x(t)] = \mathcal{L}[x(t)]|_{s=j\omega} + \mathcal{L}[x(-t)]|_{s=-j\omega} \tag{6.67}$$

In particular, if $x(t)$ is positive time and absolutely integrable, then

$$\bar{X}(\omega) := \mathcal{F}[x(t)] = \mathcal{L}[x(t)]|_{s=j\omega} =: X(s)|_{s=j\omega} = X(j\omega) \tag{6.68}$$

Using these formulas, we can compute the Fourier transform of an absolutely integrable signal from the (one-sided) Laplace transform.

EXAMPLE 6.9.1

Consider

$$x(t) = \begin{cases} 3e^{-2t} & \text{for } t \geq 0 \\ 2e^{3t} & \text{for } t < 0 \end{cases}$$

Because the signal approaches zero exponentially as $t \to \pm\infty$, it is absolutely integrable. The Laplace transform of $3e^{-2t}$ is $3/(s+2)$. The flipping of the negative-time signal is $2e^{-3t}$ whose Laplace transform is $2/(s+3)$. Thus the Fourier transform of $x(t)$ is

$$\mathcal{F}[x(t)] = \left.\frac{3}{s+2}\right|_{s=j\omega} + \left.\frac{2}{s+3}\right|_{s=-j\omega} = \frac{3}{j\omega+2} + \frac{2}{-j\omega+3} = \frac{13-j\omega}{\omega^2+j\omega+6}$$

It is the frequency spectrum of the signal.

EXAMPLE 6.9.2

Consider

$$x_1(t) = 2e^{-0.3t}\sin t \tag{6.69}$$

for $t \geq 0$, as shown in Figure 1.14(a). It is positive time. Because

$$\int_0^{\infty} |2e^{-0.3t}\sin t|\,dt < 2\int_0^{\infty} e^{-0.3t}\,dt = \frac{2}{0.3} < \infty$$

it is absolutely integrable. Its Laplace transform is, using Table 6.1,

$$X_1(s) = \frac{2}{(s+0.3)^2+1} = \frac{2}{s^2+0.6s+1.09} \tag{6.70}$$

Thus the Fourier transform or the frequency spectrum of (6.69) is

$$X_1(j\omega) = \left.\frac{2}{(s+0.3)^2+1}\right|_{s=j\omega} = \frac{2}{(j\omega+0.3)^2+1} \tag{6.71}$$

The spectrum is defined for all ω in $(-\infty, \infty)$.

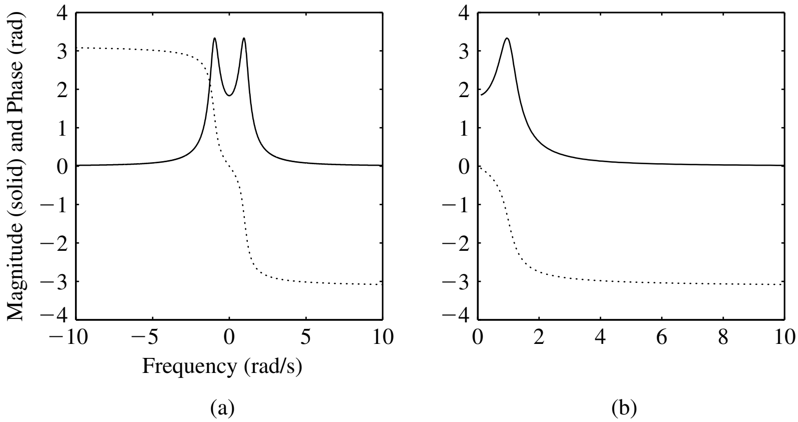


Figure 6.16 Frequency spectrum of (6.69). (**a**) Using (6.71). (**b**) Using freqs on (6.70).

We use MATLAB to compute (6.71) for ω in $[-10, 10]]$ as follows:

```
w=-10:0.01:10;
X=2./((j*w+0.3).^2+1);
plot(w,abs(X),w,angle(X),':')
```

It yields the magnitude spectrum (solid line) and phase spectrum (dotted line) in Figure 6.16(a). Its magnitude spectrum is even and its phase spectrum is odd. The magnitude spectrum has peaks at $\pm\omega = 1$ rad/s. It is reasonable in view of the factor $\sin t$ in the signal. Because of symmetric of the magnitude spectrum and antisymmetric of the phase spectrum, we often plot them only for $\omega \geq 0$.

The MATLAB function freqs computes frequency responses of transfer functions. Because of (6.68), it can also be used to compute frequency spectra of signals. Using the numerator's and denominator's coefficients of $X_1(s)$ in (6.70), we type

```
n=2;d=[1 0.6 1.09];
[H,w]=freqs(n,d);
plot(w,abs(H),w,angle(H),':')
```

It generates the plot in Figure 6.16(b). It is the magnitude and phase spectra of (6.69). Recall that the magnitude spectrum of (6.69) can also be directly computed using FFT as we did in Example 5.6.3. The plot in Figure 5.15(a) is close to the magnitude spectrum in Figure 6.16(b).

To conclude this section, we mention that (6.67) and (6.68) hold only if $x(t)$ is absolutely integrable. For example, the signal $x(t) = e^{2t}$, for $t \geq 0$, is not absolutely integrable. Its Laplace transform is $1/(s-2)$, and its Fourier transform is not defined and cannot equal $1/(j\omega-2)$. The Laplace transform of the step function $q(t)$ is $Q(s) = 1/s$. Its Fourier transform, as discussed in (4.56), is

$$\bar{Q}(\omega) = \pi\delta(\omega) + \frac{1}{j\omega} \tag{6.72}$$

which differs from $Q(j\omega)$. Note that the step function is not absolutely integrable. Thus the condition of absolute integrability is essential in using (6.67) and (6.68).

6.9.1 Why Fourier Transform Is Not Used in System Analysis

This subsection compares the Laplace and Fourier transforms in system analysis. Consider an LTIL system described by

$$y(t) = \int_{\tau=0}^{t} h(t-\tau)u(\tau)\,d\tau$$

Its Laplace transform is

$$Y(s) = H(s)U(s) \tag{6.73}$$

and its Fourier transform is, as derived in (4.46),

$$\bar{Y}(\omega) = \bar{H}(\omega)\bar{U}(\omega) \tag{6.74}$$

Note the use of different notations for the Laplace and Fourier transforms; otherwise confusion may arise. Thus system analysis can be carried out using the Laplace and Fourier transforms. However, the Fourier transform is, as we will show, less general, less revealing, and more complicated. Thus we did not discuss in Chapter 4 its application to system analysis.

Equation (6.73) is applicable whether the system is stable or not and whether the frequency spectrum of the input is defined or not. For example, consider $h(t) = u(t) = e^t$, for $t \geq 0$. Their Laplace transforms both equal $1/(s-1)$. Thus the output is given by

$$Y(s) = H(s)U(s) = \frac{1}{s-1}\frac{1}{s-1} = \frac{1}{(s-1)^2}$$

and, using Table 6.1, $y(t) = te^t$. The output grows unbounded. For this example, the Fourier transforms of $h(t)$, $u(t)$, and $y(t)$ are not defined and (6.74) is not applicable. Thus the Laplace transform is more general.

Next we consider a system with transfer function

$$H(s) = \frac{s}{(s+1)(s+2)} \tag{6.75}$$

It is stable and its impulse response (inverse Laplace transform of $H(s)$) is positive time and absolutely integrable. Thus its Fourier transform is, using (6.68),

$$\bar{H}(\omega) = H(j\omega) = \frac{j\omega}{(j\omega+1)(j\omega+2)} = \frac{j\omega}{2-\omega^2+3j\omega} \tag{6.76}$$

From the poles of $H(s)$, we can determine immediately the general form of its time response, its time constant, and, consequently, its speed of response. Although the poles and zero of the system are embedded in its Fourier transform $\bar{H}(\omega)$ in (6.76), it requires some effort to find its poles and zero (see Section 7.7). Thus $\bar{H}(\omega)$ is less revealing than $H(s)$.

We next compute the step responses of (6.75) and (6.76). The step response of (6.75) is given by

$$Y(s) = \frac{s}{(s+1)(s+2)}\frac{1}{s} = \frac{1}{(s+1)(s+2)} = \frac{1}{s+1} - \frac{1}{s+2}$$

which implies $y(t) = e^{-t} - e^{-2t}$, for $t \geq 0$. If we use (6.76), then the output is given by, using (6.74) and (6.72),

$$\bar{Y}(\omega) = \frac{j\omega}{2 - \omega^2 + 3j\omega}\left[\pi\delta(\omega) + \frac{1}{j\omega}\right]$$

To find its inverse Fourier transform is complicated. Therefore there seems no reason to use the Fourier transform in system analysis.

6.9.2 Phasor Analysis

Phasors are often used in introductory network analysis to carry out *sinusoidal steady-state analysis*. Consider the cosine function

$$A\cos(\omega t + \theta) = \mathrm{Re}[Ae^{j\theta}e^{j\omega t}]$$

where $A \geq 0$. For a given ω, the cosine function is uniquely defined by A and θ and can be represented by the *phasor*

$$A \measuredangle \theta$$

Consider a system—in particular, a RLC network with transfer function $H(s)$. If $H(s)$ is stable, and if the input is $u(t) = a\cos\omega_0 t$, then the output approaches

$$y_{ss}(t) = a|H(j\omega_0)|\cos(\omega_0 t + \measuredangle H(j\omega_0))$$

In other words, if the input is a cosine function with phasor $a\measuredangle 0$, then the output approaches a cosine function with phasor $a|H(j\omega_0)|\measuredangle H(j\omega_0)$. Thus phasors can be used to study sinusoidal steady-state responses. However, the system must be stable. Otherwise, the result will be incorrect.

EXAMPLE 6.9.3

Consider the network shown in Figure 6.17. It consists of a capacitor with capacitance 1F and a resistor with resistance -1Ω. Note that such a negative resistance can be generated using an op-amp circuit. See Problem 2.14. The transfer function of the network is $H(s) = s/(s-1)$. If we apply $u(t) = \cos 2t$, then its output was computed in Example 6.8.4 as

$$y(t) = 0.2e^t + |H(j2)|\cos(2t + \measuredangle H(j2))$$

Even though the output contains the phasor $|H(j2)|\measuredangle H(j2)$, it will be buried or overwhelmed by the exponentially increasing function $0.2e^t$. Thus the circuit will burn out or saturate and the phasor analysis cannot be used.

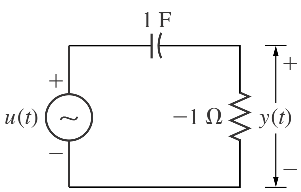


Figure 6.17 Unstable Network.

In conclusion, RLC networks with positive R, L, and C, are, as discussed earlier, automatically stable. In this case, phasor analysis can be employed to compute their sinusoidal steady-state responses. Furthermore, their impedances can be defined as R, $j\omega L$, and $1/j\omega C$. However, if a system is not stable, then phasor analysis will yield erroneous results.

6.10 FREQUENCY RESPONSES AND FREQUENCY SPECTRA

We showed in Section 6.8 that the output of a stable system with transfer function $H(s)$ excited by $u(t) = ae^{j\omega_0 t}$ approaches $aH(j\omega_0)e^{j\omega_0 t}$ as $t \to \infty$. We now extend the formula to the general case. The input and output of a system with transfer function $H(s)$ is related by

$$Y(s) = H(s)U(s) \tag{6.77}$$

This is a general equation, applicable whether the system is stable or not and whether the spectrum of the input is defined or not.

Let us replace s in (6.77) by $j\omega$ to yield

$$Y(j\omega) = H(j\omega)U(j\omega) \tag{6.78}$$

If the system is not stable or if the input grows unbounded, then one of $H(j\omega)$ and $U(j\omega)$ does not have physical meaning and, consequently, neither does their product $Y(j\omega)$. However, if the system is BIBO stable and if the input is absolutely integrable, then $H(j\omega)$ is the frequency response of the system and $U(j\omega)$ is, as shown in (6.68), the frequency spectrum of the input. Now the question is, Does their product have any physical meaning?

Before proceeding, we show that if $H(s)$ is stable (its impulse response is absolutely integrable in $[0, \infty)$) and if $u(t)$ is absolutely integrable in $[0, \infty)$, so is $y(t)$. Indeed we have

$$y(t) = \int_{\tau=0}^{t} h(t-\tau)u(\tau)\,d\tau$$

which implies

$$|y(t)| \le \int_{\tau=0}^{t} |h(t-\tau)||u(\tau)|\,d\tau \le \int_{\tau=0}^{\infty} |h(t-\tau)||u(\tau)|\,d\tau$$

Thus we have

$$\begin{aligned}
\int_{t=0}^{\infty} |y(t)|\,dt &\le \int_{t=0}^{\infty}\left(\int_{\tau=0}^{\infty} |h(t-\tau)||u(\tau)|\,d\tau\right)dt \\
&= \int_{\tau=0}^{\infty}\left(\int_{t=0}^{\infty} |h(t-\tau)|\,dt\right)|u(\tau)|\,d\tau \\
&= \int_{\tau=0}^{\infty}\left(\int_{\bar{t}=-\tau}^{\infty} |h(\bar{t})|\,d\bar{t}\right)|u(\tau)|\,d\tau \\
&= \left(\int_{\bar{t}=0}^{\infty} |h(\bar{t})|\,d\bar{t}\right)\left(\int_{\tau=0}^{\infty} |u(\tau)|\,d\tau\right)
\end{aligned}$$

where we have interchanged the order of integrations, introduced a new variable $\bar{t} = t - \tau$, and used the causality condition $h(\bar{t}) = 0$ for $\bar{t} < 0$. Thus if $h(t)$ and $u(t)$ are absolutely integrable, so is $y(t)$.

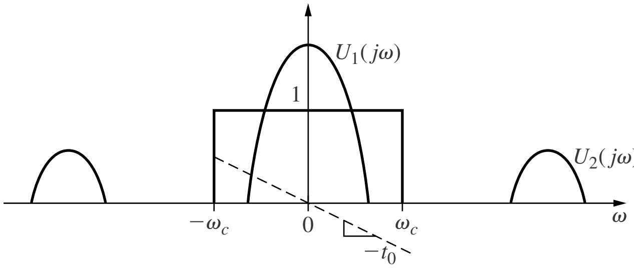


Figure 6.18 Spectra of $u_1(t)$ and $u_2(t)$.

Let us consider again (6.78). If $H(s)$ is causal and stable and if an input is positive time and absolutely integrable, then the output is positive time and absolutely integrable and, consequently, has a well-defined frequency spectrum that is given by (6.78). Thus (6.78) is of fundamental importance and filter design is based on the equation.

Let us consider the magnitude and phase responses of an ideal lowpass filter with cutoff frequency ω_c shown in Figure 6.12(a) or

$$H(j\omega) = \begin{cases} 1 \cdot e^{-j\omega t_0} & \text{for } |\omega| \le \omega_c \\ 0 & \text{for } |\omega| > \omega_c \end{cases} \tag{6.79}$$

Now if $u(t) = u_1(t) + u_2(t)$ and if the magnitude spectra of $u_1(t)$ and $u_2(t)$ are as shown in Figure 6.18, then the output frequency spectrum is given by

$$Y(j\omega) = H(j\omega)U(j\omega) = U_1(j\omega)e^{-j\omega t_0} \tag{6.80}$$

It is, as derived in (4.35), the frequency spectrum of $u_1(t - t_0)$. Thus the output of the ideal lowpass filter is

$$y(t) = u_1(t - t_0) \tag{6.81}$$

That is, the filter stops completely the signal $u_2(t)$ and passes $u_1(t)$ with only a delay of t_0 seconds. This is called a *distortionless transmission* of $u_1(t)$. Note that t_0 is the group delay of the ideal lowpass filter defined in (6.54).

To conclude this section, we mention an important implication of (6.78). If $U(j\omega) = 0$ in some frequency range, then so is $Y(j\omega) = 0$ in the same frequency range. It means that *an LTI stable system cannot generate new frequency components; it can only modify the existing frequency components of the input.* This is, however, not the case for nonlinear systems as shown in Example 4.2.5. This is a limitation of LTI systems.

6.10.1 Resonance

In this subsection we show one application of (6.78) and compare it with the corresponding time-domain equation

$$y(t) = \int_0^t h(t - \tau)u(\tau)\, d\tau \tag{6.82}$$

Consider the network in Figure 6.8(a). The network has transfer function $s/(s^2 + 1)$ and poles $\pm j$. As discussed in Section 6.7, if we apply the input $u(t) = \sin \omega_0 t$, with $\omega_0 \neq 1$, then the output will be bounded. However, if we apply $\sin t$, then the output will grow unbounded. This is due to the fact that the poles of the input $\sin t$ coincide with the poles of the network. Thus the input is amplified or reinforced by the network and grows unbounded. This is called *resonance*. We discuss this concept further in the following.

Consider a system with transfer function

$$H(s) = \frac{20}{s^2 + 0.4s + 400.04} = \frac{20}{(s + 0.2 + 20j)(s + 0.2 - 20j)} \tag{6.83}$$

It is stable and has its impulse response (inverse Laplace transform of H(s)) shown in Figure 6.19(a) and its magnitude response (magnitude of $H(j\omega)$) shown in Figure 6.19(aa). The magnitude response shows a narrow spike with peak roughly at $\omega_r = 20$ rad/s. We call ω_r the *resonance frequency*.

Let us study the outputs of the system excited by the inputs

$$u_i(t) = e^{-0.5t} \cos \omega_i t$$

with $\omega_1 = 5$, $\omega_2 = 20$, and $\omega_3 = 35$. The inputs $u_i(t)$ against t are plotted in Figures 6.19(b), 6.19(c), and 6.19(d). Their magnitude spectra $|U_i(j\omega)|$ against ω are plotted in Figures 6.19(bb), 6.19(cc), and 6.19(dd). Figures 6.19(bbb), 6.19(ccc), and 6.19(ddd) show their excited outputs $y_i(t)$. The output $y_1(t)$ in Figure 6.19(bbb) is the convolution of the impulse response in Figure 6.19(a) and the input $u_1(t)$ in Figure 6.19(b), $y_2(t)$ in Figure 6.19(ccc) is the convolution of Figures 6.19(a) and 6.19(c), and $y_3(t)$ in Figure 6.19(ddd) is the convolution of Figures 6.19(a) and 6.19(d).

The output $y_2(t)$ excited by $u_2(t)$ has larger magnitudes and a longer duration than the outputs excited by $u_1(t)$ and $u_3(t)$. This will be difficult, if not impossible, to explain in the time domain: convolutions of the impulse response and the inputs. However, it can be easily explained in the frequency domain: The nonzero magnitude spectrum of $u_2(t)$ is centered around the resonance frequency of the system. In other words, the input $u_2(t)$ resonates with the system, thus its output is the largest among the three shown in Figure 6.19.

No physical system is designed to be completely rigid. Every structure or mechanical system will vibrate when it is subjected to a shock or an oscillating force. For example, the wings of a Boeing 747 vibrate fiercely when it flies into a storm. Thus in designing a system, if its magnitude response has a narrow spike, it is important that the spike should not coincide with the most significant part of the frequency spectrum of possible external excitation. Otherwise, excessive vibration may occur and cause eventual failure of the system. The most infamous such failure was the collapse of the first Tacoma Narrows Suspension Bridge in Seattle in 1940 due to wind-induced resonance.

To conclude this section, we discuss how the plots in Figures 6.19 are obtained. The inverse Laplace transform of (6.83) is $h(t) = e^{-0.2t} \sin 20t$. Typing

```
t=0:0.05:10;h=exp(-0.2*t).*sin(20*t);plot(t,h)
```

yields the impulse response of (6.83) in Figure 6.19(a). Typing

```
n=20;d=[1 0.4 400.04];[H,w]=freqs(n,d);plot(w,abs(H))
```

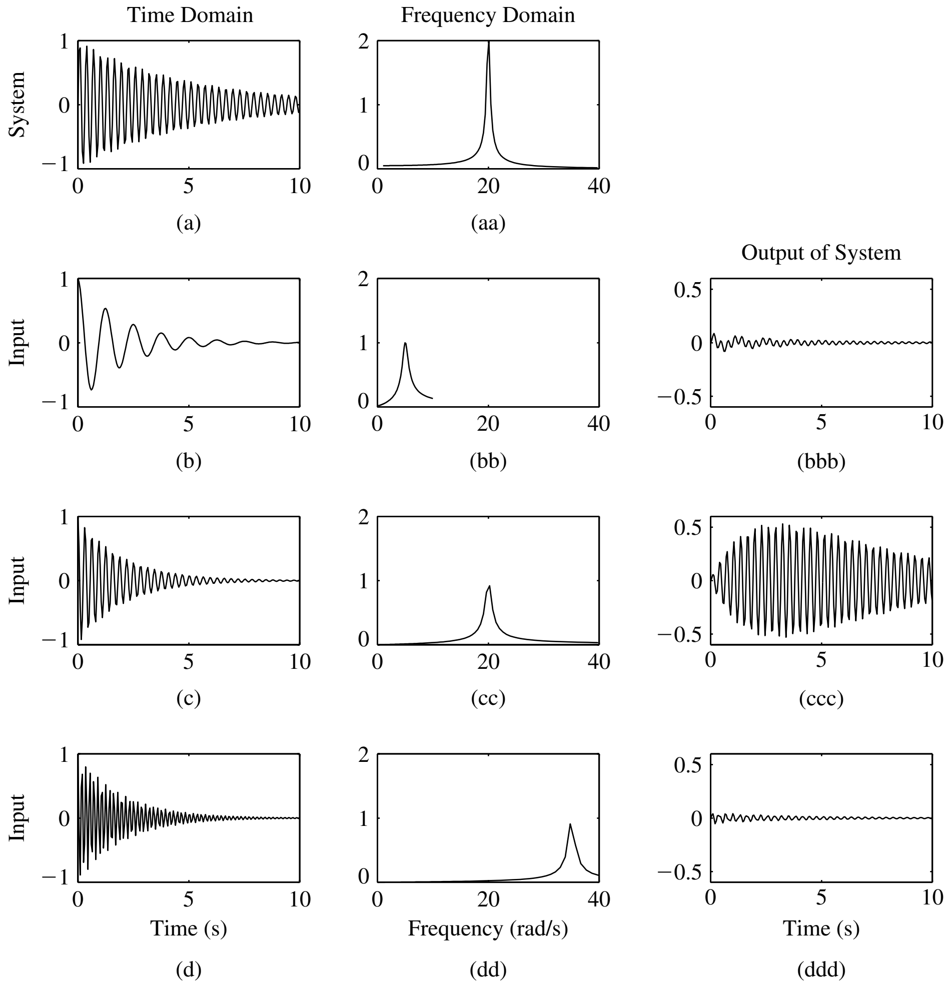


Figure 6.19 (**a**) The impulse response $h(t)$ of the system in (6.83). (**aa**) The magnitude response $|H(j\omega)|$ of the system in (6.83). (**b**) $u_1 = e^{-0.5t}\cos 5t$. (**bb**) Its magnitude spectrum. (**bbb**) Its excited output $y_1(t)$. (**c**) $u_2 = e^{-0.5t}\cos 20t$. (**cc**) Its magnitude spectrum. (**ccc**) Its excited output $y_2(t)$. (**d**) $u_3 = e^{-0.5t}\cos 35t$. (**dd**) Its magnitude spectrum. (**ddd**) Its excited output $y_3(t)$.

yields the magnitude response of (6.83) in Figure 6.19(aa). Typing

```
t=0:0.05:10;u1=exp(-0.5*t).*cos(5*t);plot(t,u1)
```

yields the input u_1 in Figure 6.19(b). Because the Laplace transform of $u_1(t)$ with $\omega_1 = 5$ is

$$U_1(s) = (s+0.5)/((s+0.5)^2+25) = (s+0.5)/(s^2+s+25.25)$$

the frequency spectrum in Figure 6.19(bb) can be obtained, as discussed in Section 6.9, by typing

```
n1=[1 0.5];d1=[1 1 25.25];[H1,w1]=freqs(n1,d1);plot(w1,abs(H1))
```

Note that the spectrum of $u_1(t)$ can also be computed using FFT as discussed in Section 5.6—in particular, Example 5.6.3. The output of (6.83) excited by $u_1(t)$, shown in Figure 6.19(bbb), is obtained by typing

```
y1=lsim(n,d,u1,t);plot(t,y1)
```

where the MATLAB function `lsim`, an acronym of *linear simulation,* computes the output of the transfer function `H=n/d` excited by `u1` at time instants denoted by `t`. The function will be discussed further in the next chapter. Replacing u_1 by u_2 and u_3 yields the remainder of Figure 6.19.

6.11 CONCLUDING REMARKS

We introduced in this chapter the concepts of transfer functions and frequency responses. They are essential in design. Design specifications can be given in the time domain as discussed in Sections 6.5.5 and 6.8.1 or in the frequency domain as shown in Figure 6.13. The former is used in control system design, whereas the latter is used in filter design. In control, we have a number of design methods. The *frequency-domain method* uses frequency responses; the *root-locus method* uses poles and zeros, and the *linear algebraic method* uses numerators' and denominators' coefficients. In filter design, we search for transfer functions whose frequency responses meet the specifications shown in Figure 6.13. See References 2, 3, and 5. Thus all concepts introduced in this chapter are fundamental in engineering.

Two general approaches are available to compute the transfer function of an LTIL system. The first approach is an analytical method that requires the knowledge of the internal structure of the system. The second approach uses measurements at the input and output terminals without using any knowledge of its internal structure. The first approach has many methods. For example, consider the network in Figure 3.3(a). We can first develop a differential equation (Example 3.8.3) and then take its Laplace transform. A simpler method is to use impedances to develop directly its transfer function (Exercise 6.3.7). Yet another method is to find, as we will discuss in the next chapter, a state-space equation to describe the network and then find its transfer function (Exercise 7.5.1). In general, the method of using impedances is the simplest for RLC networks.

The second approach again has many methods. If we measure its step response, its differentiation yields the impulse response (Example 3.7.1) whose Laplace transform is the transfer function. A better way is to take the ratio of the Laplace transforms of the step response and step function (Example 6.3.1). Yet another way is to measure the frequency response and then search a transfer function to match the measured response. This will be discussed in the next chapter. See Problem 7.21.

A system with input $u(t)$, output $y(t)$, and transfer function $H(s)$ is often denoted as shown in Figure 6.20. The diagram may lead to the writing

$$y(t) = H(s)u(t)$$

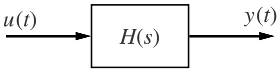


Figure 6.20 System with input $u(t)$, output $y(t)$, and transfer function $H(s)$.

This is incorrect; we cannot mix time functions and Laplace transforms in the same equation. The correct expression should be

$$Y(s) = H(s)U(s)$$

or

$$y(t) = \int_{\tau=0}^{t} h(t-\tau)u(\tau)\,d\tau$$

where $h(t)$ is the inverse Laplace transform of $H(s)$. If the system is stable and if the input frequency spectrum is defined, then we have $Y(j\omega) = H(j\omega)U(j\omega)$, which yields the output frequency spectrum. If the system is not stable, the equation $Y(j\omega) = H(j\omega)U(j\omega)$ has no physical meaning.

To conclude this chapter, we mention that the results for LTI lumped systems may not be extendible to LTI distributed systems. For example, an LTI lumped system is stable if its impulse response vanishes as $t \to \infty$. However, an LTI distributed system may be unstable even if its impulse response vanishes. An LTI lumped system is stable if its transfer function has no pole in the closed RHP. However, a distributed system can be unstable even if its transfer function is analytic (has no singular point or pole) in the closed RHP. See Ömer Morgül, "Comments on 'An unstable plant with no poles,'" *IEEE Transactions on Automatic Control,* vol. 46, p. 1843, 2001. In fact, analysis of distributed systems is not simple. Other than transmission lines and time delay, it is generally difficult to build distributed systems. On the other hand, LTI lumped systems can, as shown in this chapter, be readily analyzed. They can also, as we will show in the next chapter, be easily built. Thus we design mostly LTI lumped systems in practice.

PROBLEMS

6.1 Consider a system with impulse response

$$h(t) = 2e^{-2t} + \sin 2t$$

What is its transfer function?

6.2 Find the transfer functions of the following differential equations:

(a) $2\ddot{y}(t) + 4\dot{y}(t) + 10y(t) = \ddot{u}(t) - \dot{u}(t) - 2u(t)$

(b) $2y^{(3)}(t) + 4\ddot{y}(t) + 10\dot{y}(t) = \ddot{u}(t) + 3\dot{u}(t) + 2u(t)$

(c) $y^{(5)}(t) + 2y(t) = u^{(3)}(t) + u(t)$

6.3 Find a differential equation to describe a system with transfer function

$$H(s) = \frac{V(s)}{R(s)} = \frac{2s^2 + 5s + 3}{s^4 + 3s^3 + 10}$$

in which the input and output are denoted by $r(t)$ and $v(t)$.

6.4 Suppose the output of an LTI system excited by the input $u(t) = \sin 2t$ is measured as

$$y(t) = 4e^{-t} + 2\sin 2t - 4\cos 2t$$

What is the transfer function of the system?

6.5 Suppose the step response of an LTI system is measured as

$$y(t) = 2 + e^{-0.2t}\sin 3t - 2e^{-0.5t}$$

What is the transfer function of the system?

6.6 Find the transfer functions from u to y and from u to i of the network shown in Figure 6.21.

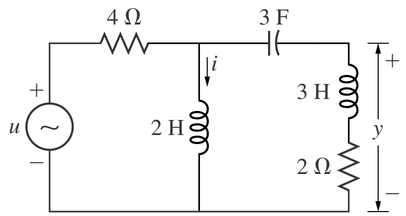


Figure 6.21

6.7 Verify that the transfer function of the op-amp circuit in Figure 6.22 is given by

$$H(s) = \frac{V_o(s)}{V_i(s)} = -\frac{Z_2(s)}{Z_1(s)}$$

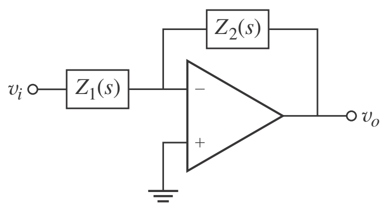


Figure 6.22

6.8 Find the poles and zeros of the following transfer functions. Also express them in zero/pole/gain form.

(a) $H_1(s) = \dfrac{3s - 6}{2s^2 + 8s + 8}$

(b) $H_2(s) = \dfrac{s^2 - s - 2}{2s^3 + 10s^2 + 18s + 10}$

Also plot their poles and zeros on a complex s-plane.

6.9 Consider the heating system shown in Figure 6.23. Let $y(t)$ be the temperature of the chamber and $u(t)$ be the amount of heat pumping into the chamber. Suppose they are related by

$$\dot{y}(t) + 0.0001y(t) = u(t)$$

If no heat is applied and if the temperature is 80°, how long will it take for the temperature to drop to 70°?

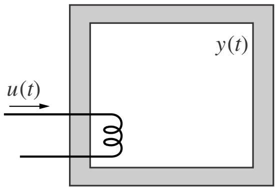


Figure 6.23

6.10 Find the impulse and step responses of a system with transfer function

$$H(s) = \frac{s^2 + 3}{(s+1)(s+2)(s-1)}$$

6.11 Verify that the impulse response of

$$H(s) = \frac{1}{s^2 + 2\sigma s + 1}$$

with $0 < \sigma < 1$, is of the form

$$h(t) = \frac{1}{\omega_r} e^{-\sigma t} \sin \omega_r t$$

What is ω_r?

6.12 Find the impulse response of a system with transfer function

$$H(s) = \frac{s - 2}{(s+2)(s^2 + 2s + 10)}$$

6.13 What is the general form of the step response of a system with transfer function

$$H(s) = \frac{10(s-1)}{(s+1)^3(s+3)}$$

6.14 Consider the transfer function

$$H(s) = \frac{N(s)}{(s+2)^4(s+0.1)(s^2+2s+10)}$$

where $N(s)$ is a polynomial of degree 7 or less and has $N(0) = 320$. What is the form of its step response?

6.15 Consider a system with transfer function

$$H(s) = \frac{1}{(s+1)(s^2+100)}$$

What is the general form of the output excited by $\sin 10t$, for $t \geq 0$? Will the output grow to infinity? Does Theorem 6.4 hold?

6.16 Repeat Problem 6.15 for the input $\sin 2t$, for $t \geq 0$.

6.17 Consider a system with transfer function

$$H(s) = \frac{1}{(s+2)(s-1.2)}$$

Compute its output excited by the input $U(s) = (s - 1.2)/[s(s + 1)]$. Is the input bounded? Does the response, $e^{1.2t}$, of the unstable pole of the system appear in the output? Is the output bounded? The only way for $e^{1.2t}$ not to appear in the output is to introduce a zero $(s - 1.2)$ in $U(s)$ to cancel the pole of $H(s)$. This type of *pole-zero cancellation* rarely occurs in practice. Thus, generally every pole of a system will be excited by an input.

6.18 Show that the positive-feedback and negative-feedback systems in Figure 3.12 are stable if and only if $|a| < 1$.

6.19 Are the systems in Problems 6.8 and 6.10 through 6.15 stable?

6.20 Use the Rough test to check the stability for each of the following polynomials.

(a) $s^5 + 3s^3 + 2s^2 + s + 1$

(b) $s^5 + 4s^4 - 3s^3 + 2s^2 + s + 1$

(c) $s^5 + 4s^4 + 3s^3 + 2s^2 + s + 1$

(d) $s^5 + 6s^4 + 23s^3 + 52s^2 + 54s + 20$

6.21 Show that a polynomial of degree 1 or 2 is a stable polynomial if and only if all coefficients are of the same sign.

6.22 Show that the polynomial

$$s^3 + a_1 s^2 + a_2 s + a_3$$

is a stable polynomial if and only if $a_1 > 0$, $a_2 > 0$, and $a_1 a_2 > a_3 > 0$.

6.23 Consider the transfer function

$$H(s) = \frac{s - 2}{s^2 + s + 100.25}$$

Compute $H(j\omega)$ at $\omega = 0$, 5, 10, 20, and 100 and then sketch roughly its magnitude and phase responses.

6.24 What is the steady-state response of the system in Problem 6.23 excited by the input

$$u(t) = 2 + \sin 5t + 3\cos 10t - 2\sin 100t$$

How long will it take to reach steady state? Is it a lowpass, bandpass, or highpass filter?

6.25 What are the time constants of the systems in Problems 6.12 through 6.14? How long will it take for their responses to reach steady state? Are the time constants of the systems in Problems 6.10 and 6.15 defined?

6.26 Consider a system described by

$$\dot{y}(t) + 0.4y(t) = 2u(t)$$

Compute its response excited by the initial condition $y(0) = 0.5$ and the input $u(t) = 1$, for $t \geq 0$. Indicate its zero-input response and zero-state response. Also indicate its transient response and steady-state response.

6.27 Consider the system in Problem 6.26. Compute its response excited by $y(0) = 0$ and $u(t) = 1$ for $t \geq 0$. Is its zero-input response identically zero? Is its transient response identically zero?

6.28 Consider the networks shown in Figures 6.24(a) and 6.24(b). Compute their transfer functions and plot their magnitude and phase responses. What are their steady-state responses excited by the input

$$u(t) = \sin 0.1t + \sin 100t$$

Are they lowpass or highpass filters? What is the 3-dB bandwidth of each filter?

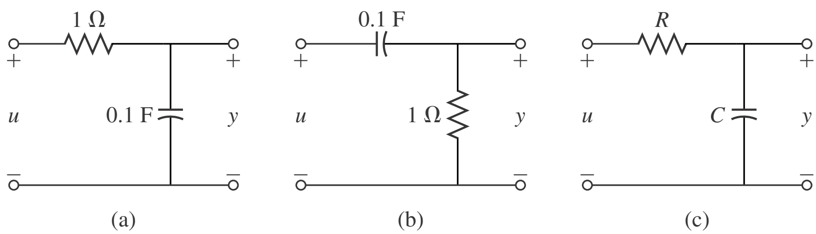


Figure 6.24

6.29 Consider the network shown in Figure 6.24(c). What are its transfer function and time constant? Show that its zero-input response is of the form

$$y_{zi}(t) = y(0)e^{-t/RC}$$

where $y(0)$ is the initial capacitor voltage.

6.30 Use

$$A \cos\theta + B \sin\theta = \alpha \cos(\theta - \phi)$$

where $\alpha = \sqrt{A^2 + B^2}$ and $\phi = \tan^{-1}(B/A)$ to show

$$0.8 \cos 2t + 0.4 \sin 2t = 0.8944 \cos(2t - 0.4636)$$

6.31 Consider the system with transfer function $1/(s + 1)^3$. Compute its step response. What is its transient response? Verify that the transient response decreases to less than 1% of its peak value in nine time constants.

6.32 Verify that for a properly designed transfer function of degree 2, a lowpass, bandpass, and highpass filter must assume, respectively, the following form:

$$H_l(s) = \frac{b}{s^2 + a_2 s + a_3}$$

$$H_b(s) = \frac{bs}{s^2 + a_2 s + a_3}$$

$$H_h(s) = \frac{bs^2}{s^2 + a_2 s + a_3}$$

6.33 Consider the transfer function

$$H(s) = \frac{k}{s+a}$$

with positive constants k and a. Show that it is a lowpass filter with 3-dB passband $[0, a]$ and bandwidth a.

6.34 Compute the Fourier transforms or spectra of

(a) $x(t) = e^{-3|t|} \qquad \text{for all } t$

(b) $x(t) = \begin{cases} e^{-3t} & \text{for all } t \geq 0 \\ e^{2t} & \text{for } t < 0 \end{cases}$

(c) $x(t) = -2e^{-t} \sin 10t \qquad \text{for } t \geq 0$

6.35 Use (6.67) to compute the spectrum of $x(t) = 1$, for all t in $(-\infty, \infty)$. Is the result correct?

6.36 Let $X(s)$ be the Laplace transform of $x(t)$ and be a proper rational function. Show that $x(t)$ approaches a constant if and only if all poles, except possibly a simple pole at $s = 0$, of $X(s)$ have negative real parts. In this case, we have

$$\lim_{t\to\infty} x(t) = \lim_{s\to 0} sX(s)$$

This is called the *final-value theorem*. Is there any difference between its derivation and the derivation of Theorem 6.4?

6.37 Verify that the formula in Problem 6.36 does not hold for $x(t) = \sin t$ and $x(t) = e^{2t}$. Give your reason.

CHAPTER 7 Realizations, Characterization, and Identification

7.1 INTRODUCTION

The transfer function is an important mathematical description of linear, time-invariant, lumped (LTIL) systems. It can be used, as discussed in the preceding chapter, to develop general properties of systems and to design frequency selective filters. It is, however, not suitable for computer computation and not convenient for physical implementation. Thus we discuss in this chapter a different mathematical description.

Every LTIL system can also be described by a state-space (ss) equation.[1] Although ss equations are not as convenient as transfer functions in developing general properties of systems and in design, they are, as we will show in this chapter, easier for computer computation. They can also be easily implemented using operational-amplifier (op-amp) circuits. Thus the ss equation is also an important mathematical description of LTIL systems.

It is possible to give a complete analytical study of ss equations; however, the study is more complex and less revealing than the corresponding study of transfer functions. Furthermore, its study has no bearing on computer computation and op-amp circuit implementation. Thus its analytical study will not be discussed in this text.

In this chapter we introduce ss equations through the *realization problem:* namely, finding ss equations from transfer functions. We then discuss basic block diagrams, op-amp circuit implementations, and computer computation of ss equations. We also compare transfer functions and ss equations and discuss their uses in determining redundancy of systems. To conclude this chapter, we discuss two identification schemes: One is based on Theorem 6.4 and the other uses linear sweep sinusoids.

The realization procedure to be discussed in this chapter is applicable to stable as well as unstable transfer functions. Thus stability and realizability are two different and independent issues. However, if an unstable transfer function is realized, the resulting ss equation or op-amp

[1]State-space equations were developed in the 1960s and thought to have the advantage over transfer functions in describing MIMO systems because all results in the SISO case can be extended to the MIMO case. In the 1970s, by considering transfer functions as ratios of two polynomials written as $H(s) = N(s)D^{-1}(s)$, the similar extension became possible. See Reference 3. Thus such advantage no longer exists.

circuit implementation will overflow, saturate, or burn out when an input is applied. Thus in practice, there is no use to realize unstable transfer functions.

Before proceeding, we mention that an ss equation is just a set of first-order differential equations. It is possible to develop an ss equation directly from a high-order differential equation. However, it is simpler to do so from a transfer function. State-space equations are expressed using matrix notations. We use boldface letters to denote vectors or matrices and use regular-face letters to denote scalars.

7.2 REALIZATIONS

Consider the following set of two equations:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \tag{7.1}$$

$$y(t) = \mathbf{c}\mathbf{x}(t) + du(t) \tag{7.2}$$

where $u(t)$ is the input, $y(t)$ the output, and $\mathbf{x}(t)$ the *state* denoting a set of initial conditions. The state $\mathbf{x}(t)$ is a column vector. If it has N components, then $\mathbf{x}$ and $\dot{\mathbf{x}}$ can be written explicitly as $\mathbf{x}(t) = [x_1(t) \;\; x_2(t) \;\; \cdots \;\; x_N(t)]'$ and $\dot{\mathbf{x}}(t) = [\dot{x}_1(t) \;\; \dot{x}_2(t) \;\; \cdots \;\; \dot{x}_N(t)]'$, where the prime denotes the transpose. We call its components *state variables*. If $\mathbf{x}$ is an $N \times 1$ vector, then $\mathbf{A}$ is an $N \times N$ square matrix, $\mathbf{b}$ is an $N \times 1$ column vector, $\mathbf{c}$ is a $1 \times N$ row vector, and d is a scalar (1×1). The matrix equation in (7.1) actually consists of N first-order differential equations; it relates the input and state and is called the *state equation*. The equation in (7.2) relates the input and output and is called the *output equation*. Note that $\mathbf{cx}$ is the product of a $1 \times N$ row vector and an $N \times 1$ column vector and is 1×1. The constant d is called the *direct transmission gain*. The set of two equations in (7.1) and (7.2) is called a CT state-space (ss) equation of dimension N.

Suppose we are given a proper transfer function $H(s)$. The problem of finding an ss equation so that its transfer function equals $H(s)$ is called the *realization* problem. The ss equation is called a *realization* of $H(s)$. We use an example to illustrate the realization procedure.

Consider the transfer function

$$H(s) = \frac{Y(s)}{U(s)} = \frac{\bar{b}_1 s^4 + \bar{b}_2 s^3 + \bar{b}_3 s^2 + \bar{b}_4 s + \bar{b}_5}{\bar{a}_1 s^4 + \bar{a}_2 s^3 + \bar{a}_3 s^2 + \bar{a}_4 s + \bar{a}_5} \tag{7.3}$$

with $\bar{a}_1 \neq 0$. The rest of the coefficients can be zero or nonzero. The transfer function is proper. We call $\bar{a}_1$ the leading coefficient. The first step in realization is to write (7.3) as

$$H(s) = \frac{N(s)}{D(s)} + d = \frac{b_1 s^3 + b_2 s^2 + b_3 s + b_4}{s^4 + a_2 s^3 + a_3 s^2 + a_4 s + a_5} + d \tag{7.4}$$

where $D(s) = s^4 + a_2 s^3 + a_3 s^2 + a_4 s + a_5$ and $N(s)/D(s)$ is strictly proper. This can be achieved by dividing the numerator and denominator of (7.3) by $\bar{a}_1$ and then carrying out a direct division as we illustrate in the next example.

EXAMPLE 7.2.1

Consider the transfer function

$$H(s) = \frac{3s^4 + 7s^3 - 2s + 10}{2s^4 + 3s^3 + 4s^2 + 7s + 5} \tag{7.5}$$

We first divide the numerator and denominator by 2 to yield

$$H(s) = \frac{1.5s^4 + 3.5s^3 - s + 5}{s^4 + 1.5s^3 + 2s^2 + 3.5s + 2.5}$$

and then carry out direct division as follows:

$$\begin{array}{r|l}
 & \qquad 1.5 \\ \hline
s^4 + 1.5s^3 + 2s^2 + 3.5s + 2.5 & 1.5s^4 + \;\; 3.5s^3 + 0\dot{s}^2 - \qquad s + \quad 5 \\
 & 1.5s^4 + 2.25s^3 + 3s^2 + 5.25s + 3.75 \\ \cline{2-2}
 & \qquad\qquad 1.25s^3 - 3s^2 - 6.25s + 1.25
\end{array}$$

Thus (7.5) can be expressed as

$$H(s) = \frac{1.25s^3 - 3s^2 - 6.25s + 1.25}{s^4 + 1.5s^3 + 2s^2 + 3.5s + 2.5} + 1.5 \tag{7.6}$$

This is in the form of (7.4).

Now we claim that the following ss equation realizes (7.4) or, equivalently, (7.3):

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -a_2 & -a_3 & -a_4 & -a_5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u(t) \tag{7.7}$$
$$y(t) = [b_1 \quad b_2 \quad b_3 \quad b_4]\mathbf{x}(t) + du(t)$$

with $\mathbf{x}(t) = [x_1(t) \;\; x_2(t) \;\; x_3(t) \;\; x_4(t)]'$. The number of state variables equals the degree of the denominator of $H(s)$. This ss equation can be obtained directly from the coefficients in (7.4). We place the denominator's coefficients, except its leading coefficient 1, with sign reversed in the first row of $\mathbf{A}$, and place the numerator's coefficients, without changing sign, directly as $\mathbf{c}$. The constant d in (7.4) is the direct transmission gain. The rest of the ss equation have fixed patterns. The second row of $\mathbf{A}$ is $[1 \;\; 0 \;\; 0 \;\; \cdots]$. The third row of $\mathbf{A}$ is $[0 \;\; 1 \;\; 0 \;\; \cdots]$ and so forth. The column vector $\mathbf{b}$ is all zero except its first entry, which is 1.

To show that (7.7) is a realization of (7.4), we must compute its transfer function. We first write the matrix equation explicitly as

$$\begin{aligned}
\dot{x}_1(t) &= -a_2x_1(t) - a_3x_2(t) - a_4x_3(t) - a_5x_4(t) + u(t) \\
\dot{x}_2(t) &= x_1(t) \\
\dot{x}_3(t) &= x_2(t) \\
\dot{x}_4(t) &= x_3(t)
\end{aligned} \tag{7.8}$$

We see that the four-dimensional state equation in (7.7) actually consists of four first-order differential equations as shown in (7.8). Applying the Laplace transform and assuming zero initial conditions yield

$$
\begin{aligned}
sX_1(s) &= -a_2X_1(s) - a_3X_2(s) - a_4X_3(s) - a_5X_4(s) + U(s) \\
sX_2(s) &= X_1(s) \\
sX_3(s) &= X_2(s) \\
sX_4(s) &= X_3(s)
\end{aligned}
$$

From the second to the last equations, we can readily obtain

$$
X_2(s) = \frac{X_1(s)}{s}, \qquad X_3(s) = \frac{X_2(s)}{s} = \frac{X_1(s)}{s^2}, \qquad X_4(s) = \frac{X_1(s)}{s^3} \tag{7.9}
$$

Substituting these into the first equation yields

$$
\left(s + a_2 + \frac{a_3}{s} + \frac{a_4}{s^2} + \frac{a_5}{s^3}\right)X_1(s) = U(s)
$$

which implies

$$
X_1(s) = \frac{s^3}{s^4 + a_2s^3 + a_3s^2 + a_4s + a_5}U(s) =: \frac{s^3}{D(s)}U(s) \tag{7.10}
$$

Substituting (7.9) and (7.10) into the following Laplace transform of the output equation in (7.7) yields

$$
\begin{aligned}
Y(s) &= b_1X_1(s) + b_2X_2(s) + b_3X_3(s) + b_4X_4(s) + dU(s) \\
&= \left(\frac{b_1s^3}{D(s)} + \frac{b_2s^3}{D(s)s} + \frac{b_3s^3}{D(s)s^2} + \frac{b_4s^3}{D(s)s^3}\right)U(s) + dU(s) \\
&= \frac{b_1s^3 + b_2s^2 + b_3s + b_4}{D(s)}U(s) + dU(s)
\end{aligned}
$$

This shows that the transfer function of (7.7) equals (7.4). Thus (7.7) is a realization of (7.4) or (7.3). The ss equation in (7.7) is said to be in the *controllable canonical form*.

We discuss a different realization. Consider the ss equation that follows:

$$
\begin{aligned}
\dot{\mathbf{x}}(t) &= \begin{bmatrix} -a_2 & 1 & 0 & 0 \\ -a_3 & 0 & 1 & 0 \\ -a_4 & 0 & 0 & 1 \\ -a_5 & 0 & 0 & 0 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}u(t) \\
y(t) &= [1 \quad 0 \quad 0 \quad 0]\mathbf{x}(t) + du(t)
\end{aligned} \tag{7.11}
$$

The $\mathbf{A}$, $\mathbf{b}$, and $\mathbf{c}$ in (7.11) are the transposes of the $\mathbf{A}$, $\mathbf{c}$, and $\mathbf{b}$ in (7.7). Thus (7.11) can be easily obtained from (7.7) or directly from the coefficients of (7.4). We will show in Section 7.5.1 that (7.11) has the same transfer function as (7.4) and thus is also a realization of (7.4). The realization in (7.11) is said to be in the *observable canonical form*.

EXAMPLE 7.2.2

Find two realizations for the transfer function in Example 7.2.1 or

$$
\begin{aligned}
H(s) &= \frac{3s^4 + 7s^3 - 2s + 10}{2s^4 + 3s^3 + 4s^2 + 7s + 5} \\
&= \frac{1.25s^3 - 3s^2 - 6.25s + 1.25}{s^4 + 1.5s^3 + 2s^2 + 3.5s + 2.5} + 1.5
\end{aligned}
\tag{7.12}
$$

Its controllable canonical form realization is, using (7.7),

$$
\begin{aligned}
\dot{\mathbf{x}}(t) &= \begin{bmatrix} -1.5 & -2 & -3.5 & -2.5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u(t) \\
y(t) &= [1.25 \quad -3 \quad -6.25 \quad 1.25]\mathbf{x}(t) + 1.5u(t)
\end{aligned}
$$

and its observable canonical form realization is, using (7.11),

$$
\begin{aligned}
\dot{\mathbf{x}}(t) &= \begin{bmatrix} -1.5 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ -3.5 & 0 & 0 & 1 \\ -2.5 & 0 & 0 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1.25 \\ -3 \\ -6.25 \\ 1.25 \end{bmatrix} u(t) \\
y(t) &= [1 \quad 0 \quad 0 \quad 0]\mathbf{x}(t) + 1.5u(t)
\end{aligned}
$$

We see that these realizations can be read out from the coefficients of the transfer function.

EXAMPLE 7.2.3

Find realizations for the transfer function

$$
H(s) = \frac{2.5s^2 - 0.87s + 0.8336}{s^2 + 0.2s + 0.82} = \frac{-1.37s - 1.1914}{s^2 + 0.2s + 0.82} + 2.5 \tag{7.13}
$$

Using (7.7) and (7.11), we can readily obtain its realizations as

$$
\begin{aligned}
\dot{\mathbf{x}}(t) &= \begin{bmatrix} -0.2 & -0.82 \\ 1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) \\
y(t) &= [-1.37 \quad -1.1914]\mathbf{x}(t) + 2.5u(t)
\end{aligned}
\tag{7.14}
$$

and

$$
\begin{aligned}
\dot{\mathbf{x}}(t) &= \begin{bmatrix} -0.2 & 1 \\ -0.82 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} -1.37 \\ -1.1914 \end{bmatrix} u(t) \\
y(t) &= [1 \quad 0]\mathbf{x}(t) + 2.5u(t)
\end{aligned}
\tag{7.15}
$$

We mention that it is possible to find infinitely many other realizations for the transfer function in (7.4), but the canonical forms in (7.7) and (7.11) are most convenient to develop and use.

EXERCISE 7.2.1

Find two realizations for each of the following transfer functions:

(a) $H_1(s) = \dfrac{2s+3}{4s+10}$

(b) $H_2(s) = \dfrac{s^2+3s-4}{2s^3+3s+1}$

(c) $H_3(s) = \dfrac{2}{s^3}$

Answers

(a) $\dot{x} = -2.5x + u$

$y = -0.5x + 0.5u$

(b) $\dot{\mathbf{x}} = \begin{bmatrix} 0 & -1.5 & -0.5 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u$

$y = [0.5 \quad 1.5 \quad -2]\mathbf{x}$

(c) $\dot{\mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u$

$y = [0 \quad 0 \quad 2]\mathbf{x}$

The MATLAB function tf2ss, an acronym for transfer function to ss equation, carries out realizations. For the transfer function in (7.12), typing

```
n=[3 7 0 -2 10];d=[2 3 4 7 5];
[a,b,c,d]=tf2ss(n,d)
```

yields

$$a = \begin{matrix} -1.5000 & -2.0000 & -3.5000 & -2.5000 \\ 1.0000 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 \\ 0 & 0 & 1.0000 & 0 \end{matrix}$$

$$b = \begin{matrix} 1 \\ 0 \\ 0 \\ 0 \end{matrix}$$

$$c = 1.2500 \;\; -3.0000 \;\; -6.2500 \;\; 1.3500$$

$$d = 1.500$$

This is the controllable canonical form realization in Example 7.2.2. In using tf2ss, there is no need to normalize the leading coefficient and to carry out direct division. Thus its use is simple and straightforward.

Improper Transfer Functions Every proper rational transfer function can be realized as in (7.1) and (7.2). Such a realization, as we will show later, can be implemented *without* using any differentiators. If a transfer function $H(s)$ is improper, then it can be expanded as

$$H(s) = H_{sp}(s) + d_0 + d_1 s + d_2 s^2 + \cdots$$

where $H_{sp}(s)$ is strictly proper. Then $H(s)$ can be realized as

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)\\ y(t) &= \mathbf{c}\mathbf{x}(t) + d_0 u(t) + d_1 \dot{u}(t) + d_2 \ddot{u}(t) + \cdots\end{aligned}$$

Its implementation requires differentiators. Thus we conclude that a rational transfer function can be implemented without using differentiators if and only if it is proper.

7.2.1 Minimal Realizations

Consider a proper rational transfer function $H(s) = N(s)/D(s)$. If $N(s)$ and $D(s)$ are coprime (have no common factors), then the degree of $H(s)$ is defined as the degree of $D(s)$. For a degree N proper transfer function, if the dimension of a realization is N, then the realization is called a *minimal-dimensional realization* or, simply, a *minimal realization*.

Consider the transfer function in (7.12). Typing `roots([3 7 0 -2 10])` in MATLAB yields the four roots of the numerator as $\{-1.78 \pm j0.47, 0.61 \pm j0.78\}$. Typing `roots([2 3 4 7 5])` yields the four roots of the denominator as $\{0.3 \pm j1.39, -1.05 \pm j0.57\}$. We see that the numerator and denominator of (7.12) have no common roots, thus the transfer function in (7.12) has degree 4. Its two realizations in Example 7.2.2 have dimension 4; thus they are minimal realizations. The transfer function in (7.13) can be shown to be coprime and thus has degree 2. Its two-dimensional realizations in (7.14) and (7.15) are all minimal realizations.

For a transfer function of degree N, it is not possible to find a realization of a dimension smaller than N. However, it is simple to find a realization of dimension $N + 1$ or higher as the next example illustrates.

EXAMPLE 7.2.4

Consider the transfer function

$$H(s) = \frac{Y(s)}{U(s)} = \frac{2}{s^2 + 3s + 2} \tag{7.16}$$

It has degree 2, and the two-dimensional ss equation

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u(t)\\ y(t) &= [0 \quad 2]\mathbf{x}(t)\end{aligned} \tag{7.17}$$

is a minimal realization.

Now if we multiply the numerator and denominator of (7.16) by $(s-1)$ to yield

$$H(s) = \frac{2(s-1)}{(s^2+3s+2)(s-1)} = \frac{2s-2}{s^3+2s^2-s-2} \tag{7.18}$$

Thus we can readily obtain, using (7.7),

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{bmatrix} -2 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}u(t) \\ y(t) &= [0 \quad 2 \quad -2]\mathbf{x}(t)\end{aligned} \tag{7.19}$$

This is a nonminimal realization of (7.16) or (7.18).

As we will see later, if we use an op-amp circuit to implement a transfer function through its nonminimal realization, the number of components used will be unnecessarily large. Furthermore, every nonminimum realization has some deficiency as we will discuss in Section 7.6. Thus there is no reason to find a nonminimal realization. When we are given a transfer function, if we first cancel out all common factors (if there are any) and then carry out realization, the resulting ss equation is automatically minimal.

7.3 BASIC BLOCK DIAGRAMS

A CT basic block diagram is a diagram that consists only of the types of elements shown in Figure 7.1. The element denoted by s^{-1} enclosed by a box in Figure 7.1(a) is an *integrator;* its output $y(t)$ equals the integration of the input $u(t)$, that is,

$$y(t) = \int_0^t u(\tau)\,d\tau + y(0) \qquad \text{(integrator)}$$

This equation is not convenient to use. Instead we assign a variable such as $x(t)$ as the output of an integrator. Then its input is $\dot{x}(t) := dx(t)/dt$. The initial condition is usually not shown and is assumed to be zero in most application. The element in Figure 7.1(b) denoted by a line with an arrow and a real number α is called a *multiplier* with gain α; its input and output are related by

$$y(t) = \alpha u(t) \qquad \text{(multiplier)}$$

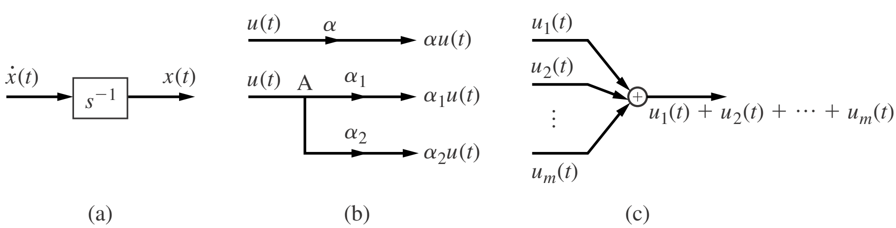


Figure 7.1 (**a**) Integrator. (**b**) Multiplier and branching point. (**c**) Adder.

Note that α can be positive or negative. If $\alpha = 1$, it is direct transmission and the arrow and α may be omitted. In addition, a signal may branch out to two or more signals with gain α_i as shown. The point A is called a *branching point*. The element denoted by a small circle with a plus sign in Figure 7.1(c) is called an *adder* or *summer*. Every adder has two or more inputs $u_i(t)$, $i = 1, 2, \ldots, m$, denoted by entering arrows, but only one output $y(t)$, denoted by a departing arrow. They are related by

$$y(t) = u_1(t) + u_2(t) + \cdots + u_m(t) \qquad \text{(adder)}$$

These elements are called *CT basic elements*. Note that a differentiator is not a basic element.

Every ss equation can be represented by a basic block diagram. We use an example to illustrate the procedure. Consider the ss equation

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \begin{bmatrix} 1.8 & 1.105 \\ -4 & -2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0.5 \\ -0.8 \end{bmatrix} u(t) \\ y(t) &= [1.1 \quad 2.4]\mathbf{x}(t) + 2.5u(t) \end{aligned} \tag{7.20}$$

with $\mathbf{x} = [x_1 \ \ x_2]'$. It consists of two first-order differential equations. In developing its basic block diagram, we must write out explicitly the two equations as

$$\dot{x}_1(t) = 1.8x_1(t) + 1.105x_2(t) + 0.5u(t) \tag{7.21}$$

and

$$\dot{x}_2(t) = -4x_1(t) - 2x_2(t) - 0.8u(t) \tag{7.22}$$

The ss equation has dimension 2; thus its block diagram needs two integrators as shown in Figure 7.2. We assign *the output of each integrator as a state variable* $x_i(t)$, for $i = 1, 2$, as shown. Then its input is $\dot{x}_i(t)$. We then use (7.21) to generate $\dot{x}_1(t)$ as shown in Figure 7.2 using dotted lines. We use (7.22) to generate $\dot{x}_2(t)$ as shown using dashed lines. The output equation in (7.20) can be generated as shown with solid lines. We see that developing a basic block diagram for an ss equation is simple and straightforward.

We plot in Figure 7.3(a) a basic block diagram for the ss equation in (7.7). The equation has dimension 4 and therefore needs four integrators. We assign the output of each integrator as a state variable x_i. Then its input is $\dot{x}_i$. Using (7.8), we can readily complete the diagram in Figure 7.3(a).

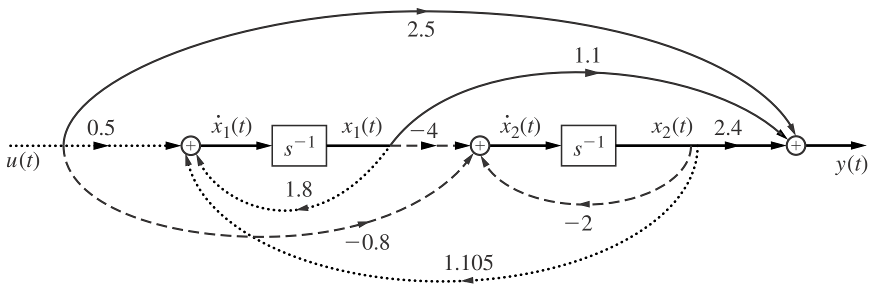


Figure 7.2 Basic block diagram of (7.20).

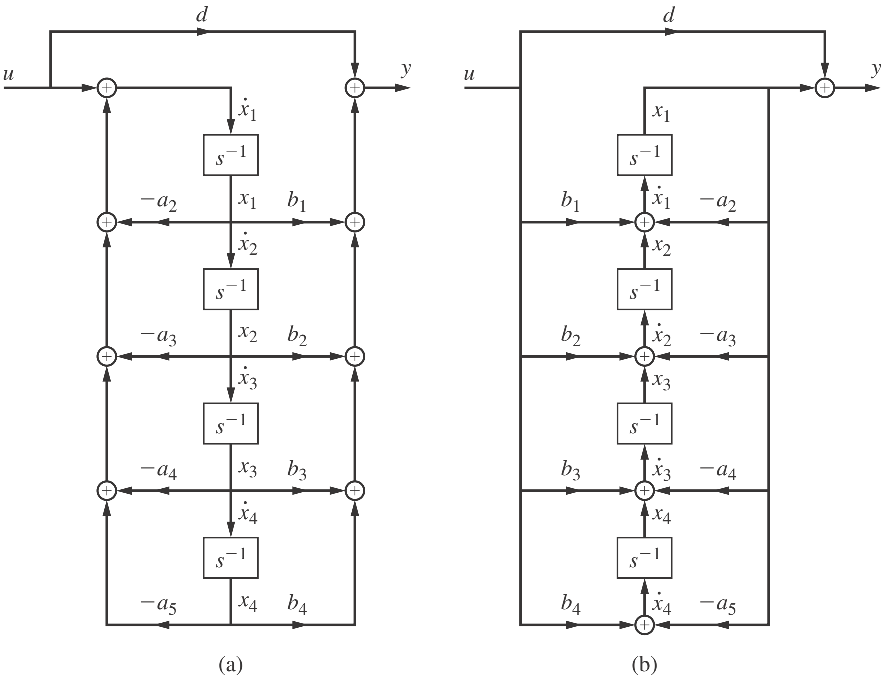


Figure 7.3 (**a**) Basic block diagram of (7.7). (**b**) Basic block diagram of (7.11).

EXERCISE 7.3.1

Verify that the basic block diagram in Figure 7.3(b) implement the observable-canonical-form ss equation in (7.11). Note that the diagram in Figure 7.3(b) can be obtained from Figure 7.3(a) by reversing every branch direction, changing every adder to a branching point, changing every branching point to an adder, and interchanging u and y. This is called transposition in graph theory.

7.3.1 Op-Amp Circuit Implementations

Every basic element can be readily implemented using an op-amp circuit as shown in Figures 7.4(a) and 7.4(b) for the integrator and adder in Figure 7.1. Thus every ss equation can be so implemented through its basic block diagram. This implementation, however, will use an unnecessarily large number of components. This section discusses how to implement an ss equation directly using the components shown in Figure 3.11 and repeated in Figures 7.4(c), 7.4(d), and 7.4(e). Note that the inverting adder in Figure 7.4(c) also carries out multiplications with gains a, b, and c. The inverting integrator in Figure 7.4(d) also carries out multiplications and additions. Note that a, b, and c must be real and positive.

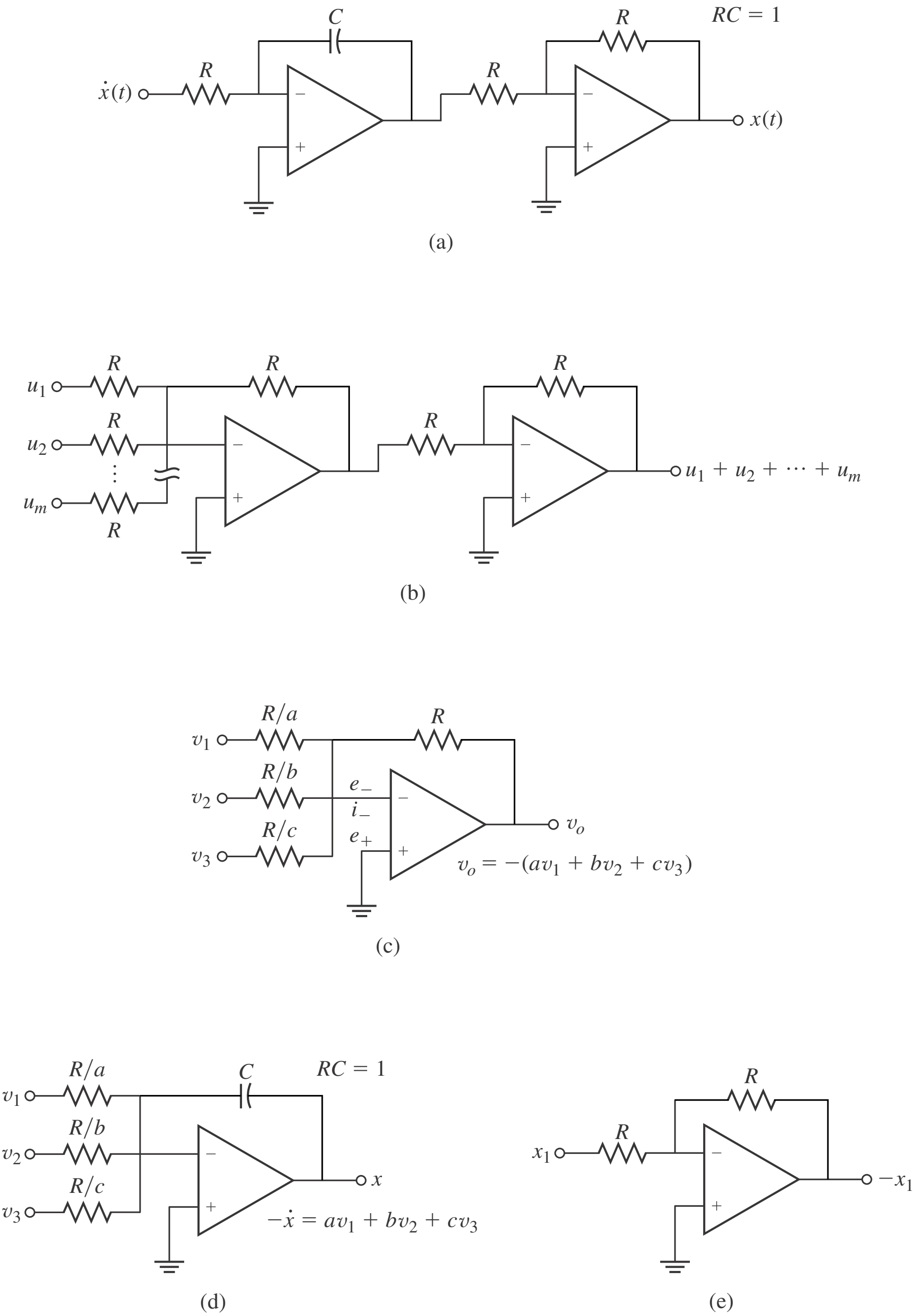


Figure 7.4 (**a**) Integrator. (**b**) Adder. (**c**) Inverting adder with multiplications. (**d**) Inverting integrator with multiplication and additions. (**e**) Inverter.

Consider the ss equation in (7.20) or

$$\begin{aligned}\dot{x}_1(t) &= 1.8x_1(t) + 1.105x_2(t) + 0.5u(t)\\ \dot{x}_2(t) &= -4x_1(t) - 2x_2(t) - 0.8u(t)\\ y(t) &= 1.1x_1(t) + 2.4x_2(t) + 2.5u(t)\end{aligned} \tag{7.23}$$

It has dimension 2, thus we need two integrators as shown in Figure 7.5(a). Let us assign the output of the left-hand-side integrator as $x_1(t)$ and the output of the right-hand-side integrator as $x_2(t)$. Then we shall generate

$$-\dot{x}_1(t) = -1.8x_1(t) - 1.105x_2(t) - 0.5u(t)$$

at the input of the left-hand-side integrator. Because all constants a, b, and c in Figure 7.4(c) are required to be positive, we must generate $-x_1(t)$ and $-x_2(t)$ from the assigned $x_1(t)$ and $x_2(t)$ by using inverters as shown in Figure 7.5(a). We also use an inverter to generate $-u(t)$ from $u(t)$. We then generate the input of the left-hand-side integrator by direct connections as shown in Figure 7.5(a). Because of the assignment of $x_2(t)$, we must generate $-\dot{x}_2(t)$ at the input of the right-hand-side integrator or

$$-\dot{x}_2(t) = 4x_1(t) + 2x_2(t) + 0.8u(t)$$

This is done as shown. Finally we assign the output of the summer as $y(t)$. Then we must generate

$$-y(t) = -1.1x_1(t) - 2.4x_2(t) - 2.5u(t)$$

at the input of the summer as shown in Figure 7.5(a). This completes the implementation of the ss equation in (7.20) or (7.23). The implementation uses 6 op amps, 2 capacitors, and 16 resistors.

We have freedom in assigning the output of each integrator as $x(t)$ or $-x(t)$. Let us check what will happen if we assign the output of the left-hand-side integrator as $x_1(t)$ and the output of the right-hand-side integrator as $-x_2(t)$ as shown in Figure 7.5(b). Using the same procedure, we can complete the op-amp circuit as shown in Figure 7.5(b). The implementation uses 5 op amps, 2 capacitors, and 14 resistors. It uses one less op amp and two less resistors than the implementation in Figure 7.5(a). The saving could be larger for higher-dimensional ss equations. Thus there is no reason to assign the outputs of all integrators as $+x_i$.

To conclude this section, we mention that if we connect two or more capacitors to an op amp as shown in Figure 3.17(b), then the number of op amps used in an implementation can be reduced. However, its implementation procedure will be less systematic. We compute the transfer function of a selected configuration with open R_i and C_i, and then we search R_i and C_i to yield the given transfer function. The implementation procedure in this section is simpler and more systematic. It is essentially an analog computer simulation diagram.

7.3.2 Stability of Op-Amp Circuits

We mentioned at the end of Section 6.7 that all practical RLC networks are stable. Thus their stability study is unnecessary. For op-amp circuits, the situation is different. We use examples to discuss the issues.

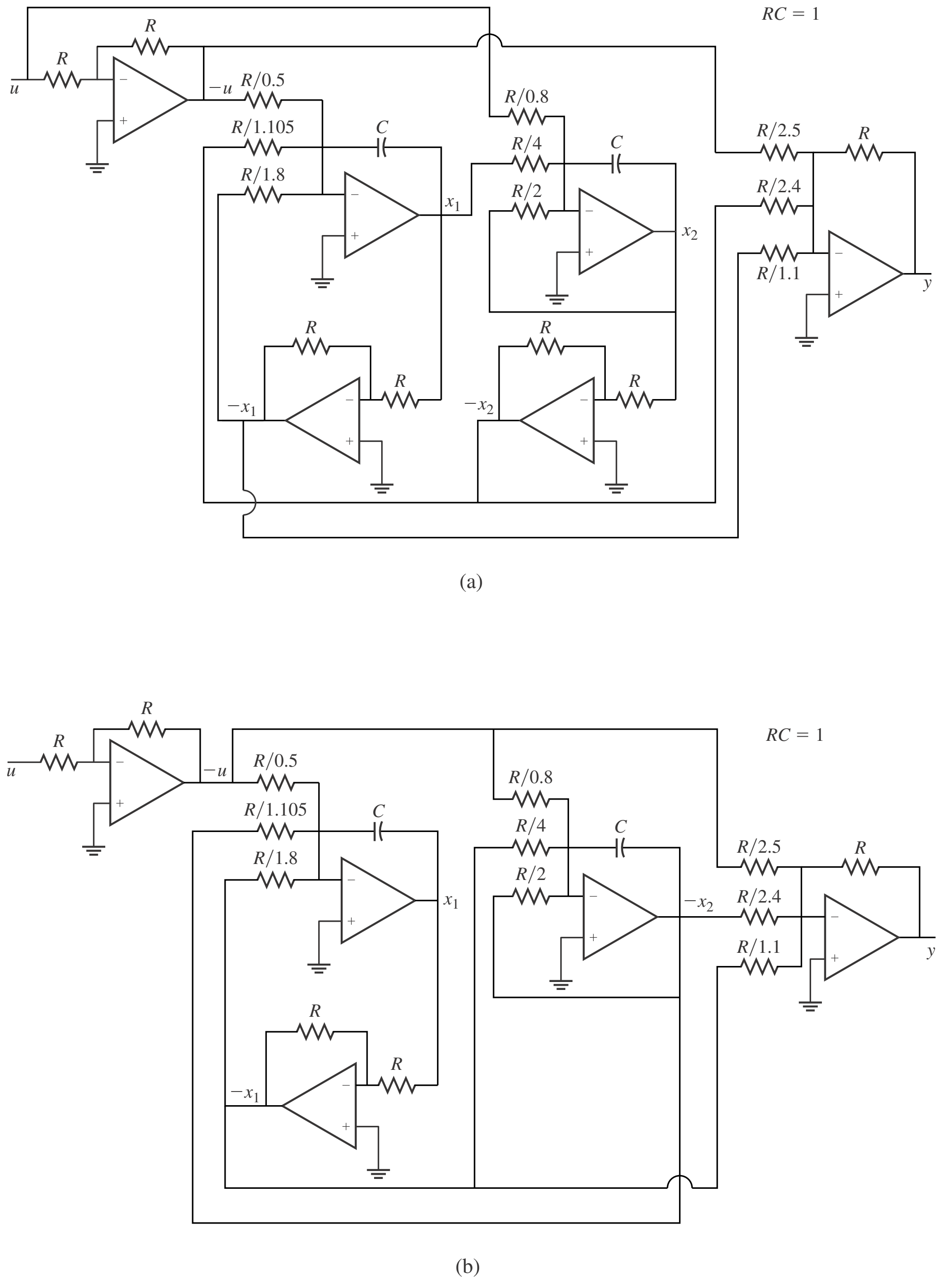


Figure 7.5 (**a**) Implementation of (7.20). (**b**) A different implementation.

Consider the transfer functions

$$H_1(s) = \frac{1}{s^2 + 2s - 3} \tag{7.24}$$

and

$$H_2(s) = \frac{1}{s^2 + 3s + 2} \tag{7.25}$$

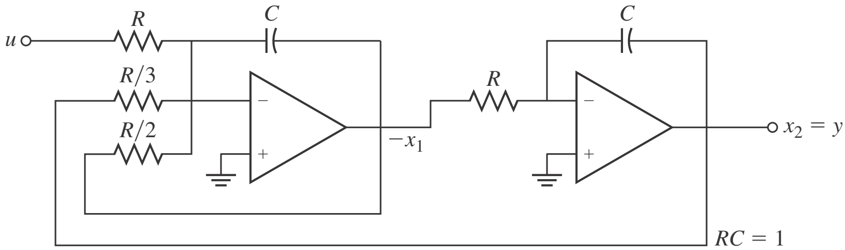
Using (7.7), we can readily obtain their ss equations as

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) \\ y(t) &= [0 \quad 1]\mathbf{x}(t) \end{aligned} \tag{7.26}$$

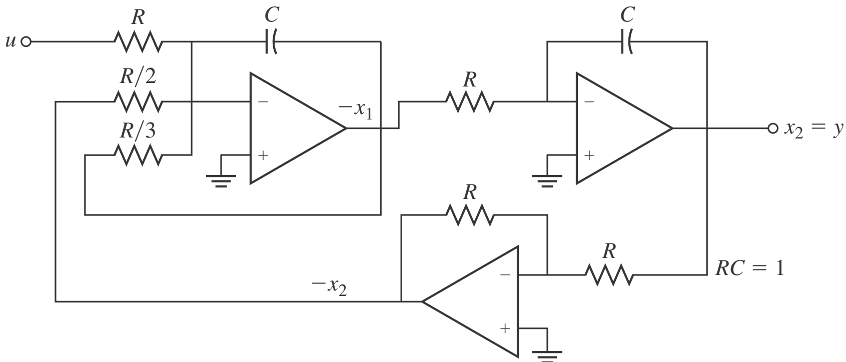
and

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) \\ y(t) &= [0 \quad 1]\mathbf{x}(t) \end{aligned} \tag{7.27}$$

The two ss equations can be implemented as shown in Figures 7.6(a) and 7.6(b).



(a)



(b)

Figure 7.6 (**a**) Unstable op-amp circuit. (**b**) Stable op-amp circuit.

We mentioned in Section 2.8.2 that negative feedback may stabilize a system and positive feedback may destabilize a system. For the circuits in Figure 7.6, all signals are fed into inverting (negative) terminals. Does this constitute as negative feedback? In fact, the situation is more complex because there is an inverter to change $x_2(t)$ into $-x_2(t)$ in Figure 7.6(b). Does this constitute a positive feedback? Thus it is difficult, if not impossible, to check stability of an op-amp circuit by inspection. To check it by measurement is simple. We apply an arbitrary input to excite it and then remove the input. Then the circuit is stable if and only if the response eventually vanishes. To check it analytically, we must compute its transfer function and then apply Theorem 6.2. The transfer function of the circuit in Figure 7.6(a) is given in (7.24). Its denominator has a negative coefficient and is therefore not a stable polynomial. Thus the circuit in Figure 7.6(a) is not stable. On the other hand, the transfer function of the op-amp circuit in Figure 7.6(b) is given in (7.25), whose denominator is a stable polynomial. Thus the circuit is stable.

To conclude this section, we mention that the stability of an op-amp circuit can also be checked directly from its ss equation. See Reference 3.

7.4 COMPUTER COMPUTATION OF STATE-SPACE EQUATIONS

Consider a system with transfer function $H(s)$ and a realization

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)\\ y(t) &= \mathbf{c}\mathbf{x}(t) + du(t)\end{aligned} \tag{7.28}$$

Although the transfer function describes only zero-state responses, its ss equation realization describes both zero-input and zero-state responses. In the following we discuss computer computation of $y(t)$ excited by some initial state $\mathbf{x}(0)$ and some input $u(t)$, for $t \geq 0$.

The first step in computer computation is to carry out discretization. The simplest way is to use the approximation

$$\dot{\mathbf{x}}(t) = \frac{\mathbf{x}(t+\Delta) - \mathbf{x}(t)}{\Delta} \tag{7.29}$$

where $\Delta > 0$ and is called the *integration step size*. Substituting (7.29) into (7.28) yields

$$\mathbf{x}(t+\Delta) = \mathbf{x}(t) + \Delta(\mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)) = (\mathbf{I} + \Delta\mathbf{A})\mathbf{x}(t) + \Delta\mathbf{b}u(t) \tag{7.30}$$

where $\mathbf{I}$ is the unit matrix with the same order as $\mathbf{A}$. Because $(1 + \Delta\mathbf{A})$ is not defined, we must use $\mathbf{x}(t) = \mathbf{I}\mathbf{x}(t)$ before combining $\mathbf{x}(t)$ and $\Delta\mathbf{A}\mathbf{x}(t)$. If we compute $\mathbf{x}(t)$ and $y(t)$ at $t = n\Delta$ for $n = 0, 1, 2, \ldots,$ then we have

$$\begin{aligned}y(n\Delta) &= \mathbf{c}\mathbf{x}(n\Delta) + du(n\Delta)\\ \mathbf{x}((n+1)\Delta) &= (\mathbf{I} + \Delta\mathbf{A})\mathbf{x}(n\Delta) + \Delta\mathbf{b}u(n\Delta)\end{aligned} \tag{7.31}$$

Using (7.31), we can compute $y(n\Delta)$, for $n = 0, 1, 2, \ldots.$ To be more specific, we use $\mathbf{x}(0)$ and $u(0)$ to compute $y(0)$ and $\mathbf{x}(\Delta)$. We then use $\mathbf{x}(\Delta)$ and $u(\Delta)$ to compute $y(\Delta)$ and $\mathbf{x}(2\Delta)$

and so forth. This can be programmed as follows:

```
Given A, b, c, d, x(0), u(t)
Select D and N, and define u(n)=u(n*D)
Compute A1=I+D*A, b1=D*b, and compute, from n=0 to N,
y(n)=c*x(n)+d*u(n)
x(n+1)=A1*x(n)+b1*u(n)
plot(n*D,y(n))
```

We see that the algorithm is simple and can be easily programmed. The actual programming depends on the language used such as FORTRAN, C, or C++ and is outside the scope of this text. The algorithm involves only repetitive multiplications and additions and thus is numerically robust. Thus computer computation of system responses is all based on ss equations.

The discretization method using (7.29) is the simplest and yields the least accurate result for the same integration step size. There are more accurate discretization methods which we will not discuss. Instead, we discuss how to use MATLAB functions to compute responses of CT systems.

7.4.1 MATLAB Computation

Consider the ss equation

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{bmatrix} 1.8 & 1.105 \\ -4 & -2 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 0.5 \\ -0.8 \end{bmatrix}u(t) \\ y(t) &= [1.1 \quad 2.4]\mathbf{x}(t) + 2.5u(t)\end{aligned} \tag{7.32}$$

In MATLAB, matrices are expressed row by row, separated by semicolons. For example, the **A**, **b**, **c**, and d in (7.32) are expressed as a=[1.8 1.105;-4 -2]; b=[0.5;-0.8]; c=[1.1 2.4];d=2.5. Note that b has a semicolon because it has two rows, whereas c has no semicolon because it has only one row. Suppose we want to find the output of the system excited by the initial state $x(0) = [2 \ \ -4]'$ and the input $u(t) = 1/(t+1)$ from $t = 0$ to $t_f = 50$, where t_f is the final time. We use the MATLAB function lsim, an acronym for *linear simulation,* to compute the output. Let us type

```
a=[1.8 1.105;-4 -2];b=[0.5;-0.8];c=[1.1 2.4];d=2.5;
cat=ss(a,b,c,d);
t=0:0.01:50;
x0=[2;-4];u=1./(t+1);
[y,t]=lsim(cat,u,t,x0);
subplot(2,1,1)
plot(t,y)
subplot(2,1,2)
lsim(cat,u,t,x0)
```

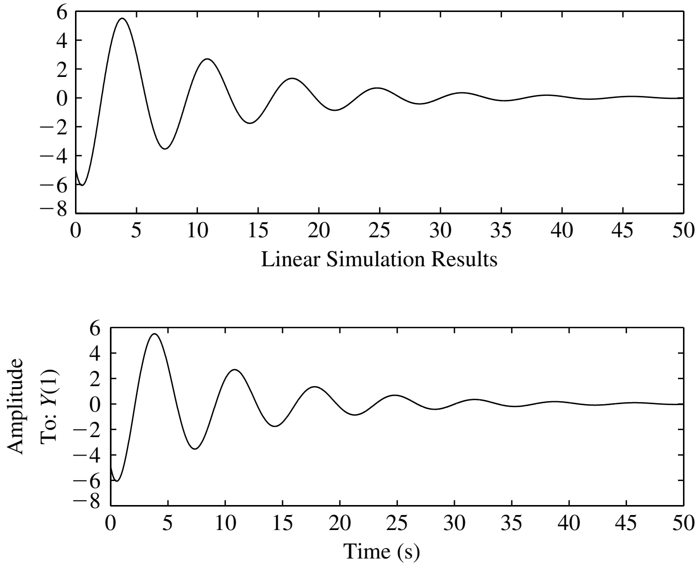


Figure 7.7 Outputs of (7.32) excited by $\mathbf{x}(0) = [2\ \ -4]'$ and $u(t) = 1/(t+1)$.

The first line expresses the matrices in MATLAB format. Note that every statement in MATLAB is executed and stored in memory. A semicolon (;) at its end suppresses its display or print. The second line defines the system. We call the system cat which is defined using the state-space model, denoted by ss. The third line indicates the time interval of computation from $t = 0$ to $t_f = 50$ with integration step size selected as 0.01. The fourth line is the initial state and the input. Note the use of dot slash (./). Most operations in MATLAB are defined for matrices; with a dot, an operation becomes element by element. We then use lsim to compute the output. The output is plotted in Figure 7.7(top) using the function plot. The function lsim generates the output at discrete time instants. Because the output of a CT system is a CT signal, we must interpolate the output between computed values. The function plot carries out the interpolation by connecting two neighboring values by a straight line, as shown in Figure 5.8(c). If the function lsim does not have the left-hand argument as shown in the last line of the preceding program, then MATLAB automatically plot the output as shown in Figure 7.7(bottom). Note that the MATLAB programs are developed for multi-input and multi-output systems. Our system has only one input and one output. Thus Figure 7.7(bottom) shows $Y(1)$ in its vertical ordinate.

Before proceeding, we mention that a CT system can also be defined using the transfer function model such as sys=tf(nu,de), where nu and de are the numerator's and denominator's coefficients. However, if we use a transfer function, all initial conditions are automatically assumed to be zero. Thus we cannot use the transfer function model to generate the plots in Figure 7.7. If we compute zero-state responses, there is no difference in using the state-space or transfer function model. If we use the latter, MATLAB first transforms the transfer function, using tf2ss, into an ss equation and then carries out computation.

Consider the ss equation in (7.32). Its transfer function can be computed as

$$H(s) = \frac{2.5s^2 - 0.87s + 0.8336}{s^2 + 0.25s + 0.82}$$

See Example 7.5.5 in a later section. To compute the impulse and step responses of (7.32), we can use the MATLAB functions impulse and step. They compute, respectively, impulse and step responses of systems. For example, typing

```
a=[1.8 1.105;-4 -2];b=[0.5;-0.8];c=[1.1 2.4];d=2.5;
[or nu=[2.5 -0.87 0.8336];de=[1 0.25 0.82];]
subplot(1,2,1)
step(a,b,c,d) [or step(nu,de)]
subplot(1,2,2)
impulse(a,b,c,d) [or impulse(nu,de)]
```

generates the step and impulse responses in Figure 7.8. In the program, we did not specify the time interval to be computed, nor did we specify the integration step size. They are selected automatically by the functions. Note that the step response approaches $0.8336/0.82 = 1.02$ and the impulse response approaches 0.

The impulse response in Figure 7.8(b) is in fact incorrect. The function impulse sets automatically the direct transmission gain to zero. Thus it yields a correct response only if $d = 0$. If $d \neq 0$, the correct impulse response should be the one in Figure 7.8(b) plus an impulse at $t = 0$ with weight $d = 2.5$.

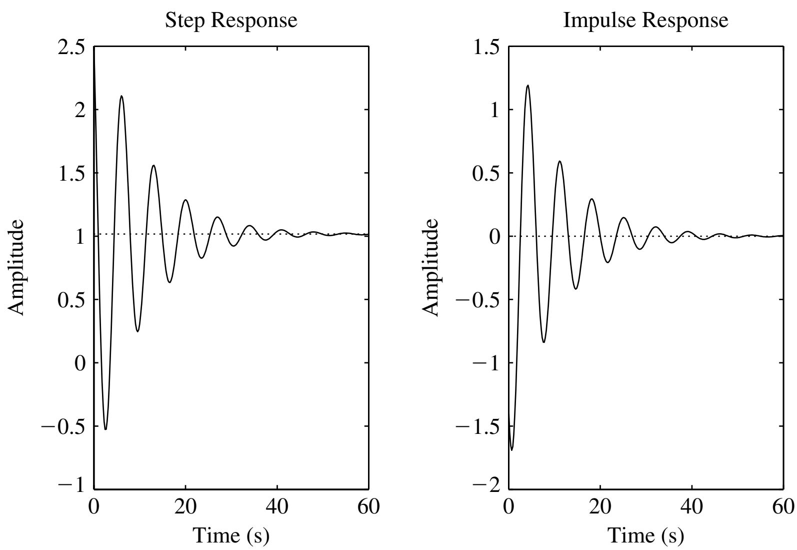


Figure 7.8 (**a**) Step response of (7.32). (**b**) Incorrect impulse response of (7.32).

Computing Impulse Responses We discuss how impulse responses are computed in MATLAB. Consider the state equation in (7.28). We show that the impulse input $u(t) = \delta(t)$ will transfer the initial state $\mathbf{x}(0) = \mathbf{0}$ to $\mathbf{x}(0+) = \mathbf{b}$. Integrating (7.28) from 0 to 0_+ yields

$$\int_0^{0_+} \dot{\mathbf{x}}(t)\,dt = \mathbf{x}(t)\Big|_{t=0}^{0_+} = \mathbf{x}(0_+) - \mathbf{x}(0) = \mathbf{x}(0_+)$$
$$= \int_0^{0_+} \mathbf{A}\mathbf{x}(t)\,dt + \int_0^{0_+} \mathbf{b}\delta(t)\,dt = \mathbf{0} + \mathbf{b} = \mathbf{b}$$

This establishes the assertion. Thus the zero-state response of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$, $y(t) = \mathbf{c}x(t)$ excited by $u(t) = \delta(t)$ (that is, impulse response) equals the *zero-input* response of $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$, $y(t) = \mathbf{c}\mathbf{x}(t)$ excited by the initial state $\mathbf{x}(0) = \mathbf{b}$. For the ss equation in (7.32) with $d = 0$, we type

```
a=[1.8 1.105;-4 -2];b=[0.5;-0.8];c=[1.1 2.4];d=0;
t=0:0.1:60;u=zeros(1,601);
cat=ss(a,b,c,d);
lsim(cat,u,t,b).
```

Note that `zeros(1,601)` is a row vector consisting of 601 zeros (1×601 zero vector), where 601 is the number of time instants in `0:0.1:60`. Thus the program computes the zero-input response excited by the initial state $\mathbf{x}(0) = \mathbf{b}$. The result is shown in Figure 7.8(b). It is the impulse response of (7.32) with $d = 0$ and is obtained without generating an impulse.

7.5 DEVELOPING STATE-SPACE EQUATIONS

Consider an LTIL system. If we compute first its transfer function, then we can obtain, using the realization procedure discussed in Section 7.2, an ss equation to describe the system. However, it is also possible to develop directly an ss equation without computing first its transfer function as we discuss in this section.

The first step in developing an ss equation is to select state variables. As discussed in Section 2.3, we can select as state variables the position and velocity of a mass for mechanical systems and all capacitor voltages and all inductor currents for RLC networks. Generally, state variables are associated with energy storage elements. The position of a mass is associated with potential energy (when it is in vertical movement) or with the energy stored in the compressed or stretched spring (when it is connected to a spring). The velocity is associated with kinetic energy. In RLC networks, capacitors and inductors are called energy storage elements because they can store energy in their electric and magnetic fields. Note that resistors cannot store energy (all energy is dissipated as heat), thus no resistor voltages nor currents can be selected as state variables. Once state variables are chosen, we then use physical laws such as Newton's law and Kirchhoff's voltage and current laws to develop ss equations. This is illustrated with examples.

EXAMPLE 7.5.1 (Mechanical System)

Consider a block with mass m connected to a wall through a spring as shown in Figure 2.15(a) and discussed in Example 3.8.4. As discussed there, the mechanical system can be described by the second-order differential equation in (3.35) or

$$m\ddot{y}(t) + f\dot{y}(t) + ky(t) = u(t) \tag{7.33}$$

Let us select the position and velocity of the mass as state variables:

$$x_1(t) = y(t) \quad \text{and} \quad x_2(t) = \dot{y}(t)$$

Then we have

$$\dot{x}_1(t) = \dot{y}(t) = x_2(t) \quad \text{and} \quad \dot{x}_2(t) = \ddot{y}(t)$$

Substituting (7.33) into $\dot{x}_2(t)$ yields

$$\begin{aligned}\dot{x}_2(t) &= (1/m)\,(-f\dot{y}(t) - ky(t) + u(t)) \\ &= (-f/m)x_2(t) + (-k/m)x_1(t) + (1/m)u(t)\end{aligned}$$

They can be expressed in matrix form as

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{bmatrix} 0 & 1 \\ -k/m & -f/m \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 0 \\ 1/m \end{bmatrix}u(t) \\ y(t) &= [1 \quad 0]\mathbf{x}(t) + 0 \cdot u(t)\end{aligned} \tag{7.34}$$

where $\mathbf{x} = [x_1 \ \ x_2]'$. This is a two-dimensional ss equation that describes the mechanical system.

We next discuss how to develop ss equations for RLC networks.

Procedure of Developing ss Equations for RLC Networks

1. Assign all capacitor voltages and all inductor currents as state variables.[2] We then write down all capacitor currents and all inductor voltages using the branch characteristics shown in Figure 7.9. If we select a capacitor current (an inductor voltage) as a state variable, then its voltage (current) is an integration and is complicated. Thus we select capacitor voltages and inductor currents. The dimension of the ss equation to be developed equals the total number of capacitors and inductors.

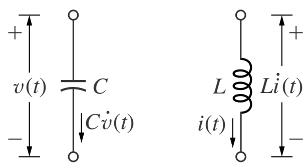


Figure 7.9 Branch characteristics of capacitor and inductor.

[2]There are some exceptions. See Problems 7.16 through 7.18.

2. Use Kirchhoff's current or voltage law to express every resistor's current and voltage in terms of state variables and, if necessary, inputs.
3. Use Kirchhoff's current or voltage law to express the derivative of each state variable in terms of state variables and inputs.

This is illustrated with examples.

EXAMPLE 7.5.2 (Electrical System)

Consider the network shown in Figures 3.8 and 6.3(a). Its transfer function was computed in Example 6.3.3 as

$$H(s) = \frac{4}{8s^2 + 25s + 7} = \frac{0.5}{s^2 + 3.125s + 0.875} \tag{7.35}$$

Its controllable-canonical-form realization is

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \begin{bmatrix} -3.125 & -0.875 \\ 1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) \\ y(t) &= [0 \quad 0.5]\mathbf{x}(t) + 0 \cdot u(t) \end{aligned} \tag{7.36}$$

We now develop directly its ss equation. We replot the network in Figure 7.10. Let us select the inductor current with the selected direction as state variable $x_1(t)$ and select the capacitor voltage with the selected polarity as state variable $x_2(t)$. Then the voltage across the inductor is $L\dot{x}_1(t) = \dot{x}_1(t)$ and the current through the capacitor is $C\dot{x}_2(t) = 2\dot{x}_2(t)$. The current through the 3-Ω resistor is $x_1(t)$; thus its voltage is $3x_1(t)$. The voltage across the 4-Ω resistor is $x_2(t)$; thus its current is $x_2(t)/4 = 0.25x_2(t)$. Applying Kirchhoff's current law to the node denoted by A yields

$$x_1(t) = 0.25x_2(t) + 2\dot{x}_2(t)$$

which implies

$$\dot{x}_2(t) = 0.5x_1(t) - 0.125x_2(t) \tag{7.37}$$

Applying Kirchhoff's voltage law around the outer loop of Figure 7.10 yields

$$u(t) = 3x_1(t) + \dot{x}_1(t) + x_2(t)$$

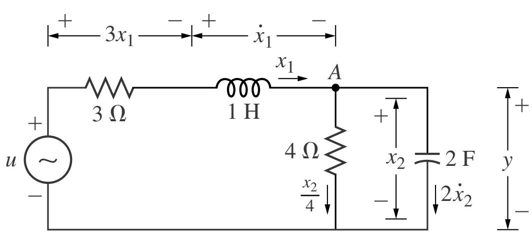


Figure 7.10 Network.

which implies

$$\dot{x}_1(t) = -3x_1(t) - x_2(t) + u(t) \tag{7.38}$$

Clearly we have $y(t) = x_2(t)$, which together with (7.37) and (7.38) can be expressed in matrix form as

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{bmatrix} -3 & -1 \\ 0.5 & -0.125 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u(t) \\ y(t) &= [0 \quad 1]\mathbf{x}(t)\end{aligned} \tag{7.39}$$

with $\mathbf{x} = [x_1 \;\; x_2]'$. This is a two-dimension ss equation that describes the network.

The network can be described by the ss equations in (7.36) and (7.39). The two equations look completely different. However, they are mathematically equivalent and either form can be used. See Reference 3. The state variables in (7.39) are associated with capacitor voltage and inductor current. The state variables in (7.36) may not have any physical meaning. This is the advantage of (7.39) over (7.36). However, in op-amp circuit implementation, the canonical form in (7.36) may use a smaller number of components than (7.39). Thus both forms are useful.

EXAMPLE 7.5.3 (Electrical System)

Consider the network shown in Figure 7.11. It consists of one resistor, one capacitor, and two inductors. The input $u(t)$ is a voltage source and the output is chosen as the voltage across the 2-H inductor as shown.

We first select the capacitor voltage $x_1(t)$ and the two inductor currents $x_2(t)$ and $x_3(t)$ as state variables. Then the capacitor current is $2\dot{x}_1(t)$ and the two inductors voltages are $\dot{x}_2(t)$ and $2\dot{x}_3(t)$. The resistor's current is clearly $x_3(t)$. Thus its voltage is $0.25x_3(t)$.

The next step is to develop $\dot{x}_i(t)$ from the network using Kirchhoff's current or voltage law. The capacitor current $2\dot{x}_1(t)$ clearly equals $x_2(t) + x_3(t)$. Thus we have

$$\dot{x}_1(t) = 0.5x_2(t) + 0.5x_3(t)$$

The voltage across the 1-H inductor is, using the left-hand-side loop of the network,

$$\dot{x}_2(t) = u(t) - x_1(t)$$

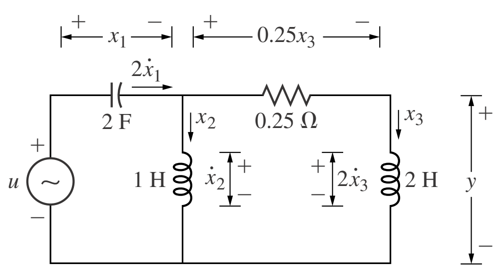


Figure 7.11 Network.

The voltage across the 2-H inductor is, using the outer loop,

$$2\dot{x}_3(t) = u(t) - x_1(t) - 0.25x_3(t)$$

or

$$\dot{x}_3(t) = -0.5x_1(t) - 0.125x_3(t) + 0.5u(t)$$

These three equations can be arranged in matrix form as

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 0.5 & 0.5 \\ -1 & 0 & 0 \\ -0.5 & 0 & -0.125 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 1 \\ 0.5 \end{bmatrix} u(t) \tag{7.40}$$

The voltage across the 2-H inductor is $2\dot{x}_3$; thus we have $y(t) = 2\dot{x}_3(t)$, which, however, is not in the form of $\mathbf{cx} + du$. From the outer loop of Figure 7.11, we can obtain

$$y(t) = -x_1(t) - 0.25x_3(t) + u(t) = [-1 \quad 0 \quad -0.25]\mathbf{x}(t) + u(t) \tag{7.41}$$

This can also be obtained by substituting $\dot{x}_3$ in the last equation of (7.40) into $y = 2\dot{x}_3$. The set of two equations in (7.40) and (7.41) describes the network. It is a three-dimensional ss equation.

EXAMPLE 7.5.4 (Electromechanical System)[3]

Consider the armature-controlled dc motor studied in Figure 3.9 and repeated in Figure 7.12. The input $u(t)$ is applied to the armature circuit with resistance R_a and inductance L_a. The generated torque $T(t)$ is proportional to armature current i_a as $T(t) = k_t i_a(t)$. The torque is used to drive the load with moment of inertia J. The viscous friction coefficient between the motor shaft and bearing is f. Let θ be the angular displacement of the load. Then we have, as discussed in (3.38),

$$k_t i_a(t) = f\dot{\theta}(t) + J\ddot{\theta}(t) \tag{7.42}$$

Now we shall develop an ss equation to describe the system. Let us select state variables of the system as

$$x_1(t) = i_a(t), \qquad x_2(t) = \theta(t), \qquad x_3(t) = \dot{\theta}(t) = d\theta(t)/dt$$

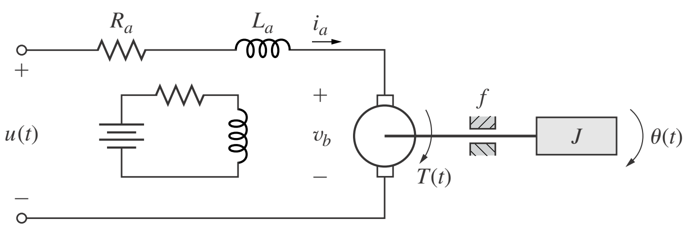


Figure 7.12 Armature-controlled dc motor.

[3]This example may be skipped without loss of continuity.

From the definition, we have $\dot{x}_2(t) = x_3(t)$. Substituting these into (7.42) yields

$$J\dot{x}_3(t) = k_t x_1(t) - f x_3(t)$$

From the armature circuit, we have, as derived in (3.39),

$$L_a \dot{x}_1(t) = u(t) - R_a x_1(t) - k_v x_3(t)$$

They can be combined as

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -R_a/L_a & 0 & -k_v/L_a \\ 0 & 0 & 1 \\ k_t/J & 0 & -f/J \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1/L_a \\ 0 \\ 0 \end{bmatrix} u(t) \tag{7.43}$$

This is the state equation. If the angular displacement θ is selected as the output, then the output equation is $y(t) = \theta(t) = x_2(t)$ or

$$y(t) = [0 \quad 1 \quad 0]\mathbf{x}(t) + 0 \cdot u(t) \tag{7.44}$$

The set of the two equations in (7.43) and (7.44) is a three-dimensional ss equation that describes the motor in Figure 7.12.

7.5.1 From State-Space Equations to Transfer Functions

We showed in the preceding section that it is possible to develop ss equations directly for systems without first computing their transfer functions. Once an ss equation is computed, it is natural to ask what is its transfer function. This subsection discusses this problem.

Consider the ss equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \tag{7.45}$$

$$y(t) = \mathbf{c}\mathbf{x}(t) + du(t) \tag{7.46}$$

Applying the vector version of (6.11) to (7.45) and assuming zero initial condition yield

$$s\mathbf{X}(s) = \mathbf{A}\mathbf{X}(s) + \mathbf{b}U(s)$$

which can be written as

$$(s\mathbf{I} - \mathbf{A})\mathbf{X}(s) = \mathbf{b}U(s)$$

where $\mathbf{I}$ is the unit matrix of the same order as $\mathbf{A}$ and has the property $\mathbf{X}(s) = \mathbf{I}\mathbf{X}(s)$. Without introducing $\mathbf{I}$, $s - \mathbf{A}$ is not defined. Premultiplying $(s\mathbf{I} - \mathbf{A})^{-1}$ to the preceding equation yields $\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}U(s)$. Substituting this into the Laplace transform of (7.46) yields

$$Y(s) = \mathbf{c}\mathbf{X}(s) + dU(s) = \mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}U(s) + dU(s) \tag{7.47}$$

Note that $\mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}$ is the product of $1 \times N$, $N \times N$, and $N \times 1$ matrices and is a scalar. Thus the transfer function of the ss equation is

$$H(s) = \mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} + d \tag{7.48}$$

We give an example.

EXAMPLE 7.5.5

Consider the ss equation in (7.20) or

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 1.8 & 1.105 \\ -4 & -2 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 0.5 \\ -0.8 \end{bmatrix}u(t)$$
$$y(t) = [1.1 \quad 2.4]\mathbf{x}(t) + 2.5u(t) \tag{7.49}$$

To compute its transfer function, we first compute the inverse of the matrix $s\mathbf{I} - \mathbf{A}$. Computing the inverse of a square matrix of order 3 or higher is complicated. However, the inverse of a 2×2 matrix is quite simple and is given by

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}^{-1} = \frac{1}{\alpha\delta - \beta\gamma}\begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} \tag{7.50}$$

We interchange the positions of the diagonal entries and change the sign of the off-diagonal entries (without changing positions), and then we divide every entry by the determinant of the matrix. The resulting matrix is the inverse. Applying the formula, we have

$$s\mathbf{I} - \mathbf{A} = s\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1.8 & 1.105 \\ -4 & -2 \end{bmatrix} = \begin{bmatrix} s-1.8 & -1.105 \\ 4 & s+2 \end{bmatrix}$$

and

$$\begin{bmatrix} s-1.8 & -1.105 \\ 4 & s+2 \end{bmatrix}^{-1} = \frac{1}{(s-1.8)(s+2)-(-1.105)(4)}\begin{bmatrix} s+2 & 1.105 \\ -4 & s-1.8 \end{bmatrix} \tag{7.51}$$

The determinant of the matrix can be simplified as $s^2 + 0.2s + 0.82$; it is a polynomial of degree 2. Substituting (7.51) into (7.48) for the ss equation in (7.49) yields

$$\begin{aligned} H(s) &= \frac{1}{s^2+0.2s+0.82}[1.1 \quad 2.4]\begin{bmatrix} s+2 & 1.105 \\ -4 & s-1.8 \end{bmatrix}\begin{bmatrix} 0.5 \\ -0.8 \end{bmatrix} + 2.5 \\ &= \frac{1}{s^2+0.2s+0.82}[1.1 \quad 2.4]\begin{bmatrix} 0.5s+0.116 \\ -0.8s-0.56 \end{bmatrix} + 2.5 \\ &= \frac{-1.37s-1.2164}{s^2+0.2s+0.82} + 2.5 = \frac{2.5s^2-0.87s+0.8336}{s^2+0.2s+0.82} \end{aligned}$$

This is the transfer function of the ss equation in (7.49).

Other than the formulation of matrices in the preceding sections, this is the only matrix operations used in this text. You may proceed even if you do not follow the steps. Computing the transfer function of an ss equation can be easily carried out using the MATLAB function ss2tf, an acronym for state space to transfer function. For the ss equation in (7.49), typing

```
a=[1.8 1.105;-4 -2];b=[0.5;-0.8];c=[1.1 2.4];d=2.5;
[n,d]=ss2tf(a,b,c,d)
```

yields n=[2.5 -0.87 0.8336],d=[1 0.2 0.82]. They are the coefficients of the numerator and denominator of the transfer function.

EXERCISE 7.5.1

Find an ss equation to describe the network in Figure 3.3(a) and then find its transfer function.

Answers

$\dot{x} = -0.5x + 0.5u$, $y = x$, $0.5/(s + 0.5)$.

We discuss a general property of (7.48). As we can see from (7.51), every entry of $(s\mathbf{I} - \mathbf{A})^{-1}$ is strictly proper. Thus $\mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}$ is also strictly proper. In other words, if an ss equation has no direct transmission gain or $d = 0$, then its transfer function is strictly proper. However, if $d \neq 0$, then its transfer function is biproper. Thus, $H(s)$ is biproper if and only if $d \neq 0$.

To conclude this subsection, we show that if $\{\mathbf{A}, \mathbf{b}, \mathbf{c}, d\}$ is a realization of $H(s)$, so is $\{\mathbf{A}', \mathbf{c}', \mathbf{b}', d\}$, where the prime denotes the transpose. Indeed, because $H(s)$ and d are scalars (1×1), we have $H'(s) = H(s)$ and $d' = d$. Using $(\mathbf{MPQ})' = \mathbf{Q}'\mathbf{P}'\mathbf{M}'$ and $\mathbf{I}' = \mathbf{I}$, we take the transpose of (7.48):

$$H'(s) = H(s) = [\mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} + d]' = \mathbf{b}'(s\mathbf{I} - \mathbf{A}')^{-1}\mathbf{c}' + d$$

This shows the assertion. Thus if (7.7) is a realization of (7.4), so is (7.11).

7.6 COMPLETE CHARACTERIZATION BY TRANSFER FUNCTIONS

If an ss equation is developed directly from a system, then the equation describes not only the relationship between its input and output but also its internal structure. Thus the ss equation is called an *internal description*. In contrast, the transfer function is called an *external description* because it does not reveal anything inside the system. Whenever we use a transfer function, all initial conditions must be assumed to be zero. Thus a transfer function describes only zero-state responses, whereas an ss equation describes not only zero-state but also zero-input responses. Thus it is natural to ask, Does the transfer function describe fully a system or is there any information of the system missing from the transfer function? This is discussed next.

Consider the network shown in Figure 7.13(a). It has an LC loop connected in series with the current source. Even though the input will excite some response in the LC loop, because the current $i(t)$ shown always equals the input $u(t)$, the network, as far as the output $y(t)$ is concerned, can be reduced to the one in Figure 7.13(aa). In other words, because the response in the LC loop will not appear at the output terminal, the two networks in Figures 7.13(a) and 7.13(aa) have the same transfer function. For the network in Figure 7.13(b), if the initial conditions of L and C are zero, the current in the LC loop will remain zero no matter what input is applied. Thus the LC loop will not be excited by any input and, consequently, will not contribute any voltage to the output, and the network can be reduced to the one in Figure 7.13(bb) in computing the transfer function. For the network in Figure 7.13(c), because the four resistors are the same, if the initial voltage across the capacitor is zero, the voltage will remain zero no matter what input is applied. Thus the transfer functions of the two networks in Figures 7.13(c)

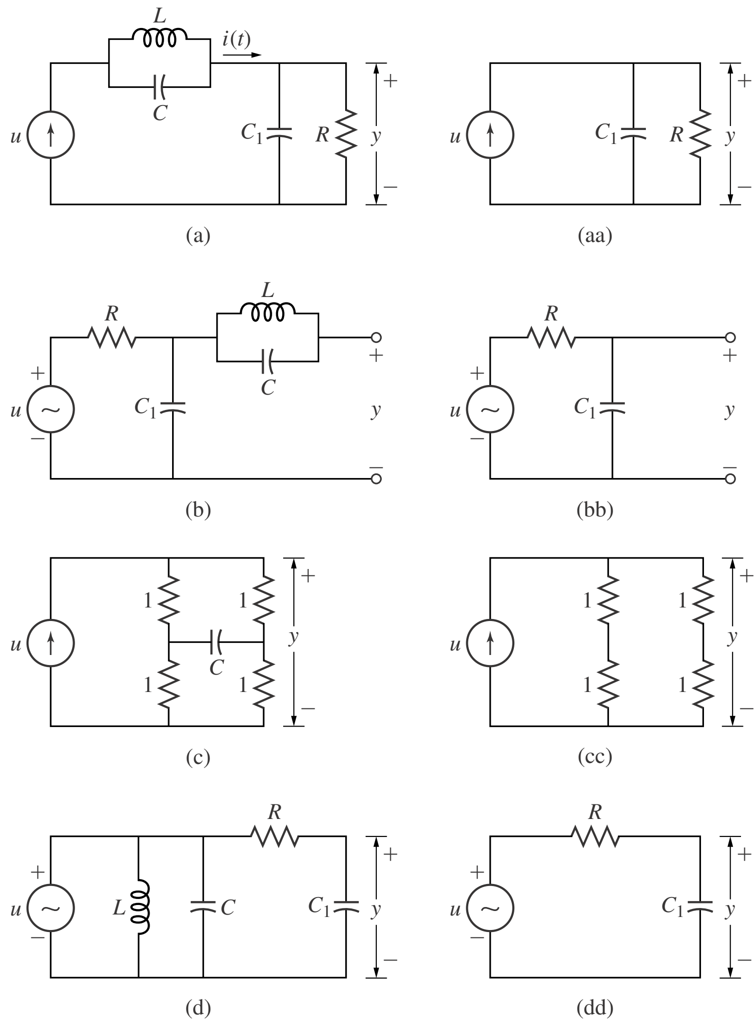


Figure 7.13 (**a**) Unobservable network. (**aa**) Reduced network. (**b**) Uncontrollable network. (**bb**) Reduced network. (**c**) Uncontrollable and unobservable network. (**cc**) Reduced network. (**d**) Unobservable network. (**dd**) Reduced network.

and 7.13(cc) are the same. The network in Figure 7.13(d) is the dual of Figure 7.13(a). The capacitor with capacitance C and the inductor with inductance L connected in parallel with the voltage source $u(t)$ can be deleted in computing its transfer function. It is clear that all networks in Figures 7.13(a) through 7.13(d) cannot be described fully by their transfer functions.

The network in Figure 7.13(a) has two capacitors and one inductor and thus has three state variables. If we develop a three-dimensional ss equation, then the network is fully described by the ss equation. A similar remark applies to the networks in Figures 7.13(c) through 7.13(d). In other words, all networks on the left-hand side of Figure 7.13 can be fully described by their ss equations but not by their transfer functions. In this sense, ss equations are more general than transfer functions.

In practice, we try to design a simplest possible system to achieve a given task. As far as the input and output are concerned, all networks on the left-hand side of Figure 7.13 can be replaced by the simpler networks on the right-hand side. In other words, the networks on the left-hand side have some redundancies or some unnecessary components, and we should not design such systems. It turns out that such redundancies can be detected from their transfer functions as well as ss equations, as we discuss next.

Consider an RLC network with proper transfer function $H(s)$. We assume that the denominator and numerator of $H(s)$ have no common factors. Then the degree of $H(s)$ is defined as the degree of its denominator. Now if the number of energy-storage elements (that is, inductors and capacitors) equals the degree of $H(s)$, then there are no redundancies, as far as energy-storage elements are concerned,[4] in the network and the network is said to be *completely characterized* by its transfer function. For example, the transfer functions of the networks in Figures 7.13(aa), 7.13(bb), and 7.13(dd) are $R/(RC_1s+1)$, $1/(RC_1s+1)$, and $1/(RC_1s+1)$, respectively. They have degree 1 and equal the number of capacitor or inductor in each network. The network in Figure 7.13(cc) has transfer function 1. It has degree 0, and the network has no capacitor or inductor. Thus all networks on the right-hand side of Figure 7.13 are completely characterized by their transfer functions. On the other hand, all networks on the left-hand side of Figure 7.13 are not completely characterized by their transfer functions.

Even though all networks on the left-hand side of Figure 7.13 cannot be fully described by their transfer functions, they can be fully described by their ss equations. However, those ss equations will show some deficiency. The ss equations describing the networks in Figures 7.13(a) and 7.13(d) are controllable but not observable. The ss equation describing the network in Figure 7.13(b) is observable but not controllable. The ss equation describing the network in Figure 7.13(c) is not controllable nor observable. In other words, all ss equations that describe fully the networks in Figures 7.13(a) through 7.13(d) do not have the so-called *controllability* and *observability* properties. This type of ss equation is not used in design.

The aforementioned properties can be checked *directly* from ss equations. However, it requires more mathematics than what we used so far. See Reference 3. Thus we discuss a different method. Consider an ss equation with dimension N. We compute its transfer function. If the transfer function has a degree less than N, then the ss equation does not have both the controllability and observability properties. It may have one but not both properties. However, if

$$\begin{aligned}\text{Dimension of ss equation} &= \text{Degree of its transfer function}\\ &= \text{No. of energy-storage elements}\end{aligned}$$

[4]We consider redundancies only in energy-storage elements in RLC circuits and number of integrators in op-amp circuits. For example, consider the network in Figure 2.11. The transfer function from u to y is 0.01, which has degree 0. The network can clearly be replaced by a simpler resistive network as far as the input and output is concerned. Thus the network in Figure 2.11 has many unnecessary resistors. We are not concerned with this type of redundancy.

then the ss equation is controllable and observable and the system described by the ss equation is also completely characterized by its transfer function. For example, the ss equations describing Figures 7.13(a), 7.13(b) and 7.13(d) have dimension 3, but their transfer functions all have degree 1. The ss equation describing Figure 7.13(c) has dimension 1, but its transfer function has degree 0. See Problem 7.14. Thus all ss equations describing the networks on the left-hand side of Figure 7.13 have some deficiency and, consequently, the networks have some redundancies and should not be used in practice.

EXAMPLE 7.6.1

Consider the state space equation in (7.19). We use the MATLAB function ss2tf to compute its transfer function. Let us type

```
a=[-2 1 2;1 0 0;0 1 0];b=[1;0;0];c=[0 2 -2];d=0;
[n,d]=ss2tf(a,b,c,d)
```

It yields n=[0 0 2 -2]; d=[1 2 -1 -2]. Thus the transfer function of (7.19) is

$$H(s) = \frac{2s-2}{s^3+2s^2-s-2} = \frac{2(s-1)}{(s+1)(s+2)(s-1)} \tag{7.52}$$

Recall that the degree of a proper rational transfer function is defined as the degree of its denominator after canceling out all common factors. Thus the transfer function in (7.52) has degree 2, which is less than the dimension of the ss equation. Thus the ss equation has some deficiency and the corresponding system should not be used in practice.

We mention that the MATLAB function minreal, an acronym for minimal realization, can also be used to check whether an ss equation is minimal—that is, both controllable and observable.

If an LTIL system has no redundancy in energy-storage elements, then it is completely characterized by its transfer function. In this case, there is no difference in using an ss equation or the transfer function to study the system. However, the former is more convenient to carry out computer computation and op-amp circuit implementation, and the latter is more convenient to carry out qualitative analysis and design.

In practice, we use transfer functions to carry out design. Once a transfer function that meets the specifications is found, we use its minimal realization to carry out implementation. Then the resulting system will not have any redundancy. Thus all practical systems, unless inadvertently implemented, are completely characterized by their transfer functions.

7.6.1 Can We Disregard Zero-Input Responses?

Transfer functions describe only zero-state responses of systems. Thus whenever we use transfer functions, all initial conditions are automatically assumed to be zero. Can we do so? The answer is affirmative if a system is completely characterized by its transfer function. We use examples to justify this assertion.

EXAMPLE 7.6.2

Consider the system with transfer function in (7.16) and ss equation in (7.17). Because the dimension of the ss equation equals the degree of the transfer function, the system is completely characterized by its transfer function $H(s) = 2/[(s+1)(s+2)]$. In order to compute its zero-input response (the response excited by initial conditions), we must use its ss equation.[5] Taking the Laplace transform of (7.45) and assuming $\mathbf{x}(0) \neq \mathbf{0}$ and $u(t) = 0$ yield

$$s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{X}(s)$$

which implies $\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0)$. Substituting this into the output equation, we obtain

$$Y_{zi}(s) = \mathbf{c}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0) \tag{7.53}$$

This gives the zero-input or natural response of the system. For the ss equation in (7.17), (7.53) can be computed as

$$Y_{zi}(s) = \frac{2x_1(0) + 2(s+3)x_2(0)}{s^2 + 3s + 2} = \frac{k_1}{s+1} + \frac{k_2}{s+2}$$

Thus its zero-input response is of the form

$$y_{zi}(t) = k_1 e^{-t} + k_2 e^{-2t} \tag{7.54}$$

for some constants k_1 and k_2. Different initial conditions yield different k_1 and k_2. However, its zero-input response is always of the form shown in (7.54).

The zero-state response of the system excited by any input, as discussed in Sections 6.5 and 6.6, is of the form

$$y(t) = \bar{k}_1 e^{-t} + \bar{k}_2 e^{-2t} + \text{(terms due to poles of input)} \tag{7.55}$$

for some constant $\bar{k}_1$ and $\bar{k}_2$. We see that the two time functions e^{-t} and e^{-2t} in (7.54) appear in the zero-state response in (7.55). Thus from the transfer function we can determine the form of natural responses of the system.

EXAMPLE 7.6.3

Consider the system with transfer function in (7.18) and ss equation in (7.19). As discussed in Example 7.6.1, the system is not completely characterized by its transfer function $H(s) = 2/[(s+1)(s+2)]$. Using (7.53), we can show that the zero-input response is of the form

$$y_{zi}(t) = k_1 e^{-t} + k_2 e^{-2t} + k_3 e^{t} \tag{7.56}$$

[5] This is also the case in using MATLAB. This is the reason we use the ss equation description in Figure 2.7.

for some constants k_i. Its zero-state response excited by any input, however, is of the form

$$y(t) = \bar{k}_1 e^{-t} + \bar{k}_2 e^{-2t} + \text{(terms due to poles of input)} \tag{7.57}$$

for some constants $\bar{k}_1$ and $\bar{k}_2$. We see that the zero-state response does not contain the time function e^t which appears in the zero-input response.

For this system, if some initial conditions are not zero, the response may grow unbounded and the system cannot be used in practice. This fact, however, cannot be detected from its transfer function. Thus for this system, we cannot disregard zero-input responses and cannot use its transfer function to study the system.

The preceding examples show that if a system is not completely characterized by its transfer function, we cannot disregard zero-input responses. On the other hand, if a system is completely characterized by its transfer function, then zero-input responses are essentially contained in zero-state responses. Thus we may simply disregard zero-input responses.

We give a different justification. If a system is completely characterized by its transfer function, then for any state $\mathbf{x}_1$, we can find an input to drive $\mathbf{x}(0) = \mathbf{0}$ to $\mathbf{x}(t_1) = \mathbf{x}_1$, at any selected time t_1. See Reference 3. We then remove the input (that is, $u(t) = 0$ for $t > t_1$), then the response of the system after t_1 is the zero-input response excited by the initial state $\mathbf{x}_1$. Thus if a system is completely characterized by its transfer function, every zero-input response can be generated as a zero-state response. Thus there is no need to study zero-input or nature responses.

7.7 IDENTIFICATION BY MEASURING FREQUENCY RESPONSES

There are a number of ways of computing the transfer function of a system.[6] If we find a differential equation or ss equation to describe the system, then its transfer function can be obtained by applying the Laplace transform. For RLC networks, we can also use transform impedances to find transfer functions (Section 6.3.2). The preceding methods are applicable only to systems with known structures and are carried out using analytical methods. If the internal structure of a system is not known, the only way to find its transfer function is by measurements at the input and output terminals. This is an *identification* problem. This method can also be used to find transfer functions of complex systems, such as op amps that contain over 20 transistors. Computing their transfer functions using analytical methods will be complicated.

Identification is not a simple topic because measured data are often corrupted by noise. See Reference 25. There are two types of identification: parametric and nonparametric. In parametric identification, we assume the system to have a transfer function of known denominator's and numerator's degrees but with unknown coefficients. We then try to determine the coefficients from measured data. In nonparametric identification, we try to identify its transfer function without prior knowledge of its form. We discuss in the following one nonparametric scheme which is based on Theorem 6.4.

---

[6]This section may be skipped without loss of continuity. However, we must accept the fact that the transfer function of an op amp in the next subsection can be obtained by measurements.

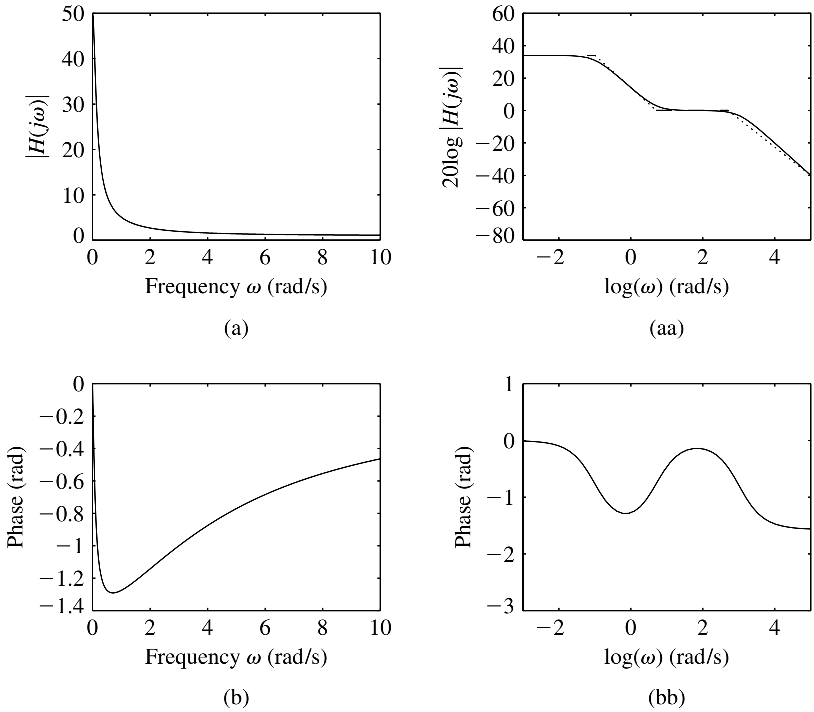


Figure 7.14 (**a**) Measured magnitude response. (**b**) Measured phase response. (**aa**) Bode gain plot. (**bb**) Bode phase plot. They are obtained using the MATLAB functions plot, loglog, and semilogx.

Consider a system with transfer function $H(s)$. If it is stable, its output excited by the input $u(t) = \sin\omega_0 t$ approaches $|H(j\omega_0)|\sin(\omega_0 t + \measuredangle H(j\omega_0))$ as $t \to \infty$. If the time constant of $H(s)$ is t_c, it takes roughly five time constants to reach steady state. Once $|H(j\omega_0)|$ and $\measuredangle H(j\omega_0)$ are measured, we then use a different ω_0 and repeat the measurement. After measuring $|H(j\omega)|$ and $\measuredangle H(j\omega)$ at a number of different ω, we can plot its magnitude and phase responses as shown in Figures 7.14(a) and 7.14(b). Clearly, the more frequencies at which we measure, the more accurate the plots. We then try to find a transfer function whose magnitude and phase responses approximate the measured ones.

It is difficult, if not impossible, to find a transfer function from the two plots in Figures 7.14(a) and 7.14(b). Now we use different coordinates to plot them. The coordinates in Figure 7.14(a) are $|H(j\omega)|$ against ω. We replot the magnitude response in Figure 7.14(aa) using $20\log|H(j\omega)|$ (in dB) and $\log\omega$. Note that the logarithms are with base 10. The coordinates in Figure 7.14(b) are $\measuredangle H(j\omega)$ against ω. We replot the phase response in Figure 7.14(bb) using $\measuredangle H(j\omega)$ and $\log\omega$. Note that $\omega = 0, 1, 10$, and ∞ become $-\infty, 0, 1$, and ∞, respectively, in $\log\omega$. The horizontal scale in Figure 7.14(aa) is for ω from $10^{-3} = 0.001$ to 10^5 or for $\log\omega$ from -3 to 5. The vertical scale from 0 to 50 in Figure 7.14(a) becomes the vertical scale from $-\infty$ to $20\log 50 = 34$ in Figure 7.14(aa). The set of two plots in Figures 7.14(aa) and 7.14(bb) is called the *Bode plot*. We call Figure 7.14(aa) the Bode gain plot and call Figure 7.14(bb) the Bode phase plot.

It turns out from the Bode gain plot, we can conclude that the transfer function must be of the form

$$H(s) = \frac{\pm 50\left(1 \pm \frac{s}{5}\right)}{\left(1 \pm \frac{s}{0.1}\right)\left(1 \pm \frac{s}{1000}\right)} \tag{7.58}$$

From the Bode phase plot, we can then conclude that the transfer function is

$$H(s) = \frac{50(1 + s/5)}{(1 + s/0.1)(1 + s/1000)} = \frac{1000(s + 5)}{(s + 0.1)(s + 1000)} \tag{7.59}$$

We call $H(0) = 50$ the dc gain. It is the gain at $s = 0$ and is different from the gain 1000 in the zero/pole/gain form. We explain in the following how the transfer function is obtained.

Using $20 \log(ab/cd) = 20 \log a + 20 \log b - 20 \log c - 20 \log d$, we can write $20 \log |H(j\omega)|$ in (7.58) as

$$\begin{aligned} 20 \log |H(j\omega)| &= 20 \log |\pm 50| + 20 \log \left|1 \pm \frac{j\omega}{5}\right| \\ &\quad - 20 \log \left|1 \pm \frac{j\omega}{0.1}\right| - 20 \log \left|1 \pm \frac{j\omega}{1000}\right| \end{aligned} \tag{7.60}$$

Thus the Bode gain plot is the sum of the four terms in (7.60). The first term is $20 \log |\pm 50| = 20 \log 50 = 34$ for all ω. Consider the third term. We call 0.1 a *corner frequency*. If ω is much smaller than 0.1, then $|1 \pm j\omega/0.1| \approx 1$. If ω is much larger than 0.1, then

$$|1 \pm j\omega/0.1| \approx |\pm j\omega/0.1| = \omega/0.1$$

Note that we consider only $\omega > 0$ in Bode plots. Thus the third term of (7.60) can be approximated as

$$-20 \log \left|1 \pm \frac{j\omega}{0.1}\right| \approx \begin{cases} -20 \log 1 = 0 & \text{for } \omega \ll 0.1 \text{ or } \log \omega \ll -1 \\ -20 \log(\omega/0.1) & \text{for } \omega \gg 0.1 \text{ or } \log \omega \gg -1 \end{cases}$$

Note that $-20 \log 1 = 0$ is a horizontal line and $-20 \log(\omega/0.1)$ is a straight line with slope -20 dB/decade[7] as shown in Figure 7.15 with dotted lines. They intersect at the corner frequency $\omega = 0.1$ or $\log 0.1 = -1$. In other words, $-20 \log |1 \pm j\omega/0.1|$ can be approximated by the horizontal line for $\omega < 0.1$ and the straight line with slope -20 dB/decade for $\omega > 0.1$ intersected at the corner frequency $\log 0.1 = -1$. Likewise, the second term $20 \log |1 \pm (j\omega/5)|$ in (7.60) can be approximated by two straight lines, one with slope 0 for $\omega < 5$ and the other with slope $+20$ dB/decade for $\omega > 5$, intersected at the corner frequency $\omega = 5$ or $\log 5 = 0.7$ as shown in Figure 7.15 with dashed lines. The fourth term can be approximated by two straight lines, one with slope 0 for $\omega < 1000$ and the other with slope -20 dB/decade for $\omega > 1000$, intersected at the corner frequency $\omega = 1000$ or $\log 1000 = 3$, as shown in Figure 7.15 with dash-and-dotted lines. Their sum is plotted in Figure 7.15 with a solid line. Shifting the solid line up by $20 \log 50 = 34$ yields the Bode gain plot in Figure 7.14(aa). From the preceding discussion, we know that every corner frequency yields a pole or zero. If the slope *decreases*

[7] It means that whenever the frequency increases 10 times, the dB gain decreases by 20.

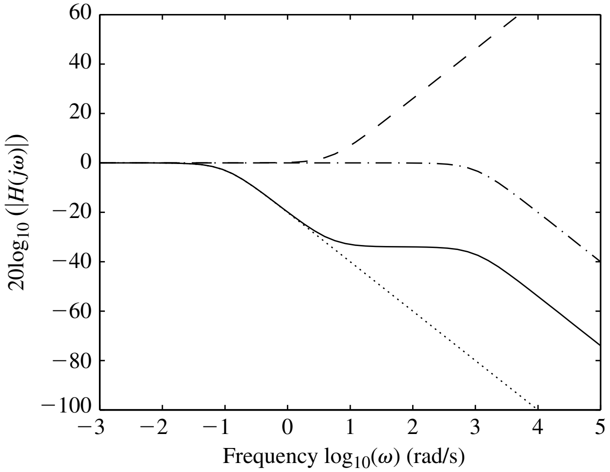


Figure 7.15 Bode gain plots of the second term (dashed line), third term (dotted line), and fourth term (dash-and-dotted line) of (7.60) and their sum (solid line).

by 20 dB/decade as ω increases, the corner frequency is a pole. If the slope *increases* by 20 db/decade, it is a zero. Note that $\pm k$ have the same gain plot. So have $(1 \pm j\omega/a)$. Thus we conclude that the transfer function which yields the Bode gain plot in Figure 7.14(aa) must be of the form shown in (7.58).

Every term in (7.58) has $\pm$ signs. Which sign to take is determined from the Bode phase plot shown in Figure 7.14(bb). We consider the phase as ω increases. For ω very small, the measured phase is 0. For ω very small, the phase of (7.58) is dictated by $k = \pm 50$. The phase is 180° or $-180°$ if $k = -50$, and it is 0 if $k = 50$. Thus we take the positive sign for the dc gain. In the neighborhood of each pole, if the phase decreases as ω increases, we take the positive sign. Otherwise, we take the negative sign. In the neighborhood of each zero, we take the positive sign if the phase increases but take the negative sign if the phase decreases. Thus from the phase plot in Figure 7.14(bb), we conclude that the transfer function equals (7.59).

In conclusion, if the magnitude response of a system plotted in the log–log coordinates can be approximated by sections of straight lines, then the intersections of those straight lines yield poles and zeros of its transfer function. If the frequency response is obtained by measurement, the system must be stable. Thus all poles must be in the LHP. Whether the dc gain is positive or negative and whether a zero is in the LHP or RHP must be determined from the phase plot.

Instruments called network analyzers or spectrum analyzers are available to measure automatically frequency responses and then generate transfer functions. We mention that for the transfer function in (7.59), if we type `n=[1000 5000]; d=[1 1000.01 100]; freqs(n,d)`, then MATLAB will yield the Bode gain and phase plots. The function `freqs` automatically selects 200 frequencies in $[0, \infty)$ and then carries out the computation. There is no need to specify the frequencies to be computed.

Limitations The preceding identification scheme is based on the assumption that the Bode gain plot can be approximated by sections of straight lines. This is so if poles and zeros are real and far apart as in (7.58). If poles and zeros are clustered together and/or complex, then it will be difficult to find straight-line approximations. In this case, it will not be simple to find a transfer function to approximate the measured Bode gain plot.

We also mention that Bode plots can be used to design control systems to have some phase margin and gain margin. The method, however, is limited to certain types of simple transfer functions. For complicated transfer functions, phase and gain margins may not have physical meaning. See Reference 5.

7.7.1 Models of Op Amps

Consider the op-amp model shown in Figure 2.16. If it is modeled as an ideal op amp (LTI, memoryless, and infinite open-loop gain), then we have

$$i_- = -i_+ = 0 \quad \text{and} \quad e_- = e_+$$

Using these two properties, we can carry out analysis and design of most op-amp circuits.

A more realistic model is to model it as an LTI memoryless system with a finite open-loop gain A. In this case, even though we still have $i_- = -i_+ = 0$ (because of a large input resistance), we no longer have $e_- = e_+$. Instead it must be replaced by

$$v_o(t) = A[e_+(t) - e_-(t)] \tag{7.61}$$

An even more realistic model is to model it as an LTI lumped system with memory. Such a system can be described by a convolution, a differential equation, or a transfer function. Among the three descriptions, the transfer function is the easiest to obtain. Thus we generalize (7.61) to

$$V_o(s) = A(s)[E_+(s) - E_-(s)] \tag{7.62}$$

where a capital letter is the Laplace transform of the corresponding lowercase letter, and $A(s)$ is a transfer function. In this model, we still have $I_-(s) = -I_+(s) = 0$. Figure 7.16 shows a typical frequency response of a Fairchild $\mu A741$ op amp copied from a spec sheet. The magnitude response is plotted using $\log |H(j\omega)|$ versus $\log(\omega/2\pi)$, and the phase response is plotted using $\measuredangle H(j\omega)$ versus $\log(\omega/2\pi)$. The numbers on the scales are the actual numbers, not their logarithms. Although the magnitude is plotted using $\log |H(j\omega)|$ instead of $20 \log |H(j\omega)|$, all discussion in the preceding section is still applicable. The only difference is that the slope will decrease by 1 per decade, instead of 20 dB per decade at each pole and increase by 1

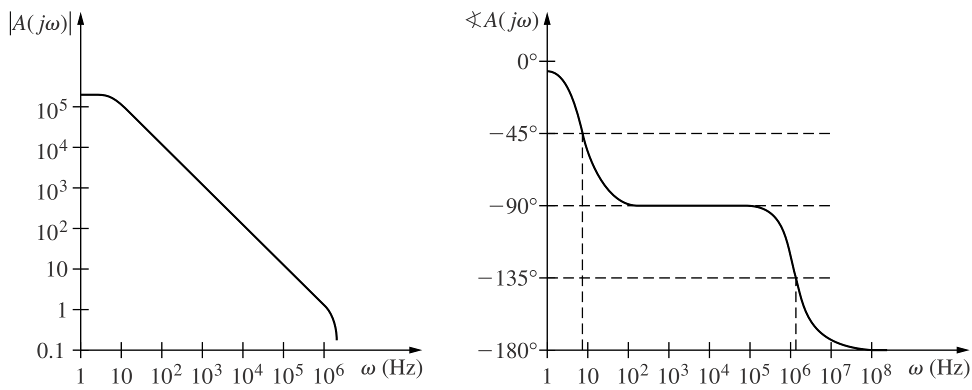


Figure 7.16 Typical magnitude and phase responses of Fairchild $\mu A741$ op amp.

per decade at each zero. The plot in Figure 7.16(a) can be approximated by three sections of straight lines. They intersect roughly at 8 and 3×10^6 in Hz or 16π and $6\pi \times 10^6$ in rad/s. At the corner frequency 16π, the slope decreases by 1/decade, thus it is a pole. At the corner frequency $6\pi \times 10^6$, the slope decreases by 1/decade (from -1 to -2), thus it is also a pole. There is no zero. Because the Bode plot is obtained by measurement, the two poles must be stable poles. It is also consistent with the phase plot in Figure 7.16. The gain for ω small is roughly 2×10^5. Thus we have $k = \pm 2 \times 10^5$. Because the phase approaches 0 for ω very small, we take the positive sign. Thus the transfer function of the op amp is

$$A(s) = \frac{2 \times 10^5}{\left(1 + \dfrac{s}{16\pi}\right)\left(1 + \dfrac{s}{6\pi 10^6}\right)} \tag{7.63}$$

It is a transfer function of degree 2.

The two poles of (7.63) are widely apart, one at -16π and the other at $-6\pi \times 10^6$. The response due to the pole $-6\pi \times 10^6$ will vanish much faster than the response due to the pole -16π, thus the response of (7.63) is dominated by the pole -16π and the transfer function in (7.63) can be reduced to

$$A(s) = \frac{2 \times 10^5}{\left(1 + \dfrac{s}{16\pi}\right)} = \frac{32\pi 10^5}{s + 16\pi} \approx \frac{10^7}{s + 50.3} \tag{7.64}$$

This is called a *single-pole* or *dominant-pole* model of the op amp in References 8, 11, and 23. This simplification or model reduction is widely used in practice. This will be discussed further in the next chapter.

7.7.2 Measuring Frequency Responses Using Sweep Sinusoids

The measurement of frequency responses discussed in Section 7.7 is based on Theorem 6.4.[8] We apply a sinusoid with frequency ω_0 and measure its steady-state response, which yields the frequency response at ω_0. If we measure the frequency responses at 50 different frequencies, then we must repeat the measurements 50 times. Is it possible to reduce the number of measurements? This is a question of practical importance.

We discuss in the following the response of a system excited by a sinusoid whose frequency changes linearly with time. Consider $u(t) = \sin \omega(t)t$, with $\omega(t) = at$, where a is a constant. This is a sinusoidal function with time-varying frequency and will be called a *linear sweep sinusoid*. Its Laplace transform cannot be expressed in close form and is complicated. Thus it is not possible to compute analytically the output of a system excited by such an input. We use the MATLAB function lsim to compute its response.

EXAMPLE 7.7.1

Consider two systems with transfer functions $H_1(s) = 5/(s+5)$ and $H_2(s) = 9/(s^2 + 0.6s + 9)$, respectively. We apply the input $u(t) = \sin \omega(t)t$ with $\omega(t) = 0.1t$ and use the MATLAB function lsim to compute their responses. The results are shown

[8] This subsection may be skipped without loss of continuity.

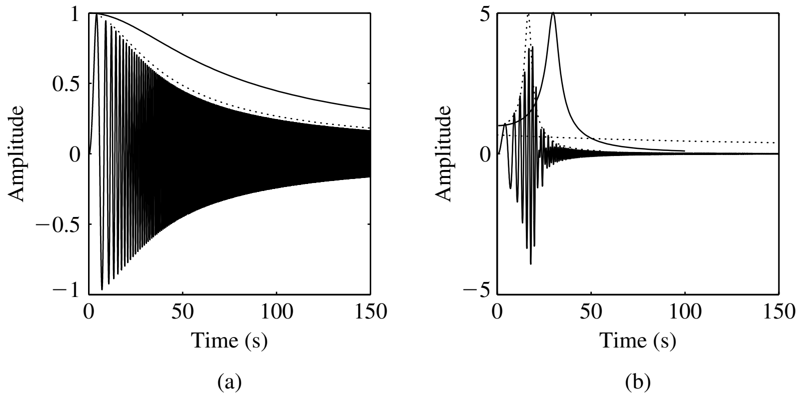


Figure 7.17 (**a**) Output of $H_1(s)$. (**b**) Output of $H_2(s)$.

in Figures 7.17(a) and 7.17(b). The horizontal ordinate is time in seconds. Because $\omega(t) = 0.1t$, we use solid lines to superpose on them the magnitude responses of $H_i(s)$ using $\omega/0.1$ as the horizontal scale. They are not the envelopes of the time responses of the systems. However, if we select $\omega/0.18$ as the horizontal scale, then the corresponding magnitude responses plotted with dotted lines are almost identical to the envelopes of the time responses. Thus one may be tempted to use a sweep sinusoid to measure the magnitude response of a system. By so doing, we need only one measurement.

For the two systems in this example, because we know the transfer functions, we can adjust b in ω/b so that their magnitude responses match the envelopes of their time responses. If we know the transfer function, then this measurement is not needed. If we do not know the transfer function, it is not clear how to obtain the magnitude response. In addition, it is difficult to obtain the phase response from the time response. Thus the method of using sweep sinusoids requires more study before it can be used to measure frequency responses of systems.

PROBLEMS

7.1 Find two realizations for each of the following transfer functions.

(a) $H_1(s) = \dfrac{6s + 3}{3s + 9}$

(b) $H_2(s) = \dfrac{s^2 + 2}{s^2 + 3s + 2}$

(c) $H_3(s) = \dfrac{1}{2s^4 + 3s^3 + 2s^2 + 5s + 10}$

7.2 Find three- and four-dimensional realizations for the transfer function

$$H(s) = \frac{s^2 - 4}{(s^2 + 2s + 1)(s + 2)}$$

What is the degree of the transfer function? Find also its minimal realization.

7.3 Find basic block diagrams for the systems in Problem 7.1.

7.4 Consider the basic block diagram in Figure 7.18. Assign the output of each integrator as a state variable and then develop an ss equation to describe it.

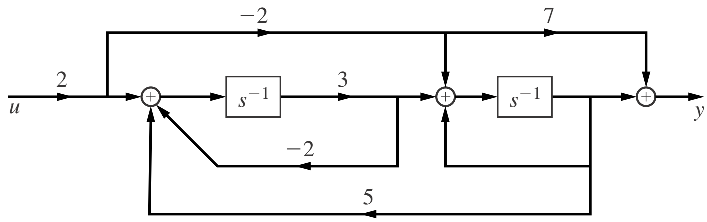


Figure 7.18

7.5 Consider $x(t) = 2\sin t + 0.1\sin 100t$, where $2\sin t$ represents a desired signal and $0.1\sin 100t$ represents noise. Show that differentiation of $x(t)$ will amplify the noise and integration of $x(t)$ will suppress the noise.

7.6 Find op-amp circuit implementations for the systems in Problem 7.1.

7.7 Implement (7.20) or (7.23) by assigning the outputs of the integrators in Figure 7.5 as $-x_1(t)$ and $x_2(t)$. What are the numbers of various components used in the implementation?

7.8 Use MATLAB to compute the step responses of the ss equations in (7.17) and (7.19). Are the results the same?

7.9 Use MATLAB to compute the response of (7.20) excited by the initial state $\mathbf{x}(0) = [1 \ \ -2]'$ and the input $u(t) = \sin\omega(t)t$ with $\omega(t) = 0.1t$ for t in $[0, 40]$ with increment 0.01.

7.10 Use MATLAB function `impulse` to compute the impulse response of (7.36). Also use `lsim` to compute its zero-input response excited by the initial state $\mathbf{x}(0) = [1 \ \ 0]'$. Are the results the same?

7.11 Consider the network shown in Figure 7.19(a) where the input $u(t)$ is a current source and where $Z_1(s)$ and $Z_2(s)$ are two arbitrary impedances of degree one or higher and have no common pole. If we consider the current $i(t)$ passing through $Z_2(s)$ as the output, what is its transfer function? If we consider the voltage $y(t)$ across $Z_2(s)$ as the output, what is its the transfer function? Is the network completely characterized by the two transfer functions?

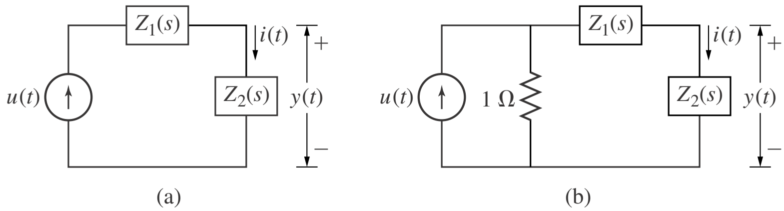


Figure 7.19

7.12 Repeat Problem 7.11 for the network shown in Figure 7.19(b).

7.13 Find the transfer function from u to y of the network shown in Figure 7.20.

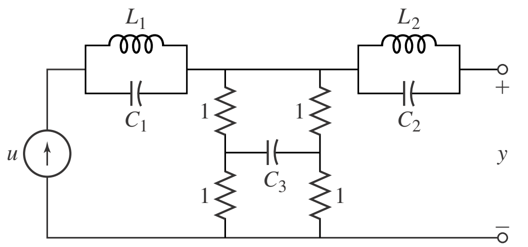


Figure 7.20

7.14 Verify that the network in Figure 7.13(c) is described by the one-dimensional ss equation

$$\dot{x}(t) = (-1/C)x(t)$$
$$y(t) = u(t)$$

where $x(t)$ is the voltage across the capacitor. Compute its transfer function. Is the network completely described by its transfer function?

7.15 Find ss equations to describe the networks shown in Figure 7.21.

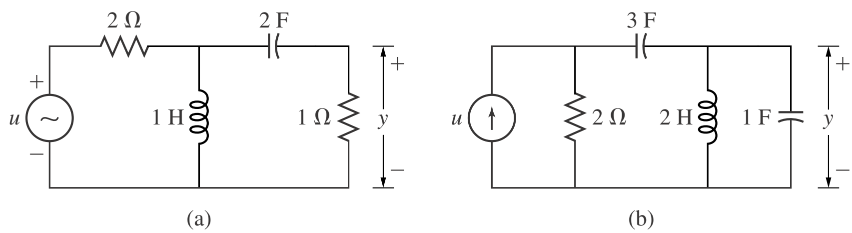


Figure 7.21

7.16 Consider the network shown in Figure 7.22(a). Because the two capacitor voltages are identical, assign only one capacitor voltage as a state variable and then develop a one-dimensional ss equation to describe the network. More generally, if a network has a loop that consists of only capacitors and if we assign all capacitor voltages as state variables, then one of them can be expressed as a linear combination of the rest. In other words, their voltages are not linearly independent. Thus there is no need to assign all capacitor voltages as state variables.

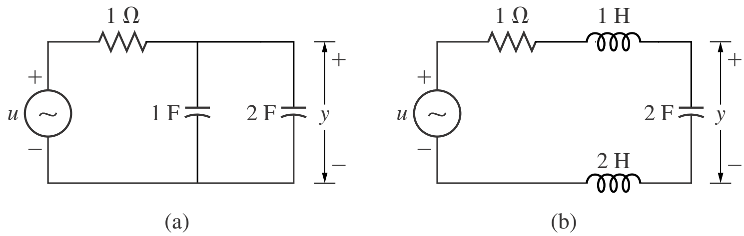


Figure 7.22

7.17 Consider the network shown in Figure 7.22(b). Because the two inductor currents are identical, assign only one inductor current as a state variable and then develop a two-dimensional ss equation to describe the network. More generally, if a network has a node or, more generally, a *cutset* that consists of only inductors, then there is no need to assign all inductor currents as state variables. See Reference 3.

7.18 Consider the network shown in Figure 7.23. Verify that if we assign the 1-F capacitor voltage as $x_1(t)$, the inductor current as $x_2(t)$, and the 2-F capacitor voltage as $x_3(t)$, then the circuit can be described by

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1/3 & 0 \\ 0 & -2 & 1 \\ 0 & -1/3 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 2/3 \\ 0 \\ 1/3 \end{bmatrix} \dot{u}(t)$$
$$y(t) = [0 \quad -2 \quad 1]\mathbf{x}(t)$$

Verify that if we assign only $x_1(t)$ and $x_2(t)$ (without assigning x_3), then the network can be described by

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1/3 \\ -1 & -2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) + \begin{bmatrix} 2/3 \\ 0 \end{bmatrix} \dot{u}(t)$$
$$y(t) = [-1 \quad -2]\mathbf{x}(t) + u(t)$$

Are the two equations in the standard ss equation form?

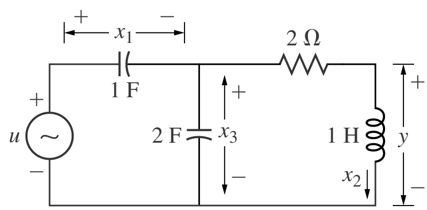


Figure 7.23

7.19 Use (7.48) to compute the transfer function of (7.39) which describes the network in Figure 7.10. The transfer function of the network can be computed using a differential equation (Example 3.8.5), an ss equation (Example 7.5.2), and transform impedances (Example 6.3.3). Which method is the simplest?

7.20 The roots $a + jb$ and $-a + jb$ are called *reciprocal roots*. If they are plotted on the complex s-plane, they are symmetric with respect to the $j\omega$-axis. Show

$$|j\omega - a - jb| = |j\omega + a - jb|$$

for all ω. Can you conclude that the transfer functions

$$H_1(s) = \frac{s+2}{(s-1+j2)(s-1-j2)}$$

and

$$H_2(s) = \frac{s-2}{(s+1+j2)(s+1-j2)}$$

have the same magnitude response? Do they have the same phase response? Which system is stable?

7.21 Consider the network shown in Figure 3.3(a). Suppose its Bode plot is obtained by measurement as shown in Figure 7.24. Can you find its transfer function?

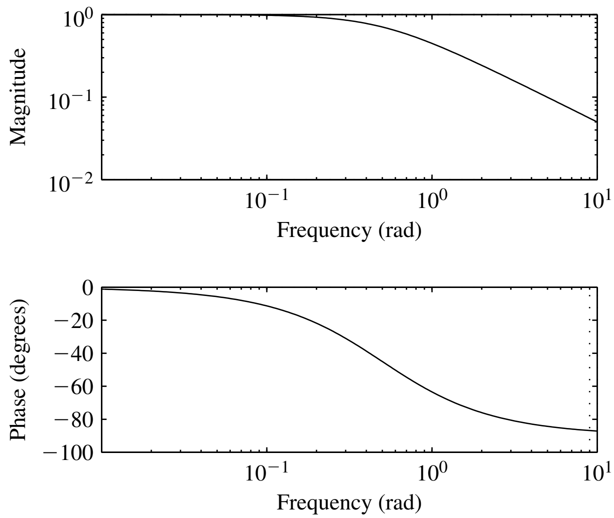


Figure 7.24

7.22 Consider the network shown in Figure 7.25(a), where T is a tunnel diode with the characteristic shown in Figure 7.25(b). Verify that the network can be described by

$$\dot{x}_1(t) = \frac{-h(x_1(t))}{C} + \frac{1}{C}x_2(t)$$
$$\dot{x}_2(t) = \frac{-x_1(t) - Rx_2(t)}{L} + \frac{1}{L}u(t)$$

This is a nonlinear ss equation. Now if v is known to be limited to the range $[a, b]$ shown, develop a linear state equation to describe the network. This problem also shows that ss equations can easily be extended to describe nonlinear systems.

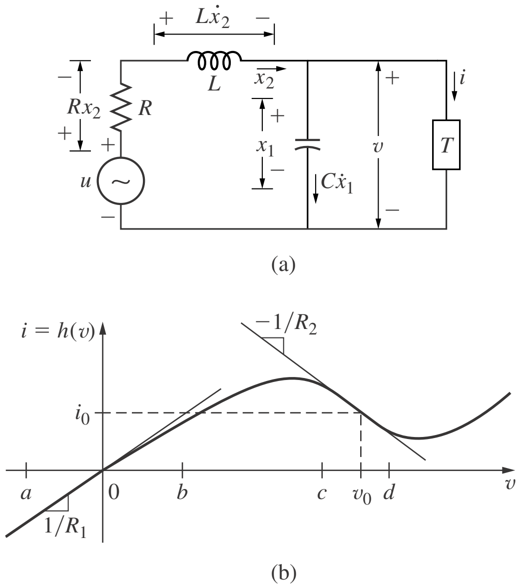


Figure 7.25

7.23 In Problem 7.22, if v is known to be limited to the range $[c, d]$ shown in Figure 7.25(b), verify that the network can be described by the linear state equation

$$\begin{bmatrix} \dot{\bar{x}}_1(t) \\ \dot{\bar{x}}_2(t) \end{bmatrix} = \begin{bmatrix} 1/CR_2 & 1/C \\ -1/L & -R/L \end{bmatrix} \begin{bmatrix} \bar{x}_1(t) \\ \bar{x}_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1/L \end{bmatrix} \bar{u}(t)$$

where $\bar{x}_1(t) = x_1(t) - v_0$, $\bar{x}_2(t) = x_2(t) - i_0$, and $\bar{u}(t) = u(t) - v_0 - Ri_0$. This is obtained by shifting the operating point from (0, 0) to (v_0, i_0), achieved by *biasing,* and then carrying out linearization. The equation is called a *small signal* linear model. This is a widely used technique in electronic circuits to develop linear models.

CHAPTER 8 Model Reduction, Feedback, and Modulation

8.1 INTRODUCTION

This chapter introduces three independent topics.[1] The first topic, consisting of Sections 8.2 and 8.3, discusses model reduction. This is widely used in engineering and yet rarely discussed in most texts. Because model reduction is based on systems' frequency responses and signals' frequency spectra, its introduction in this text is believed to be appropriate. We introduce the concept of operational frequency ranges of devices and apply it to op-amp circuits, seismometers, and accelerometers. The second topic, consisting of Sections 8.4 through 8.6, discusses feedback. We saw feedback in all op-amp circuits in the preceding chapters. We now discuss it in a more general setting. Feedback is used in refrigerators, ovens, and homes to maintain a set temperature and in auto-cruise of automobiles to maintain a set speed. A properly designed feedback system is less sensitive to external disturbances (uphill or downhill in the case of auto-cruise) and parameter variations (no passengers to four passengers). Thus feedback is widely used in practice. We first discuss three basic connections of two systems and the loading problem. We then use an example to demonstrate the advantage of using feedback. We also develop feedback models for op-amp circuits and show that the characteristic of an op-amp circuit is insensitive to its open-loop gain. To conclude this part, we introduce the Wien-bridge oscillator. We design it directly and by using a feedback model. The last topic, consisting of Sections 8.7 and 8.8, discusses two types of modulations and their demodulations. This is basic in communication.

8.2 OP-AMP CIRCUITS USING A SINGLE-POLE MODEL

We introduce model reduction by way of answering the question posed in Chapter 2. Consider the op-amp circuits shown in Figures 2.17(a) and 2.17(b) and repeated in Figure 8.1. We showed in Section 2.8.2 that the two circuits have, using the ideal model, the property $v_o(t) = v_i(t)$. However, the circuit in Figure 8.1(a) can be used as a voltage follower or buffer, but not the one in Figure 8.1(b). We now give the reason.

[1]The order of their studies can be changed (except that Section 8.4.4 requires Section 8.2.1).

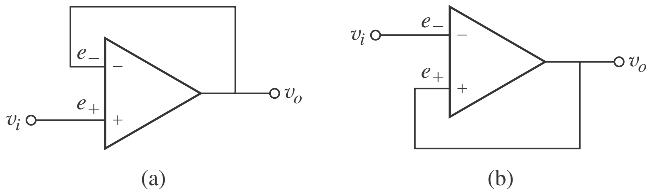


Figure 8.1 (**a**) Voltage follower. (**b**) Unstable system.

If an op amp is modeled as memoryless and operates in its linear region, then its inputs and output can be related by

$$v_o(t) = A[e_+(t) - e_-(t)] \tag{8.1}$$

where A is a constant. Now if the op amp is modeled to have memory and has the transfer function

$$A(s) = \frac{10^7}{s + 50.3} \tag{8.2}$$

as discussed in Section 7.7.1, then (8.1) must be modified as

$$V_o(s) = A(s)[E_+(s) - E_-(s)] \tag{8.3}$$

Now we use this more realistic model to study the stability of the op-amp circuits.

Consider the circuit in Figure 8.1(a). Substituting $V_i(s) = E_+(s)$ and $V_o(s) = E_-(s)$ into (8.3) yields

$$V_o(s) = A(s)(V_i(s) - V_o(s))$$

which implies

$$(1 + A(s))V_o(s) = A(s)V_i(s)$$

Thus the transfer function of the circuit in Figure 8.1(a) is

$$H(s) = \frac{V_o(s)}{V_i(s)} = \frac{A(s)}{A(s) + 1} \tag{8.4}$$

or, substituting (8.2),

$$H(s) = \frac{\dfrac{10^7}{s + 50.3}}{1 + \dfrac{10^7}{s + 50.3}} = \frac{10^7}{s + 50.3 + 10^7} \approx \frac{10^7}{s + 10^7} \tag{8.5}$$

It is stable. If we apply a step input, the output $v_o(t)$, as shown in Theorem 6.4, will approach $H(0) = 1$. If we apply $v_i(t) = \sin 10t$, then, because

$$H(j10) = \frac{10^7}{j10 + 10^7} \approx 1e^{j \cdot 0}$$

the output will approach $\sin 10t$. Furthermore, because the time constant is $1/10^7$, it takes roughly 5×10^{-7} second to reach steady state. In other words, the output will follow the input almost instantaneously. Thus the device is a very good voltage follower.

Now we consider the op-amp circuit in Figure 8.1(b). Substituting $V_i(s) = E_-(s)$ and $V_o(s) = E_+(s)$ into (8.3) yields

$$V_o(s) = A(s)(V_o(s) - V_i(s))$$

which implies

$$(1 - A(s))V_o(s) = -A(s)V_i(s)$$

Thus the transfer function of the circuit is

$$H(s) = \frac{V_o(s)}{V_i(s)} = \frac{-A(s)}{1 - A(s)} = \frac{A(s)}{A(s) - 1} \tag{8.6}$$

or, substituting (8.2),

$$H(s) = \frac{\dfrac{10^7}{s + 50.3}}{\dfrac{10^7}{s + 50.3} - 1} = \frac{10^7}{10^7 - s - 50.3} \approx \frac{10^7}{10^7 - s} \tag{8.7}$$

It has a pole in the RHP. Thus the circuit is not stable and its output will grow unbounded when an input is applied. Thus the circuit will either burn out or run into a saturation region and cannot be used in practice.

To conclude this section, we mention that the instability of the circuit in Figure 8.1(b) can also be established using a memoryless but nonlinear model of the op amp. See Reference 6 (pp. 190–193).

8.2.1 Model Reduction—Operational Frequency Range

In order to facilitate analysis and design, engineers often employ simplification and approximation in developing mathematical equations for systems. We use the voltage follower in Figure 8.1(a) as an example to discuss the issue. If we use the ideal model ($A = \infty$), then we obtain immediately $v_o(t) = v_i(t)$. Thus its transfer function is 1. If we model the op amp as a memoryless system with constant gain $A = 2 \times 10^5$, then the transfer function of the voltage follower is, following (8.4),

$$H(s) = \frac{A}{A + 1} = \frac{2 \times 10^5}{2 \times 10^5 + 1} = 0.99999 \tag{8.8}$$

which is practically 1. If we model the op amp as a system with memory and with the transfer function in (8.2), then the transfer function of the voltage follower is

$$H_1(s) = \frac{A(s)}{A(s) + 1} \approx \frac{10^7}{s + 10^7} \tag{8.9}$$

Its dc gain is $H_1(0) = 1$. Because its time constant is $1/10^7$, its step response approaches 1 almost instantaneously. Thus there seems no reason to use the transfer function in (8.9).

Is this true for any input? Let us apply $u(t) = \sin \omega_0 t$ with $\omega_0 = 10$ rad/s. Then the output approaches $\sin \omega_0 t$ immediately if we use $H(s) = 1$ or almost immediately if we use the transfer function in (8.9). However, if $\omega_0 = 10^{20}$, then the results will be different. If the transfer function

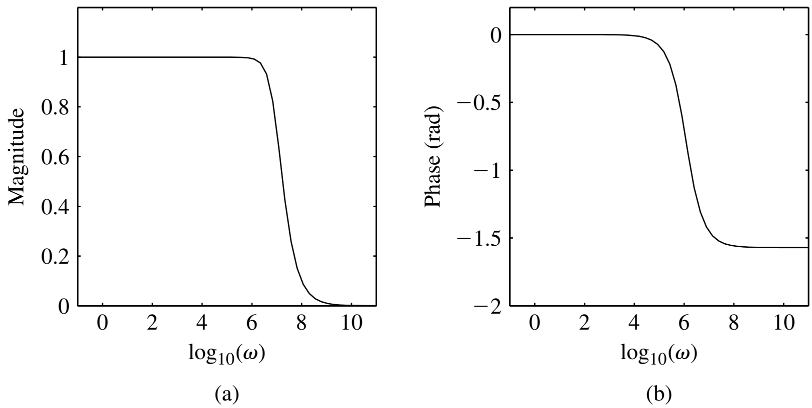


Figure 8.2 (**a**) Magnitude response of (8.9). (**b**) Phase response.

is 1, then the output is $\sin 10^{20}t$. To find the output of (8.9), we compute

$$H_1(j10^{20}) = \frac{10^7}{j10^{20} + 10^7} \approx \frac{10^7}{j10^{20}} \approx 0 \cdot e^{-j\pi/2}$$

Thus the output approaches 0. In other words, the voltage follower will not follow the input $\sin 10^{20}t$.

In practice, every system or device is designed to operate only for a class of inputs. In other words, no physical system can function identically for all signals. For this reason, there are low-frequency amplifiers, RF (radio frequency) amplifiers, and microwave amplifiers. In order to find the class of inputs that the voltage follower can follow, we must plot the frequency response of (8.9). Figure 8.2(a) plots the magnitude response of (8.9) against $\log_{10} \omega$ for $\omega = 0.1$ to 10^{11}.[2] Figure 8.2(b) shows the corresponding phase response. We see that the magnitude response is 1 and the phase response is practically 0 in the frequency range $[0, 10^6]$. Thus if an input frequency spectrum lies inside the range, then the transfer function in (8.9) can be reduced or simplified as 1. Equivalently, if $|j\omega| \leq 10^6$, then $j\omega + 10^7 \approx 10^7$ and (8.9) can be reduced as

$$H_{1r}(s) = \frac{10^7}{s + 10^7} \approx \frac{10^7}{10^7} = 1$$

We call this the *simplified* or *reduced model* of (8.9). This model is valid only if the input frequency spectrum lies inside the frequency range $[0, 10^6]$. For example, the frequency spectrum of $\sin 10^{20}t$ lies outside the range, and the simplified transfer function cannot be used. We call $[0, 10^6]$ the *operational frequency range* of the voltage follower.

As we will see in the next section, what is the exact operational frequency range of a system or device is not important. What is important is the fact that every device will yield the intended result only in a limited frequency range. Because the voltage follower in (8.9) is actually a lowpass filter with 3-dB passband $[0, 10^7]$ (see Section 6.8.2 and Problem 6.33), we may use the passband as its operational frequency range. In this frequency range the transfer function can be

[2]Because of the wide range of ω, it is difficult to plot the magnitude response against linear frequency.

reduced to 1. It is important to mention that stability is essential in every model reduction. For example, the magnitude and phase responses of (8.7) are identical to those in Figure 8.2 in the frequency range $[0, 10^6]$. But (8.7) cannot be reduced to $H_r(s) = 1$ for any signal. For example, the output of (8.7) excited by $u(t) = 1$ or $\sin 2t$ grows unbounded, which is entirely different from the corresponding output of $H_r(s) = 1$. We give one more example.

PID Controllers Consider the transfer function

$$C(s) = k_p + \frac{k_i}{s} + k_d s$$

where k_p is a gain or a *proportional* parameter, k_i is the parameter associated with an *integrator,* and k_d is the parameter associated with a *differentiator*. The transfer function is called a *PID controller* and is widely used in control systems. See Reference 5.

The differentiator $k_d s$ is actually a simplified model. In reality, it is designed as

$$H(s) = \frac{k_d s}{1 + s/N} \tag{8.10}$$

where N is a constant, called a *taming factor*. The transfer function is biproper and can be easily implemented. See Problem 8.4. We plot in Figure 8.3(a) with a solid line the magnitude response of $k_d s$ with $k_d = 2$ and with a dotted line the magnitude response of (8.10) with $N = 20$. We see that they are close for low frequencies. Figure 8.3(b) shows the corresponding outputs excited by $\sin 1.5t$; they are very close. Thus the biproper transfer function in (8.10) acts as a differentiator. Control signals are usually of low frequencies. For ω small so that $1 + j\omega/N \approx 1$, then (8.10) can be reduced as $H_r(s) = k_d s/1$. Thus the PID controller is actually a reduced transfer function and is applicable only for low-frequency signals.

The preceding discussion can be stated formally as follows. Consider a device or a system with transfer function $H(s)$. If $H(s)$ is stable and if its frequency response equals the frequency response of a simpler transfer function $H_r(s)$ in some frequency range B, then $H(s)$ can be reduced as $H_r(s)$ for input signals whose frequency spectra lie inside B. We call B the operational

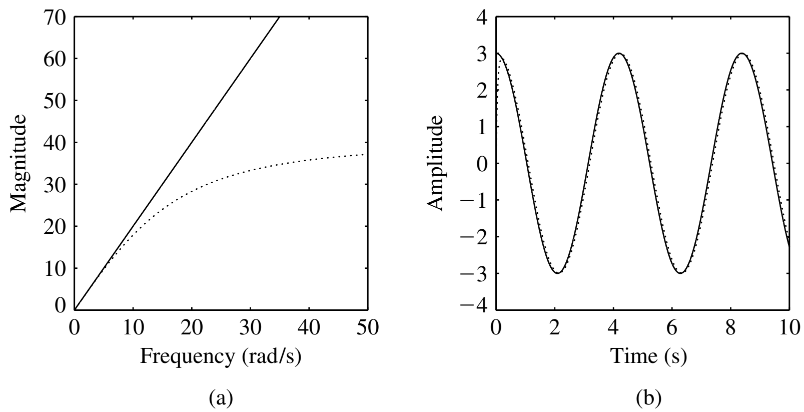


Figure 8.3 (**a**) Magnitude responses of $2s$ (solid line) and (8.10) with $N = 20$ (dotted line). (**b**) Outputs of $2s$ (solid line) and (8.10) (dotted line) excited by $\sin 1.5t$.

frequency range of the device. We use (6.78) to justify the assertion. Let $y(t)$ and $y_r(t)$ be the outputs of $H(s)$ and $H_r(s)$ excited by $u(t)$. Then we have

$$Y(j\omega) = H(j\omega)U(j\omega)$$

and

$$Y_r(j\omega) = H_r(j\omega)U(j\omega)$$

If $H(j\omega) = H_r(j\omega)$ for ω in B, then $Y(j\omega) = Y_r(j\omega)$ for ω in B. If the spectrum of $u(t)$ is limited to B or, equivalently, $U(j\omega) = 0$ for ω outside B, then $Y(j\omega) = Y_r(j\omega) = 0$ for ω outside B. Thus we have $Y(j\omega) = Y_r(j\omega)$ for all ω. Consequently we conclude $y(t) = y_r(t)$ for all t. This is the basis of model reduction. If $H(s)$ is not stable, the equation $Y(j\omega) = H(j\omega)U(j\omega)$ has no meaning and $H(s)$ cannot be reduced.

8.3 SEISMOMETERS AND ACCELEROMETERS

A seismometer[3] is a device to measure and record vibratory movements of the ground caused by earthquakes or man-made explosion. There are many types of seismometers. We consider the one based on the model shown in Figure 8.4(a). A block with mass m, called a *seismic mass,* is supported inside a case through a spring with spring constant k and a dashpot as shown. The dashpot generates a viscous friction with viscous friction coefficient f. The case is rigidly attached to the ground. Let u and z be, respectively, the displacements of the case and seismic mass relative to the inertia space. They are measured from the equilibrium position. By this, we mean that if $u = 0$ and $z = 0$, then the gravity of the seismic mass and the spring force cancel out and the mass remains stationary. We mention that the model in Figure 8.4(a) is essentially the same as the one in Figure 3.7(b). The only difference is that the input in Figure 3.7(b) is a shock or a sudden force and the ground is motionless. The input in Figure 8.4(a) is the movement of the ground.

Now if the ground vibrates and $u(t)$ becomes nonzero, the spring will insert a force on the seismic mass and cause it to move. If the mass is rigidly attached to the case (or $k = \infty$), then

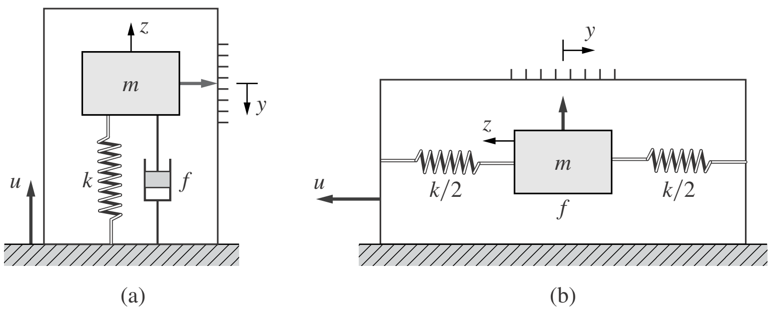


Figure 8.4 (**a**) Seismometer. (**b**) Accelerometer.

[3]This section may be skipped without loss of continuity.

$z(t) = u(t)$. Otherwise, we generally have $z(t) < u(t)$. Let us define $y(t) := u(t) - z(t)$. It is the displacement of the mass with respect to the case and can be read from the scale on the case as shown. It can also be transformed into a voltage signal using, for example, a potentiometer. Now the acceleration force of m must be balanced out by the spring force $ky(t)$ and the viscous friction force $f\dot{y}(t)$. Thus we have

$$ky(t) + f\dot{y}(t) = m\ddot{z}(t) = m(\ddot{u}(t) - \ddot{y}(t))$$

or

$$m\ddot{y}(t) + f\dot{y}(t) + ky(t) = m\ddot{u}(t) \tag{8.11}$$

Applying the Laplace transform and assuming zero initial conditions, we obtain the transfer function from u to y of the seismometer as

$$H(s) = \frac{Y(s)}{U(s)} = \frac{ms^2}{ms^2 + fs + k} \tag{8.12}$$

This system is always stable for any positive m, f, and k (see Problem 6.21).

If $k = \infty$ or the seismic mass is rigidly attached to the case, then $z(t) = u(t)$, $y(t) = 0$, and the system cannot be used to measure the movements of the ground. If the seismic mass remains stationary ($z(t) = 0$) when the case moves, then we have $y(t) = u(t)$ and the system will be a good seismometer. Unfortunately, it is impossible to detach the mass completely from the case, thus we should select a small spring constant. Arbitrarily, we select $m = 1$ and $k = 0.2$. We then plot in Figure 8.5 the magnitude responses of (8.12) for $f = 0.2, 0.4, 0.63$, and 0.8. It is

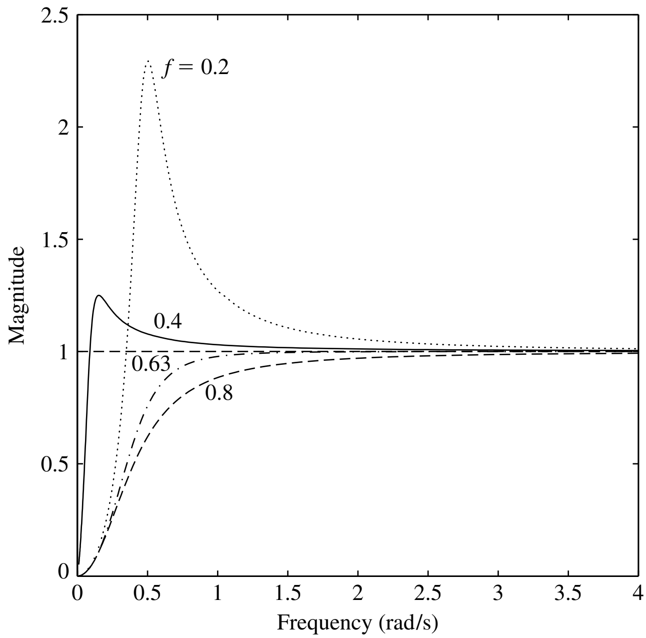


Figure 8.5 Magnitude responses of (8.12) for $m = 1$, $k = 0.2$, and $f = 0.2, 0.4, 0.63, 0.8$.

obtained in MATLAB by typing

```
n=[1 0 0];
d1=[1 0.2 0.2];[H1, w]=freqs(n,d1);
d2=[1 0.4 0.2];[H2, w]=freqs(n,d2);
d3=[1 0.63 0.2];[H3, w]=freqs(n,d3);
d4=[1 0.8 0.2];[H4, w]=freqs(n,d4);
plot(w,abs(H1),':',w,abs(H2),w,abs(H3),'--',w,abs(H4),'-.')
```

We see that for $f = 0.63$, the magnitude response approaches 1 at the smallest ω or has the largest frequency range $\omega > 1.25$ in which the magnitude response roughly equals 1.[4] Thus we conclude that (8.12) with $m = 1$, $f = 0.63$, and $k = 0.2$ is a good seismometer. However, it will yield accurate results *only* for signals whose frequency spectra lie inside the range $[1.25, \infty]$. In other words, in this operational frequency range or

$$|jf\omega + k| \ll m\omega^2$$

the transfer function in (8.12) can be simplified or reduced as

$$H_r(s) = \frac{Y(s)}{U(s)} = \frac{ms^2}{ms^2} = 1$$

which implies $y(t) = u(t)$.

Figure 8.6 shows the response of (8.12) excited by some $u(t)$. Figure 8.6(a) shows the signal $u_1(t) = e^{-0.3t}\sin t \cos 20t$, for $t \geq 0$. Its frequency spectrum, similar to the one in Figure 5.15(b), is centered around $\omega = 20$ rad/s. It lies inside the operational frequency range [1.25, 00] of the seismometer. The reading of the seismometer (output excited by $u_1(t)$) is shown in Figure 8.6(aa). It is obtained in MATLAB as follows:

```
n=[1 0 0];d=[1 0.63 0.2];t=0:0.1:10;
u=exp(-0.3*t).*sin(t).*cos(20*t);
y=lsim(n,d,u,t);plot(t,y)
```

The reading $y(t)$ follows closely the input.

Figure 8.6(b) shows the signal $u_2(t) = e^{-0.3t}\sin t$. Its spectrum, similar to the one in Figure 5.15(a), is centered around $\omega = 0$ with peaks at $\omega = \pm 1$. It is outside the operational frequency range [1.25, 00] of the seismometer. Thus the seismometer will not follow the signal as shown in Figure 8.6(bb). This shows the importance of the concepts of operational frequency ranges of systems and frequency spectra of signals.

Consider next the system shown in Figure 8.4(b). It is a model of a type of accelerometers. An accelerometer is a device that measures the acceleration of an object to which the device is attached. By integrating the acceleration twice, we obtain the velocity and distance (position) of the object. It is widely used in inertial navigation systems on airplanes and ships.

---

[4]Here we define the operational frequency range to be the range in which $|H(j\omega) - 1| < 0.02$. If we define it differently as $|H(j\omega) - 1| < 0.1$, then $f = 0.4$ yields a larger operational frequency range.

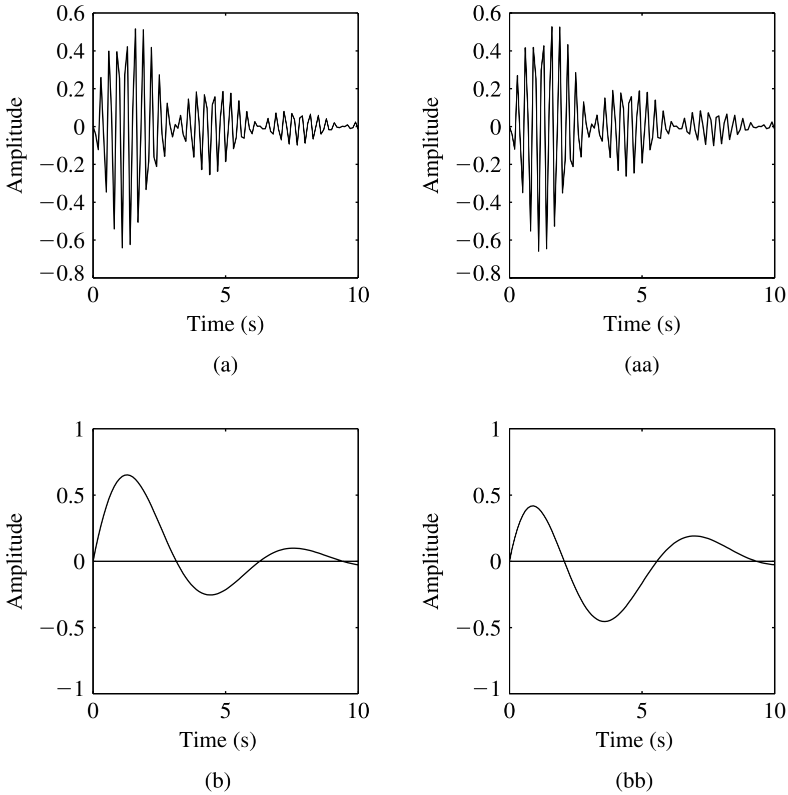


Figure 8.6 (**a**) $u_1(t) = e^{-0.3t} \sin t \cos 20t$. (**aa**) Output of the seismometer excited by u_1. (**b**) $u_2(t) = e^{-0.3t} \sin t$. (**bb**) Output of the same seismometer excited by u_2.

The model consists of a block with seismic mass m attached to a case through two springs as shown in Figure 8.4(b). The case is filled with oil to create viscous friction and is attached rigidly to an object such as an airplane. Let u and z be, respectively, the displacements of the case and mass with respect to the inertia space. Because the mass is floating inside the case, u may not equal z. We define $y := u - z$. It is the displacement of the mass with respect to the case and can be transformed into a voltage signal. The input of the accelerometer is u and the output is y. Let the spring constant of each spring be $k/2$ and let the viscous friction coefficient be f. Then we have, as in (8.11) and (8.12),

$$ky(t) + f\dot{y}(t) = m\ddot{z}(t) = m\ddot{u}(t) - m\ddot{y}(t)$$

and

$$H(s) = \frac{Y(s)}{U(s)} = \frac{ms^2}{ms^2 + fs + k} \tag{8.13}$$

It is stable for any positive m, f, and k.

If the movement of an object, such as a commercial airplane or big ship, is relatively smooth, then the frequency spectrum of the input signal is of low frequency. In this frequency range, if

$$|m(j\omega)^2 + f \times j\omega| = |-m\omega^2 + jf\omega| \ll k$$

then, as far as the frequency response is concerned, the transfer function can be approximated by

$$H_r(s) = \frac{Y(s)}{U(s)} = \frac{ms^2}{k} \tag{8.14}$$

or, in the time domain,

$$y(t) = \frac{m}{k}\frac{d^2u(t)}{dt^2} \tag{8.15}$$

Thus the output y is proportional to the acceleration. We plot in Figure 8.7 the magnitude responses of (8.13) with $m = 1, k = 10$, and $f = 3, 3.75, 4$, and 4.75. We also plot in Figure 8.7 the magnitude response of (8.14) with the solid line that grows outside the vertical range. We see that the magnitude response of (8.13) with $m = 0, k = 10$, and $f = 3.7$ or 4 is very close to the magnitude response of (8.14) in the frequency range roughly $[0, 2.3]$. Thus in this frequency range, the device in Figure 8.4(b) described by (8.13) with $m = 1,\ f = 4$, and $k = 10$ can function as an accelerometer. We use examples to verify this assertion.

If the input is

$$u(t) = 0.5t^2 + (\cos\omega_0 t - 1)/\omega_0{}^2 \qquad \text{for } 0 \le t \le 2\pi/\omega_0 \tag{8.16}$$

with $\omega_0 = 0.1$, then the output of (8.13) is as shown in Figure 8.8(a) with a solid line. We plot in Figure 8.8(a) with a dotted line the second derivative of $u(t)$ or

$$\frac{d^2u(t)}{dt} = 1 - \cos\omega_0 t \tag{8.17}$$

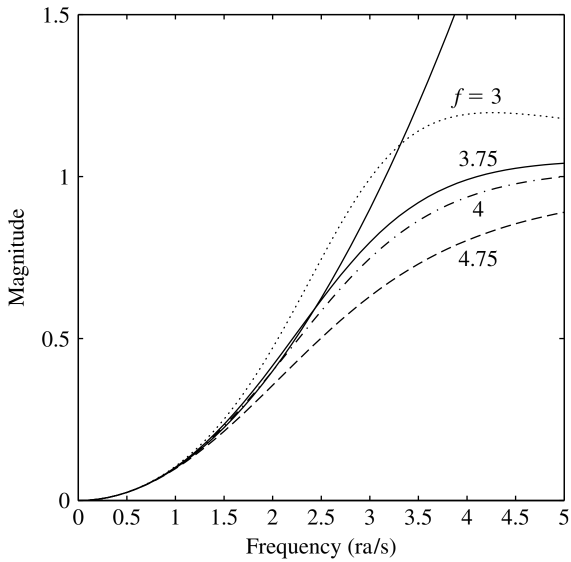


Figure 8.7 Magnitude responses of (8.13) for $m = 1, k = 10$, and $f = 3, 3, 75, 4, 4.75$.

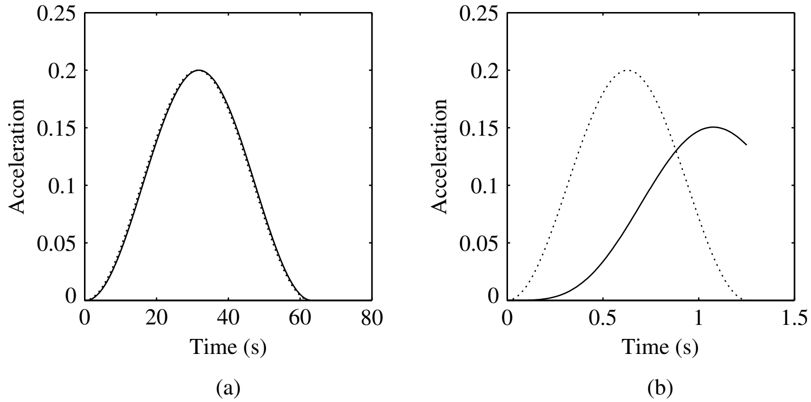


Figure 8.8 (**a**) The output (solid line) of the accelerometer excited by $u(t)$ in (8.16) with $\omega_0 = 0.1$ and the actual acceleration of $u(t)$ multiplied by 0.1 (dotted line). (**b**) The output (solid line) of the accelerometer excited by $u(t)$ in (8.16) with $\omega_0 = 5$ and the actual acceleration of $u(t)$ multiplied by 0.1 (dotted line).

multiplied by $m/k = 0.1$. It is indistinguishable from the output of the accelerometer. Note that the factor 0.1 can be easily taken care of by calibration or amplification. Thus the accelerometer gives a very accurate reading of the acceleration of $u(t)$ in (8.16) with $\omega_0 = 0.1$.

If the input is (8.16) with $\omega_0 = 5$, then the output is as shown in Figure 8.8(b) with solid line. The actual acceleration (multiplied by 0.1) is as shown with a dotted line. The output is quite different from the actual acceleration. Thus the accelerometer will not yield an accurate acceleration of the signal in (8.16) with $\omega_0 = 5$.

To find the reason that the accelerometer gives an accurate acceleration of $u(t)$ in (8.16) with $\omega_0 = 0.1$ but an incorrect acceleration with $\omega_0 = 5$, we must compute the frequency spectra of $u(t)$. They were computed in Section 5.6 using FFT and plotted in Figures 5.13 and 5.14. They are repeated in Figure 8.9. Figure 8.9(a) shows the magnitude spectrum of $u(t)$ with $\omega_0 = 0.1$. Even though not all of the spectrum lies inside the range [0, 2.3], its most significant

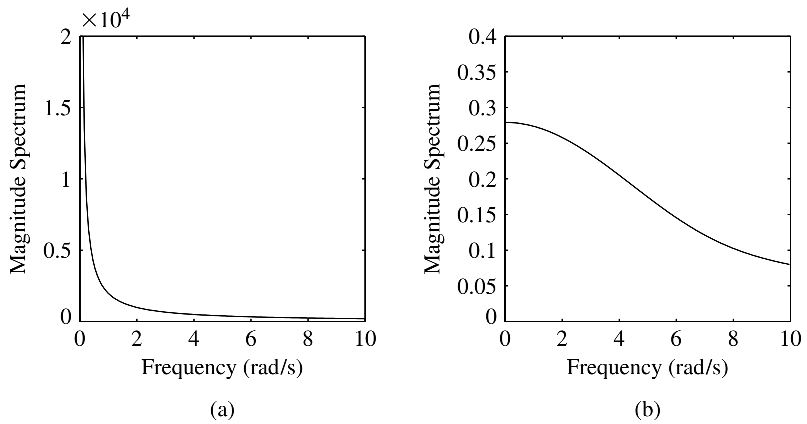


Figure 8.9 (**a**) Magnitude spectrum of (8.16) with $\omega_0 = 0.1$. (**b**) Magnitude spectrum of (8.16) with $\omega_0 = 5$.

part does. Thus the accelerometer yields an accurate acceleration of the signal. Figure 8.9(b) shows the magnitude spectrum of $u(t)$ with $\omega_0 = 5$. The spectrum is not concentrated inside the operational frequency range [0, 2.3] of the accelerometer. Thus the reading is quite different from the actual acceleration. This example shows once again the importance of systems' frequency responses and signals' frequency spectra. This example also shows that there is no need to define an operational frequency range precisely because most signals encountered in practice are not exactly band-limited. Furthermore, in the case of seismometers, relative magnitudes of different earthquakes are more revealing than precise movement of each earthquake.

The transfer function in (8.12) or (8.13) can be designed to function as a seismometer or as an accelerometer. One design criterion is to find m, f, and k to have the largest operational frequency range. For example, for the seismometer with $m = 1$ and $k = 0.2$, its operational frequency range is the largest for $f = 0.63$. For a different set of m and k, we will find a different f to yield the largest operational frequency range. See Problems 8.6 and 8.7. Thus the design is not unique and there is a large leeway in designing an actual system.

For the transfer function in (8.12) or (8.13), $\sqrt{k/m}$ is often called *the natural frequency*. It equals $\sqrt{k/m} = \sqrt{0.2/1} = 0.45$ for the seismometer and $\sqrt{10/1} = 3.16$ for the accelerometer. In the literature, it is suggested that the lowest frequency of the input signal should be larger than the natural frequency of the seismometer, and the highest frequency of the input signal should be smaller than the natural frequency of the accelerometer. In practice, no input is a pure sinusoid; therefore what is its highest or lowest frequency is not clear. In view of our discussion, we can replace the preceding statements as follows: If most of the input frequency spectrum lies inside a device's operational frequency range, then the device will yield a good result. Of course, what is "most of" is open for argument. Fortunately, this is engineering, not mathematics.

In application, an accelerometer will always give out a reading, even if it is not accurate. Consequently, its velocity and distance (or position) may not be accurate. Fortunately, the reading of the position of an airplane can be reset or corrected using GPS (global positioning system) signals or signals from ground control towers.

To conclude this section, we discuss a miniaturized accelerometer that is used in automobiles to trigger an air bag. First, we modify the arrangement in Figure 8.4(b) to the one shown in Figure 8.10(a). Because the spring is not in line with the direction of movement, the equation

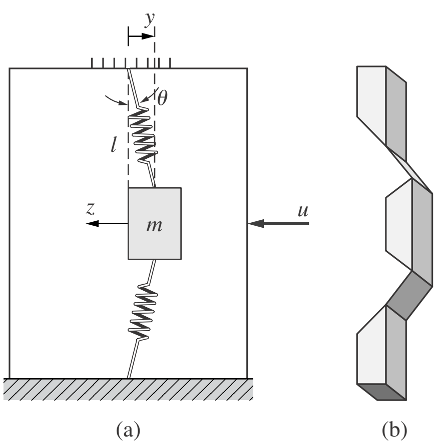


Figure 8.10 **(a)** Nonlinear accelerometer. **(b)** Silicon piezoresistive accelerometer.

describing the spring force is not simple. For example, if the spring has length l and if the seismic mass moves y, then the spring is stretched to have length $y/\sin\theta$ and the spring force is

$$k(y/\sin\theta - l)$$

where θ is the angle shown in Figure 8.10(a). We then multiply this spring force by $\sin\theta$ to have the force in line with the movement. Thus the equation describing the system is nonlinear and is complicated. Fortunately, the device is used only to trigger an air bag when the deceleration during a collision is larger than a set value, thus the design can be used. Figure 8.10(b) shows a design in which the spring is replaced by silicon piezoresistive material. The material will generate a voltage when it is stressed. It implements essentially the setup in Figure 8.10(a). The device can be miniaturized and mass produced. It is inexpensive and can function reliably. Thus it is used in every automobile. The purpose of this discussion is to point out that understanding the basic idea is only a first step in developing a device. It takes months or years of development and engineering ingenuity to develop a workable and commercially successful device.

8.4 COMPOSITE SYSTEMS

A system is often built by connecting two or more subsystems. This section discusses three basic connections. Consider the connections shown in Figure 8.11. The two subsystems are described by

$$Y_1(s) = H_1(s)U_1(s), \qquad Y_2(s) = H_2(s)U_2(s) \tag{8.18}$$

Let $u(t)$ and $y(t)$ be the input and output of the overall system. In the *tandem* or *cascade* connection shown in Figure 8.11(a), we have $u(t) = u_1(t)$, $y_1(t) = u_2(t)$, and $y(t) = y_2(t)$.

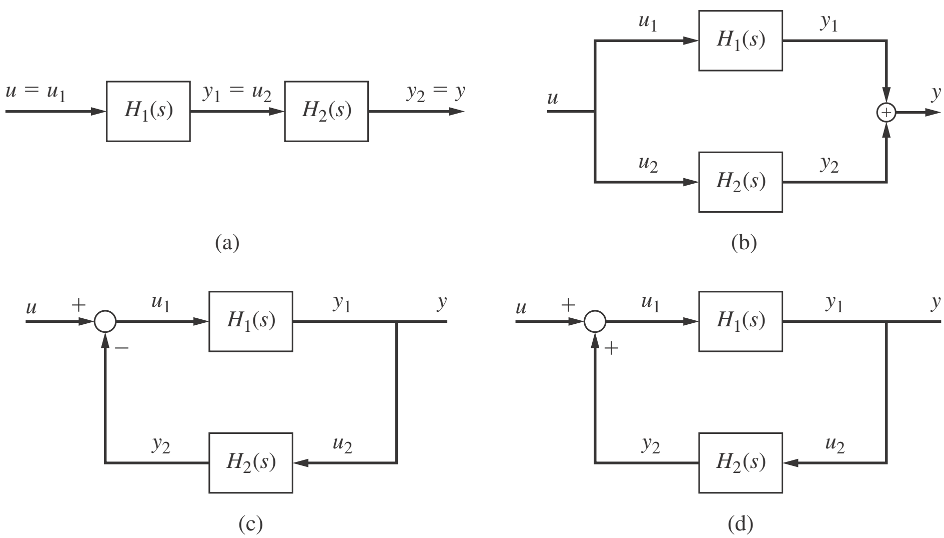


Figure 8.11 (**a**) Tandem connection. (**b**) Parallel connection. (**c**) Negative feedback connection. (**d**) Positive feedback connection.

Using these and (8.18), we have

$$Y(s) = Y_2(s) = H_2(s)U_2(s) = H_2(s)Y_1(s) = H_2(s)H_1(s)U(s)$$

Thus the transfer function of the tandem connection is $H_2(s)H_1(s)$, the product of the two individual transfer functions. In the parallel connection shown in Figure 8.11(b), we have $u(t) = u_1(t) = u_2(t)$ and $y(t) = y_1(t) + y_2(t)$ which imply

$$Y(s) = Y_1(s) + Y_2(s) = (H_1(s) + H_2(s))U(s)$$

Thus the transfer function of the parallel connection is $H_1(s)+H_2(s)$. For the negative feedback connection shown in Figure 8.11(c), we have $u_1(t) = u(t) - y_2(t)$, $y(t) = y_1(t) = u_2(t)$, or

$$U_1(s) = U(s) - H_2(s)U_2(s) = U(s) - H_2(s)H_1(s)U_1(s)$$

which implies

$$(1 + H_2(s)H_1(s))U_1(s) = U(s) \quad \text{or} \quad U_1(s) = \frac{U(s)}{1 + H_1(s)H_2(s)}$$

Thus we have

$$Y(s) = H_1(s)U_1(s) = \frac{H_1(s)}{1 + H_1(s)H_2(s)}U(s)$$

and the transfer function of the negative feedback system is

$$H(s) = \frac{H_1(s)}{1 + H_1(s)H_2(s)} \tag{8.19}$$

EXERCISE 8.4.1

Verify that the transfer function of the positive feedback system shown in Figure 8.11(d) is

$$H(s) = \frac{H_1(s)}{1 - H_1(s)H_2(s)} \tag{8.20}$$

The preceding derivations involve only algebraic manipulations; thus, transfer functions of various connections can be easily obtained. If we use convolutions or ss equations, then their derivations are much more complex, especially in feedback connections. Thus the use of transfer functions to study feedback systems is much simpler.

8.4.1 Loading Problem

The manipulation in the preceding section is correct mathematically, but may not be so in engineering. The next example illustrates this fact.

EXAMPLE 8.4.1

Consider the two RC networks shown in Figure 8.12(a). Using impedances, we can find their transfer functions as

$$H_1(s) = \frac{1/s}{1+1/s} = \frac{1}{s+1} \quad \text{and} \quad H_2(s) = \frac{1}{1+1/2s} = \frac{2s}{2s+1} \tag{8.21}$$

Let us connect them in tandem as shown with dashed lines. We compute its transfer function from u to y. The impedance of the parallel connection of $1/s$ and $1+1/2s$ is

$$Z_1(s) = \frac{(1/s)(1+1/2s)}{(1/s)+(1+1/2s)} = \frac{2s+1}{s(2s+3)}$$

Thus the voltage $V_1(s)$ shown is given by

$$V_1(s) = \frac{Z_1(s)}{1+Z_1(s)}U(s) = \frac{2s+1}{2s^2+5s+1}U(s)$$

and the output $Y(s)$ is given by

$$Y(s) = \frac{1}{1+1/2s}V_1(s) = \frac{2s}{2s+1}\frac{2s+1}{2s^2+5s+1}U(s)$$

Thus the transfer function from u to y of the tandem connection is

$$H(s) = \frac{2s}{2s^2+5s+1} \tag{8.22}$$

Does this transfer function equal the product of $H_1(s)$ and $H_2(s)$? We compute

$$H_1(s)H_2(s) = \frac{1}{s+1}\frac{2s}{2s+1} = \frac{2s}{2s^2+3s+1}$$

It is different from the transfer function in (8.22). Thus the transfer function of the tandem connection of the two networks does not equal the product of the two individual transfer functions.

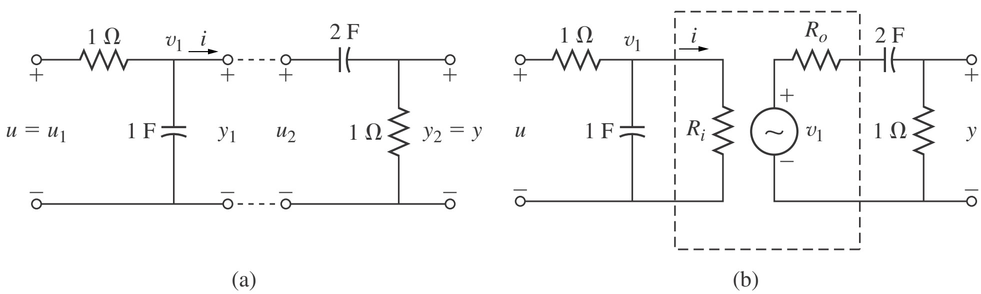


Figure 8.12 (**a**) Two networks. (**b**) Inserting isolating amplifier.

The preceding example shows that $H(s) = H_1(s)H_2(s)$ may not hold in engineering. If this happens, we say that the connection has a loading problem. The loading in Figure 8.12(a) is due to the fact that the current $i(t)$ shown is zero before connection and becomes nonzero after connection. This provides a method of eliminating the loading.

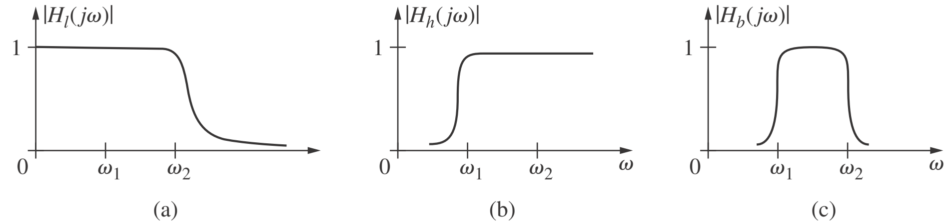


Figure 8.13 (**a**) Lowpass filter. (**b**) Highpass filter. (**c**) Bandpass filter obtained by tandem connection of (a) and (b).

Let us insert an amplifier with gain 1 or a voltage follower between the two circuit as shown in Figure 8.12(b). If the amplifier has a very large input resistance R_i, the current $i(t)$ will remain practically zero. If the output resistance R_o is zero or very small, then there will be no internal voltage drop in the amplifier. Thus the amplifier is called an *isolating amplifier* or *buffer* and can eliminate or reduce the loading problem. In conclusion, in electrical systems, loading can often be eliminated by inserting voltage followers as shown in Figure 2.18(b). For op-amp circuits, there is no loading problem because op amps have large input resistances and small output resistances. Thus if we design a lowpass op-amp filter with frequency response shown in Figure 8.13(a) and a highpass op-amp circuit with frequency response shown in Figure 8.13(b), then their tandem connection will yield a bandpass filter with frequency response shown in Figure 8.13(c). See References 6 (pp. 795–800) and 8.

Unlike electrical systems, there is no way to eliminate the loading problem in mechanical systems. For example, consider the armature-controlled dc motor driving a load shown in Figure 3.9. It is not possible to develop a transfer function for the motor and a transfer function for the load and then connect them together. We must develop a transfer function for the motor and load as a unit. If a transducer such as a potentiometer or a tachometer is connected to the motor shaft, its moment of inertia must be included in the load. Thus for mechanical systems, care must be exercised in developing their transfer functions. See Reference 5.

8.4.2 Why Feedback?

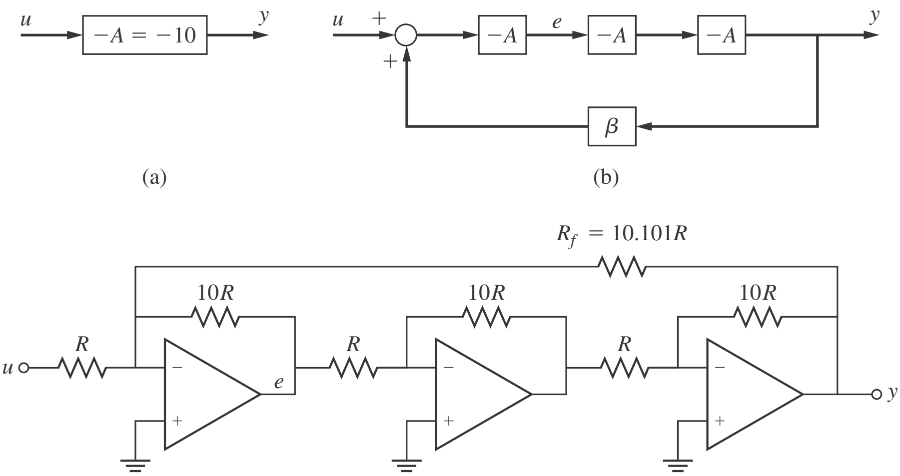
We use an example to show the main reason for using feedback. Suppose we are asked to design an inverting amplifier with gain $A = 10$. Such an amplifier can be easily built using the circuit in Figure 2.29(a) with $R_2 = 10R_1$ and is denoted by the box in Figure 8.14(a). Now consider the arrangement shown in Figure 8.14(b). It consists of three boxes, each with gain $-A = -10$ and a positive feedback with gain β. The arrangement can be implemented as shown in Figure 8.14(c).

The system in Figure 8.14(b) is memoryless. Let $-A_f$ be its gain from u to y. Then we have, using (8.20),

$$-A_f = \frac{(-A)^3}{1 - \beta(-A)^3} = -\frac{A^3}{1 + \beta A^3}$$

or

$$A_f = \frac{A^3}{1 + \beta A^3} \tag{8.23}$$



(a) (b) (c)

Figure 8.14 (**a**) Inverting amplifier. (**b**) Inverting feedback amplifier. (**c**) Op-amp implementation of (b).

Now we will find a β so that $A_f = 10$. We solve

$$10 = \frac{10^3}{1 + 10^3\beta}$$

which implies $10 + 10^4\beta = 10^3$ and

$$\beta = \frac{10^3 - 10}{10^4} = 0.099$$

In other words, if $\beta = 0.099$, then the feedback system in Figure 8.14(b) is also an inverting amplifier with gain 10. The feedback gain β can be implemented as shown in Figure 8.14(c) with $R_f = R/\beta = 10.101R$ (Problem 8.10).

Even though the inverting feedback amplifier in Figure 8.14(b) uses three times more components than the one in Figure 8.14(a), it is the preferred one. We give the reason. To *dramatize* the effect of feedback, we assume that A decreases 10% each year due to aging or whatever reason. In other words, A is 10 in the first year, 9 in the second year, and 8.1 in the third year as listed in Table 8.1. Next we compute A_f from (8.23) with $\beta = 0.099$ and $A = 9$:

$$A_f = \frac{9^3}{1 + 0.099 \times 9^3} = 9.963$$

We see that even though A decreases 10%, A_f decreases only $(10 - 9.963)/10 = 0.0037$ or less than 0.4%. If $A = 8.1$, then

$$A_f = \frac{8.1^3}{1 + 0.099 \times 8.1^3} = 9.913$$

and so forth. They are listed in Table 8.1. We see that the inverting feedback amplifier is much less sensitive to variations of A.

TABLE 8.1 Gains for Inverting Open and Feedback Amplifiers

n	1	2	3	4	5	6	7	8	9	10
A	10	9.0	8.1	7.29	6.56	5.9	5.3	4.78	4.3	3.87
A_f	10	9.96	9.91	9.84	9.75	9.63	9.46	9.25	8.96	8.6

If an amplifier is to be replaced when its gain decreases to 9 or less, then the open-loop amplifier in Figure 9.14(a) lasts only one year, whereas the feedback amplifier in Figure 9.14(b) lasts almost 9 years. Thus even though the feedback amplifier uses three times more components, it lasts nine times longer. Thus it is more cost effective. Not to mention the inconvenience and cost of replacing the open-loop amplifier every year. In conclusion, a properly designed feedback system is much less sensitive to parameter variations and external disturbances. However feedback may cause instability of systems. Thus stability is an important issue in feedback systems as we discuss next.

8.4.3 Stability of Feedback Systems

This subsection discusses stability of composite systems—in particular, feedback systems. All composite systems shown in Figure 8.11 are linear, time-invariant, and lumped. Thus if we compute their overall transfer functions, then their stability can be easily checked. However, there is one problem, as we discuss next.

EXAMPLE 8.4.2

Consider the system shown in Figure 8.15(a). It is the tandem connection of an unstable system with transfer function $H_1(s) = 1/(s-1)$ and a stable system with transfer function $H_2(s) = (s-1)/(s+1)$. The overall transfer function is

$$H(s) = H_2(s)H_1(s) = \frac{s-1}{s+1}\frac{1}{s-1} = \frac{1}{s+1}$$

We see that there is a *pole-zero cancellation*. Thus the degree of $H(s)$ is less than the sum of the degrees of $H_1(s)$ and $H_2(s)$. The overall transfer function is stable.

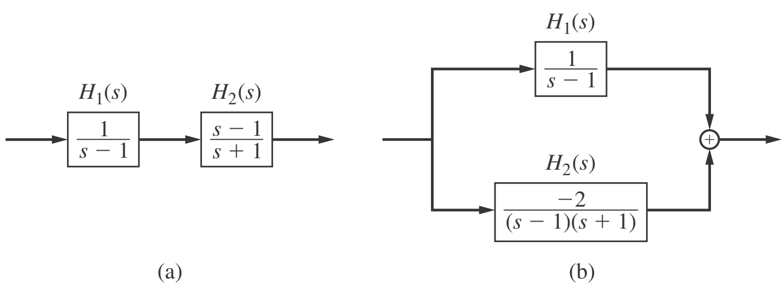


Figure 8.15 (**a**) Tandem connection of two systems. (**b**) Parallel connection of two systems.

Consider next the system shown in Figure 8.15(b). It is the parallel connection of two unstable systems with transfer functions $H_1(s) = 1/(s-1)$ and $H_2(s) = -2/(s-1)(s+1)$. Its overall transfer function is

$$H(s) = H_1(s) + H_2(s) = \frac{1}{s-1} + \frac{-2}{(s-1)(s+1)}$$

$$= \frac{s+1-2}{(s-1)(s+1)} = \frac{s-1}{(s-1)(s+1)} = \frac{1}{s+1}$$

which is stable. Note that the degree of $H(s)$ is less than the sum of the degrees of $H_1(s)$ and $H_2(s)$.

Although the two overall systems in the preceding example are stable, they cannot be used in practice because they involve unstable pole-zero cancellations. Such exact cancellations are difficult to achieve in practice because all physical components have some tolerances. For example, a 10 kΩ resistor with 10% tolerance may not have a resistance exactly equal to 10 kΩ; its actual resistance could be any value between 990 and 1100 Ω. Thus exact pole-zero cancellations rarely occur in practice. If no such cancellation occurs, then we have

$$\text{Degree of } H(s) = \text{Degree of } H_1(s) + \text{Degree of } H_2(s)$$

In this case, the overall system will be completely characterized by its overall transfer function. The two composite systems in Figure 8.15 are not completely characterized by their transfer functions. Thus the instability of their subsystems do not show up in their overall transfer functions.

We assume in the remainder of this subsection that every overall system is completely characterized by its overall transfer function. Under this assumption, *every tandem or parallel connection of two subsystems is stable if and only if the two subsystems are both stable*. For feedback connection, the situation is quite different. Consider the negative-feedback system shown in Figure 8.16(a). The two subsystems have transfer functions

$$H_1(s) = \frac{s-10}{s+1} \quad \text{and} \quad H_2(s) = 1$$

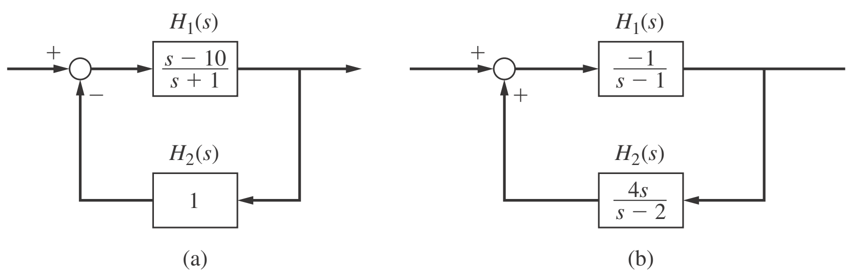


Figure 8.16 (**a**) Unstable feedback system with stable subsystems. (**b**) Stable feedback system with unstable subsystems.

They are both stable. Its overall transfer function is, using (8.19),

$$H(s) = \frac{H_1(s)}{1 + H_1(s)H_2(s)} = \frac{\dfrac{s-10}{s+1}}{1 + \dfrac{s-10}{s+1}}$$

$$= \frac{s-10}{s+1+s-10} = \frac{s-10}{2s-9}$$

Its degree equals the sum of the degrees of $H_1(s)$ and $H_2(s)$, thus the feedback system is completely characterized by $H(s)$. The transfer function $H(s)$ has an unstable pole at 4.5. Thus the negative-feedback system is unstable even though its subsystems are both stable.

Next we consider the positive-feedback system shown in Figure 8.16(b). The two subsystems have transfer functions

$$H_1(s) = \frac{-1}{s-1} \quad \text{and} \quad H_2(s) = \frac{4s}{s-2}$$

They are both unstable. Its overall transfer function is, using (8.20),

$$H(s) = \frac{H_1(s)}{1 - H_1(s)H_2(s)} = \frac{\dfrac{-1}{s-1}}{1 - \dfrac{-1}{s-1}\dfrac{4s}{s-2}}$$

$$= \frac{-(s-2)}{(s-1)(s-2)+4s} = \frac{-(s-2)}{s^2+s+2}$$

which is stable. Thus the positive-feedback system is stable even though its subsystems are both unstable. We see that stability of a positive- or negative-feedback system is independent of the stability of its subsystems. Thus it is important to check the stability of every feedback system.

The easiest method of checking the stability of any composite system, including any positive- or negative-feedback system, is to compute first its overall transfer function and then to check whether its denominator is a stable polynomial. We mention that the *Nyquist stability criterion* can be used to check the stability of the negative-feedback system shown in Figure 8.11(c) *without* computing its overall transfer function. We plot the so-called Nyquist plot and then check its encirclement around a critical point. The plot is not simple, especially if $H_1(s)$ or $H_2(s)$ has poles on the imaginary axis. For simple $H_1(s)$ and $H_2(s)$, the gain and phase margins of the plot will provide some information about the relative stability of the feedback system. However, the two margins may not be useful for complicated systems. See Reference 5. If we compute its overall transfer function, then we can determine not only its stability but also every thing about the system.

8.4.4 Inverse Systems

Consider a system with transfer function $H(s)$. If we know its input $u(t)$, the output $y(t)$ is simply the inverse Laplace transform of $H(s)U(s)$. Now if we know the output $y(t)$, can we compute its input $u(t)$?

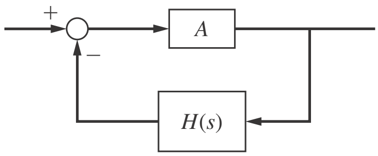


Figure 8.17 Feedback implementation of inverse system.

Consider a system with transfer function $H(s)$. We call the system with transfer function $H_{in}(s) := H^{-1}(s)$ the *inverse system* of $H(s)$. If the inverse $H_{in}(s)$ of $H(s)$ exists, we apply $y(t)$ to $H_{in}(s)$. Then its output $\bar{y}(t)$ equals $u(t)$ for

$$\bar{y}(t) = \mathcal{L}^{-1}[H_{in}(s)Y(s)] = \mathcal{L}^{-1}[H^{-1}(s)H(s)U(s)] = \mathcal{L}^{-1}[U(s)] = u(t)$$

This problem may find application in oil exploration: To find the source from measured data.

Every system, including every inverse system, involves two issues: realizability and stability. If the transfer function of a system has a proper rational function, then it can be implemented using an op-amp circuit. If the transfer function is not stable, then its realization is useless because it will saturate or burn out. Thus we require every inverse system to have a proper rational transfer function and to be stable.

Consider a system with transfer function $H(s)$. If $H(s)$ is biproper, then its inverse $H_{in}(s) = H^{-1}(s)$ is also biproper. If all zeros of $H(s)$ lie inside the open left half s-plane,[5] then $H^{-1}(s)$ is stable. Thus if a stable biproper transfer function has all its zeros inside the open LHP, then its inverse exists and can be easily built. If it has one or more zeros inside the closed RHP, then its inverse system is not stable and cannot be used.

If $H(s)$ is strictly proper, then $H^{-1}(s)$ is improper and cannot be realized. In this case, we may try to implement its inverse system as shown in Figure 8.17 where A is a very large positive gain. The overall transfer function of the feedback system in Figure 8.17 is, using (8.19),

$$H_o(s) = \frac{A}{1 + AH(s)}$$

Now if A is very large such that $|AH(j\omega)| \gg 1$, then the preceding equation can be approximated by

$$H_o(s) = \frac{A}{1 + AH(s)} \approx \frac{A}{AH(s)} = H^{-1}(s) \tag{8.24}$$

Thus the feedback system can be used to implement approximately an inverse system. Furthermore, for A very large, $H_o(s)$ is practically independent on A, thus the feedback system is insensitive to the variations of A. Thus it is often suggested in the literature to implement an inverse system as shown in Figure 8.17.

The problem is actually not so simple. In order for the approximation in (8.24) to be valid, the overall system $H_o(s)$ must be stable. This may not always be the case as we show next.

[5]Such a transfer function is called a *minimum-phase transfer function*. See Reference 5.

EXAMPLE 8.4.3

Consider $H(s) = 1/(s^2 + s + 3)$. Because $H^{-1}(s) = s^2 + s + 3$ is improper, it cannot be implemented without using differentiators. If we implement its inverse system as shown in Figure 8.17, then the overall transfer function is

$$H_o(s) = \frac{A}{1 + AH(s)} = \frac{A(s^2 + s + 3)}{s^2 + s + 3 + A}$$

It is stable for any $A \geq -3$. Thus if A is very large, $H_o(s)$ can be reduced as

$$H_{or}(s) \approx \frac{A(s^2 + s + 3)}{A} = s^2 + s + 3 = H^{-1}(s)$$

in the frequency range $|-\omega^2 + j\omega + 1| \ll A$. Thus the feedback system can be used to implement approximately the inverse system of $H(s)$. The approximation, however, is valid only for low-frequency signals as demonstrated in the following.

Consider the signal $u_1(t) = e^{-0.3t} \sin t$ shown in Figure 8.18(a). Its spectrum, as shown in Figure 5.15(a), has peak at $\omega = 1$ and is practically zero for $\omega > 10$. If we select $A = 1000$, then we have $|-\omega^2 + j\omega + 1| \ll A$. The output $y_1(t)$ of $H(s)$ excited by $u_1(t)$ is shown in Figure 8.18(aa). The output of $H_o(s)$ with $A = 1000$ excited by $y_1(t)$ is shown in Figure 8.18(aaa) with a solid line. It is very close to $u_1(t)$ shown with a dotted line. Thus the feedback system in Figure 8.17 implements the inverse of $H(s)$ for $u_1(t)$.

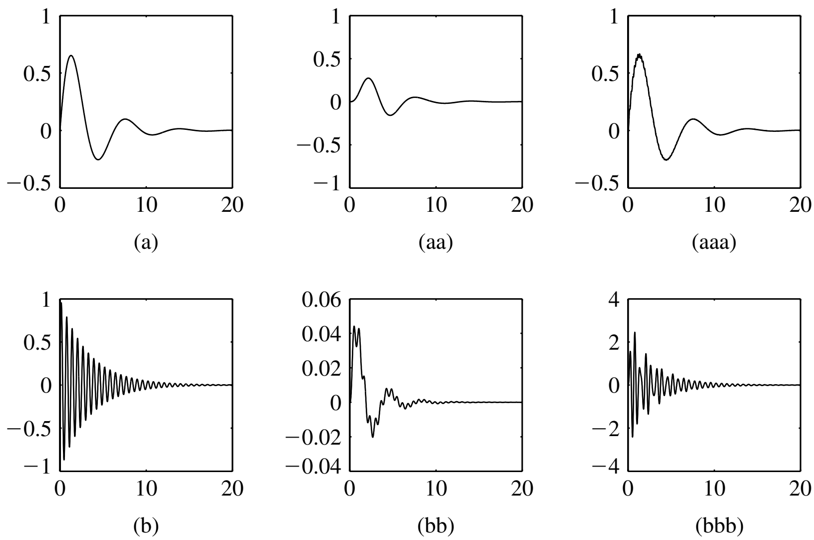


Figure 8.18 (**a**) Input $u_1(t)$ of $H(s)$. (**aa**) Output $y_1(t)$ of $H(s)$. (**aaa**) Output (solid line) of $H_o(s)$ excited by $y_1(t)$ and $u(t)$ (dotted line). (**b**) Input $u_2(t)$ of $H(s)$. (**bb**) Output $y_2(t)$ of $H(s)$. (**bbb**) Output of $H_o(s)$ excited by $y_2(t)$.

Next we consider $u_2(t) = e^{-0.3t} \sin 10t$ and select $A = 200$. The corresponding results are shown in Figures 8.18(b), 8.18(bb), and 8.18(bbb). The output of the feedback system is different from $u_2(t)$. This is so because we do not have $|-\omega^2 + 1 + j\omega| \ll A = 200$. Thus the feedback system is not an inverse system of $H(s)$ for high-frequency signals.

EXAMPLE 8.4.4

Consider $H(s) = 1/(s^3 + 2s^2 + s + 1)$. Because $H^{-1}(s) = s^3 + 2s^2 + s + 1$ is improper, it cannot be realized without using differentiators. Consider the feedback system in Figure 8.17. Its transfer function is

$$H_o(s) = \frac{A}{1 + AH(s)} = \frac{A(s^3 + 2s^2 + s + 1)}{s^3 + 2s^2 + s + 1 + A}$$

Using the Routh test, we can show that $s^3 + 2s^2 + s + 1 + A$ is not a stable polynomial for any $A \geq 1$ (Problem 8.16). Thus $H_o(s)$ is not stable for A large and cannot be reduced to $H^{-1}(s)$. To verify this assertion, we apply the output (Figure 8.19(b)) of $H(s)$ excited by $u(t) = e^{-0.3t} \sin t$ (Figure 8.19(a)) to $H_o(s)$ with $A = 1000$. The output of $H_o(s)$ shown in Figure 8.19(c) grows unbounded and does not resemble $u(t)$. Thus the feedback system in Figure 8.17 cannot implement the inverse system of $H(s)$.

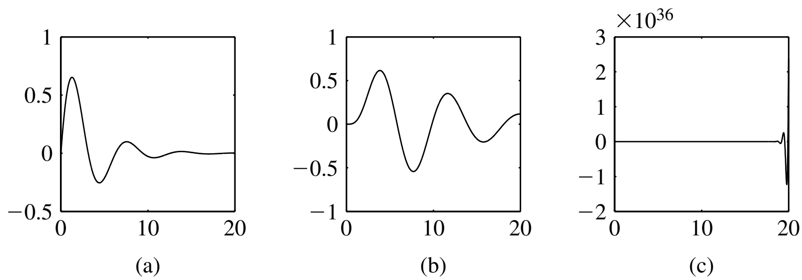


Figure 8.19 (**a**) Input $u(t)$ of $H(s)$. (**b**) Its output $y(t)$. (**c**) Output of $H_o(s)$ excited by $y(t)$.

Even though feedback implementation of inverse systems is often suggested in the literature, such implementation is not always possible, as shown in the preceding examples. In fact, it is basically a model reduction problem and involves two issues: stability and operational frequency range. Without considering these two issues, the implementation is simply a mathematical exercise with no engineering bearing.

8.5 WIEN-BRIDGE OSCILLATOR

Sinusoidal generators are devices that maintain sinusoidal oscillations once they are excited. We discuss one simple such circuit in this section.

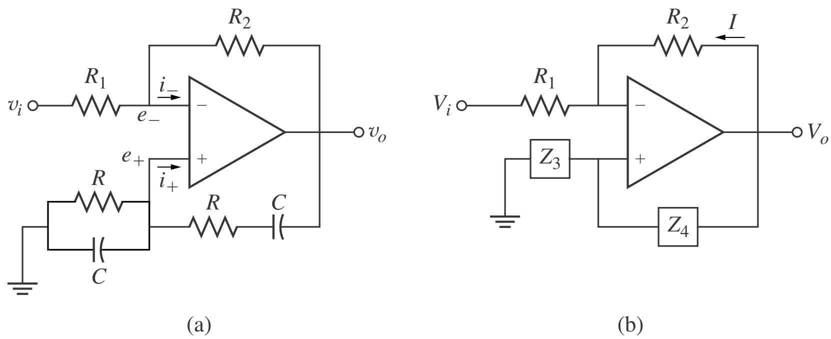


Figure 8.20 (**a**) Wien-bridge oscillator. (**b**) Using impedances.

Consider the op-amp circuit, called the *Wien-bridge oscillator,* shown in Figure 8.20(a) where the op amp is modeled as ideal. We use transform impedances to compute its transfer function from the input v_i to the output v_o. Let $Z_3(s)$ and $Z_4(z)$ be, respectively, the impedances of the parallel and series connections of R and C as shown in Figure 8.20(b). Then we have

$$Z_3 = \frac{R(1/Cs)}{R + (1/Cs)} = \frac{R}{RCs + 1} \tag{8.25}$$

and

$$Z_4 = R + (1/Cs) \tag{8.26}$$

The voltage $E_+(s)$ at the noninverting terminal is, using $I_+(s) = 0$,

$$E_+(s) = \frac{Z_3}{Z_3 + Z_4} V_o(s)$$

The current flowing from v_o to v_i is, using $I_-(s) = 0$,

$$I(s) = \frac{V_o(s) - V_i(s)}{R_1 + R_2}$$

Thus the voltage $E_-(s)$ at the inverting terminal is

$$\begin{aligned} E_-(s) &= V_i(s) + I(s)R_1 = V_i(s) + \frac{R_1}{R_1 + R_2}(V_o(s) - V_i(s)) \\ &= \frac{R_1 V_o(s) + R_2 V_i(s)}{R_1 + R_2} \end{aligned}$$

Equating $E_+(s) = E_-(s)$ yields

$$\frac{Z_3}{Z_3 + Z_4} V_o(s) = \frac{R_1 V_o(s) + R_2 V_i(s)}{R_1 + R_2}$$

which implies

$$(R_1 + R_2) Z_3 V_o(s) = (Z_3 + Z_4) R_1 V_o(s) + (Z_3 + Z_4) R_2 V_i(s)$$

and

$$(R_2 Z_3 - R_1 Z_4) V_o(s) = (Z_3 + Z_4) R_2 V_i(s)$$

Thus the transfer function from v_i to v_o is

$$H(s) = \frac{V_o(s)}{V_i(s)} = \frac{(Z_3 + Z_4)R_2}{R_2 Z_3 - R_1 Z_4} \tag{8.27}$$

Substituting (8.25) and (8.26) into (8.27) and after simple manipulation, we finally obtain the transfer function as

$$H(s) = \frac{-((RCs)^2 + 3RCs + 1)R_2}{R_1(RCs)^2 + (2R_1 - R_2)RCs + R_1} \tag{8.28}$$

It has two poles and two zeros.

We now discuss the condition for the circuit to maintain a sustained oscillation once it is excited. If $H(s)$ has one or two poles inside the open RHP, its output will grow unbounded once the circuit is excited. If the two poles are inside the open LHP, then the output will eventually vanish once the input is removed. Thus the condition for the circuit to maintain a sustain oscillation is that the two poles are on the $j\omega$-axis. This is the case if

$$2R_1 = R_2$$

Under this condition, the denominator of (8.28) becomes $R_1[(RC)^2 s^2 + 1]$ and the two poles are located at $\pm j\omega_0$ with

$$\omega_0 := \frac{1}{RC}$$

They are pure imaginary poles. After the circuit is excited and the input is removed, the output $v_o(t)$ is of the form

$$v_o(t) = k_1 \sin(\omega_0 t + k_2)$$

for some constants k_1 and k_2. It is a sustained oscillation with frequency $1/RC$ rad/s. Because of its simplicity in structure and design, the Wien-bridge oscillator is widely used.[6]

Although the frequency of oscillation is fixed by the circuit, the amplitude k_1 of the oscillation depends on how it is excited. Different excitation will yield different amplitude. In order to have a fixed amplitude, the circuit must be modified. First we select a R_2 slightly larger than $2R_1$. Then the transfer function in (8.28) becomes unstable. In this case, even though no input is applied, once the power is turned on, the circuit will start to oscillate due to thermal noise or power-supply transient. The amplitude of oscillation will increase with time. When it reaches A, a value predetermined by a nonlinear circuit, called a limiter, the circuit will maintain $A \sin \omega_r t$, where ω_r is the imaginary part of the unstable poles and is very close to ω_0. See References 8, 11, and 23. Thus an actual Wien-bridge oscillator is more complex than the one shown in Figure 8.20. However, the preceding linear analysis does illustrate the basic design of the oscillator.

[6]Once an input is removed, responses of the Wien-bridge oscillator can be considered to be excited by nonzero initial conditions. The Wien-bridge oscillator is said to be *marginally stable* because its zero-input responses will not approach zero and will not grow unbounded. Note that the oscillator has a pair of pure imaginary poles and is therefore not BIBO stable.

8.6 FEEDBACK MODEL OF OP-AMP CIRCUITS

The Wien-bridge oscillator was designed directly in the preceding section. Most electronics texts carry out its design using a feedback model. Thus we develop first a feedback model for a general op-amp circuit.

Consider the op-amp circuit shown in Figure 8.21(a) in which the op amp is modeled as $I_-(s) = -I_+(s) = 0$ and $V_o(s) = A(s)[E_+(s) - E_-(s)]$. Let $Z_i(s)$, for $i = 1:4$, be any transform impedances. We develop a block diagram for the circuit.

The current I_1 passing through the impedance Z_1 is $(V_1 - E_-)/Z_1$, and the current I_2 passing through Z_2 is $(V_o - E_-)/Z_2$. Because $I_- = 0$, we have $I_1 = -I_2$ or

$$\frac{V_1 - E_-}{Z_1} = -\frac{V_o - E_-}{Z_2}$$

which implies

$$E_- = \frac{Z_2 V_1 + Z_1 V_o}{Z_1 + Z_2} = \frac{Z_2}{Z_1 + Z_2} V_1 + \frac{Z_1}{Z_1 + Z_2} V_o \tag{8.29}$$

This equation can also be obtained using a linearity property. If $V_o = 0$, the voltage E_- at the inverting terminal excited by V_1 is $[Z_2/(Z_1 + Z_2)]V_1$. If $V_1 = 0$, the voltage E_- excited by V_o is $[Z_1/(Z_1 + Z_2)]V_o$. Thus we have (8.29).

Likewise, because $I_+ = 0$, we have, at the noninverting terminal,

$$\frac{V_2 - E_+}{Z_3} = -\frac{V_o - E_+}{Z_4}$$

which implies

$$E_+ = \frac{Z_4 V_2 + Z_3 V_o}{Z_3 + Z_4} = \frac{Z_4}{Z_3 + Z_4} V_2 + \frac{Z_3}{Z_3 + Z_4} V_o \tag{8.30}$$

This equation can also be obtained using a linearity property as discussed above. Using (8.29), (8.30), and $V_o = A(E_+ - E_-)$, we can obtain the block diagram in Figure 8.21(b). It is a feedback model of the op-amp circuit in Figure 8.19(a). If $V_1 = 0$ and $V_2 = 0$, the feedback model reduces to the one in Reference 11 (p. 847).

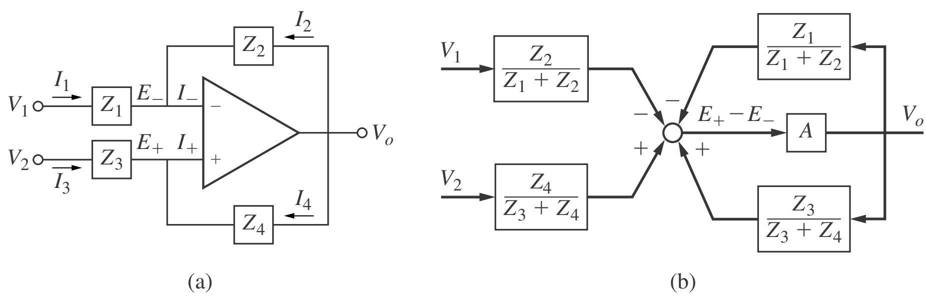


Figure 8.21 (**a**) Op-amp circuit. (**b**) Its block diagram.

8.6.1 Feedback Model of Wien-Bridge Oscillator

Applying Figure 8.21(b) to Figure 8.20(b), we can obtain for the Wien-bridge oscillator the feedback model shown in Figure 8.22(a) with Z_3 and Z_4 given in (8.25) and (8.26). Note that the summer in Figure 8.21(b) has been plotted as two summers in Figure 8.22(a). This is permitted because the signals at the input of the gain A are the same. If A is very large or infinite, the negative-feedback loop implements the inverse of $R_1/(R_1+R_2)$. Note that the negative-feedback system is memoryless and is always stable for any finite A, albeit very large. Thus Figure 8.22(a) can be reduced as shown in Figure 8.22(b). Let us define

$$H_1(s) := \frac{R_1 + R_2}{R_1} \tag{8.31}$$

and

$$H_2(s) := \frac{Z_3}{Z_3 + Z_4} = \frac{RCs}{(RCs)^2 + 3RCs + 1} \tag{8.32}$$

where we have substituted (8.25) and (8.26). Because the input V_i enters into the inverting terminal and feedback enters into the noninverting terminal, we have the negative and positive signs shown. This is essentially the feedback model used in References 8 and 23. Note that the system in Figure 8.22(b) is a positive feedback system.

The feedback system in Figure 8.22 is slightly different from the ones in Figures 8.11(c) and (d). However, using the same procedure, we can obtain its transfer function as

$$H(s) = \frac{V_o(s)}{V_i(s)} = \frac{-R_2}{R_1 + R_2} \times \frac{H_1(s)}{1 - H_1(s)H_2(s)} \tag{8.33}$$

Substituting (8.31) and (8.32) into (8.33), we will obtain the same transfer function in (8.28) (Problem 8.19).

We discuss an oscillation condition, called the *Barkhausen criterion*. The criterion states that if there exists an ω_0 such that

$$H_1(j\omega_0)H_2(j\omega_0) = 1 \tag{8.34}$$

then the feedback system in Figure 8.22(b) can maintain a sinusoidal oscillation with frequency ω_0. Note that once the circuit is excited, the input is removed. Thus the oscillation condition depends only on $H_1(s)$ and $H_2(s)$. Their product $H_1(s)H_2(s)$ is called the *loop gain*. It is the product of the transfer functions along the loop.

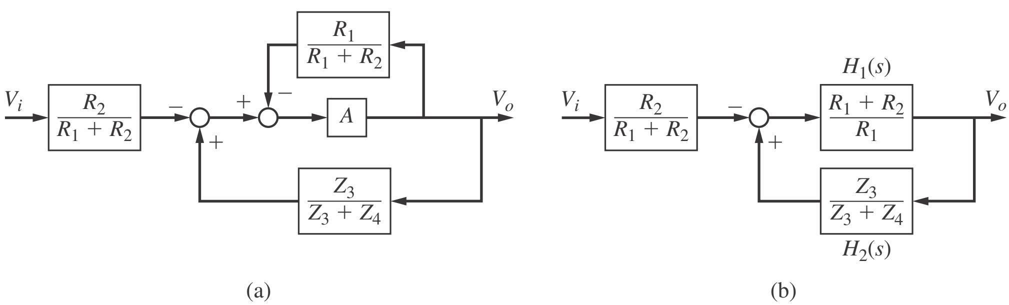


Figure 8.22 (**a**) Feedback model of Wien-bridge oscillator with finite A. (**b**) With $A = \infty$.

The criterion in (8.34) can be established as follows. Suppose the feedback system in Figure 8.22(b) maintains the steady-state oscillation $\text{Re}(Ae^{j\omega_0 t})$ at the output after the input is removed ($v_i = 0$). Then the output of $H_2(s)$ is $\text{Re}(AH_2(j\omega_0)e^{j\omega_0 t})$ (See (6.50).) This will be the input of $H_1(s)$ because $v_i = 0$. Thus the output of $H_1(s)$ is

$$\text{Re}(H_1(j\omega_0)AH_2(j\omega_0)e^{j\omega_0 t})$$

If the condition in (8.34) is met, then this output reduces to $\text{Re}(Ae^{j\omega_0 t})$, the sustained oscillation. This establishes (8.34). The criterion is widely used in designing Wien-bridge oscillators. See References 8, 11, and 23.

We mention that the criterion is the same as checking whether or not $H(s)$ in (8.33) has a pole at $j\omega_0$. We write (8.34) as

$$1 - H_1(j\omega_0)H_2(j\omega_0) = 0$$

which implies $H_1(j\omega_0) \neq 0$ and $|H(j\omega_0)| = \infty$. Thus $j\omega_0$ is a pole of $H(s)$ in (8.33). In other words, the criterion is the condition for $H(s)$ to have a pole at $j\omega_0$. Because $H(s)$ has only real coefficients, if $j\omega_0$ is a pole, so is $-j\omega_0$. Thus the condition in (8.34) checks the existence of a pair of complex conjugate poles at $\pm j\omega_0$. This is the condition used in Section 8.5. In conclusion, the Wien-bridge oscillator can be designed directly as in Section 8.5 or using a feedback model as in this subsection.

8.7 MODULATION

Every device is designed to operate only in a limited frequency range. For example, motors are classified as dc and ac motors. A dc motor can be driven by dc signals or signals with frequency spectra centered around $\omega = 0$. A 60-Hz ac motor cannot be driven by a dc signal; it will operate most efficiently if the applied signal has frequency spectrum centered around 60 Hz. Now suppose we have a signal $x(t)$ with frequency spectrum $X(\omega)$ as shown in Figure 8.23. Because the spectrum does not cover the operational frequency range of the motor, the signal cannot drive the motor. This problem can be resolved if we shift the spectrum of the signal to center around 60 Hz by multiplying $x(t)$ with $\cos 120\pi t$. See (4.40) and (4.41). Thus modulation can be used to match up the frequency spectrum of a signal with the operational frequency range of a device. This is used in control systems.

Modulation is important in communication. The frequency spectrum of human voices lies generally between 200 Hz and 4 kHz. Such signals are not transmitted directly for three reasons. First, the wavelength of 200-Hz signals is $(3 \times 10^8)/200 = 1.5 \times 10^6$ meters. In order to transmit such signals efficiently, the size of an antenna must be at least one-tenth of their wavelengths, or

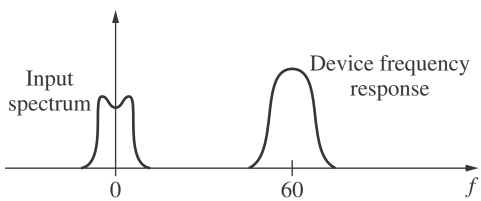


Figure 8.23 Matching operational frequency range.

1.5×10^5 meters. This is impractical. Second, such signals will attenuate rapidly in atmosphere. And third, interference will occur if two or three signals are transmitted simultaneously. For these reasons, such signals are modulated before transmission.

There are many modulation schemes. We discuss only two of them. Consider a signal $x(t)$ with frequency spectrum $X(\omega)$ band-limited to W, that is,

$$X(\omega) = 0 \qquad \text{for } |\omega| > W$$

Let us multiply $x(t)$ by $\cos \omega_c t$ to yield

$$x_m(t) = x(t) \cos \omega_c t \tag{8.35}$$

The process can be represented as shown in Figure 8.24(a). If $x(t) = 2e^{-0.3t} \sin t$ as shown in Figure 1.14(a), then $x_m(t)$ with $\omega_c = 20$ is as shown in Figure 8.24(aa). We call $\cos \omega_c t$ the *carrier signal,* and we call ω_c the *carrier frequency*. The signal $x(t)$ is called the *modulation signal,* and $x_m(t)$ is called the *modulated signal*. If the spectrum of $x(t)$ is as shown in Figure 8.25(a), then the spectrum $X_m(\omega)$ of $x_m(t)$ is, using (4.41), as shown in Figure 8.25(b).

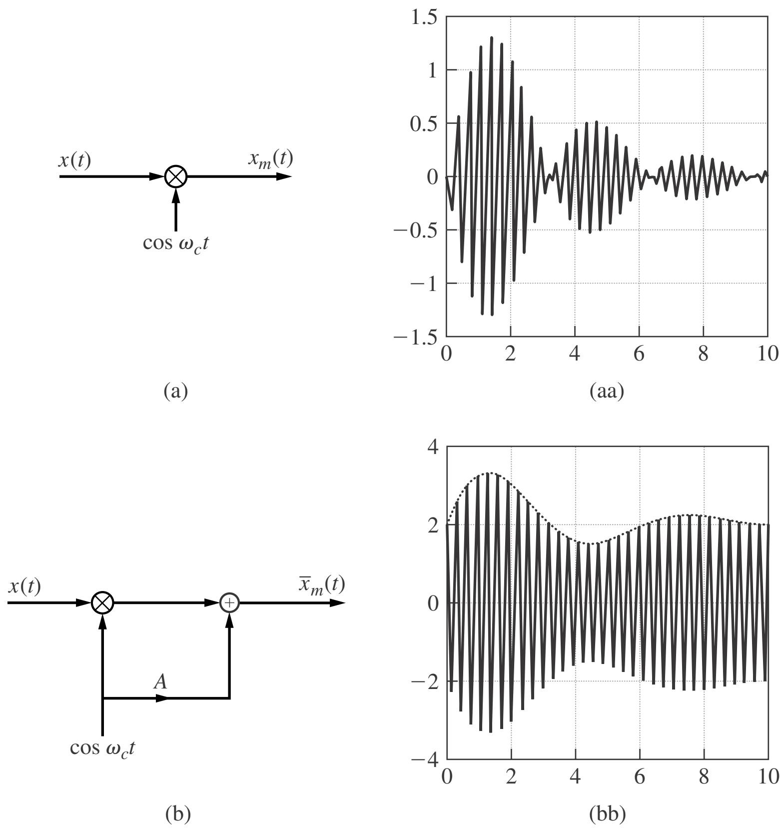


Figure 8.24 (**a**) DSB-SC amplitude modulation. (**aa**) $x_m(t) = x(t) \cos \omega_c t$. (**b**) Amplitude modulation. (**bb**) $\bar{x}_m(t) = [A + x(t)] \cos \omega_c t$.

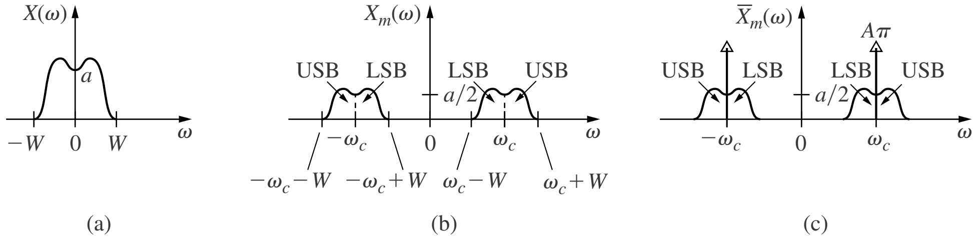


Figure 8.25 (**a**) Spectrum of $x(t)$. (**b**) Spectrum of $x_m(t)$. (**c**) Spectrum of $\bar{x}_m(t)$.

Next we consider the modulation scheme shown in Figure 8.24(b). The modulated signal is given by

$$\bar{x}_m(t) = [A + x(t)] \cos \omega_c t \tag{8.36}$$

where A is selected so that $A + x(t) > 0$ for all t. We plot $\bar{x}_m(t)$ in Figure 8.24(bb) for $x(t) = 2e^{-0.3t} \sin t$, $A = 2$, and $\omega_c = 20$. We see that the upper envelope of $\bar{x}_m(t)$, denoted by a dotted line, equals $x(t)$ shifted up by $A = 2$.

The frequency spectrum of $\bar{x}_m(t)$ is, using (4.54) and (4.41),

$$\begin{aligned}\bar{X}_m(\omega) &= \mathcal{F}[A \cos \omega_c t] + \mathcal{F}[x(t) \cos \omega_c t] \\ &= A\pi[\delta(\omega - \omega_c) + \delta(\omega + \omega_c)] \\ &\quad + 0.5[X(\omega - \omega_c) + X(\omega + \omega_c)]\end{aligned} \tag{8.37}$$

and is plotted in Figure 8.25(c). This modulation scheme is called the *amplitude modulation (AM)* or, more specifically, the *double-sideband AM*. Double sidebands refer to the fact that the spectrum $\bar{X}_m(\omega)$ contains both upper sideband (USB) and lower sideband (LSB) as shown in Figure 8.23(c). It is possible to design a single sideband AM that contains only an upper or lower sideband. See Reference 15.

The spectrum in (8.37) contains two impulses at $\pm\omega_c$ which are the spectrum of the carrier signal $\cos \omega_c t$. The two impulses due to the carrier signal do not appear in $X_m(\omega)$. Thus the modulation scheme in (8.35) is called the *double sideband-suppressed carrier (DSB-SC) modulation*.

8.7.1 Filtering and Synchronous Demodulation

Recovering $x(t)$ from $x_m(t)$ or $\bar{x}_m(t)$ is called *demodulation*. This section discusses recovering $x(t)$ from $x_m(t)$. Let us multiply $x_m(t)$ by $\cos \omega_c t$. Then we have, using $\cos^2 \theta = 0.5(1 + \cos 2\theta)$,

$$y(t) := x_m(t) \cos \omega_c t = x(t)(\cos \omega_c t)^2 = 0.5x(t)[1 + \cos 2\omega_c t]$$

and, using (4.41),

$$Y(\omega) := \mathcal{F}[y(t)] = 0.5X(\omega) + 0.25X(\omega - 2\omega_c) + 0.25X(\omega + 2\omega_c)$$

which is plotted in Figure 8.26. We see that its spectrum consists of three parts: the spectrum of $x(t)/2$ and the spectrum of $x(t)/4$ shifted to $\pm 2\omega_c$. Thus if we design a lowpass filter with gain 2 and cutoff frequency larger than W but less than $2\omega_c - W$, then we can recover $x(t)$.

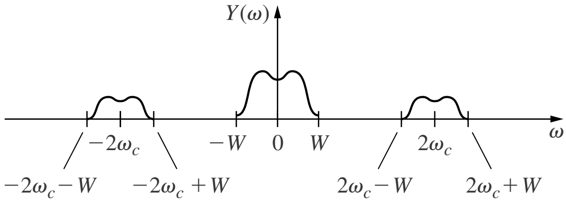


Figure 8.26 Spectrum of $y(t)$.

We use the problem posed in Section 1.4.3 to illustrate the procedure. Consider the signal $x_1(t) = 2e^{-0.3t}\sin t$ and $x_2(t) = -x_1(t)$ shown in Figures 1.14(a) and 1.14(b). If we transmit $x_1(t) + x_2(t) = 0$ directly, there is no way to recover $x_1(t)$ and $x_2(t)$. We claimed there that if we transmit $x_1(t) + x_2(t)\cos 20t$, then it is possible to recover $x_1(t)$ and $x_2(t)$. We now show how to achieve this.

The magnitude spectrum of

$$x(t) = x_1(t) + x_2(t)\cos 20t = 2e^{-0.3t}\sin t - 2e^{-0.3t}\sin t\cos 20t \tag{8.38}$$

for $\omega \geq 0$, is plotted in Figure 8.27 with a solid line. It is the sum of the magnitude spectra in Figures 5.15(a) and 5.15(b). The spectrum of $x_1(t)$, which is centered around $\omega = 0$, and the spectrum of $x_{2m}(t) = x_2(t)\cos 20t$, which is centered around $\omega = 20$, are widely apart. Thus if we design a lowpass filter that passes the spectrum of $x_1(t)$ and stops the spectrum of $x_{2m}(t)$, then the output of the filter excited by $x_1(t) + x_{2m}(t)$ will yield $x_1(t)$. We demonstrate this by using the lowpass filter[7]

$$H_1(s) = \frac{61.4}{s^3 + 4.9s^2 + 31s + 61.4} \tag{8.39}$$

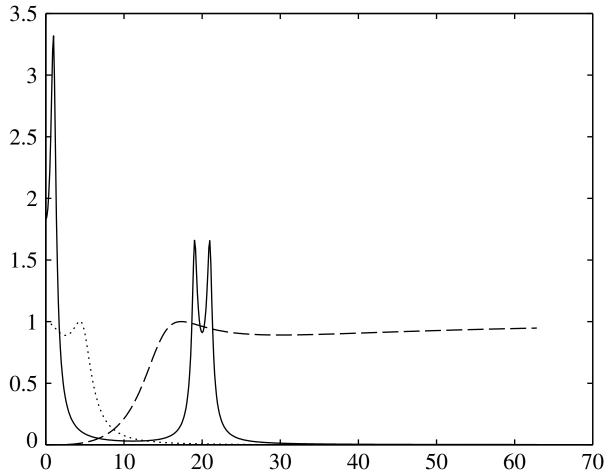


Figure 8.27 Magnitude spectrum of $x(t)$ (solid line). Magnitude response of the lowpass filter in (8.39) (dotted line). Magnitude response of the highpass filter in (8.40) (dashed line).

[7]It is the order 3 Chebyshev type I lowpass filter with passband cutoff frequency 5 rad/s and passband tolerance 1 dB, and it is obtained in MATLAB by typing [b,a]=cheby1(3,1,5,'s'). See Reference 2.

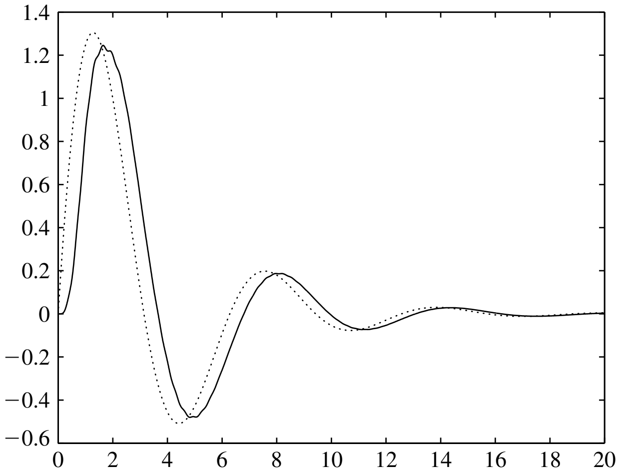


Figure 8.28 Output of the lowpass filter (solid line) and signal $x_1(t)$ (dotted line).

Its magnitude response is shown in Figure 8.27 with a dotted line. We see that the spectrum of $x_{2m}(t)$ is completely blocked by the filter, and the output of the filter should be close to $x_1(t)$. To verify this, we carry out the following computer simulation:

```
t=0:0.01:20;
x=2*exp(-0.3*t).*sin(t)-2*exp(-0.3*t).*sin(t).*cos(20*t);
b=61.4;a=[1 4.9 31 61.4];
sys=tf(b,a);
y=lsim(sys,x,t);
x1=2*exp(-0.3*t).*sin(t);
plot(t,y,t,x1,':')
```

The first line specifies the time interval to be computed: from 0 to 20 seconds with time increment 0.01. The second line is the input signal in (8.38). The third line lists the coefficients of the lowpass filter in (8.39). The fourth line defines the system, called sys, using transfer function. The fifth line computes the output of the filter. In order to compare this output with $x_1(t)$, we plot in Figure 8.28 the output of the filter (solid line) and $x_1(t)$ (dotted line) using the last two lines of the program. We see that the output of the filter is close to $x_1(t)$ except for a small time delay and a small attenuation.

To recover $x_2(t)$ from $x(t) = x_1(t) + x_2(t)\cos 20t$, we use the highpass filter[8]

$$H_2(s) = \frac{s^3}{s^3 + 37.8s^2 + 452.6s + 6869.4} \tag{8.40}$$

Its magnitude response is plotted in Figure 8.27 with a dashed line. We see that the spectrum of $x_1(t)$ is blocked by the highpass filter. Thus the output of the filter will be close to $x_{2m}(t) = x_2(t)\cos 20t$.

[8]It is the order 3 Chebyshev type I highpass filter with passband cutoff frequency 15 rad/s and passband tolerance 1 dB, and it is obtained in MATLAB by typing `[b,a]=cheby1(3,1,15,'high','s')`. See Reference 2.

Let $y_m(t)$ be the output of the highpass filter in (8.40) and be assumed of the form $y_m(t) = y(t)\cos(20t+\theta)$, for some θ. It turns out that whether or not we can recover $x_2(t)$ depends on the θ selected. First we use Theorem 6.4 to select $\theta = \measuredangle H_2(j20) = 1.69$; that is, we assume that the output of the highpass filter is of the form $y_m(t) = y(t)\cos(20t+1.67)$. Then $y(t)$ should be close to $x_2(t)$. To recover $y(t)$ from $y_m(t)$, we multiply $y_m(t)$ by $\cos(20t+1.69)$ to yield

$$\begin{aligned} y_{mm} &= y(t)\cos(20t+\theta)\cos(20t+\theta) = y(t)\,[0.5(1+\cos(40t+2\theta))] \\ &= 0.5y(t) + 0.5y(t)\cos(40t+2\theta) \end{aligned} \tag{8.41}$$

It consists of the signal $y(t)$ itself and its modulated signal with carrier frequency 40 rad/s. If $y_{mm}(t)$ passes through the lowpass filter in (8.39) with gain doubled, then the output of the filter should be close to $x_2(t)$. The program that follows implements the preceding procedure:

```
t=0:0.01:20;
x=2*exp(-0.3*t).*sin(t)-2*exp(-0.3*t).*sin(t).*cos(20*t);
b=[1 0 0 0];a=[1 37.8 452.6 6869.4];
sys=tf(b,a);
[ym,t]=lsim(sys,x,t);
ymm=ym.*cos(20*t+1.69);
b1=2*61.4;a1=[1 4.9 31 61.4];
sys1=tf(b1,a1);
y=lsim(sys1,ymm,t);
x2=-2*exp(-0.3*t).*sin(3*t);
plot(t,y,t,x2,':')
```

The result is plotted in Figure 8.29 with a solid line. It is indeed close to $x_2(t)$ (dotted line) except for a small time delay and a small attenuation. This verifies the assertion that the signals $x_1(t)$ and $x_2(t)$ can be recovered from their sum if one of them is modulated with a properly selected carrier frequency.

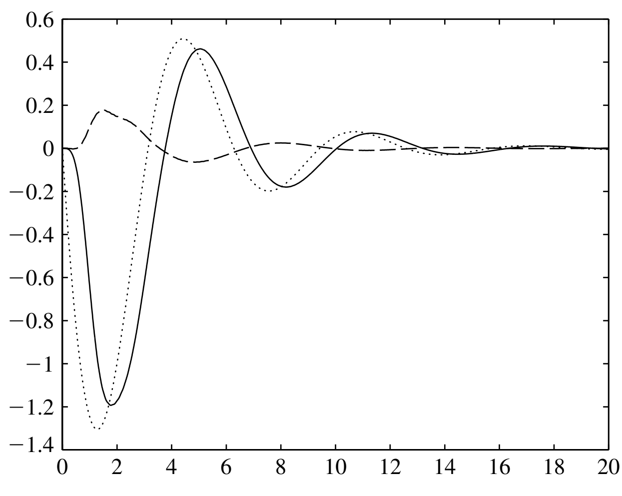


Figure 8.29 Result of highpass filtering and demodulation with a correctly selected phase (solid line), with an incorrectly selected phase (dashed line), and the original signal $x_2(t)$ (dotted line).

In this demodulation process, if θ is selected as zero—that is, we multiply y_m in the program by $\cos 20t$ instead of $\cos(20t + 1.67)$—then the program yields the dashed line in Figure 8.29. The result is quite different from the original $x_2(t)$. Thus this demodulating scheme requires not only to synchronize the two $\cos \omega_c t$ at the demodulator and modulator but also to include the phase introduced by the highpass filter. This is difficult to achieve in practice. Thus the DSB-SC modulation is not used in radio transmission. However, it can be used in control systems where the modulation is used only to match up the operational frequency range and there is no need to recover the original signal.

8.8 AM MODULATION AND ASYNCHRONOUS DEMODULATION

There are many ways to implement AM modulation and demodulation. See References 18 and 28. We discuss only the ones that are based on the material in this text. Consider the arrangement shown in Figure 8.30. It is called a *switching modulator* where the switching is provided by the diode.[9] The input is the sum of $x(t)$ and $c \cos \omega_c t$, that is, $x(t) + c \cos \omega_c t$. We require $c \gg x(t)$ so that $x(t) + c \cos \omega_c t > 0$ if $\cos \omega_c t > 0$ and $x(t) + c \cos \omega_c t < 0$ if $\cos \omega_c t < 0$. Thus the diode is forward biased if $\cos \omega_c > 0$ and reverse biased if $\cos \omega_c < 0$. Consequently, the voltage v_o across the resistor is the product of the input and the switching function $p(t)$ in (4.10) with $P = 2\pi/\omega_c$ or, equivalently, with ω_0 replaced by ω_c. Thus we have

$$\begin{aligned} v_o(t) &= [x(t) + c \cos \omega_c t] p(t) \\ &= [x(t) + c \cos \omega_c t] \left[\frac{1}{2} + \frac{2}{\pi} \cos \omega_c t - \frac{2}{3\pi} \cos 3\omega_c t + \cdots \right] \end{aligned} \tag{8.42}$$

Using $\cos\theta \cos\phi = 0.5[\cos(\theta - \phi) + \cos(\theta + \phi)]$, we can write (8.42) as

$$\begin{aligned} v_o(t) = \frac{1}{2} x(t) + \frac{c}{\pi} + \left[\frac{c}{2} + \frac{2}{\pi} x(t) \right] \cos \omega_c t \\ + \text{terms containing } \cos k\omega_c t \text{ with } k \geq 2 \end{aligned} \tag{8.43}$$

See Example 4.2.5. Let $x(t)$ be band-limited to W and let ω_c be larger than $2W$. Now if we designed an ideal *bandpass* filter with lower and upper cutoff frequencies $\omega_c - W$ and $\omega_c + W$, then the output of the filter will be

$$y(t) = \left[\frac{c}{2} + \frac{2}{\pi} x(t) \right] \cos \omega_c t \tag{8.44}$$

This is essentially the AM signal in (8.36). This is a simple way of implementing AM modulation. We mention that after a bandpass filter is connected, the resistor shown in Figure 8.28 is no longer needed.

[9] The output of a modulator contains frequencies other than those contained in the input signal. The output of an LTI system cannot generate new frequencies other than those contained in the input signal. See the discussion at the end of Section 6.10. Thus a modulator cannot be linear and time invariant. It must be nonlinear and/or time-varying. The switching modulator in Figure 8.30 is nonlinear but time-invariant.

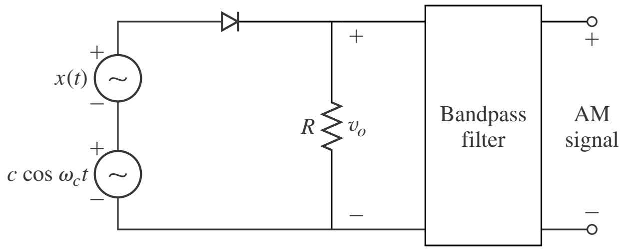


Figure 8.30 Switching modulator.

Next we discuss the recovering of $x(t)$ from the AM signal $\bar{x}_m(t) = (A + x(t)) \cos \omega_c t$. We use the fact that the upper envelope of $\bar{x}_m(t)$ equals $A + x(t)$ as shown in Figure 8.24(bb). Consider the network enclosed by the dashed lines shown in Figure 8.31(a). It consists of a diode, a resistor, and a capacitor as shown. If the capacitor is absent or $C = 0$, then the circuit is a half-wave rectifier. See Section 2.9. Thus if the input is $\bar{x}_m(t)$, its output voltage will be as shown in Figure 8.31(b) with dotted lines. It equals the part of $\bar{x}_m(t) \geq 0$. If the resistor is absent or $R = \infty$, then the output voltage will be as shown with the horizontal dashed line. In this case, the input voltage will charge the capacitor whenever it is larger than $v_o(t)$ and the diode will block the capacitor to discharge when the input voltage is smaller than $v_o(t)$. Thus if there is no resistor, the circuit is a *peak rectifier*. See Reference 23.

We now discuss the case where both R and C are present. When the diode is forward biased, the input voltage charges up the capacitor rapidly. When the diode is reverse biased, the capacitor discharges through the resistor in the form $e^{-t/RC}$ and with time constant RC. See Problem 6.29. If there is no R or $R = \infty$, the time constant is infinity and the capacitor will not discharge as shown in Figure 8.31(b) with the horizontal dashed line. If $C = 0$, the time constant is zero, thus the capacitor discharges immediately and the voltage $v_o(t)$ follows the input voltage when the diode conducts as shown in Figure 8.31(b) with dotted lines. This is what we discussed in the preceding paragraph.

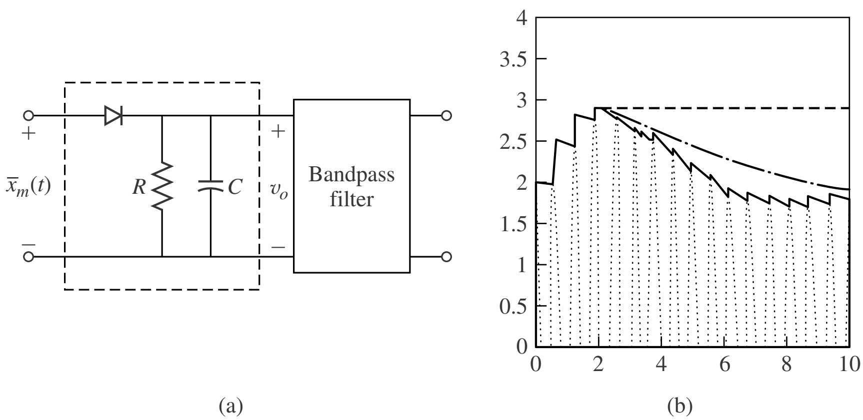


Figure 8.31 (**a**) Envelope detector and bandpass filter. (**b**) $v_o(t)$: Dotted line if $C = 0$, dashed line if $R = \infty$, dash-and-dotted line if RC is too large, and solid line for properly selected RC.

If R and C are selected properly, then the output will be as shown in Figure 8.31(b) with a solid line which is close to the upper envelope of $\bar{x}_m(t)$ or $A + x(t)$. Thus the network enclosed by the dashed lines in Figure 8.31(a) with such R and C is called an *envelope detector*. Note that if R and C are not selected properly, the output may be as shown with the dot-and-dashed line. Thus the selection of R and C is not a simple task.

The output of the envelope detector contains the dc gain A and ripples. Because A has zero frequency and ripples have high frequencies, they can be eliminated using a bandpass filter. If the spectrum of $x(t)$ is band-limited to W and if we design a bandpass filter with upper cutoff frequency W and lower cutoff frequency close to 0, then the output of the bandpass filter will be close to $x(t)$. This is one way of recovering $x(t)$ from $\bar{x}_m(t)$. If the bandpass filter is built with an op-amp circuit, then there will be no loading problem. If the filter is built with an RLC network, then we must consider the loading problem. In fact, the RC inside the dotted lines in Figure 8.31(a) is a lowpass filter and can be combined with the bandpass filter.

We give a different interpretation of the demodulation scheme in Figure 8.31(a). Suppose there are no R and C, then the voltage $v_o(t)$ is, as in (8.42), the product of the input $\bar{x}_m(t) = [A + x(t)]\cos\omega_c t$ and $p(t)$, that is,

$$\begin{aligned} v_o(t) &= \{[A + x(t)]\cos\omega_c t\}\, p(t) \\ &= [A + x(t)]\cos\omega_c t \left[\frac{1}{2} + \frac{2}{\pi}\cos\omega_c t - \frac{2}{3\pi}\cos 3\omega_c t + \cdots\right] \end{aligned} \tag{8.45}$$

which can be written as

$$v_o(t) = \frac{1}{\pi}[A + x(t)] + \text{terms containing } \cos k\omega_c t \text{ with } k \geq 1 \tag{8.46}$$

See (8.42), (8.43), and Example 4.2.5. Thus if we design a bandpass filter with upper cutoff frequency W and lower cutoff frequency close to zero and if $\omega_c > 2W$, then the output of the filter will yield a signal close to $x(t)/\pi$. In conclusion, using a diode and a bandpass filter, we can carry out amplitude modulation and amplitude demodulation. The passband of the bandpass filter is $[\omega_c - W, \omega_c + W]$ for modulation and $(0, W]$ for demodulation. If $x(t)$ is a voice signal, then its frequency spectrum is generally limited to 200 to 4 kHz. In this case, the passband of the bandpass filter for demodulation can be selected as [200, 4000] in Hz.

In communication, different applications are assigned to different frequency bands as shown in Figure 8.32. For example, the frequency band is limited from 540 to 1600 kHz for AM radio transmission, and from 87.5 to 108 MHz for FM (frequency modulation) transmission. In AM transmission, radio stations are assigned carrier frequencies from 540 to 1600 kHz with 10 kHz increments as shown in Figure 8.33. This is called *frequency-division multiplexing*.

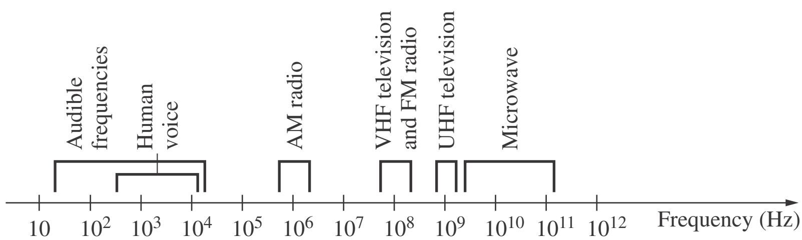


Figure 8.32 Various frequency bands.

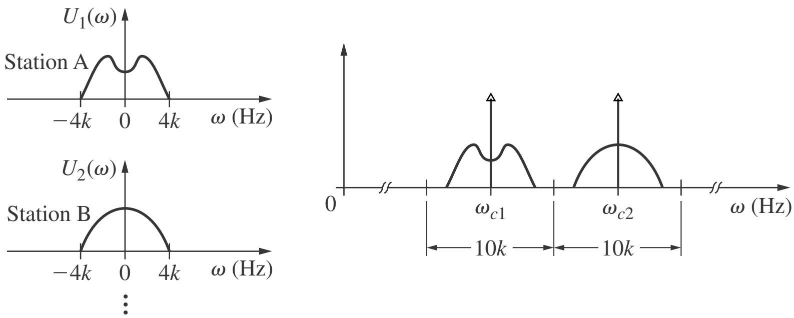


Figure 8.33 Frequency-division multiplexing.

PROBLEMS

8.1 Consider the voltage follower shown in Figure 8.1(a). What is its overall transfer function if the op amp is modeled to have the transfer function

$$A(s) = \frac{10^5}{s + 100}$$

Find the frequency range in which the overall transfer function has magnitude 0.707 or larger. We may call the range the operational frequency range of the voltage follower.

8.2 Consider the op-amp circuit shown in Figure 2.30(a) with $R_2 = 10R_1$. Show that it is an inverting amplifier with gain 10 if the op amp is ideal. Suppose the op amp is modeled to have the transfer function in (8.2). Show that the circuit is not stable. What is its output excited by a step input? Can it be used as an inverting amplifier?

8.3 Consider the noninverting amplifier shown in Figure 2.30(b) with $R_2 = 10R_1$. What is its transfer function if the op amp is modeled as ideal? What is its transfer function if the op amp is modeled to have the transfer function in (8.2)? Compare their step responses? What is its operational frequency range in which the op amp can be modeled as ideal?

8.4 Find a realization of (8.10) with $k_d = 2$ and $N = 20$ and then implement it using an op-amp circuit. Also show that the circuit in Figure 8.34 can implement (8.10) by selecting C, R_1, and R_2.

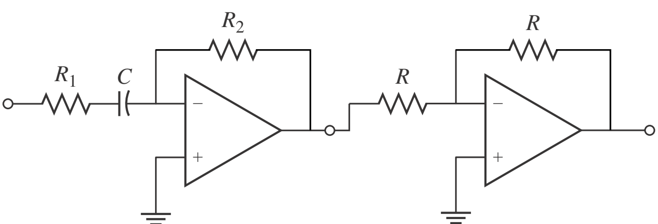


Figure 8.34

8.5 The device shown in Figure 8.35 can be used to measure pressure and is called a *manometer*. Suppose the transfer function from p to y is

$$H(s) = \frac{Y(s)}{P(s)} = \frac{0.1}{0.01s^2 + 0.1s + 1}$$

(a) If $p(t) = a$, what is its steady-state response? Is the steady-state height proportional to the applied pressure? (Calibrations are carried out for steady-state response in most instruments.)

(b) What is the time constant of $H(s)$? Roughly, how many seconds will it take for its step response to reach steady state?

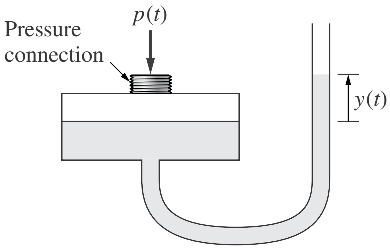


Figure 8.35

8.6 Consider the transfer function in (8.12). If $m = 1$ and $k = 2$, find the f so that the operational frequency range for (8.12) to act as a seismometer is the largest. Is this operational frequency range larger or smaller than $[1.25, \infty)$ computed for $m = 1$, $k = 0.2$, and $f = 0.63$?

8.7 Repeat Problem 8.6 for $m = 1$ and $k = 0.02$.

8.8 Verify that the tandem connection of two LTI single-input single-output systems has the commutative property; that is, the order of the two systems can be interchanged. Does this property hold for multi-input multi-output systems?

8.9 What is the output excited by $\sin 2t$ of the tandem connection of a differentiator followed by a multiplier with time-varying gain t? What is the output if the order of the differentiator and multiplier is interchanged? Does the commutative property hold for time-varying systems?

8.10 Verify that if $R_f = R/\beta$, then Figure 8.14(c) implements the β in Figure 8.14(b).

8.11 Consider the negative feedback system shown in Figure 8.11(c) with

$$H_1(s) = \frac{-10}{s+1} \quad \text{and} \quad H_2(s) = 2$$

Check the stability of $H_1(s)$, $H_2(s)$, and the feedback system. Is it true that if all subsystems are stable, then the negative feedback system is stable?

8.12 Consider the negative feedback system shown in Figures 8.11(c) with direct negative feedback ($H_2(s) = 1$) and

$$H_1(s) = \frac{2}{s-1}$$

It is called a unity negative feedback system. Check the stability of $H_1(s)$ and the feedback system. Is it true that if $H_1(s)$ is unstable, so is the negative feedback system?

8.13 Consider a unity negative feedback system with

$$H_1(s) = \frac{s^3 + 2s^2 - 2}{s^4 + s^2 + 2s + 3}$$

Is $H_1(s)$ stable? Is the feedback system stable?

8.14 Consider the op-amp circuit shown in Figure 8.36 where the op amp is modeled as $V_o(s) = A(s)[E_+(s) - E_-(s)]$ and $I_- = -I_+ = 0$. Verify that the transfer function from $V_i(s)$ to $V_o(s)$ is

$$H(s) = \frac{V_o(s)}{V_i(s)} = \frac{-A(s)Z_2(s)}{Z_1(s) + Z_2(s) + A(s)Z_1(s)}$$

Also derive it using the block diagram in Figure 8.21(b).

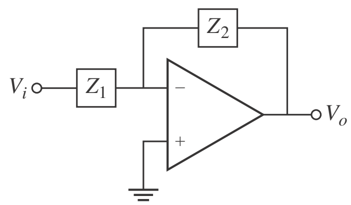


Figure 8.36

8.15 Consider the op-amp circuit in Figure 8.36 with $Z_1 = R$ and $Z_2 = 10R$. Find its transfer functions if $A = 10^5$ and $A = 2 \times 10^5$. The two open-loop gains differ by 100%. How much of a difference is there between the two transfer functions? Is the transfer function sensitive to the variation of A?

8.16 Use the Routh test to verify that the polynomial $s^3 + 2s^2 + s + 1 + A$ is a stable polynomial if and only if $-1 < A < 1$.

8.17 Can you use Figure 8.17 to implement the inverse system of $H(s) = (s+1)/(s^2 + 2s + 5)$? How about $H(s) = (s-1)/(s^2 + 2s + 5)$?

8.18 Consider the stable biproper transfer function $H(s) = (s-2)/(s+1)$. Is its inverse system stable? Can its inverse system be implemented as shown in Figure 8.17?

8.19 Verify that the transfer function in (8.33) equals the one in (8.28).

8.20 Find the transfer function from v_i to v_o of the op-amp circuit in Figure 8.37. What is the condition for the circuit to maintain a sinusoidal oscillation once it is excited, and what is its frequency of oscillation? Are the results the same as those obtained in Section 8.5?

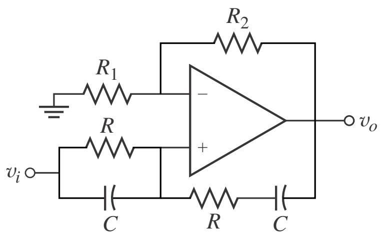


Figure 8.37

8.21 Use Figure 8.21(b) to develop a feedback block diagram for the circuit in Figure 8.37, and then compute its transfer function. Is the result the same as the one computed in Problem 8.20?

8.22 Consider $x(t) = \cos 10t$. What is its frequency spectrum. What are the spectra of $x_m(t) = x(t) \cos 100t$ and $\bar{x}_m(t) = (x(t) + 2) \cos 100t$.

8.23 Repeat Problem 8.22 for $x(t) = \cos 10t \cos 20t$.

8.24 Develop a procedure to recover $x(t)$ from $x_m(t)$ in Problem 8.22.

8.25 Repeat Problem 8.24 for Problem 8.23.

8.26 Consider a signal $x(t)$ band-limited to W and consider its modulated signal $x_m(t) = x(t) \cos \omega_c t$. What is the condition on ω_c in order to recover $x(t)$ from $x_m(t)$? Give your reasons.

CHAPTER 9 DT Transfer Functions—z-Transform

9.1 INTRODUCTION

Every DT LTI lumped and causal system, as discussed in Chapter 3, can be described by an advanced-form difference equation such as

$$
\begin{aligned}
&a_1 y[n+N] + a_2 y[n+N-1] + \cdots + a_{N+1} y[n] \\
&\quad = b_1 u[n+M] + b_2 u[n+M-1] + \cdots + b_{M+1} u[n]
\end{aligned}
\tag{9.1}
$$

with $a_1 \neq 0, b_1 \neq 0$, and $N \geq M$. Other than at least one of a_{N+1} and b_{M+1} being different from 0, all other a_i and b_i can be zero or nonzero. The equation is said to have order $\max(N, M) = N$. The system can also be described by a delayed-form difference equation such as

$$
\begin{aligned}
&\bar{a}_1 y[n] + \bar{a}_2 y[n-1] + \cdots + \bar{a}_{\bar{N}+1} y[n-\bar{N}] \\
&\quad = \bar{b}_1 u[n] + \bar{b}_2 u[n-1] + \cdots + \bar{b}_{\bar{M}+1} u[n-\bar{M}]
\end{aligned}
\tag{9.2}
$$

with $\bar{a}_1 \neq 0$, $\bar{a}_{\bar{N}+1} \neq 0$, and $\bar{b}_{\bar{M}+1} \neq 0$. The rest of $\bar{a}_i$ and $\bar{b}_i$, however, can be zero or nonzero. There is no condition imposed on the integers $\bar{N}$ and $\bar{M}$. The equation is said to have order $\max(\bar{N}, \bar{M})$. A delayed form can be easily obtained from an advanced form and vice versa. Thus both forms will be used in this chapter.

Once a mathematical description is available, the next step is to carry out analysis. There are two types of analyses: quantitative and qualitative. It turns out that the high-order difference equations in (9.1) and (9.2) are not suitable for either analysis. For quantitative analysis or computer computation, we will transform the high-order difference equations into a set of advanced-form first-order difference equations, called a DT state-space (ss) equation, and then carry out computation. This will be discussed in the next chapter. This chapter studies only qualitative analysis. We first introduce the z-transform to transform the difference equations into algebraic equations, called *transfer functions*. We then use them to develop general properties of DT systems.

All DT signals in this chapter are expressed in terms of time index or, equivalently, are assumed to have sampling period 1 unless stated otherwise.

9.2 z-TRANSFORM

Consider a DT signal $x[n]$. Its z-transform is defined as

$$X(z) := \mathcal{Z}[x[n]] := \sum_{n=0}^{\infty} x[n]z^{-n} \tag{9.3}$$

and its inverse z-transform is given by

$$x[n] := \mathcal{Z}^{-1}[X(z)] := \frac{1}{2\pi j} \oint X(z)z^{n-1}\,dz \tag{9.4}$$

where z is a complex variable, called the z-transform variable. The integration in (9.4) will be discussed shortly. The function $X(z)$ is called the z-transform of $x[n]$, and $x[n]$ is the inverse z-transform of $X(z)$. The z-transform as defined is often called the one-sided z-transform. If the lower summation limit 0 in (9.3) is replaced by $-\infty$, then (9.3) is called the two-sided z-transform. The two-sided z-transform is rarely used in practice, thus its discussion is omitted.

The z-transform of $x[n]$ is defined as a power series of z^{-1}. In this series, z^{-n} can be interpreted as indicating the nth sampling instant. In other words, z^0 indicates the initial time instant $n = 0$; z^{-1} indicates time instant $n = 1$ and so forth. In this sense, there is not much difference between a sequence and its z-transform. Before proceeding, we mention that multiplying a number by z^{-1} shifts the number to the right by one sample or delays the number by one sample. Thus z^{-1} is called the *unit delay element*. Multiplying a number by z shifts the number to the left by one sample or advances the number by one sample. Thus z is called the *unit advance element*.

The z-transforms we encounter in this chapter, however, can all be expressed in closed form as rational functions of z. They will be developed using the formula

$$1 + r + r^2 + r^3 + \cdots = \sum_{n=0}^{\infty} r^n = \frac{1}{1-r} \tag{9.5}$$

where r is a real or complex constant with magnitude less than 1, that is, $|r| < 1$. See (5.1) and (5.2). We give an example.

EXAMPLE 9.2.1

Consider the DT signal $x[n] = 1.3^n$, defined for all integers n. This signal grows exponentially to ∞ as $n \to \infty$. Its z-transform is, using (9.5),

$$\begin{aligned} X(z) &= \sum_{n=0}^{\infty} 1.3^n z^{-n} = \sum_{n=0}^{\infty} (1.3z^{-1})^n \\ &= \frac{1}{1-1.3z^{-1}} = \frac{z}{z-1.3} \end{aligned} \tag{9.6}$$

which requires $|1.3z^{-1}| < 1$ or $1.3 < |z|$. For example, if $z = 1$, the infinite sum in (9.6) diverges (approaches ∞) and does not equal $1/(1-1.3) = -10/3$. Thus (9.6)

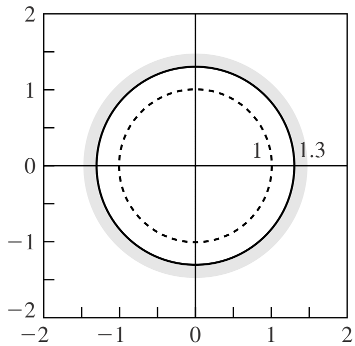


Figure 9.1 Region of convergence.

holds only if $|z| > 1.3$. The region outside the circle with radius 1.3 shown in Figure 9.1 is called the *region of convergence*. The region of convergence is needed in using (9.4) to compute the inverse z-transform of $X(z)$. The integration in (9.4) is to be carried out, in the counterclockwise direction, around a circular contour lying inside the region of convergence. In particular, we can select $z = ce^{j\omega}$ in (9.4). Then (9.4) becomes, using $dz = cje^{j\omega}\, d\omega$,

$$x[n] := \mathcal{Z}^{-1}[X(z)] = \frac{1}{2\pi j}\int_{\omega=0}^{2\pi} X(ce^{j\omega})(ce^{j\omega})^{n-1} cje^{j\omega}\, d\omega$$
$$= \frac{c^n}{2\pi}\int_{\omega=0}^{2\pi} X(ce^{j\omega})e^{jn\omega}\, d\omega \tag{9.7}$$

For our example, if we select $c = 2$, then the contour lies inside the region of convergence and (9.7) will yield

$$x[n] = \begin{cases} 1.3^n & \text{for } n \geq 0 \\ 0 & \text{for } n < 0 \end{cases}$$

a positive-time signal. Thus we often say that the z-transform is defined for positive-time signals. Because we encounter mostly positive-time signals in practice, the z-transform as introduced suits us well.

For the preceding example, if c is selected as 1, outside the region of convergence, then its inverse z-transform will yield a negative-time signal. More generally, let $X(z)$ be the z-transform of a positive-time signal. Then its inverse z-transform may yield a positive-time, negative-time, or two-sided signal depending on the c selected. In other words, without specifying the region of convergence, the inverse z-transform in (9.7) may not yield the original $x[n]$. Fortunately, in our application, we deal exclusively with positive-time signals and will not use (9.7) to compute the inverse z-transform. Furthermore, even though (9.6) was developed under the assumption $|z| > 1.3$, we may consider the z-transform $1/(1 - 1.3z^{-1}) = z/(z - 1.3)$ to be defined for all z except at $z = 1.3$. There is no need to pay any attention to its region of convergence. Thus we will not be concerned further with the region of convergence.

We give two more examples to conclude this section.

EXAMPLE 9.2.2

Find the z-transform of the impulse sequence $\delta[n-n_0]$ with $n_0 \geq 0$. The sequence is 1 at $n = n_0$ and 0 for $n \neq n_0$. By definition, we have

$$\mathcal{Z}[\delta[n-n_0]] = \sum_{n=0}^{\infty} \delta[n-n_0]z^{-n} = z^{-n_0} \tag{9.8}$$

Note that if $n_0 < 0$, then the sequence $\delta[n-n_0]$ is a negative-time sequence and its z-transform is 0.

EXAMPLE 9.2.3

Find the z-transform of $x[n] = 2 \cdot b^n$, where b is a real or complex constant. By definition, we have

$$X(z) = \mathcal{Z}[2 \cdot b^n] = \sum_{n=0}^{\infty} 2 \cdot b^n z^{-n} = 2\sum_{0}^{\infty} \left(bz^{-1}\right)^n = \frac{2}{1-bz^{-1}} = \frac{2z}{z-b} \tag{9.9}$$

We discuss two special cases of $\mathcal{Z}[b^n] = 1/(1-bz^{-1}) = z/(z-b)$. If $b = 0$, then $x[n] = \delta[n]$, the impulse sequence and its z-transform is $z/z = 1$. If $b = 1$, then $x[n] = q[n]$, a step sequence. Thus we have

$$\mathcal{Z}[q[n]] = \frac{1}{1-z^{-1}} = \frac{z}{z-1}$$

If $b = e^{j\omega_0}$, then

$$\mathcal{Z}[e^{jn\omega_0}] = \frac{1}{1-e^{j\omega_0}z^{-1}} = \frac{z}{z-e^{j\omega_0}}$$

EXERCISE 9.2.1

Find the z-transforms of $2\delta[n]$, 2, 2^n, and $(-2)^n$.

Answers

2, $2z/(z-1)$, $z/(z-2)$, $z/(z+2)$

The z-transforms of the DT signals in the preceding examples are all rational functions of z. This is, however, not always the case. In fact, given a DT signal, the following situations may occur:

1. Its z-transform does not exist. In other words, there is no region of convergence. The sequences e^{n^2} and e^{e^n}, for $n \geq 0$, are such examples. These two sequences are mathematically contrived and do not arise in practice. Thus it is fair to say that all DT signals encountered in practice are z-transformable.

2. Its z-transform exists but cannot be expressed in closed form. Most, if not all, randomly generated sequences of infinite length belong to this type.
3. Its z-transform exists and can be expressed in closed form but not as a rational function of z. For example, the z-transform of $x[0] = 0$ and $x[n] = 1/n$, for $n > 0$, is $-\ln(1 - z^{-1})$. It is an irrational function of z.
4. Its z-transform exists and is a rational function of z.

We study in this chapter only z-transforms that belong to the last class.

9.2.1 From Laplace Transform to z-Transform

The Laplace transform is defined for CT positive-time signals, whereas the z-transform is defined for DT positive-time sequences. If we apply the Laplace transform directly to a DT sequence $x[n]$, the result will be identically zero. Thus we must modify the sequence before proceeding. Let $x[n]$ be a positive sequence or

$$x[n] = \sum_{n=0}^{\infty} x[k]\delta[n-k]$$

where $\delta[n]$ is the impulse sequence defined in (1.17). As in (5.5) with $T = 1$, we modify it as

$$x_d(t) = \sum_{k=0}^{\infty} x[k]\delta(t-k) = \sum_{n=0}^{\infty} x[n]\delta(t-n) \tag{9.10}$$

where $\delta(t)$ is the impulse defined in Section 1.5. This is a CT representation of the DT sequence $x[n]$. Applying the Laplace transform to (9.10) yields

$$\begin{aligned} X_d(s) = \mathcal{L}[x_d(t)] &= \int_{t=0}^{\infty} \left(\sum_{n=0}^{\infty} x[n]\delta(t-n) \right) e^{-st}\, dt \\ &= \sum_{n=0}^{\infty} x[n] \left(\int_{t=0}^{\infty} \delta(t-n)e^{-st}\, dt \right) = \sum_{n=0}^{\infty} x[n]e^{-sn} \end{aligned} \tag{9.11}$$

where we have interchanged the order of summation and integration and used the sifting property of impulses. We see that if we replace e^s by z, then (9.11) becomes the z-transform of $x[n]$ defined in (9.3). Thus we have

$$\mathcal{Z}[x[n]] = \mathcal{L}[x_d(t)]|_{z=e^s} \tag{9.12}$$

This establishes the relationship between the Laplace transform and z-transform.

Let us discuss the implication of $z = e^s$. We show that it maps the $j\omega$-axis on the s-plane into the unit circle on the z-plane, and it maps the open left half s-plane into the interior of the unit circle on the z-plane as shown in Figure 9.2. Indeed, because

$$z = e^s = e^{\sigma + j\omega} \tag{9.13}$$

where $s = \sigma + j\omega$, if $\sigma = 0$ (the imaginary axis of the s-plane), then $|z| = |e^{j\omega}| = 1$ (the unit circle on the z-plane). Thus the $j\omega$-axis is mapped into the unit circle. Note that the mapping is not one-to-one. If fact, all $s = 0, \pm j2\pi, \pm j4\pi, \ldots$ are mapped into $z = 1$. If $\sigma < 0$

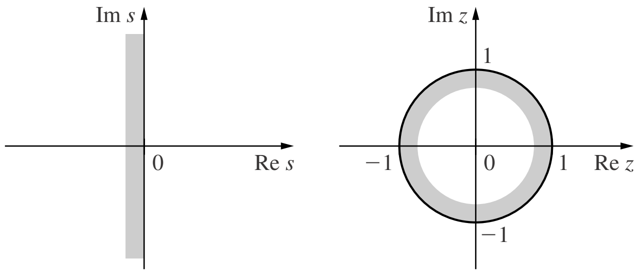


Figure 9.2 Mapping between the s-plane and z-plane.

(left half s-plane), then

$$|z| = |e^{\sigma}||e^{j\omega}| = |e^{\sigma}| < 1$$

(interior of the unit circle on the z-plane). This establishes the mapping shown in Figure 9.2.

9.3 DT TRANSFER FUNCTIONS

Consider a DT LTI system with input $u[n]$ and output $y[n]$. If it is initially relaxed at $n = 0$ or, equivalently, $u[n] = y[n] = 0$, for $n < 0$, then the input and output can be described by the convolution

$$y[n] = \sum_{k=0}^{\infty} h[n-k]u[k] \tag{9.14}$$

where $h[n]$ is the impulse response of the system. If the system is causal, then $h[n] = 0$, for $n < 0$. Thus all sequences in (9.14) are positive-time. Applying the z-transform to $y[n]$ yields

$$Y(z) = \mathcal{Z}[y[n]] = \sum_{n=0}^{\infty} y[n]z^{-n} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\infty} h[n-k]u[k] \right) z^{-(n-k)} z^{-k}$$

Interchanging the order of summations, introducing a new index $\bar{n} := n - k$, and using the causality property $h[\bar{n}] = 0$, for $\bar{n} < 0$, we obtain

$$\begin{aligned} Y(z) &= \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} h[n-k]z^{-(n-k)} \right) u[k]z^{-k} = \sum_{k=0}^{\infty} \left(\sum_{\bar{n}=-k}^{\infty} h[\bar{n}]z^{-\bar{n}} \right) u[k]z^{-k} \\ &= \left(\sum_{\bar{n}=0}^{\infty} h[\bar{n}]z^{-\bar{n}} \right) \sum_{k=0}^{\infty} u[k]z^{-k} \end{aligned}$$

or

$$Y(z) = H(z)U(z) \tag{9.15}$$

where $Y(z)$ and $U(z)$ are the z-transform of the input and output and

$$H(z) = \mathcal{Z}[h[n]] = \sum_{n=0}^{\infty} h[n]z^{-n} \tag{9.16}$$

is called the *(discrete-time or digital) transfer function*. It is the z-transform of the impulse response. Thus the z-transform transforms the convolution in (9.14) into the multiplication in (9.15). Because the convolution in (9.14) describes only zero-state responses, so does the transfer function. Thus whenever we use (9.15), the system is implicitly assumed to be initially relaxed.

Using (9.15), we can also define the transfer function as

$$H(z) = \left.\frac{\mathcal{Z}[\text{output}]}{\mathcal{Z}[\text{input}]}\right|_{\text{all initial conditions zero}} \tag{9.17}$$

This definition includes as a special case the previous definition. Indeed, if the input is an impulse sequence, then its z-transform is 1 and the transfer function simply equals the z-transform of the corresponding output which, by definition, is the impulse response. Because the transfer function can be obtained from any input–output pair, this establishes formally the assertion in Chapter 2 that the characteristics of an LTI system can be determined from any single pair of input and output.

Every DT LTI system that is initially relaxed can be described by a DT transfer function. If the system is distributed, then its transfer function either is an irrational function of z or cannot be expressed in closed form. If the system is lumped, then its transfer function is a rational function of z as we derive in the next subsection.

9.3.1 From Difference Equations to Rational Transfer Functions

Every DT LTI lumped system can be described by, in addition to a convolution, a differential equation such as the one in (9.1) or (9.2). In order to apply the z-transform to (9.1) and (9.2), we first develop some formulas. Let $X(z)$ be the z-transform of $x[n]$, where $x[n]$ is a two-sided sequence. Then we have

$$\mathcal{Z}[x[n-1]] = z^{-1}X(z) + x[-1] \tag{9.18}$$

$$\mathcal{Z}[x[n-2]] = z^{-2}X(z) + x[-1]z^{-1} + z[-2] \tag{9.19}$$

$$\mathcal{Z}[x[n+1]] = z[X(z) - x[0]] \tag{9.20}$$

$$\mathcal{Z}[x[n+2]] = z^{2}[X(z) - x[0] - x[1]z^{-1}] \tag{9.21}$$

Indeed, by definition, we have

$$X(z) = \sum_{n=0}^{\infty} x[n]z^{-n}$$

and

$$X_{-1}(z) := \mathcal{Z}[x[n-1]] = \sum_{n=0}^{\infty} x[n-1]z^{-n} = x[-1] + \sum_{n=1}^{\infty} x[n-1]z^{-(n-1)-1}$$

$$= x[-1] + z^{-1}\sum_{\bar{n}=0}^{\infty} x[\bar{n}]z^{-\bar{n}} = x[-1] + z^{-1}X(z)$$

where we have introduced $\bar{n} = n - 1$. This shows (9.18). This can be explained as follows. The z-transform is defined only for the positive-time part of $x[n]$. The sequence $x[n-1]$ shifts $x[n]$ one sample to the right. Thus the z-transform of $x[n-1]$ is $z^{-1}X(z)$. Furthermore, $x[-1]$ is the entry at $n = 0$ in $x[n-1]$ but is not included in $X(z)$. Thus the z-transform of $x[n-1]$ is as shown in (9.18). Using the same argument, we can establish (9.19). In general, for any positive integer k, we have

$$\mathcal{Z}[x[n-k]] = z^{-k}X(z) + \sum_{l=1}^{k} x[-l]z^{l-k} \tag{9.22}$$

Note that if $x[n]$ is positive time, then $x[n] = 0$, for $n < 0$. In this case, (9.22) reduces to $\mathcal{Z}[x[n-k]] = z^{-k}X(z)$, for any positive integer k. This is consistent with the fact that multiplying by z^{-1} delays one sample. Thus, multiplying by z^{-k} delays k samples.

To show (9.20), we compute

$$\begin{aligned} X_1(z) := \mathcal{Z}[x[n+1]] &= \sum_{n=0}^{\infty} x[n+1]z^{-n} = \sum_{n=0}^{\infty} x[n+1]z^{-(n+1)+1} \\ &= z\sum_{\bar{n}=1}^{\infty} x[\bar{n}]z^{-\bar{n}} = z\left[\sum_{\bar{n}=0}^{\infty} x[\bar{n}]z^{-\bar{n}} - x[0]\right] \\ &= z[X(z) - x[0]] \end{aligned}$$

This can be explained as follows. The sequence $x[n+1]$ shifts $x[n]$ one sample to the left. Thus $x[0]$ is located at $n = -1$ in the shifted sequence $x[n+1]$ and will not appear in the z-transform of $x[n+1]$. Thus we subtract $x[0]$ from $X(z)$ and then multiply it by z to yield (9.20). Using the same argument, we can establish (9.21). In general, for any positive integer k, we have

$$\mathcal{Z}[x[n+k]] = z^k\left[X(z) - \sum_{l=0}^{k-1} x[l]z^{-l}\right] \tag{9.23}$$

We use examples to show the use of the formulas.

EXAMPLE 9.3.1 (Delayed Form)

Consider a DT system described by the second-order difference equation

$$2y[n] + 3y[n-1] + 5y[n-2] = u[n-1] - 2u[n-2] \tag{9.24}$$

Applying the z-transform yields

$$\begin{aligned} &2Y(z) + 3\left(z^{-1}Y(z) + y[-1]\right) + 5\left(z^{-2}Y(z) + y[-1]z^{-1} + y[-2]\right) \\ &\quad = z^{-1}U(z) + u[-1] - 2\left(z^{-2}U(z) + u[-1]z^{-1} + u[-2]\right) \end{aligned}$$

which can be combined as

$$(2 + 3z^{-1} + 5z^{-2})Y(z) = (z^{-1} - 2z^{-2})U(z) + a + bz^{-1}$$

with $a := -3y[-1] - 5y[-2] + u[-1] - 2u[-2]$ and $b := -5y[-1] - 2u[-1]$. Thus we have

$$Y(z) = \frac{z^{-1} - 2z^{-2}}{2 + 3z^{-1} + 5z^{-2}}U(z) + \frac{a + bz^{-1}}{2 + 3z^{-1} + 5z^{-2}} \tag{9.25}$$

We see that the output in the transform domain consists of two parts: One part is excited by the input, and the other part is excited by the initial conditions. The former is the zero-state or forced response, and the latter is the zero-input or natural response. This confirms the fact that the response of every linear system can always be decomposed as the zero-state response and zero-input response.

If all initial conditions are zero or $y[n] = u[n] = 0$, for $n < 0$, then (9.25) reduces to

$$Y(z) = \frac{z^{-1} - 2z^{-2}}{2 + 3z^{-1} + 5z^{-2}} U(z)$$

Thus the transfer function of the system described by (9.24) is

$$H(z) = \frac{z^{-1} - 2z^{-2}}{2 + 3z^{-1} + 5z^{-2}} \tag{9.26}$$

EXAMPLE 9.3.2 (Advanced Form)

Consider a DT system described by the second-order difference equation

$$2y[n+2] + 3y[n+1] + 5y[n] = u[n+1] - 2u[n] \tag{9.27}$$

which is obtained from (9.24) by adding 2 to all indices. Applying the z-transform to (9.27), we obtain

$$2z^2(Y(z) - y[1]z^{-1} - y[0]) + 3z\,(Y(z) - y[0]) + 5Y(z) = z\,(U(z) - u[0]) - 2U(z)$$

which can be combined as

$$(2z^2 + 3z + 5)Y(z) = (z - 2)U(z) + (2y[1] + 3y[0] - u[0])z + 2y[0]z^2 \tag{9.28}$$

This equation consists of $y[0]$, $y[1]$, and $u[0]$ which are not initial conditions of (9.24) or (9.27). For the second-order difference equation in (9.27), the initial conditions are $y[n]$ and $u[n]$, for $n = -1$ and -2. Now we use (9.27) to express $y[1]$ and $y[0]$ in terms of the initial conditions. For $n = -2$ and -1, (9.27) becomes, respectively,

$$2y[0] + 3y[-1] + 5y[-2] = u[-1] - 2u[-2]$$

and

$$2y[1] + 3y[0] + 5y[-1] = u[0] - 2u[-1]$$

Substituting these into (9.28) yields

$$(2z^2 + 3z + 5)Y(z) = (z - 2)U(z) + bz + az^2$$

where b and a are defined in the preceding example. Thus we have

$$Y(z) = \frac{z - 2}{2z^2 + 3z + 5} U(z) + \frac{az^2 + bz}{2z^2 + 3z + 5} \tag{9.29}$$

This equation is identical to (9.25) if its numerator and denominator are multiplied by z^{-2}. If all initial conditions are zero, then the equation reduces to

$$Y(z) = \frac{z-2}{2z^2+3z+5}U(z)$$

Thus the transfer function is

$$H(z) = \frac{z-2}{2z^2+3z+5} \tag{9.30}$$

If we multiply its numerator and denominator by z^{-2}, then (9.30) becomes (9.26). Thus (9.26) and (9.30) are identical.

Let us compute the transfer function of the difference equations in (9.1) and (9.2). If a system is initially relaxed or, equivalently, $y[n] = u[n] = 0$, for $n < 0$, we can use, as demonstrated in the preceding example, $\mathcal{Z}[x[n+k]] = z^k X(z)$, instead of (9.23), in computing the transfer function. Applying $\mathcal{Z}[x[n+k]] = z^k X(z)$ to (9.1) yields

$$a_1 z^N Y(z) + a_2 z^{N-1} Y(z) + \cdots + a_{N+1} Y(z) = b_1 z^M U(z) + b_2 z^{M-1} U(z) + \cdots + b_{M+1} U(z)$$

or

$$(a_1 z^N + a_2 z^{N-1} + \cdots + a_N z + a_{N+1}) Y(z) = (b_1 z^M + b_2 z^{M-1} + \cdots + b_M z + b_{M+1}) U(z)$$

Thus the transfer function of (9.1) is

$$H(z) = \frac{b_1 z^M + b_2 z^{M-1} + \cdots + b_M z + b_{M+1}}{a_1 z^N + a_2 z^{N-1} + \cdots + a_N z + a_{N+1}} \tag{9.31}$$

Applying $\mathcal{Z}[x[n-k]] = z^{-k} X(z)$ to (9.2) yields

$$\bar{a}_1 Y(z) + \bar{a}_2 z^{-1} Y(z) + \cdots + \bar{a}_{\bar{N}+1} z^{-\bar{N}} Y(z) = \bar{b}_1 U(z) + \bar{b}_2 z^{-1} U(z) + \cdots + \bar{b}_{\bar{M}+1} z^{-\bar{M}} U(z)$$

or

$$(\bar{a}_1 + \bar{a}_2 z^{-1} + \cdots + \bar{a}_{\bar{N}+1} z^{-\bar{N}}) Y(z) = (\bar{b}_1 + \bar{b}_2 z^{-1} + \cdots + \bar{b}_{\bar{M}+1} z^{-\bar{M}}) U(z)$$

Thus the transfer function of (9.2) is

$$\bar{H}(z) = \frac{\bar{b}_1 + \bar{b}_2 z^{-1} + \cdots + \bar{b}_{\bar{M}+1} z^{-\bar{M}}}{\bar{a}_1 + \bar{a}_2 z^{-1} + \cdots + \bar{a}_{\bar{N}+1} z^{-\bar{N}}} \tag{9.32}$$

In conclusion, delayed- and advanced-form difference equations can be easily transformed to transfer functions. We call (9.31) a *positive-power transfer function* and call (9.32) a *negative-power transfer function*. Either form can be easily transformed into the other form.

Unlike CT systems where we use exclusively positive-power transfer functions, we may encounter both forms in DT systems. Thus we discuss further the two forms. For the transfer function in (9.31), we always assume $a_1 \neq 0$, $b_1 \neq 0$, and at least one of b_{M+1} and a_{N+1} to be different from zero. If both b_{M+1} and a_{N+1} are zero, then the numerator and denominator of (9.31) have the common factor z and both M and N can be reduced by 1. As in the CT case, the transfer function in (9.31) is defined to be improper if $M > N$, proper if $N \geq M$, strictly proper if $N > M$, and biproper if $N = M$. If the denominator and numerator of $H(z)$ has no common factors or are coprime, then $H(z)$ has degree $\max(N, M)$. If $H(z)$ is proper, then its degree is N, the degree of its denominator.

For the negative-power transfer function in (9.32), we always assume $\bar{a}_{\bar{N}+1} \neq 0$, $\bar{b}_{\bar{M}+1} \neq 0$, and at least one of $\bar{a}_1$ and $\bar{b}_1$ to be different from zero. We now discuss its properness condition. Because the properness of a rational function has been defined for positive-power form, we must translate the condition into (9.32). Recall from Section 6.3.3 that the properness of $H(z)$ can also be determined from $H(\infty)$. If $z = \infty$, then $z^{-1} = 0$ and $H(\infty)$ in (9.32) reduces to $\bar{b}_1/\bar{a}_1$. Thus (9.32) is improper if $\bar{a}_1 = 0$ and $\bar{b}_1 \neq 0$, proper if $\bar{a}_1 \neq 0$, biproper if $\bar{a}_1 \neq 0$ and $\bar{b}_1 \neq 0$, and strictly proper if $\bar{a}_1 \neq 0$ and $\bar{b}_1 = 0$. These conditions are independent of $\bar{N}$ and $\bar{M}$. The integer $\bar{N}$ can be larger than, equal to, or smaller than $\bar{M}$. Thus the properness condition for negative-power transfer functions is less natural. This is one advantage of using positive-power transfer functions. If the denominator and numerator of $\bar{H}(z)$ in (9.32) has no common factors or are coprime, then $\bar{H}(z)$ has degree $\max(\bar{N}, \bar{M})$. If $\bar{H}(z)$ is proper, its degree is still $\max(\bar{N}, \bar{M})$, not necessarily the degree of its denominator.

We study only causal systems. The condition for the positive-power transfer function in (9.31) to describe a causal system is that (9.31) is proper or $N \geq M$. The condition for the negative-power transfer function in (9.32) to describe a causal system is $\bar{a}_1 \neq 0$.

The response of every linear system can be decomposed into zero-state and zero-input responses. As in the CT case, zero-input responses are generally parts of zero-state responses. Thus in discussing general properties of systems, we study only zero-state responses that can be described by transfer functions. Thus in the remainder of this chapter we study mostly transfer functions.

EXERCISE 9.3.1

Find the transfer function of the advanced-form difference equation

$$y[n+3] + 3y[n+1] + 2y[n] = 2u[n+2] - 0.5u[n]$$

Is it proper? What is its degree?

Answers

$(2z^2 - 0.5)/(z^3 + 3z + 2)$. Yes. 3.

EXERCISE 9.3.2

Find the negative-power transfer function of the difference equation

$$y[n] + y[n-2] = u[n] + u[n-1] - 2u[n-3]$$

Is it proper? What is its degree? Transform it into a positive-power transfer function. Is it proper? What is its degree?

Answers

$(1 + z^{-1} - 2z^{-3})/(1 + z^{-2})$. Yes. 3. $(z^3 + z^2 - 1)/(z^3 + z)$. Yes. 3.

It is simple to obtain transfer functions from difference equations. The converse is also simple.

EXERCISE 9.3.3

Transform the transfer function

$$H(z) = \frac{Y(z)}{U(z)} = \frac{2z^2 - 3z + 1}{3z^2 + 4z + 2}$$

into the delayed-form and advanced-form difference equations.

Answers

$3y[n] + 4y[n-1] + 2y[n-2] = 2u[n] - 3u[n-1] + u[n-2]$
$3y[n+2] + 4y[n+1] + 2y[n] = 2u[n+2] - 3u[n+1] + u[n]$

9.3.2 Poles and Zeros

This section introduces the concepts of poles and zeros for DT transfer functions. All discussion in Section 6.3.4 is directly applicable to positive-power transfer functions, thus the discussion here will be brief.

Consider a proper rational transfer function $H(z) = N(z)/D(z)$, where $D(z)$ and $N(z)$ are polynomials of z with real coefficients. A finite real or complex number λ is called a zero of $H(z)$ if $H(\lambda) = 0$. It is a pole if $|H(\lambda)| = \infty$. If $N(z)$ and $D(z)$ are coprime, then all roots of $N(z)$ are the zeros of $H(z)$ and all roots of $D(z)$ are the poles of $H(z)$. For example, consider

$$H(z) = \frac{8z^3 - 24z - 16}{2z^5 + 20z^4 + 98z^3 + 268z^2 + 376z + 208} \tag{9.33}$$

To find its zeros and poles, we can apply the MATLAB function roots to its numerator and denominator. We can also apply the function tf2zp as

```
n=[8 0 -24 -16];d=[2 20 98 268 376 208];
[z,p,k]=tf2zp(n,d)
```

which yields z=[-1 -1 2]; p=[-2 -2 -2 -2-3j -2+3j]; k=4. Thus $H(z)$ can be expressed in the zero/pole/gain form as

$$H(z) = \frac{4(z+1)^2(z-2)}{(z+2)^3(z+2+3j)(z+2-3j)} \tag{9.34}$$

It has simple zero at 2 and simple poles at $-2 \pm j3$. It has a repeated zero at -1 with multiplicity 2 and has a repeated pole at -2 with multiplicity 3. If all coefficients of $H(z)$ are real, complex conjugate poles and zeros must appear in pairs. Note that $H(\infty) = 0$. But $z = \infty$ is not a zero because we consider only finite zeros. We see that the discussion for CT transfer function $H(s)$ can be directly applied to DT transfer function $H(z)$ without any modification.

We next consider negative-power transfer functions such as

$$H(z) = \frac{4z^{-1} + 6z^{-2}}{2 + 5z^{-1} + 3z^{-2}} = \frac{4z^{-1}(1 + 1.5z^{-1})}{(2 + z^{-1})(1 + 3z^{-1})} \tag{9.35}$$

Setting $1 + 1.5z^{-1} = 0$, $z^{-1} = -1/1.5$, or $z = -1.5$, we have $H(-1.5) = 0$. Thus -1.5 is a zero of (9.35). Likewise, we can verify that $-1/2$ and -3 are poles. Note that setting $z^{-1} = 0$ or $z = 1/0 = \infty$, we have $H(\infty) = 0$. But $z = \infty$ is not a zero because we consider only finite zeros. Thus zeros and poles can also be obtained from negative-power transfer functions.

Although the poles and zeros of (9.35) can be obtained directly as discussed above, it is simpler to transform (9.35) into positive power form by multiplying its numerator and denominator by z^2 to yield

$$H(z) = \frac{4z + 6}{2z^2 + 5z + 3} = \frac{4(z + 1.5)}{(2z + 1)(z + 3)} = \frac{2(z + 1.5)}{(z + 0.5)(z + 3)}$$

Then it has zero at -1.5, and it has poles at -0.5 and -3.

9.3.3 Transfer Functions of FIR and IIR Systems

As discussed in Section 3.2.1, DT LTIL systems can be classified as FIR (finite impulse response) and IIR (infinite impulse response). We discuss their transfer functions. Note that all CT LTIL systems, excluding memoryless systems, are IIR. Thus we do not classify CT LTIL systems as FIR or IIR.

Consider a DT Nth-order FIR system with impulse response $h[n]$ for $n = 0, 1, \ldots, N$, and $h[N] \neq 0$. Note that it has length $N + 1$. Its transfer function is

$$\begin{aligned} H(z) &= h[0] + h[1]z^{-1} + \cdots + h[N]z^{-N} \\ &= \frac{h[0]z^N + h[1]z^{N-1} + \cdots + h[N]}{z^N} \end{aligned}$$

and has degree N. All its N poles are located at $z = 0$. Thus every FIR filter has poles only at the origin of the z-plane. Conversely, if a transfer function has poles only at $z = 0$, then it describes an FIR filter.

EXAMPLE 9.3.3

Consider the 30-day moving average filter discussed in Example 3.6.1. It has impulse response $h[n] = 1/30$, for $n = 0 : 29$. Thus its transfer function is

$$H(z) = \frac{1}{30}\sum_{n=0}^{29} z^{-n} = \frac{\sum_{n=0}^{29} z^n}{30z^{29}}$$

It is a biproper rational function with degree 29.

Using (5.1), we can write the transfer function as

$$H(z) = \frac{Y(z)}{U(z)} = \frac{1 - z^{30}}{30z^{29}(1 - z)} = \frac{0.033(z^{30} - 1)}{z^{29}(z - 1)} \tag{9.36}$$

Because its numerator and denominator have the common factor $z - 1$, $z = 1$ is not a pole of the transfer function. Thus the assertion that an FIR system has poles only at $z = 0$ is still valid.

From (9.36), we have

$$z^{30}Y(z) - z^{29}Y(z) = 0.033[z^{30}U(z) - U(z)]$$

In finding its difference equation, we can use $z^k X(z) = \mathcal{Z}[x[n+k]]$ instead of (9.23). Thus its time-domain description is

$$y[n+30] - y[n+29] = 0.033(u[n+30] - u[n])$$

It can be transformed into the delayed form as

$$y[n] - y[n-1] = 0.033(u[n] - u[n-30])$$

which is the recursive difference equation in (3.27).

While every DT FIR system has poles only at $z = 0$, every DT IIR system has at least one pole other than $z = 0$. The reason is as follows. The inverse z-transform of $z/(z-a)$ is a^n, for $n \geq 0$, which has, if $a \neq 0$, infinitely many nonzero entries. If $H(z)$ has one or more poles other than $z = 0$, then its impulse response (inverse z-transform) has infinitely many nonzero entries. Thus $H(z)$ describes an IIR filter.

9.4 PROPERTIES OF z-TRANSFORM

We discuss some properties of the z-transform and develop a z-transform table.

Linearity The z-transform is a linear operator. That is, if $X_1(z) = \mathcal{Z}[x_1[n]]$ and $X_2(z) = \mathcal{Z}[x_2[n]]$, then for any constants α_1 and α_2, we have

$$\mathcal{Z}[\alpha_1 x_1[n] + \alpha_2 x_2[n]] = \alpha_1 \mathcal{Z}[x_1[n]] + \alpha_2 \mathcal{Z}[x_2[n]] = \alpha_1 X_1(z) + \alpha_2 X_2(z)$$

This can directly be verified from the definition of the z-transform.

EXAMPLE 9.4.1

We use $\mathcal{Z}[b^n] = z/(z-b)$ to compute the z-transform of $\sin \omega_0 n$. Using Euler's identity $\sin \omega_0 n = (e^{j\omega_0 n} - e^{-j\omega_0 n})/2j$, we have

$$\begin{aligned}
\mathcal{Z}[\sin \omega_0 n] &= \frac{1}{2j}[\mathcal{Z}[e^{j\omega_0 n}] - \mathcal{Z}[e^{-j\omega_0 n}]] = \frac{1}{2j}\left[\frac{z}{z - e^{j\omega_0}} - \frac{z}{z - e^{-j\omega_0}}\right] \\
&= \frac{(e^{j\omega_0} - e^{-j\omega_0})z}{2j[z^2 - (e^{j\omega_0} + e^{-j\omega_0})z + 1]} = \frac{(\sin \omega_0)z}{z^2 - 2(\cos \omega_0)z + 1} \\
&= \frac{(\sin \omega_0)z^{-1}}{1 - 2(\cos \omega_0)z^{-1} + z^{-2}}
\end{aligned}$$

EXERCISE 9.4.1

Verify

$$\mathcal{Z}[\cos \omega_0 t] = \frac{z(z - \cos \omega_0)}{z^2 - 2(\cos \omega_0)z + 1} = \frac{1 - (\cos \omega_0)z^{-1}}{1 - 2(\cos \omega_0)z^{-1} + z^{-2}}$$

EXERCISE 9.4.2

Compute the z-transforms of

(a) $e^{2n} - (-1.5)^n$

(b) $2 + 0.5 \sin 3n$

Answers

(a) $8.89z/[(z - 7.39)(z + 1.5)]$

(b) $z(2z^2 + 4.03z + 1.93)/[(z - 1)(z^2 + 1.98z + 1)]$

Multiplication by b^n in the Time Domain If $X(z) = \mathcal{Z}[x[n]]$, then

$$\mathcal{Z}[b^n x[n]] = X(z/b)$$

By definition, we have

$$\mathcal{Z}[b^n x[n]] = \sum_{0}^{\infty} b^n x[n] z^{-n} = \sum_{0}^{\infty} x[n](z/b)^{-n} = X(z/b)$$

This establishes the formula. Using the formula, we have

$$\mathcal{Z}[b^n \sin \omega_0 n] = \frac{(\sin \omega_0)(z/b)}{(z/b)^2 - 2(\cos \omega_0)(z/b) + 1} = \frac{b(\sin \omega_0)z}{z^2 - 2b(\cos \omega_0)z + b^2}$$

and

$$\mathcal{Z}[(\cos \omega_0 n)x[n]] = \mathcal{Z}\left[\frac{e^{j\omega_0 n} + e^{-j\omega_0 n}}{2} x[n]\right] = 0.5[X(e^{-j\omega_0}z) + X(e^{j\omega_0}z)]$$

EXERCISE 9.4.3

Verify

$$\mathcal{Z}[b^n \cos \omega_0 n] = \frac{(z - b\cos \omega_0)z}{z^2 - 2b(\cos \omega_0)z + b^2}$$

Multiplication by n in the Time Domain If $X(z) = \mathcal{Z}[x[n]]$, then

$$\mathcal{Z}[nx[n]] = -z\frac{dX(z)}{dz}$$

The differentiation of (9.3) with respect to z yields

$$\frac{d}{dz}X(z) = \sum_{n=0}^{\infty}(-n)x[n]z^{-n-1}$$

which becomes, after multiplying its both sides by $-z$,

$$-z\frac{d}{dz}X(z) = \sum_{n=0}^{\infty} nx[n]z^{-n}$$

Its right-hand side, by definition, is the z-transform of $nx[n]$. This establishes the formula.

Using the formula and $\mathcal{Z}[b^n] = z/(z-b)$, we can establish

$$\mathcal{Z}[nb^n] = -z\frac{d}{dz}\left[\frac{z}{z-b}\right] = -z\frac{(z-b)-z}{(z-b)^2} = \frac{bz}{(z-b)^2} = \frac{bz^{-1}}{(1-bz^{-1})^2}$$

EXERCISE 9.4.4

Verify

$$\mathcal{Z}[n^2b^n] = \frac{b(z+b)z}{(z-b)^3}$$

To conclude this section, we list in Table 9.1 some z-transform pairs. Unlike the Laplace transform where we use exclusively positive-power form, we may encounter both negative-power and positive-power forms. Thus we list both forms in the table.

The z-transforms listed in Table 9.1 are all proper rational functions. If $x[0] \neq 0$, such as $\delta[n]$, 1, b^n, and $b^n \cos\omega_0 n$, then its z-transform is biproper. If $x[0] = 0$, such as nb^n, n^2b^n, and $b^n \sin\omega_0 n$, then its z-transform is strictly proper. The Laplace transforms in Table 6.1 are all strictly proper except the one of $\delta(t)$.

TABLE 9.1 z-Transform Pairs

$x[n],\ n \geq 0$	$X(z)$	$X(z)$
$\delta[n]$	1	1
$\delta[n-n_0]$	z^{-n_0}	z^{-n_0}
1 or $q[n]$	$\dfrac{z}{z-1}$	$\dfrac{1}{1-z^{-1}}$
b^n	$\dfrac{z}{z-b}$	$\dfrac{1}{1-bz^{-1}}$
nb^n	$\dfrac{bz}{(z-b)^2}$	$\dfrac{bz^{-1}}{(1-bz^{-1})^2}$
n^2b^n	$\dfrac{(z+b)bz}{(z-b)^3}$	$\dfrac{(1+bz^{-1})bz^{-1}}{(1-bz^{-1})^3}$
$b^n \sin\omega_0 n$	$\dfrac{(\sin\omega_0)bz}{z^2-2(\cos\omega_0)bz+b^2}$	$\dfrac{(\sin\omega_0)bz^{-1}}{1-2(\cos\omega_0)bz^{-1}+b^2z^{-2}}$
$b^n \cos\omega_0 n$	$\dfrac{(z-b\cos\omega_0)z}{z^2-2(\cos\omega_0)bz+b^2}$	$\dfrac{(1-(\cos\omega_0)bz^{-1})}{1-2(\cos\omega_0)bz^{-1}+b^2z^{-2}}$

9.5 INVERSE z-TRANSFORM

This section discusses the inverse z-transform, that is, to compute the time sequence of a proper rational z-transform $X(z)$. As in the CT case, the z-transform is not used in computer computation. It is used mainly in developing general properties of DT systems. Thus our discussion concentrates on developing general forms of inverse z-transforms. The basic procedure is to express $X(z)$ as a sum of terms whose inverse z-transforms are available in a table such as Table 9.1, and then use the table to find the inverse z-transform. We can use either positive-power or negative-power z-transforms. We discuss mainly the former because we are more familiar with positive-power polynomials than negative-power polynomials. We use examples to illustrate the procedure. The reader is assumed to be familiar with all the discussion in Sections 6.5 through 6.5.3.

EXAMPLE 9.5.1

Consider the proper rational z-transform

$$X(z) = \frac{2z^2 + 10}{(z+1)(z-2)} \tag{9.37}$$

Instead of $X(z)$, we expand $X(z)/z$ into a partial fraction expansion. The reasons for this expansion will be given in the next example. We expand $X(z)/z$ as

$$\bar{X}(z) := \frac{X(z)}{z} = \frac{2z^2 + 10}{z(z+1)(z-2)} = k_0 + \frac{k_1}{z} + \frac{k_2}{z+1} + \frac{k_3}{z-2} \tag{9.38}$$

In this expansion, the procedure discussed in Section 6.5 can be directly applied. Thus we have

$$\begin{aligned}
k_0 &= \bar{X}(\infty) = 0 \\
k_1 &= \left.\frac{2z^2 + 10}{(z+1)(z-2)}\right|_{z=0} = \frac{10}{-2} = -5 \\
k_2 &= \left.\frac{2z^2 + 10}{z(z-2)}\right|_{z=-1} = \frac{12}{3} = 4 \\
k_3 &= \left.\frac{2z^2 + 10}{z(z+1)}\right|_{z=2} = \frac{18}{2 \times 3} = 3
\end{aligned}$$

and

$$\frac{X(z)}{z} = -5\frac{1}{z} + 4\frac{1}{z+1} + 3\frac{1}{z-2}$$

This becomes, after multiplying the whole equation by z,

$$X(z) = -5 + 4\frac{z}{z+1} + 3\frac{z}{z-2} \tag{9.39}$$

Using Table 9.1, we have

$$x[n] = -5\delta[n] + 4 \times (-1)^n + 3 \times 2^n \tag{9.40}$$

for $n \geq 0$. We see that if we expand $X(z)/z$, then the procedure in Section 6.5 can be directly applied. After the expansion, we then multiply the whole equation by z. Then every term of the resulting equation is in Table 9.1, and the inverse z-transform of $X(z)$ can be read out from the table.

EXAMPLE 9.5.2

Consider the z-transform in Example 9.5.1. If we expand it directly as

$$X(z) = \frac{2z^2 + 10}{(z+1)(z-2)} = k_0 + \frac{k_1}{z+1} + \frac{k_2}{z-2}$$

then the k_i can be easily computed as in the preceding example. However the last two terms $1/(z+1)$ and $1/(z-2)$ are not in Table 9.1. Thus we need additional manipulation to find its inverse z-transform. Thus the expansion is not desirable. If we expand it as

$$X(z) = k_0 + \frac{k_1 z}{z+1} + \frac{k_2 z}{z-2} \tag{9.41}$$

then every term is in Table 9.1, and its inverse z-transform is

$$x[n] = k_0\delta[n] + k_1(-1)^n + k_2 2^n$$

for $n \geq 0$. Thus the remaining problem is to compute k_i. Their computation however requires some modification from the procedure in Section 6.5. If we multiply (9.41) by $(z+1)/z$ and then set $z = -1$, we obtain

$$k_1 = X(z)\frac{z+1}{z}\bigg|_{z=-1} = \frac{2z^2+10}{z(z-2)}\bigg|_{z=-1} = \frac{12}{3} = 4$$

If we multiply (9.41) by $(z-2)/z$ and then set $z = 2$, we obtain

$$k_2 = X(z)\frac{z-2}{z}\bigg|_{z=2} = \frac{2z^2+10}{z(z+1)}\bigg|_{z=2} = \frac{18}{2 \times 3} = 3$$

To find k_0, we set $z = \infty$ in (9.41) to yield

$$X(\infty) = 2 = k_0 + k_1 + k_2$$

which implies $k_0 = 2 - 3 - 4 = -5$. The result is the same as the one in (9.40).

From the preceding two examples, we see that there are many ways to carry out partial fraction expansions. However, if we expand $X(z)/z$, then all discussion in Section 6.5 is directly applicable.

EXAMPLE 9.5.3

Compute the step response of a DT system with transfer function $H(z) = 10/[(z-1)(z+2)]$. If $u[n] = q[n]$, then $U(z) = z/(z-1)$. Thus the step response in the z-transform domain is

$$Y(z) = H(z)U(z) = \frac{10}{(z-1)(z+2)}\frac{z}{z-1} = \frac{10z}{(z-1)^2(z+2)} \tag{9.42}$$

Instead of expanding $Y(z)$, we expand $Y(z)/z$ as

$$\bar{Y}(z) := \frac{Y(z)}{z} = \frac{10z}{(z-1)^2(z+2)} \times \frac{1}{z} = \frac{10}{(z-1)^2(z+2)}$$
$$= k_0 + r_1\frac{1}{z-1} + r_2\frac{1}{(z-1)^2} + k_1\frac{1}{z+2} \tag{9.43}$$

The parameters k_0, r_2, and k_1 can be computed using simple formulas as

$$k_0 = \bar{Y}(\infty) = 0$$
$$r_2 = \left.\frac{10}{z+2}\right|_{z=1} = \frac{10}{3}$$
$$k_1 = \left.\frac{10}{(z-1)^2}\right|_{z=-2} = \frac{10}{(-3)^2} = \frac{10}{9}$$

See Section 6.5.2. The remaining parameter r_1 can be computed from (9.43) by selecting an arbitrary z other than the poles of $\bar{Y}(z)$. Let us select $z = 0$. Then (9.43) becomes

$$\frac{10}{2} = k_0 - r_1 + r_2 + k_1\frac{1}{2}$$

which implies $r_1 = -5 + (10/3) + (10/18) = -10/9$. Thus we have

$$\frac{Y(z)}{z} = \frac{-10}{9}\frac{1}{z-1} + \frac{10}{3}\frac{1}{(z-1)^2} + \frac{10}{9}\frac{1}{z+2}$$

and

$$Y(z) = \frac{-10}{9}\frac{z}{z-1} + \frac{10}{3}\frac{z}{(z-1)^2} + \frac{10}{9}\frac{z}{z+2}$$

Thus the inverse z-transform of $Y(z)$, using Table 9.1, is

$$y[n] = \frac{-10}{9} + \frac{10}{3}n + \frac{10}{9}(-2)^n$$

for $n \geq 0$. This is the step response of the system.

Finding inverse z-transform involving complex poles is more complex. We discuss in the following only the general form of its inverse z-transform. Consider the z-transform

$$X(z) = \frac{2z+5}{z^2+1.2z+0.85} \tag{9.44}$$

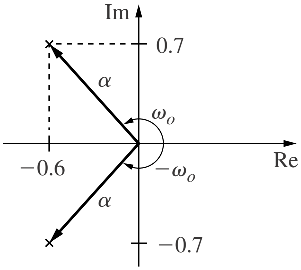


Figure 9.3 Complex poles expressed in polar form.

We first compute its poles or the roots of its denominator as

$$z^2 + 1.2z + 0.85 = (z + 0.6)^2 - 0.36 + 0.85 = (z + 0.6)^2 + 0.49$$
$$= (z + 0.6)^2 - (j0.7)^2 = (z + 0.6 - j0.7)(z + 0.6 + j0.7)$$

Thus the transfer function has a pair of complex conjugate poles at $-0.6 \pm j0.7$ as shown in Figure 9.3. Unlike the CT case where the real part of complex-conjugate poles yields the envelope of its inverse Laplace transform and the imaginary part yields the frequency of oscillation, the real and imaginary parts of complex-conjugate poles in the DT system have no physical meaning. Now we express the poles in polar form as

$$-0.6 \pm j0.7 = 0.92e^{\pm j2.28} =: \alpha e^{\pm j\omega_0}$$

Then the denominator of (9.44) can be expressed as

$$z^2 + 1.2z + 0.85 = (z - (-0.6 + j0.7))(z - (-0.6 - j0.7)) = (z - 0.92e^{j2.28})(z - 0.92e^{-j2.28})$$

Let us expand $X(z)/z$ as

$$\frac{X(z)}{z} = k_0 + \frac{k_1}{z} + \frac{\bar{k}_2}{z - 0.92e^{j2.28}} + \frac{\bar{k}_2^*}{z - 0.92e^{-j2.28}}$$

where $k_0 = 0$, $k_1 = 5/0.85$, and $\bar{k}_2^*$ is the complex conjugate of $\bar{k}_2$. Then we have

$$X(z) = k_1 + \frac{\bar{k}_2 z}{z - 0.92e^{j2.28}} + \frac{\bar{k}_2^* z}{z - 0.92e^{-j2.28}}$$

Its inverse z-transform is, writing $\bar{k}_2 = \beta e^{j\gamma}$ and using Table 9.1,

$$\begin{aligned}
x[n] &= k_1\delta[n] + \beta e^{j\gamma}(0.92e^{j2.28})^n + \beta e^{-j\gamma}(0.92e^{-j2.28})^n \\
&= k_1\delta[n] + \beta(0.92)^n e^{j(2.28n+\gamma)} + \beta(0.92)^n e^{-j(2.28n+\gamma)} \\
&= k_1\delta[n] + 2\beta(0.92)^n \cos(2.28n + \gamma) \\
&= k_1\delta[n] + k_2(0.92)^n \cos(2.28n + k_3) \\
&= k_1\delta[n] + k_2(0.92)^n \sin(2.28n + \bar{k}_3)
\end{aligned}$$

for some real constants k_1, k_2, k_3, and $\bar{k}_3$. In conclusion, a pair of complex conjugate poles $\alpha e^{\pm j\omega_0}$ will generate a response of the form

$$k_2\alpha^n \cos(\omega_0 n + k_3) \quad \text{or} \quad k_2\alpha^n \sin(\omega_0 n + \bar{k}_3)$$

Its envelope is determined by the magnitude of the poles and its frequency by the phase.

We mention that the inverse z-transform of $X(z) = N(z)/D(z)$ can also be obtained by direct division of $N(z)$ by $D(z)$. The division can be used to compute the first few terms of the inverse z-transform. However, it is difficult to obtain its general form. Thus its discussion is omitted.

9.6 SIGNIFICANCE OF POLES AND ZEROS

Just as the CT case, poles and zeros of a DT transfer function are plotted on the complex z-plane with crosses and circles. However, unlike the CT case where the s-plane is divided into the left half-plane, the right half-plane, and the $j\omega$-axis, the complex z-plane is divided into the unit circle, its interior, and its exterior. This is expected in view of the mapping discussed in Figure 9.2.

We discuss responses of some poles. We consider first real poles. Let α be real and positive. Then the inverse z-transform of $z/(z-\alpha)$ is α^n. The time response α^n grows unbounded if $\alpha > 1$, is a step sequence if $\alpha = 1$, and vanishes if $\alpha < 1$. If α is a repeated pole such as $z/(z-\alpha)^2$, then its inverse z-transform is $n\alpha^n$. It grows unbounded if $\alpha \geq 1$. If $\alpha < 1$, $n\alpha^n$ approaches 0 because

$$\lim_{n\to\infty} \frac{(n+1)\alpha^{n+1}}{n\alpha^n} = \alpha < 1$$

following Cauchy's ratio test. See Reference 20. Similarly, we can show that $n^k\alpha^n$ approaches 0, as $n \to \infty$, for any $\alpha < 1$ and any positive integer k.

Next we consider the pole at $\alpha e^{j\pi} = -\alpha$, where α is real and positive. The pole is located on the negative real axis. The inverse z-transform of $z/(z+\alpha)$ is $(-\alpha)^n$. Thus its value changes sign at every sampling instant, and the time sequence contains the highest-frequency component π. Recall that we have assumed $T = 1$ in our discussion. Thus the Nyquist frequency range is $(-\pi/T, \pi/T] = (-\pi, \pi]$, and π is the highest frequency. The time sequence $(-\alpha)^n$ grows alternatively to infinity if $\alpha > 1$, is 1 or -1 if $\alpha = 1$, and vanishes alternatively if $\alpha < 1$ as shown in Figure 9.4(a). If the real and negative pole is repeated with multiplicity 2, then its time response $n(-\alpha)^n$ grows unbounded if $\alpha \geq 1$ and vanishes if $\alpha < 1$.

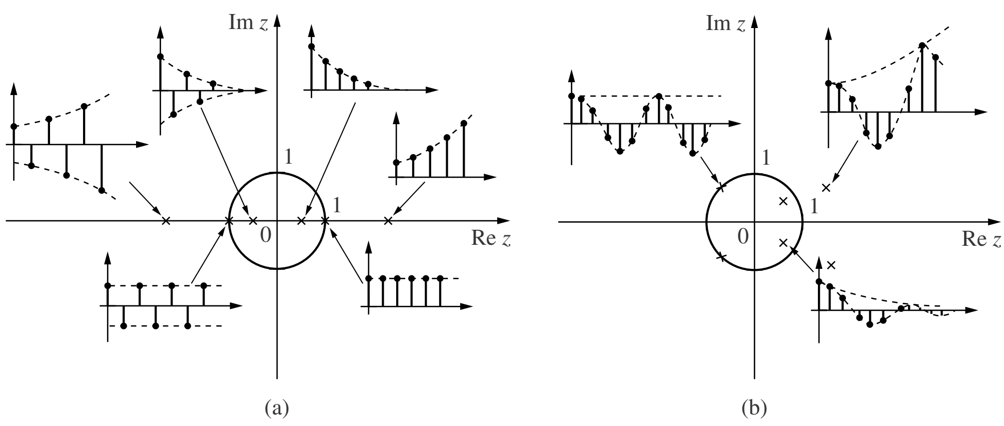


Figure 9.4 (**a**) Responses of real poles. (**b**) Responses of complex poles.

Let us consider $\alpha e^{j\omega_0}$, where α is real and positive and $0 < \omega_0 < \pi$. The complex pole together with its complex conjugate $\alpha e^{-j\omega_0}$ generate the response $\alpha^n \sin \omega_0 n$ as shown in Figure 9.4(b). The response grows unbounded if $\alpha > 1$, is a sinusoidal sequence with frequency ω_0 if $\alpha = 1$, and vanishes if $\alpha < 1$. If the pair of complex conjugate poles is repeated, then its time response grows unbounded if $\alpha \geq 1$ and vanishes if $\alpha < 1$. In conclusion, the time response of a pole, simple or repeated, real or complex, approaches zero if and only if the pole lies inside the unit circle.

We summarize the preceding discussion in the following:

Pole Location	Response as $n \to \infty$
Interior of the unit circle, simple or repeated	0
Exterior of the unit circle, simple or repeated	∞ or $-\infty$
Unit circle, simple	Constant or sustained oscillation
Unit circle, repeated	∞ or $-\infty$

Next we consider the following DT transfer functions:

$$H_1(z) = \frac{1.9712}{(z-0.6)^2(z-0.9e^{j2.354})(z-0.9e^{-j2.354})}$$

$$= \frac{1.9712}{z^4 + 1.27z^3 + 0.45z^2 - 0.4572z - 0.2916} \tag{9.45}$$

$$H_2(z) = \frac{19.712(z-0.9)}{z^4 + 1.27z^3 + 0.45z^2 - 0.4572z - 0.2916} \tag{9.46}$$

$$H_3(z) = \frac{0.219(z^2+4z+4)}{z^4 + 1.27z^3 + 0.45z^2 - 0.4572z - 0.2916} \tag{9.47}$$

$$H_4(z) = \frac{1.9712(z^3+2z^2+3z-5)}{z^4 + 1.27z^3 + 0.45z^2 - 0.4572z - 0.2916} \tag{9.48}$$

They all have the same set of poles and the property $H_i(1) = 1$. Even though they have different zeros, their step responses can all be expressed as

$$Y(z) = H_i(z)U(z) = H_i(z)\frac{z}{z-1}$$

which can be expanded as

$$\frac{Y(z)}{z} = \frac{k_1}{z-0.6} + \frac{k_2}{(z-0.6)^2} + \frac{\bar{k}_3}{z-0.9e^{j2.354}} + \frac{\bar{k}_3^*}{z-0.9e^{-j2.354}} + \frac{k_5}{z-1}$$

or

$$Y(z) = \frac{k_1 z}{z-0.6} + \frac{k_2 z}{(z-0.6)^2} + \frac{\bar{k}_3 z}{z-0.9e^{j2.354}} + \frac{\bar{k}_3^* z}{z-0.9e^{-j2.354}} + \frac{k_5 z}{z-1}$$

with $k_5 = H_i(1) = 1$. Thus their step responses in the time domain are all of the form

$$y[n] = k_1 0.6^n + k_2 n(0.6)^n + k_3 0.9^n \cos(2.354n + k_4) + k_5 \tag{9.49}$$

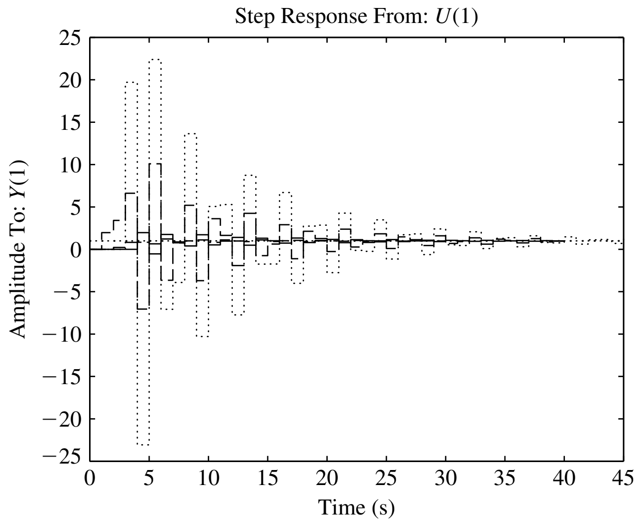


Figure 9.5 Step responses of the transfer functions in (9.45) through (9.48).

for $n \geq 0$. This form is determined solely by the poles of $H_i(z)$ and $U(z)$. The transfer function $H_i(z)$ has a repeated real pole at 0.6, which yields the response $k_1 0.6^n + k_2 n 0.6^n$, and a pair of complex conjugate poles $0.9e^{\pm j2.354}$, which yields $k_3 0.9^n \cos(2.354n + k_4)$. The z-transform of the step input is $z/(z-1)$. Its pole at $z = 1$ yields the response $H_i(1) \times 1^n = H_i(1)$. We plot in Figure 9.5 the step responses of the transfer functions in (9.45) through (9.48). They are obtained in MATLAB by typing

```
n1=1.9712;d=[1 1.27 0.45 -0.4572 -0.2916];
n2=19.712*[1 -0.9];n3=0.219*[1 4 4];
n4=1.9712*[1 2 3 -5];
dstep(n1,d)
hold on
dstep(n2,d,':')
dstep(n3,d,'--')
dstep(n4,d,'-.')
```

We see that even though their step responses are all of the form in (9.49), the responses right after the application of the step input are all different. This is due to different set of k_i. In conclusion, poles dictate the general form of responses; zeros affect only the parameters k_i. Thus we conclude that zeros play a lesser role than poles in determining responses of systems.

To conclude this section, we mention that the response generated by dstep(ni,d) is not computed using the transfer function ni/d for the same reasons discussed in Section 6.5.4. As in the CT case, the transfer function is first transformed into an ss equation, and then we use the ss equation to carry out the computation. This will be discussed in the next chapter. We also mention that dstep(ni,d) can be replaced by

```
cat=tf(ni,d,T);
step(cat)
```

The first line uses transfer function (tf) to define the DT system with sampling period T. Without the third argument T, the system is a CT system. Therefore, it is important to include the third argument which can be selected as any positive number, in particular, $T = 1$. The second line computes the step response of the system. If we type step(cat, tf), then it generates the step response up to the sampling instant equal to or less than t_f. Without the second argument t_f, the function will automatically select a t_f.

9.7 STABILITY

This section introduces the concept of stability for DT systems. If a DT system is not stable, its response excited by any input generally will grow unbounded. Thus every DT system designed to process signals must be stable. Let us give a formal definition.

DEFINITION 9.1 A DT system is BIBO (bounded-input bounded-output) stable or, simply, stable, if *every* bounded input sequence excites a bounded output sequence. Otherwise, the system is said to be unstable.

A signal is bounded if it does not grow to ∞ or $-\infty$. In other words, a signal $u[n]$ is bounded if there exists a constant M_1 such that $|u[n]| \leq M_1 < \infty$ for all n. As in the CT case, Definition 9.1 cannot be used to conclude the stability of a system because there are infinitely many bounded inputs to be checked. However, if we can find a bounded input that excites an unbounded output, then we can conclude that the system is not stable. In fact, the stability of a DT system can be determined from its mathematical descriptions without applying any input. In other words, stability is a property of a system and is independent of applied inputs. The output of a stable system excited by any bounded input must be bounded; its output excited by an unbounded input is generally unbounded.

THEOREM 9.1 A DT LTI system with impulse response $h[n]$ is BIBO stable if and only if $h[n]$ is absolutely summable in $[0, \infty)$, that is,

$$\sum_{n=0}^{\infty} |h[n]| \leq M < \infty$$

for some constant M.

Proof We first show that the system is BIBO stable under the condition. Indeed, the input $u[n]$ and output $y[n]$ of the system are related by

$$y[n] = \sum_{k=0}^{n} h[k]u[n-k]$$

If $u[k]$ is bounded or $|u[n]| \le M_1$ for all $n \ge 0$, then we have

$$\begin{aligned}|y[n]| &= \left|\sum_{k=0}^{n} h[k]u[n-k]\right| \le \sum_{k=0}^{n} |h[k]||u[n-k]| \\ &\le M_1 \sum_{k=0}^{n} |h[k]| \le M_1 \sum_{k=0}^{\infty} |h[k]| \le M_1 M\end{aligned}$$

for all $n \ge 0$. This shows that if $h[n]$ is absolutely summable, then the system is BIBO stable.

Next we show that if $h[n]$ is not absolutely summable, then there exists a bounded input that will excite an unbounded output. If $h[n]$ is not absolutely summable, then for any arbitrarily large M_2, there exists an n_1 such that

$$\sum_{k=0}^{n_1} |h[k]| \ge M_2$$

Let us select an input as, for $k = 0 : n_1$,

$$u[n_1 - k] = \begin{cases} 1 & \text{if } h[k] \ge 0 \\ -1 & \text{if } h[k] < 0 \end{cases}$$

For this bounded input, the output $y[n]$ at $n = n_1$ is

$$y[n_1] = \sum_{k=0}^{n_1} h[k]u[n_1 - k] = \sum_{k=0}^{n_1} |h[k]| \ge M_2$$

This shows that if $h[n]$ is not absolutely summable, then there exists a bounded input that will excite an output with an arbitrarily large magnitude. Thus the system is not stable. This establishes the theorem.

We mention that a necessary condition for $h[n]$ to be absolutely summable is that $h[n] \to 0$ as $n \to \infty$.[1] However, the condition is not sufficient. For example, consider $h[0] = 0$ and $h[n] = 1/n$, for $n \ge 1$. It approaches 0 as $n \to \infty$. Let us compute

$$\begin{aligned}S := \sum_{n=0}^{\infty} |h[n]| = \sum_{n=1}^{\infty} \frac{1}{n} &= (1) + \left(\frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) \\ &+ \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \left(\frac{1}{9} + \cdots + \frac{1}{16}\right) + \cdots\end{aligned}$$

The sum inside every pair of parentheses is 1/2 or larger, thus we have

$$S \ge 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots = \infty$$

[1]In the CT case, $h(t) \to 0$ as $t \to \infty$ is not a necessary condition for $h(t)$ to be absolutely integrable. See the footnote on page 224. See also the footnote on page 154.

and the sequence is not absolutely summable. Thus a DT system with impulse sequence $1/n$, for $n \geq 1$, is not stable even though its impulse response approaches zero. Note that the z-transform of $1/n$ is an irrational function of z, thus the DT system is not a lumped system.

The sequence $1/n$ is not absolutely summable because it does not approach zero fast enough. If a sequence approaches 0 sufficiently fast, such as b^n or $n^k b^n$, for $|b| < 1$ and any positive integer k, then it is absolutely summable. Indeed, for $|b| < 1$, we have

$$\sum_{n=0}^{\infty} |b|^n = \frac{1}{1 - |b|} < \infty$$

and

$$\sum_{n=0}^{\infty} n^k |b|^n < \infty$$

following Cauchy's ratio test

$$\lim_{n\to\infty} \frac{(n+1)^k |b|^{n+1}}{n^k |b|^n} = |b| < 1$$

See Reference 20. This is so because as $n \to \infty$, $n^k |b|^n$ decreases in the rate of $|b|^n$, which decreases *exponentially* to zero for $|b| < 1$. See the discussion leading to (1.22). It approaches zero much faster than $1/n$.

THEOREM 9.2 A DT LTI lumped system with proper rational transfer function $H(z)$ is stable if and only if every pole of $H(z)$ has a magnitude less than 1 or, equivalently, all poles of $H(z)$ lie inside the unit circle on the z-plane.

If $H(z)$ has one or more poles lying outside the unit circle, then its impulse response grows unbounded and is not absolutely summable. If it has poles on the unit circle, then its impulse response will not approach 0 and is not absolutely summable. Thus if $H(z)$ has one or more poles on or outside the unit circle, then the system is not stable. On the other hand, if every pole of $H(z)$ has a magnitude less than 1, then its time response approaches zero exponentially and is absolutely summable. Thus we conclude that the system is stable. Note that the stability of a system is independent of the zeros of its transfer function. The zeros can be located outside, on, or inside the unit circle.

COROLLARY 9.2 A DT LTI lumped system with impulse response $h[n]$ is BIBO stable if and only if $h[n] \to 0$ as $n \to \infty$.

The difference between Theorem 9.1 and Corollary 9.2 is that the former is applicable to distributed and lumped systems whereas the latter is applicable only to lumped systems. For example, the system with impulse response $h[n] = 1/n$ is a distributed system, thus Theorem 9.2

and its corollary are not applicable. We study only LTI lumped systems, thus we use mostly Theorem 9.2 and its corollary to check stability of systems.

EXAMPLE 9.7.1

Consider a DT system with transfer function

$$H(z) = \frac{(z+2)(z-10)}{(z-0.9)(z+0.95)(z+0.9+j0.7)(z+0.9-j0.7)}$$

Its two real poles 0.9 and -0.95 have magnitudes less than 1. We compute the magnitude of the complex poles:

$$\sqrt{(0.9)^2+(0.7)^2} = \sqrt{0.81+0.49} = \sqrt{1.3} = 1.14$$

It is larger than 1. Thus the system is not stable.

EXAMPLE 9.7.2

Consider an Nth-order FIR system. Such a filter has impulse response $h[n]$, for $n = 0 : N$, and $h[n] = 0$, for $n > N$. Clearly, we have

$$\sum_{n=0}^{\infty} |h[n]| = \sum_{n=0}^{N} |h[n]| < \infty$$

Thus every FIR filter is stable. This can also be concluded from its transfer function

$$H(z) = \sum_{n=0}^{N} h[n]z^{-n} = \frac{h[0]z^N + h[1]z^{N-1} + \cdots + h[N-1]z + h[N]}{z^N}$$

All its N poles are located at $z = 0$. They all lie inside the unit circle. Thus it is stable.

EXERCISE 9.7.1

Determine the stability of the following systems:

(a) $\dfrac{z+1}{(z-0.6)^2(z+0.8+j0.6)(z+0.8-j0.6)}$

(b) $\dfrac{3z-6}{(z-2)(z+0.2)(z-0.6+j0.7)(z-0.6-j0.7)}$

(c) $\dfrac{z-10}{z^2(z+0.95)}$

Answers

No. Yes. Yes.

If a DT system is known to be LTI and lumped, then its stability can easily be checked by measurement. We apply $u[0] = 1$ and $u[n] = 0$ for $n > 0$ (an impulse sequence), then the output is the impulse response. Thus the system is stable if and only if the response approaches zero as $n \to \infty$.

9.7.1 Jury Test

Consider a system with proper rational transfer function $H(z) = N(z)/D(z)$. We assume that $N(s)$ and $D(s)$ are coprime. Then the poles of $H(z)$ are the roots of $D(z)$. If $D(z)$ has degree three or higher, computing its roots by hand is not simple. This, however, can be carried out using a computer. For example, consider

$$D(z) = 2z^3 - 0.2z^2 - 0.24z - 0.08 \tag{9.50}$$

Typing

```
d=[2 -0.2 -0.24 -0.08];
roots(d)
```

in MATLAB yields -0.5, $-0.2 \pm j0.2$. It has a real root and one pair of complex conjugate roots. They all have magnitudes less than 1. Thus any transfer function with (9.50) as its denominator is BIBO stable. Of course, if we transform $H(z)$ into the zero/pole/gain form by calling the MATLAB function tf2zp, we can also determine the stability of $H(z)$ from its pole locations.

We discuss in the following a method of checking the stability of $H(z)$ without computing its poles or, equivalently, the roots of its denominator.[2] We first define a polynomial to be *DT stable* if all its roots have magnitudes less than 1. The Jury test to be introduced checks whether or not a polynomial is DT stable without computing its roots. We use the following polynomial of degree 5 to illustrate the procedure:

$$D(z) = a_0z^5 + a_1z^4 + a_2z^3 + a_3z^2 + a_4z + a_5 \qquad \text{with } a_0 > 0 \tag{9.51}$$

We call a_0 the leading coefficient. If the leading coefficient is negative, we apply the procedure to $-D(z)$. Because $D(z)$ and $-D(z)$ have the same set of roots, if $-D(z)$ is DT stable, so is $D(z)$. The polynomial $D(z)$ has degree 5 and six coefficients a_i, $i = 0:5$. We form Table 9.2, called the Jury table. The first row is simply the coefficients of $D(z)$ arranged in the descending power of z. The second row is the reversal of the first row. We compute $k_1 = a_5/a_0$, the ratio of the last entries of the first two rows. The first b_i row is obtained by subtracting from the first a_i row the product of the second a_i row and k_1. Note that the last entry of the first b_i row is automatically zero and is discarded in the subsequent discussion. We then reverse the order of b_i to form the second b_i row and compute $k_2 = b_4/b_0$. The first b_i row subtracting the product of the second b_i row and k_2 yields the first c_i row. We repeat the process until the table is completed

[2]This subsection may be skipped without loss of continuity.

TABLE 9.2 The Jury Table

a_0	a_1	a_2	a_3	a_4	a_5	
a_5	a_4	a_3	a_2	a_1	a_0	$k_1 = a_5/a_0$
b_0	b_1	b_2	b_3	b_4	0	(1st a_i row) $-k_1$(2nd a_i row)
b_4	b_3	b_2	b_1	b_0		$k_2 = b_4/b_0$
c_0	c_1	c_2	c_3	0		(1st b_i row) $-k_2$(2nd b_i row)
c_3	c_2	c_1	c_0			$k_3 = c_3/c_0$
d_0	d_1	d_2	0			(1st c_i row) $-k_3$(2nd c_i row)
d_2	d_1	d_0				$k_4 = d_2/d_0$
e_0	e_1	0				(1st d_i row) $-k_4$(2nd d_i row)
e_1	e_0					$k_5 = e_1/e_0$
f_0	0					(1st e_i row) $-k_5$(2nd e_i row)

as shown. We call b_0, c_0, d_0, e_0, and f_0 the *subsequent leading coefficients*. If $D(z)$ has degree N, then the table has N subsequent leading coefficients.

THEOREM 9.3 A polynomial with a positive leading coefficient is DT stable if and only if every subsequent leading coefficient is positive. If any subsequent leading coefficient is 0 or negative, then the polynomial is not DT stable.

The proof of this theorem is beyond the scope of this text. We discuss only its employment.

EXAMPLE 9.7.3

Consider

$$D(z) = z^3 - 2z^2 - 0.8$$

Note that the polynomial has a missing term. We form

1	−2	0	−0.8	
−0.8	0	−2	1	$k_1 = -0.8/1 = -0.8$
0.36	−2	−1.6	0	
−1.6	−2	0.36		$k_2 = -1.6/0.36 = -4.44$
−6.74				

A negative subsequent leading coefficient appears. Thus the polynomial is not DT stable.

EXAMPLE 9.7.4

Consider

$$D(z) = 2z^3 - 0.2z^2 - 0.24z - 0.08$$

We form

2	−0.2	−0.24	−0.08	
−0.08	−0.24	−0.2	2	$k_1 = -0.08/2 = -0.04$
1.9968	−0.2096	−0.248	0	
−0.248	−0.2096	1.9968		$k_2 = -0.248/1.9968 = -0.124$
1.966	−0.236	0		
−0.236	1.966			$k_3 = -0.236/1.966 = -0.12$
1.938	0			

The three subsequent leading coefficients are all positive. Thus the polynomial is DT stable.

EXERCISE 9.7.2

Check the DT stability of

(a) $2z^3 - 0.2z^2 - 0.24z$

(b) $6z^4 + 4z^3 - 1.5z^2 - z + 0.1$

Answers

(a) Yes.

(b) Yes.

We defined in Section 6.7.1 a polynomial to be stable if all its roots have negative real parts. A stable polynomial cannot have missing terms or negative coefficients. This is not the case for DT stable polynomials (all its roots have magnitudes less than 1). For example, the two polynomials in the preceding Exercise have missing term or negative coefficients and are still DT stable. Thus the conditions for a polynomial to be stable or DT stable are different and have nothing to do with each other.

EXERCISE 9.7.3

Check the stability and DT stability of the polynomials:

(a) $z^3 + 2.1z^2 + 1.93z + 0.765$

(b) $z^3 + 0.3z^2 - 0.22z - 0.765$

(c) $z^3 + 3.2z^2 + 3.25z + 1.7$

(d) $z^2 - 2z + 2$

Answers

(a) Yes, yes.

(b) No, yes.

(c) Yes, no.

(d) No, no.

9.8 FREQUENCY RESPONSES

Before discussing the implication of stability, we introduce the concept of frequency responses. Consider a DT transfer function $H(z)$. Its values along the unit circle on the z-plane—that is, $H(e^{j\omega})$, for all ω—is called the *frequency response*. In general, $H(e^{j\omega})$ is complex-valued and can be expressed in polar form as

$$H(e^{j\omega}) = A(\omega)e^{j\theta(\omega)} \tag{9.52}$$

where $A(\omega)$ and $\theta(\omega)$ are real-valued functions of ω and $A(\omega) \geq 0$. We call $A(\omega)$ the *magnitude response,* and we call $\theta(\omega)$ the *phase response*. We first give an example.

EXAMPLE 9.8.1

Consider the DT transfer function

$$H(z) = \frac{z+1}{10z-8} \tag{9.53}$$

Its frequency response is

$$H(e^{j\omega}) = \frac{e^{j\omega}+1}{10e^{j\omega}-8} \tag{9.54}$$

We compute

$$\omega = 0: \qquad H(1) = \frac{1+1}{10-8} = 1 = 1 \cdot e^{j0}$$

$$\omega = \pi/2 = 1.57: \qquad H(j1) = \frac{j1+1}{j10-8} = \frac{1.4e^{j\pi/4}}{12.8e^{j2.245}} = 0.11e^{-j1.46}$$

$$\omega = \pi = 3.14: \qquad H(-1) = \frac{-1+1}{-10-8} = 0$$

Other than these ω, its computation is quite complex. Fortunately, the MATLAB function freqz, where the last character z stands for the z-transform, carries out the computation. To compute the frequency response of (9.53) from $\omega = -10$ to 10 with increment 0.01, we type

```
n=[1 1];d=[10 -8];
w=-10:0.01:10;
H=freqz(n,d,w);
plot(w,abs(H),w,angle(H),':')
```

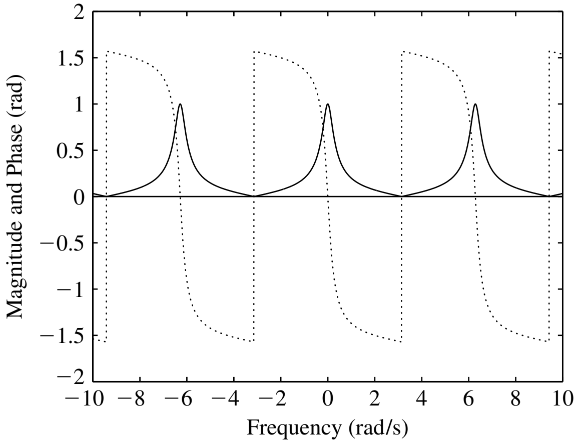


Figure 9.6 Magnitude response (solid line) and phase response (dotted line) of (9.53).

The result is shown in Figure 9.6. We see that the frequency response is periodic with period 2π. This is expected because we have assumed $T = 1$ and the Nyquist frequency range is $[-\pi, \pi)$ or $(-\pi, \pi]$. It also follows from

$$e^{j\omega} = e^{j(\omega+2\pi)}$$

If all coefficients of $H(z)$ are real, then we have $H(e^{j\omega}) = [H(e^{-j\omega})]^*$, which imples

$$A(\omega)e^{j\theta(\omega)} = [A(-\omega)e^{j\theta(-\omega)}]^* = A(-\omega)e^{-j\theta(-\omega)}$$

Thus we have

$$A(\omega) = A(-\omega) \qquad \text{(even)} \tag{9.55}$$

and

$$\theta(\omega) = -\theta(-\omega) \qquad \text{(odd)} \tag{9.56}$$

In other words, if $H(z)$ has only real coefficients, then its magnitude response is even and its phase response is odd. Thus we often plot frequency responses only in the positive frequency range $[0, \pi]$. If we do not specify ω in using freqz(n,d), then freqz selects automatically 512 points in $[0, \pi]$.[3] Thus typing

```
n=[1 1];d=[10 -8];
[H,w]=freqz(n,d);
subplot(1,2,1)
plot(w,abs(H))
subplot(1,2,2)
plot(w,angle(H))
```

[3]The function freqz is based on FFT, thus its number of frequencies is selected to be a power of 2 or $2^9 = 512$. However, the function freqs for computing frequency responses of CT systems has nothing to do with FFT, and its default uses 200 frequencies.

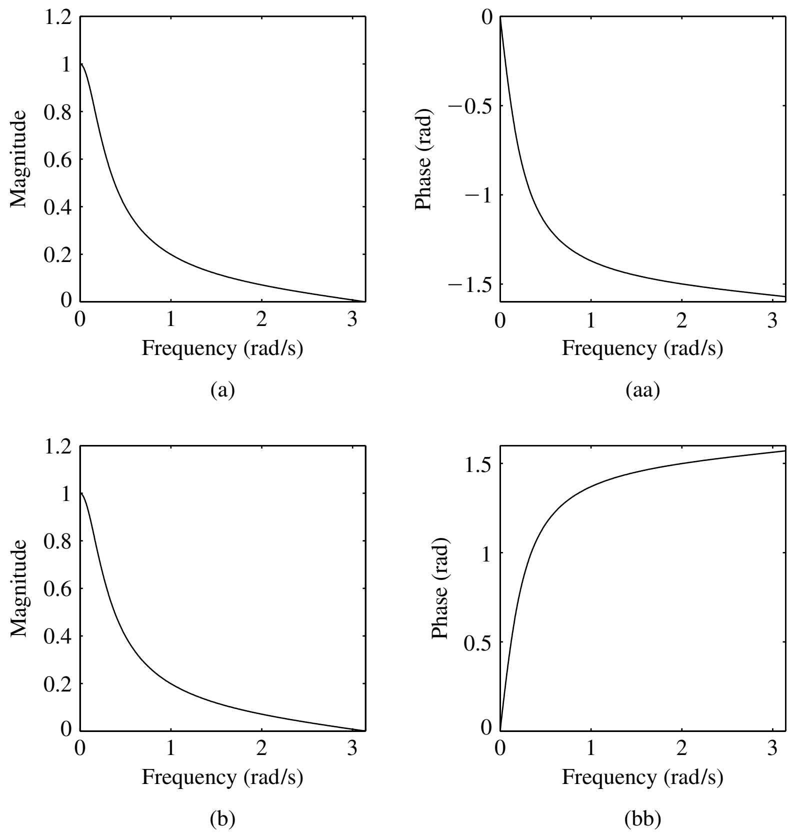


Figure 9.7 (**a**) Magnitude response of $(z+1)/(10z-8)$. (**aa**) Its phase response. (**b**) Magnitude response of $(z+1)/(-8z+10)$. (**bb**) Its phase response.

generates the magnitude and phase responses of (9.53) in Figures 9.7(a) and 9.7(aa). Note that we did not specify the frequencies in the preceding program.

Next we consider the DT transfer function

$$H_1(z) = \frac{z+1}{-8z+10} \tag{9.57}$$

Its pole is $10/8 = 1.25$ and is outside the unit circle. Thus the system is not stable. Replacing `d=[10 -8]` by `d=[-8 10]`, the preceding program generates the magnitude and phase responses of (9.57) in Figures 9.7(b) and 9.7(bb). We see that the magnitude responses of (9.53) and (9.57) are identical, but their phase responses are different. See Problem 9.12.

We now discuss the implication of stability and physical meaning of frequency responses. Consider a DT system with transfer function $H(z)$. Let us apply to it the input $u[n] = ae^{j\omega_0 n}$. If the system is not stable, then the output will grow unbounded or maintain an oscillation with a frequency different from ω_0. See Problem 9.13. However, if the system is stable, then the output approaches, as we show next, $aH(e^{j\omega_0})e^{j\omega_0 n}$ as $n \to \infty$.

The z-transform of $u[n] = ae^{j\omega_0 n}$ is $az/(z - e^{j\omega_0})$. Thus the output of $H(z)$ is

$$Y(z) = H(z)U(z) = H(z)\frac{az}{z - e^{j\omega_0}}$$

We expand $Y(z)/z$ as

$$\frac{Y(z)}{z} = \frac{k_1}{z - e^{j\omega_0}} + \text{terms due to poles of } H(z)$$

with

$$k_1 = aH(z)|_{z=e^{j\omega_0}} = aH(e^{j\omega_0})$$

Thus we have

$$Y(z) = aH(e^{j\omega_0})\frac{z}{z - e^{j\omega_0}} + \text{terms due to poles of } H(z)$$

which implies

$$y[n] = aH(e^{j\omega_0})e^{j\omega_0 n} + \text{responses due to poles of } H(z)$$

If $H(z)$ is stable, then all responses due to its poles approach 0 as $n \to \infty$. Thus we have

$$y_{ss}[n] := \lim_{n\to\infty} y[n] = aH(e^{j\omega_0})e^{j\omega_0 n} \tag{9.58}$$

or, substituting (9.52),

$$\begin{aligned} y_{ss}[n] &= aA(\omega_0)e^{j\theta(\omega_0)}e^{j\omega_0 n} = aA(\omega_0)e^{j[\omega_0 n + \theta(\omega_0)]} \\ &= aA(\omega_0)\left[\cos(\omega_0 n + \theta(\omega_0)) + j\sin(\omega_0 n + \theta(\omega_0))\right] \end{aligned} \tag{9.59}$$

They are the DT counterparts of (6.50) and (6.51). We call $y_{ss}[n]$ the *steady-state response*. We list some special cases of (9.58) or (9.59) as a theorem.

THEOREM 9.4 Consider a DT system with proper rational transfer function $H(z)$. If the system is BIBO stable, then

$u[n] = a$ for $n \geq 0 \to y_{ss}[n] = aH(1)$
$u[n] = a\sin\omega_0 n$ for $n \geq 0 \to y_{ss}[n] = a|H(e^{j\omega_0})|\sin(\omega_0 n + \measuredangle H(e^{j\omega_0}))$
$u[n] = a\cos\omega_0 n$ for $n \geq 0 \to y_{ss}[n] = a|H(e^{j\omega_0})|\cos(\omega_0 n + \measuredangle H(e^{j\omega_0}))$

The steady-state response in (9.58) is excited by the input $u[n] = ae^{j\omega_0 n}$. If $\omega_0 = 0$, the input is a step sequence with amplitude a, and the output approaches a step sequence with amplitude $aH(1)$. If $u[n] = a\sin\omega_0 n = \text{Im}\, ae^{j\omega_0 n}$, where Im stands for the imaginary part, then the output approaches the imaginary part of (9.59) or $aA(\omega_0)\sin(\omega_0 n + \theta(\omega_0))$. Using the real part of $ae^{j\omega_0 n}$, we will obtain the next equation. In conclusion, if we apply a sinusoidal input to a system, then the output approaches a sinusoidal signal with the same frequency, but its amplitude will be modified by $A(\omega_0) = |H(e^{j\omega_0})|$ and its phase by $\measuredangle H(e^{j\omega_0}) = \theta(\omega_0)$. We give an example.

EXAMPLE 9.8.2

Consider a system with transfer function $H(z) = (z + 1)/(10z + 8)$. The system has pole $-8/10 = -0.8$ which has a magnitude less than 1, thus it is stable. We compute the steady-state response of the system excited by

$$u[n] = 2 + \sin 6.38n + 0.2 \cos 3n = 2 + \sin 0.1n + 0.2 \cos 3n \tag{9.60}$$

where $2 + \sin 0.1n$ will be considered the desired signal and $0.2 \cos 3n$ will be considered noise. In this problem, the sampling period is implicitly assumed to be 1. Thus the Nyquist frequency range is $(-\pi, \pi]$. Because the frequency 6.38 is outside the range, we have subtracted from it $2\pi = 6.28$ to bring it inside the range as $6.38 - 6.28 = 0.1$.

In order to apply Theorem 9.4, we read from Figures 9.7(a) and 9.7(aa) $H(1) = 1$, $H(e^{j0.1}) = 0.9e^{-j0.4}$, and $H(e^{j3}) = 0.008e^{-j1.6}$. Clearly, the reading cannot be very accurate. Then Theorem 9.4 implies

$$y_{ss}[n] = 2 \cdot 1 + 0.9 \sin(0.1n - 0.4) + 0.2 \cdot 0.008 \cos(3n - 1.6) \tag{9.61}$$

We see that the noise is essentially eliminated by the system. The system passes low-frequency signals and stops high-frequency signals and is therefore called a lowpass filter.

We mention that the condition of stability is essential in Theorem 9.4. If a system is not stable, then the theorem is not applicable as demonstrated in the next example.

EXAMPLE 9.8.3

Consider a system with transfer function $H_1(z) = (z + 1)/(-8z + 10)$. It has pole at 1.25 and is not stable. Its magnitude and phase responses are shown in Figures 9.7(b) and 9.7(bb). We can also read out $H_1(e^{j0.1}) = 0.91e^{j0.42}$. If we apply the input $u[n] = \sin 0.1n$, the output can be computed as

$$y[n] = -0.374(1.25)^n + 0.91 \sin(0.1n + 0.42)$$

for $n \geq 0$ (Problem 9.7). Although the output contains $|H_1(e^{j0.1})| \sin(0.1n + \measuredangle H_1(e^{j0.1}))$, the term is buried by $-0.374(1.25)^n$ as $n \to \infty$. Thus the output grows unbounded and Theorem 9.4 is not applicable. Furthermore, $H_1(e^{j\omega})$ has no physical meaning.

In view of Theorem 9.4, if we can design a stable system with the magnitude response (solid line) and the phase response (dotted line) shown in Figure 9.8(a), then the system will pass sinusoids with frequency $|\omega| < \omega_c$ and stop sinusoids with frequency $|\omega| > \omega_c$. We require the phase response to be linear as shown to avoid distortion as we will discuss later. We call such a system an *ideal lowpass filter* with cutoff frequency ω_c. Figures 9.8(b) and 9.8(c) show the characteristics of ideal bandpass and highpass filters. They are essentially the same as those in

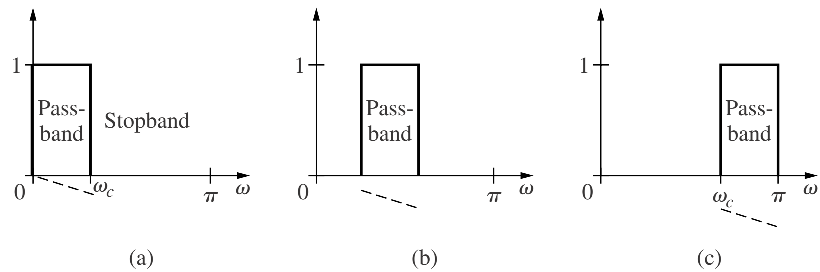


Figure 9.8 (**a**) Ideal lowpass filter. (**b**) Ideal bandpass filter. (**c**) Ideal highpass filter.

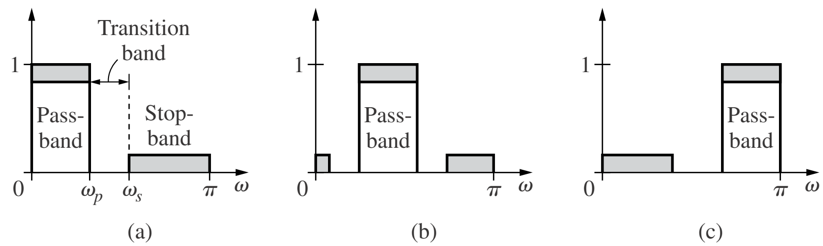


Figure 9.9 Specifications for practical (**a**) lowpass filter, (**b**) bandpass filter, and (**c**) highpass filter.

Figure 6.12. The only difference is the frequency range $[0, \infty)$ for the CT case and $[0, \pi]$ for the DT case. The ideal filters are not causal (their impulse responses are not identically zero for $n \leq 0$) and cannot be built in the real world. Thus the specifications in Figure 9.8 are modified as shown in Figure 9.9, which corresponds to Figure 6.13 for the CT case. The design problem then becomes the search of stable proper rational transfer functions whose magnitude responses lie inside the specified regions. See, for example, Reference 2.

9.8.1 Speed of Response—Time Constant

Consider the transfer function $H_1(z)$ in (9.45). Its step response was computed in (9.49) as

$$y[n] = k_1 0.6^n + k_2 n(0.6)^n + k_3 0.9^n \cos(2.354n + k_4) + H_i(1)$$

Because the system is stable, the response approaches the steady state

$$y_{ss}[n] = \lim_{n\to\infty} y[n] = H_i(1) = 1 \tag{9.62}$$

as $n \to \infty$. Mathematically speaking, it takes an infinite amount of time for the response to reach steady state. In practice, we often consider the response to have reached steady state if the response reaches and remains within $\pm 1\%$ of its steady state. One may then wonder: How fast will the response reach steady state? This question is important in designing practical systems.

Let us call $y_{tr}[n] := y[n] - y_{ss}[n]$ the *transient response*. For the transfer function in (9.45) with step input, the transient response is

$$y_{tr}[n] = k_1 0.6^n + k_2 n(0.6)^n + k_3 0.9^n \cos(2.35n + k_4) \tag{9.63}$$

Clearly the faster the transient response approaches zero, the faster the total response reaches steady state.

The form of the transient response in (9.63) is dictated only by the poles of $H_1(z)$. Because $H_1(z)$ is stable, all terms in (9.63) approach zero as $n \to \infty$. The smaller the magnitude of a pole, the faster its time response approaches zero. Thus the time for the transient response to approach zero is dictated by the pole that has the largest magnitude.

We introduced in Section 1.6 the concept of time constant for $b^n = \mathcal{Z}^{-1}[z/(z-b)]$ with $|b| < 1$. We showed there that the function b^n decreases to less than 1% of its peak magnitude in five time constants. We now extend the concept to the general case. Let $H(z)$ be a stable proper rational function and let $|b|$ be the largest magnitude of all poles. Because $H(z)$ is stable, we have $|b| < 1$. We define the *time constant* of $H(z)$ as

$$t_c = \frac{-1}{\ln |b|} \tag{9.64}$$

Note that t_c is positive because $|b| < 1$. Then generally, the transient response of $H(z)$ will decrease to less than 1% of its peak magnitude in n_c samples where

$$n_c := \text{round}(5t_c) = \text{round}(-5/\ln |b|) \tag{9.65}$$

Thus in practice, we may consider the response to have reached steady state in roughly n_c samples or $n_c T$ seconds. For example, all transfer functions in (9.45) through (9.48) have the same time constant $-1/\ln 0.9 = 9.47$, and their responses, as shown in Figure 9.5, all reach steady state in round $(5 \times 9.47 = 47.35) = 47$ samples.

It is important to mention that the rule of five time constants should be used only as a guide. It is possible to construct examples whose transient responses will not decrease to less than 1% of their peak magnitudes in five time constants. However, it is generally true that the smaller the time constant, the faster the system responds.

EXERCISE 9.8.1

What are the time constants of the following transfer functions:

(a) $\dfrac{1}{z - 0.7}$

(b) $\dfrac{z - 5}{(z+1)(z^2 + 0.8z + 0.41)}$

(c) $\dfrac{3(z - 5)}{(z + 0.9)(z^2 + 1.2z + 0.9225)}$

Answers

(a) 2.88.

(b) Not defined.

(c) 24.8.

9.9 FREQUENCY RESPONSES AND FREQUENCY SPECTRA

We showed in the preceding section that the output of a stable system with transfer function $H(z)$ excited by $u[n] = ae^{j\omega_0 n}$ approaches $aH(e^{j\omega_0})e^{j\omega_0 n}$ as $n \to \infty$. We now extend the formula to the general case.

Consider a DT positive-time signal $x[n]$ with sampling period 1. If it is absolutely summable, its frequency spectrum is defined in (5.7), with $T = 1$, as

$$X_d(\omega) = \sum_{n=0}^{\infty} x[n]e^{-jn\omega}$$

Let $X(z)$ be the z-transform of $x[n]$, that is,

$$X(z) = \sum_{n=0}^{\infty} x[n]z^{-n}$$

We see that replacing $z = e^{j\omega}$, the z-transform of $x[n]$ becomes the frequency spectrum of $x[n]$, that is,

$$X_d(\omega) = \mathcal{Z}[x[n]]\big|_{z=e^{j\omega}} = X(e^{j\omega}) \tag{9.66}$$

This is similar to the CT case where the frequency spectrum of a positive-time and absolutely integrable $x(t)$ equals its Laplace transform with s replaced by $j\omega$. It is important to mention that (9.66) holds only if $x[n]$ is positive time and absolutely summable.

The input and output of a DT system with transfer function $H(z)$ is related by

$$Y(z) = H(z)U(z) \tag{9.67}$$

This equation is applicable whether the system is stable or not and whether the frequency spectrum of the input signal is defined or not. For example, consider $H(z) = z/(z-2)$, which is unstable, and $u[n] = 1.2^n$, which grows unbounded and its frequency spectrum is not defined. The output of the system is

$$Y(z) = H(z)U(z) = \frac{z}{z-2}\frac{z}{z-1.2} = \frac{2.5z}{z-2} - \frac{1.5z}{z-1.2}$$

which implies $y[n] = 2.5 \times 2^n - 1.5 \times 1.2^n$, for $n \geq 0$. The output grows unbounded and its frequency spectrum is not defined.

Before proceeding, we show that if $H(z)$ is stable (its impulse response $h[n]$ is absolutely summable) and if $u[n]$ is absolutely summable, then so is the output. Indeed, we have

$$y[n] = \sum_{k=0}^{n} h[n-k]u[k]$$

which implies

$$|y[n]| \leq \sum_{k=0}^{n} |h[n-k]||u[k]| \leq \sum_{k=0}^{\infty} |h[n-k]||u[k]|$$

Thus we have

$$\sum_{n=0}^{\infty}|y[n]| \le \sum_{n=0}^{\infty}\left(\sum_{k=0}^{\infty}|h[n-k]||u[k]|\right) = \sum_{k=0}^{\infty}\left(\sum_{n=0}^{\infty}|h[n-k]|\right)|u[k]|$$

$$= \sum_{k=0}^{\infty}\left(\sum_{\bar{n}=-k}^{\infty}|h(\bar{n})|\right)|u[k]| = \left(\sum_{\bar{n}=0}^{\infty}|h[\bar{n}]|\right)\left(\sum_{k=0}^{\infty}|u[k]|\right)$$

where we have interchanged the order of summations, introduced a new index $\bar{n} = n - k$, and used the causality condition $h[\bar{n}] = 0$ for $\bar{n} < 0$. Thus if $h[n]$ and $u[n]$ are absolutely summable, so is $y[n]$.

Let us substitute $z = e^{j\omega}$ into (9.67) to yield

$$Y(e^{j\omega}) = H(e^{j\omega})U(e^{j\omega}) \tag{9.68}$$

The equation is meaningless if the system is not stable or if the input frequency spectrum is not defined. However, if the system is BIBO stable and if the input is absolutely summable, then the output is absolutely summable and its frequency spectrum is well defined and equals the product of the frequency response $H(e^{j\omega})$ of the system and the frequency spectrum $U(e^{j\omega})$ of the input. This is a fundamental equation and is the basis of digital filter design.

Let us consider the magnitude and phase responses of an ideal lowpass filter with cutoff frequency ω_c shown in Figure 9.8(a) or

$$H(e^{j\omega}) = \begin{cases} 1 \cdot e^{-j\omega n_0} & \text{for } |\omega| \le \omega_c \\ 0 & \text{for } \omega_c < |\omega| \le \pi \end{cases} \tag{9.69}$$

with $n_0 > 0$. Now if $u[n] = u_1[n] + u_2[n]$ and if the magnitude spectra of $u_1[n]$ and $u_2[n]$ are as shown in Figure 9.10, then the output frequency spectrum is given by

$$Y(e^{j\omega}) = H(e^{j\omega})U(e^{j\omega}) = U_1(e^{j\omega})e^{-j\omega n_0} \tag{9.70}$$

If the z-transform of $u_1[n]$ is $U_1(z)$, then the z-transform of $u_1[n-n_0]$ is $z^{-n_0}U_1(z)$ as derived in (9.22). Thus the spectrum in (9.70) is the spectrum of $u_1[n-n_0]$. In other words, the output of the ideal lowpass filter is

$$y[n] = u_1[n-n_0] \tag{9.71}$$

That is, the filter stops completely the signal $u_2[n]$ and passes $u_1[n]$ with only a delay of n_0 samples. This is called a *distortionless transmission* of $u_1[n]$.

We stress once again that the equation in (9.67) is more general than the equation in (9.68). Equation (9.68) is applicable only if the system is stable and the input frequency spectrum is defined.

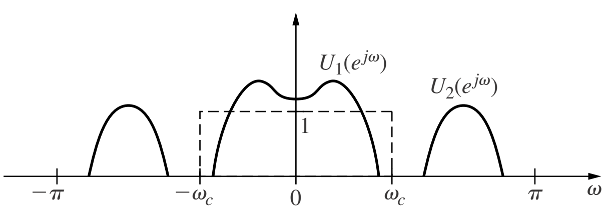


Figure 9.10 Spectra of $u_1(t)$ and $u_2(t)$.

9.10 DIGITAL PROCESSING OF CT SIGNALS

Because of many advantages of DT techniques, CT signals are now widely processed using DT systems. We use an example to illustrate how this is carried out. Consider the CT signal in (6.52) or

$$u(t) = 1 + \sin 0.1t + 0.2\cos 20t \tag{9.72}$$

where $0.2\cos 20t$ is noise. In Example 6.8.3, we used the CT lowpass filter $2/(s+2)$ to eliminate the noise as shown in Figure 9.11(a). We plot in Figures 9.12(a) and 9.12(b) the input and output of the analog filter. They are actually the plots in Figure 6.11. The noise $0.2\cos 20t$ is essentially eliminated in Figure 9.12(b).

We now discuss digital processing of the CT signal. In order to process $u(t)$ digitally, we must first select a sampling period. The Nyquist sampling theorem requires $T < \pi/20 = 0.157$. Let us select $T = 0.15$. Then the output of the ADC in Figure 9.11(b) is

$$u(nT) = 1 + \sin 0.1nT + 0.2\cos 20nT \tag{9.73}$$

with $T = 0.15$. Its positive Nyquist frequency range is $[0, \pi/T = 20.94]$.

In designing DT filters, we may normalize the sampling period to 1. This can be achieved by multiplying all frequencies by $T = 0.15$. Thus the frequencies

$$0 \quad 0.1 \quad 20$$

in the positive Nyquist frequency range $[0, \pi/T = 20.94]$ become

$$0 \times 0.15 \quad 0.1 \times 0.15 \quad 20 \times 0.15$$

or

$$0 \quad 0.015 \quad 3$$

in the normalized positive Nyquist frequency range $[0, \pi]$. Thus we must design a DT filter to pass sinusoids with frequencies 0 and 0.015 and to stop sinusoid with frequency 3. The DT filter

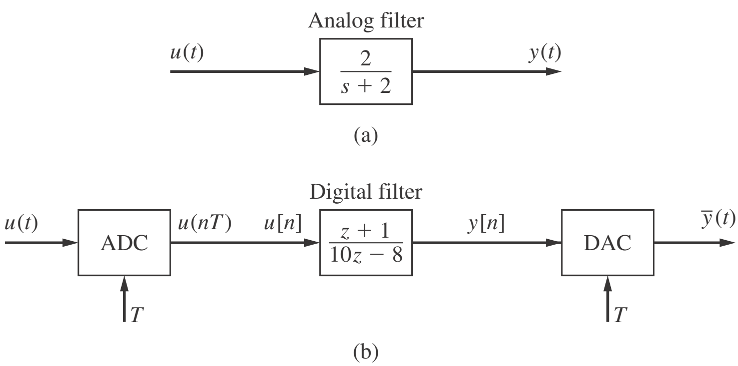


Figure 9.11 (**a**) Analog procession of CT signal. (**b**) Digital processing of CT signal.

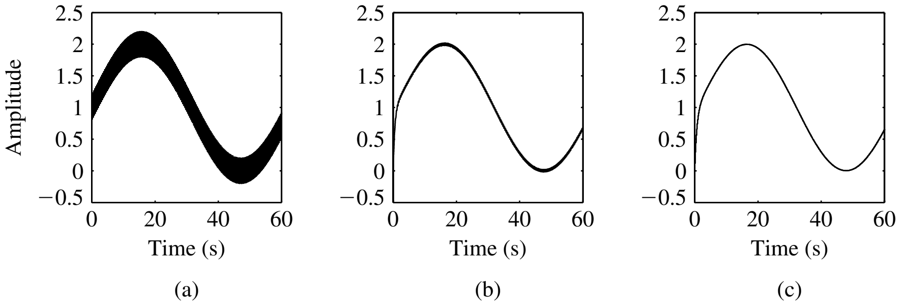


Figure 9.12 (**a**) The CT signal in (9.72). (**b**) The output $y(t)$ (solid line) of the analog filter $2/(s+2)$ in Figure 9.11(a) and the desired signal $1 + \sin 0.1t$ (dotted line). (**c**) The output $\bar{y}(t)$ (solid line) in Figure 9.11(b), that is, the output $y[n]$ of the digital filter $(z+1)/(10z-8)$ passing through a zero-order hold, and the desired signal (dotted line).

$H(z) = (z+1)/(10z-8)$, whose magnitude response is shown in Figure 9.7(a), can achieve this. Let us type

```
n=0:400;T=0.15;
uT=1+sin(0.1*n*T)+0.2.*cos(20*n*T);
dsys=tf([1 1],[10 -8],T);
yT=lsim(dsys,uT)
stairs(n*T,yT)
```

in MATLAB. The first line is the number of samples to be used and the selected sampling period. The second line is the sampled input or, equivalently, the output of the ADC in Figure 9.11(b). The third line uses transfer function (tf) to define the DT system by including the sampling period T. It is important to mention that in designing a DT filter, we may assume $T = 1$. However the resulting filter is applicable for any $T > 0$. The function lsim, an acronym for linear simulation, computes the output of the digital filter $(z+1)/(10z-8)$. It is the sequence of numbers $y[n]$ shown in Figure 9.11(b). The MATLAB function stairs carries out zero-order hold to yield the output $\bar{y}(t)$. See Section 5.3.2. It is the output of the DAC shown in Figure 9.11(b) and is plotted in Figure 9.12(c). The result is comparable to, if not better than, the one obtained using an analog filter.

PROBLEMS

9.1 Consider a DT system with impulse response

$$h[n] = 1 + 2 \cdot 0.8^n - 3 \cdot (-0.7)^n$$

What is its transfer function?

9.2 Find the transfer function for each of the following difference equations:

(a) $2y[n+2] + 4y[n+1] + 10y[n] = u[n+2] - u[n+1] - 2u[n]$

(b) $2y[n] + 4y[n-1] + 10y[n-2] = u[n-1] + 3u[n-2] + 2u[n-3]$

(c) $y[n] + y[n-5] = u[n-2] + u[n-5]$

Compare your results with those in Problem 6.2. Are they proper?

9.3 Find a delayed-form and an advanced-form difference equation for the DT transfer function

$$H(z) = \frac{V(z)}{R(z)} = \frac{2z^2 + 5z + 3}{z^4 + 3z^3 + 10}$$

9.4 Suppose the step response of a DT system is measured as $y_q[n] = h[n]$, where $h[n]$ is given in Problem 9.1. What is the transfer function of the system?

9.5 Find the poles and zeros for each of the following transfer functions:

(a) $H(z) = \dfrac{3z + 6}{2z^2 + 2z + 1}$

(b) $H(z) = \dfrac{z^{-1} - z^{-2} - 6z^{-3}}{1 + 2z^{-1} + z^{-2}}$

9.6 Find the impulse and step responses of a DT system with transfer function

$$H(z) = \frac{0.9z}{(z+1)(z-0.8)}$$

9.7 Verify that the output of $H(z) = (z+1)/(-8z+10)$ excited by $u[n] = \sin 0.1n$ is given by

$$y[n] = -0.374(1.25)^n + 0.91\sin(0.1n + 0.42)$$

for $n \geq 0$.

9.8 What is the general form of the response of

$$H(z) = \frac{z^2 + 2z + 1}{(z-1)(z-0.5+j0.6)(z-0.5-j0.6)}$$

excited by a step sequence?

9.9 Consider the DT transfer function

$$H(z) = \frac{N(z)}{(z+0.6)^3(z-0.5)(z^2+z+0.61)}$$

where $N(z)$ is a polynomial of degree 6 or less. What is the general form of its step response?

9.10 Use Jury's test to check the DT stability of the polynomials

(a) $z^3 + 4z^2 + 2$

(b) $z^3 - z^2 + 2z - 0.7$

(c) $2z^4 + 1.6z^3 + 1.92z^2 + 0.64z + 0.32$

9.11 Find polynomials of degree 1 which are, respectively, stable and DT stable, stable but not DT stable, DT stable but not stable, and neither stable nor DT stable.

9.12 Verify analytically that the magnitude responses of

$$H_1(z) = \frac{z+1}{10z-8} \quad \text{and} \quad H_2(z) = \frac{z+1}{-8z+10}$$

are identical. Do they have the same phase responses?

9.13 Consider the system in Problem 9.6. Is the system stable? What is the form of the steady-state response of the system excited by $u[n] = \sin 2n$? Does the response contain any sinusoid with frequency other than 2 rad/s? Does Theorem 9.4 hold?

9.14 Consider the system in Problem 9.9. If $N(1) = 1.6^3 \times 0.5 \times 0.61$, what is its steady-state response excited by a unit step input? How many samples will it take to reach steady state?

9.15 Compute the values of

$$H(z) = \frac{z-1}{z+0.8}$$

at $z^{j\omega}$ with $\omega = 0, \pi/4, \pi/2, 3\pi/4,$ and π. Use the values to plot roughly the magnitude and phase responses of $H(z)$.

9.16 Find the steady-state response of the DT system in Problem 9.15 excited by

$$u[n] = 2 + \sin 0.1n + \cos 3n$$

How many samples will it take to reach steady state? Is the system a lowpass or highpass system?

9.17 Let $X(z)$ be the z-transform of $x[n]$. Show that if all poles, except possibly a simple pole at $z = 1$, of $X(z)$ have magnitude less than 1, then $x[n]$ approaches a constant (zero or nonzero) as $n \to \infty$ and

$$\lim_{n\to\infty} x[n] = \lim_{z\to 1}(z-1)X(z)$$

This is called the *final-value theorem* of the z-transform. This is the DT counterpart of the one in Problem 6.35.

9.18 Consider $x[n] = 2^n$, for $n \geq 0$, and its z-transform $X(z) = z/(z-2)$. What is the final value of $x[n]$? What is the value of $\lim_{z\to 1}(z-1)X(z)$? Does the final-value theorem hold for this sequence? Why?

CHAPTER 10

DT State-Space Equations and Realizations

10.1 INTRODUCTION

We showed in the preceding chapter that every DT linear, time-invariant, and lumped (LTIL) system can be described by a rational transfer function. From the transfer function, we can develop general properties of the system. However, transfer functions, for the same reasons discussed in Section 6.5.4, are not used in computing responses of systems. For computer computation, transfer functions are transformed, as we discuss in this chapter, into state-space (ss) equations. State-space equations also provide structures for implementing DT systems. Thus the DT ss equation is an important mathematical description of DT systems. However, its analytical study will not be discussed in this text because it plays no role in computer computation and is more complex and less revealing than the corresponding study of DT transfer functions.

Two approaches are available to develop DT ss equations. The first approach is to develop DT ss equations, as in the CT case, from transfer functions. The procedure is identical to the CT case and will be discussed in Section 10.3. The second approach is to develop ss equations directly from high-order difference equations. Because this approach is widely used in DSP texts, we discuss it in the next section. All the discussion in this chapter is applicable to stable as well as unstable systems. Thus realization and stability are two independent issues.

In the design of DT systems, the sampling period can be assumed to be 1. Thus the sampling period is suppressed in this chapter and all DT signals are plotted against time index. However, the resulting systems can be used to process DT signals of any sampling period $T > 0$.

10.2 FROM DIFFERENCE EQUATIONS TO BASIC BLOCK DIAGRAMS

A DT basic block diagram is a diagram that consists only of the types of elements shown in Figure 10.1. The element denoted by z^{-1} enclosed by a box in Figure 10.1(a) is a unit-sample delay element or, simply, a *unit delay element;* its output $y[n]$ equals the input $u[n]$ delayed by one sample, that is,

$$y[n] = u[n-1] \quad \text{or} \quad y[n+1] = u[n] \qquad \text{(unit delay)}$$

A unit delay element is simply a memory location. We store a number in the location at time instant n and then fetch it in the next sample.

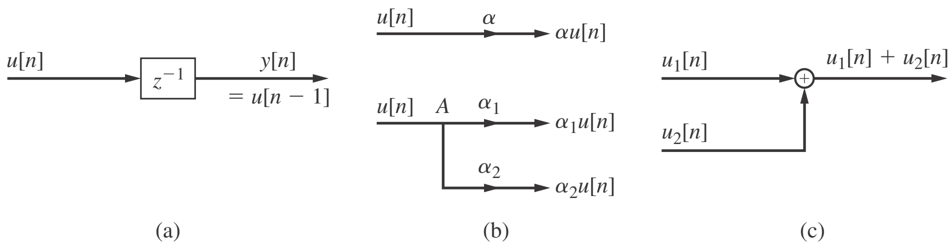


Figure 10.1 (**a**) Unit delay element. (**b**) Multiplier. (**c**) Adder.

The element in Figure 10.1(b) denoted by a line with an arrow and a real number α is called a multiplier with gain α; its input and output are related by

$$y[n] = \alpha u[n] \qquad \text{(multiplier)}$$

If $\alpha = 1$, it is direct transmission and the arrow and α may be omitted. In addition, a signal may branch out to two or more signals with gain α_i as shown. The point denoted by A is called a *branching point*. The element denoted by a small circle with a plus sign in Figure 10.1(c) is called an adder or summer. Every adder has two inputs $u_i[n]$, $i = 1, 2$, denoted by two entering arrows, and only one output $y(t)$, denoted by a departing arrow. They are related by

$$y[n] = u_1[n] + u_2[n] \qquad \text{(adder)}$$

These elements are called *DT basic elements*. These are basic operations in every PC or specialized digital hardware. Note that a unit advance element is noncausal and is not a DT basic element.

Every DT LTIL system can be described by an advanced-form or delayed-form difference equation. If we use delayed-form, then we can readily obtain a basic block diagram. For example, consider the following fourth-order difference equation

$$\begin{aligned} y[n] &+ a_2 y[n-1] + a_3 y[n-2] + a_4 y[n-3] + a_5 y[n-4] \\ &= \hat{b}_1 u[n] + \hat{b}_2 u[n-1] + \hat{b}_3 u[n-2] + \hat{b}_4 u[n-3] + \hat{b}_5 u[n-4] \end{aligned} \tag{10.1}$$

where we have normalized a_1 to 1. Let us write (10.1) as

$$\begin{aligned} y[n] = {} & \hat{b}_1 u[n] + \hat{b}_2 u[n-1] + \hat{b}_3 u[n-2] + \hat{b}_4 u[n-3] + \hat{b}_5 u[n-4] \\ & - a_2 y[n-1] - a_3 y[n-2] - a_4 y[n-3] - a_5 y[n-4] \end{aligned} \tag{10.2}$$

If we use a chain of four unit delay elements to generate $y[n-i]$, for $i = 1 : 4$ and another chain of four unit delay elements to generate $u[n-i]$, for $i = 1 : 4$, then we can readily verify that the basic block diagram in Figure 10.2 implements (10.2) and, consequently, (10.1). It is called a *direct-form* structure. The structure consists of the tandem connection of two subsystems shown.

For single-input single-output LTIL systems that are initially relaxed, we may interchange the order of their tandem connection as shown in Figure 10.3(a). This follows from the fact

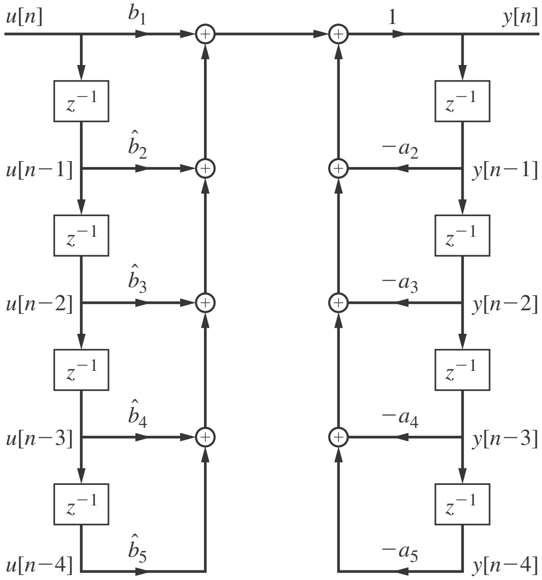


Figure 10.2 Direct-form structure of (10.1).

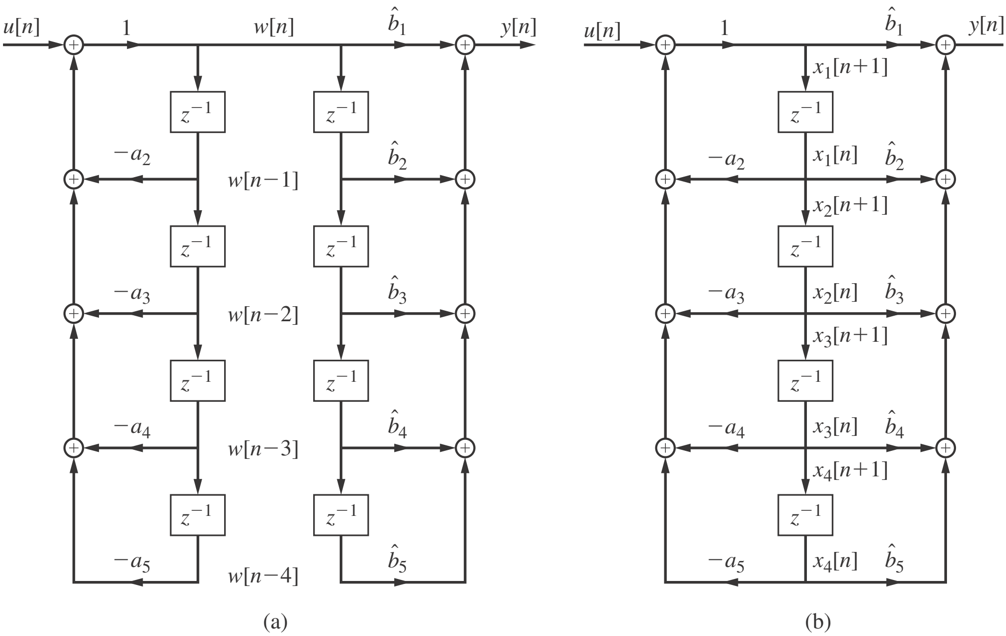


Figure 10.3 Direct-form II structure of (10.1).

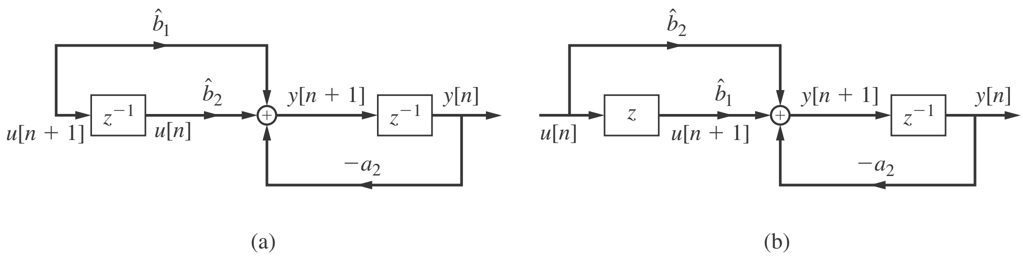


Figure 10.4 (**a**) Invalid block diagram of (10.3). (**b**) Block diagram of (10.3) that contains a unit advance element.

that $H_1(z)H_2(z) = H_2(z)H_1(z)$.[1] If we assign $w[n]$ as shown, then the variables along the two chains of unit delay elements are the same. Thus we can combine the two chains as shown in Figure 10.3(b). This basic block diagram is called the *direct-form II* in texts on digital signal processing.

The basic block diagrams in Figures 10.2 and 10.3 can be readily developed from delayed-form difference equations. The procedure, however, cannot be used for advanced-form difference equations. For example, consider the first-order advanced-form difference equation:

$$y[n+1] + a_2 y[n] = \hat{b}_1 u[n+1] + \hat{b}_2 u[n] \tag{10.3}$$

We can develop the block diagrams in Figure 10.4 to realize (10.3). The response of Figure 10.4(a) is excited by the input $u[n]$. Because there is no route for the input $u[n]$ to go to $u[n+1]$, the left-hand side unit delay element is not in operation and the diagram in Figure 10.4(a) is not valid. Although the diagram in Figure 10.4(b) implements (10.3), it is not a basic block diagram because it contains a unit advance element. Thus it seems not possible to develop a direct-form basic block diagram *directly* from an advanced-form difference equation.[2] Because differential equations for CT systems correspond to advanced-form difference equations, we don't encounter direct-form basic block diagrams in CT systems.

10.2.1 Basic Block Diagrams to State-Space Equations

Consider the following set of two equations:

$$\mathbf{x}[n+1] = \mathbf{A}\mathbf{x}[n] + \mathbf{b}u[n] \tag{10.4}$$

$$y[n] = \mathbf{c}\mathbf{x}[n] + du[n] \tag{10.5}$$

where $u[n]$ is the input, $y[n]$ is the output, and $\mathbf{x}[n]$ is the *state* denoting a set of initial conditions. The state $\mathbf{x}[n]$ is a column vector. If it has N components, then it can be written explicitly as $\mathbf{x}[n] = [x_1[n]\ x_2[n] \cdots x_N[n]]'$, where the prime denotes the transpose. If $\mathbf{x}$ is $N \times 1$, then $\mathbf{A}$ is an $N \times N$ square matrix, $\mathbf{b}$ is an $N \times 1$ column vector, $\mathbf{c}$ is a $1 \times N$ row vector, and d is scalar

[1] This commutative property does not hold for multi-input multi-output systems, nor for time-varying systems. See Problems 8.8 and 8.9.

[2] If we introduce an intermediate variable, then it is possible to develop a direct-form II block diagram. This will not be discussed.

(1×1). The matrix equation in (10.4) actually consists of N first-order difference equations; it relates the input and state and is called the *state equation*. The equation in (10.5) relates the input and output and is called the *output equation*. Note that $\mathbf{c}$ is a $1 \times N$ row vector and $\mathbf{cx}$ is the product of $1 \times N$ and $N \times 1$ vectors and is 1×1. The constant d is called the *direct transmission gain*. The set of two equations in (10.4) and (10.5) is called a DT state-space (ss) equation of dimension N.

Consider a basic block diagram. If we assign the *output of each unit delay element* as a state variable $x_i[n]$, then we can develop an ss equation to describe the diagram. For example, if we assign $x_i[n]$, for $i = 1:4$ as shown in Figure 10.3(b), then the input of each unit delay element is $x_i[n+1]$. From the diagram, we can read out

$$
\begin{aligned}
x_1[n+1] &= -a_2x_1[n] - a_3x_2[n] - a_4x_3[n] - a_5x_4[n] + u[n] \\
x_2[n+1] &= x_1[n] \\
x_3[n+1] &= x_2[n] \\
x_4[n+1] &= x_3[n] \\
y[n] &= \hat{b}_1x_1[n+1] + \hat{b}_2x_1[n] + \hat{b}_3x_2[n] + \hat{b}_4x_3[n] + \hat{b}_5x_4[n]
\end{aligned}
\tag{10.6}
$$

The output $y[n]$ contains $\hat{b}_1x_1[n+1]$ which is not permitted in (10.5). This can be resolved by substituting the first equation in (10.6) into the last equation to yield

$$
\begin{aligned}
y[n] &= \hat{b}_1(-a_2x_1[n] - a_3x_2[n] - a_4x_3[n] - a_5x_4[n] + u[n]) \\
&\quad + \hat{b}_2x_1[n] + \hat{b}_3x_2[n] + \hat{b}_4x_3[n] + \hat{b}_5x_4[n] \\
&= (\hat{b}_2 - a_2\hat{b}_1)x_1[n] + (\hat{b}_3 - a_3\hat{b}_1)x_2[n] + (\hat{b}_4 - a_4\hat{b}_1)x_3[n] \\
&\quad + (\hat{b}_5 - a_5\hat{b}_1)x_4[n] + \hat{b}_1u[n]
\end{aligned}
\tag{10.7}
$$

The first four equations in (10.6) and (10.7) can be expressed in matrix form as

$$
\begin{aligned}
\mathbf{x}[n+1] &= \begin{bmatrix} -a_2 & -a_3 & -a_4 & -a_5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u[n] \\
y[n] &= [\hat{b}_2 - a_2\hat{b}_1 \quad \hat{b}_3 - a_3\hat{b}_1 \quad \hat{b}_4 - a_4\hat{b}_1 \quad \hat{b}_5 - a_5\hat{b}_1]\mathbf{x}[n] + \hat{b}_1u[n]
\end{aligned}
\tag{10.8}
$$

This is of the form shown in (7.7) and is said to be in the *controllable canonical form*. This ss equation is developed from a basic block diagram that is obtained directly from a delayed-form difference equation. We discuss in the next section a different way of developing ss equations.

10.3 REALIZATIONS

By applying the z-transform and assuming zero initial conditions, a difference equation can be transformed into a transfer function. We encounter both positive-power and negative-power transfer functions. If we use positive-power transfer functions, then all the discussion in Section 7.2 for CT systems can be directly applied.

Suppose we are given a DT proper transfer function $H(z)$. The problem of finding an ss equation so that its transfer function equals $H(z)$ is called the *realization* problem. The ss equation is called a *realization* of $H(z)$. Let $X(z)$ be the z-transform of $x[n]$. Then we have, as derived in (9.20), $\mathcal{Z}[x[n+1]] = z(X(z) - x[0])$. Using its vector version, applying the z-transform to (10.4), and assuming zero initial state, we obtain

$$z\mathbf{X}(z) = \mathbf{A}\mathbf{X}(z) + \mathbf{b}U(z)$$

where $U(z)$ is the z-transform of $u[n]$. Using $z\mathbf{X} = z\mathbf{I}\mathbf{X}$, where $\mathbf{I}$ is a unit matrix, we can write the preceding equation as

$$(z\mathbf{I} - \mathbf{A})\mathbf{X}(z) = \mathbf{b}U(z)$$

Thus we have $\mathbf{X}(z) = (z\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}U(z)$. Substituting this into the z-transform of (10.5), we obtain

$$Y(z) = \mathbf{c}\mathbf{X}(z) + dU(z) = [\mathbf{c}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} + d]U(z)$$

Thus the transfer function of the ss equation in (10.4) and (10.5) is

$$H(z) = \mathbf{c}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} + d \tag{10.9}$$

This equation is identical to (7.48) if z is replaced by s. Thus the realization problem is the same as, given $H(z)$, finding $\{\mathbf{A}, \mathbf{b}, \mathbf{c}, d\}$ to meet (10.9).

Consider the DT transfer function

$$H(z) = \frac{Y(z)}{U(z)} = \frac{\bar{b}_1 z^4 + \bar{b}_2 z^3 + \bar{b}_3 z^2 + \bar{b}_4 z + \bar{b}_5}{\bar{a}_1 z^4 + \bar{a}_2 z^3 + \bar{a}_3 z^2 + \bar{a}_4 z + \bar{a}_5} \tag{10.10}$$

with $\bar{a}_1 \neq 0$. We call $\bar{a}_1$ the leading coefficient. The rest of the coefficients can be zero or nonzero. The transfer function is proper and describes a causal system. If a transfer function is improper, then it cannot be realized by any ss equation of the form shown in (10.4) and (10.5). The first step in realization is to write (10.10) as

$$H(z) = \frac{N(z)}{D(z)} + d = \frac{b_1 z^3 + b_2 z^2 + b_3 z + b_4}{z^4 + a_2 z^3 + a_3 z^2 + a_4 z + a_5} + d \tag{10.11}$$

with $D(z) = z^4 + a_2 z^3 + a_3 z^2 + a_4 z + a_5$ and $N(z)/D(z)$ is strictly proper. This can be achieved by dividing the numerator and denominator of (10.10) by $\bar{a}_1$ and then carrying out a direct division. The procedure is identical to the one in Example 7.2.1 and will not be repeated.

Now we claim that the following ss equation realizes (10.11):

$$\mathbf{x}[n+1] = \begin{bmatrix} -a_2 & -a_3 & -a_4 & -a_5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u[n] \tag{10.12}$$
$$y[n] = [b_1 \quad b_2 \quad b_3 \quad b_4]\mathbf{x}[n] + du[n]$$

with $\mathbf{x}[n] = [x_1[n] \;\; x_2[n] \;\; x_3[n] \;\; x_4[n]]'$. This ss equation can be obtained directly from the coefficients in (10.11). We place the denominator's coefficients, except its leading coefficient 1, with sign reversed in the first row of **A**, and we place the numerator's coefficients, without changing sign, directly as **c**. The constant d in (10.11) is the direct transmission gain. The rest of the ss equation have fixed patterns. The second row of **A** is $[1 \ \ 0 \ \ 0 \ \ \cdots]$. The third row of **A** is $[0 \ \ 1 \ \ 0 \ \ \cdots]$ and so forth. The column vector **b** is all zero except its first entry which is 1.

To show that (10.12) is a realization of (10.11), we must compute its transfer function. Following the procedure in Section 7.2, we write (10.12) explicitly as

$$\begin{aligned} x_1[n+1] &= -a_2x_1[n] - a_3x_2[n] - a_4x_3[n] - a_5x_4[n] + u[n] \\ x_2[n+1] &= x_1[n] \\ x_3[n+1] &= x_2[n] \\ x_4[n+1] &= x_3[n] \end{aligned}$$

Applying the z-transform and assuming zero initial conditions yield

$$\begin{aligned} zX_1(z) &= -a_2X_1(z) - a_3X_2(z) - a_4X_3(z) - a_5X_4(z) + U(z) \\ zX_2(z) &= X_1(z) \\ zX_3(z) &= X_2(z) \\ zX_4(z) &= X_3(z) \end{aligned} \tag{10.13}$$

From the second to the last equation, we can readily obtain

$$X_2(z) = \frac{X_1(z)}{z}, \quad X_3(z) = \frac{X_2(z)}{z} = \frac{X_1(z)}{z^2}, \quad X_4(z) = \frac{X_1(z)}{z^3} \tag{10.14}$$

Substituting these into the first equation of (10.13) yields

$$\left(z + a_2 + \frac{a_3}{z} + \frac{a_4}{z^2} + \frac{a_5}{z^3}\right) X_1(z) = U(z)$$

which implies

$$X_1(z) = \frac{z^3}{z^4 + a_2z^3 + a_3z^2 + a_4z + a_5} U(z) =: \frac{z^3}{D(z)} U(z) \tag{10.15}$$

where $D(z) = z^4 + a_2z^3 + a_3z^2 + a_4z + a_5$. Substituting (10.15) into (10.14) and then into the following z-transform of the output equation in (10.12) yields

$$\begin{aligned} Y(z) &= b_1X_1(z) + b_2X_2(z) + b_3X_3(z) + b_4X_4(z) + dU(z) \\ &= \left(\frac{b_1z^3}{D(z)} + \frac{b_2z^2}{D(z)} + \frac{b_3z}{D(z)} + \frac{b_4}{D(z)}\right) U(z) + dU(z) \\ &= \left(\frac{b_1z^3 + b_2z^2 + b_3z + b_4}{D(z)} + d\right) U(z) \end{aligned} \tag{10.16}$$

This shows that the transfer function of (10.12) equals (10.11). Thus (10.12) is a realization of (10.11) or (10.10). The realization in (10.12) is said to be in the *controllable canonical form*. We mention that the realization in (10.12) is identical to that in (10.8). This can be verified by showing $b_{i-1} = \hat{b}_i - a_i\hat{b}_1$, for $i = 2 : 5$ and $d = \hat{b}_1$. We also mention that the preceding derivation is identical to the one in Section 7.2.

We discuss a different realization. Let $\{\mathbf{A}, \mathbf{b}, \mathbf{c}, d\}$ be a realization of $H(z)$, that is, $H(z) = \mathbf{c}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} + d$. Using $H'(z) = H(z)$, $d' = d$, $\mathbf{I}' = \mathbf{I}$, and $(\mathbf{MPQ})' = \mathbf{Q}'\mathbf{P}'\mathbf{M}'$, we have

$$H'(z) = H(z) = \left(\mathbf{c}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} + d\right)' = \mathbf{b}'(z\mathbf{I} - \mathbf{A}')^{-1}\mathbf{c}' + d \tag{10.17}$$

Thus from (10.12), we can obtain another realization of (10.11) as

$$\begin{aligned} \mathbf{x}[n+1] &= \begin{bmatrix} -a_2 & 1 & 0 & 0 \\ -a_3 & 0 & 1 & 0 \\ -a_4 & 0 & 0 & 1 \\ -a_5 & 0 & 0 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} u[n] \\ y[n] &= [1 \quad 0 \quad 0 \quad 0]\mathbf{x}[n] + du[n] \end{aligned} \tag{10.18}$$

The $\mathbf{A}$, $\mathbf{b}$, and $\mathbf{c}$ in (10.18) are the transposes of the $\mathbf{A}$, $\mathbf{c}$, and $\mathbf{b}$ in (10.12). Thus (10.18) can be easily obtained from (10.12) or directly from the coefficients of (10.11). The realization in (10.18) is said to be in the *observable canonical form*.

We plot in Figure 10.5(a) the basic block diagram of the controllable-canonical-form ss equation in (10.12), and we plot in Figure 10.5(b) the basic block diagram of the observable-canonical-form ss equation in (10.18). The left-hand side or feedback part of Figure 10.3(b)

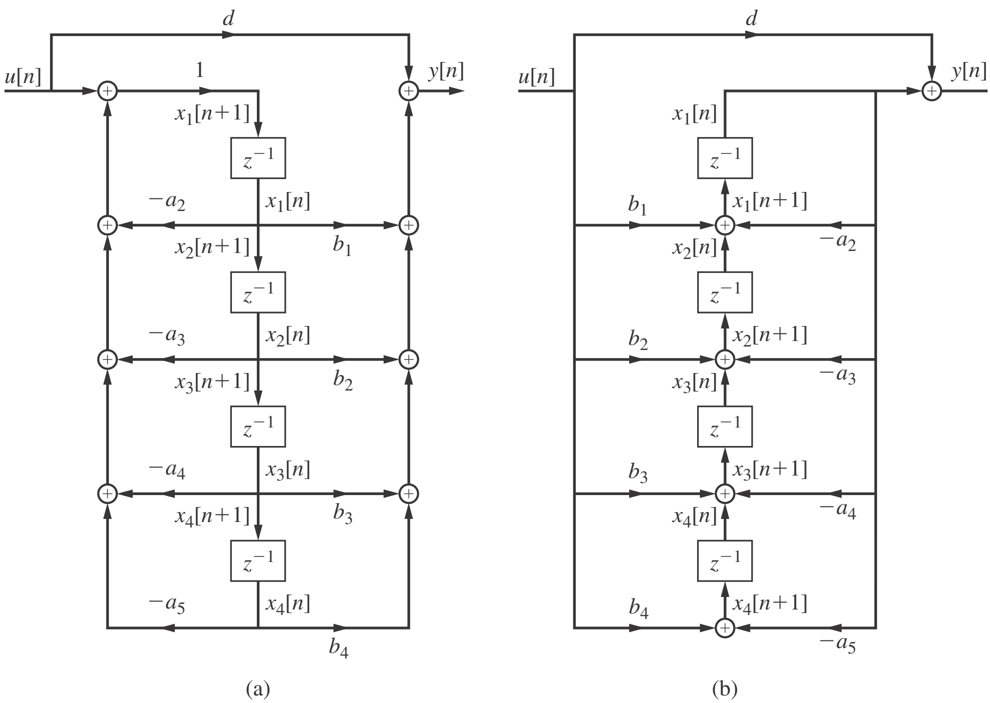


Figure 10.5 Basic block diagrams of (**a**) controllable canonical form (direct form II) and (**b**) observable canonical form (direct form II transposed).

is identical to the left-hand side or feedback part of Figure 10.5(a); their right-hand sides or feedforward parts are different. Their ss equations, however, are identical. The basic block diagram in Figure 10.5(a) developed from a controllable-form ss equation is called *direct form II* in DSP texts. The one in Figure 10.5(b) developed from an observable-form ss equation is called *direct form II transposed* in DSP texts. Note that observable form is obtained from controllable form by transposition as shown in (10.17). See also Exercise 7.3.1.

EXAMPLE 10.3.1

Find two realizations for the transfer function

$$\begin{aligned} H(z) &= \frac{3z^4 + 7z^3 - 2z + 10}{2z^4 + 3z^3 + 4z^2 + 7z + 5} \\ &= \frac{1.25z^3 - 3z^2 - 6.25z + 1.25}{z^4 + 1.5z^3 + 2z^2 + 3.5z + 2.5} + 1.5 \end{aligned} \tag{10.19}$$

Note that (10.19) is the same as (7.5) if z is replaced by s. Its controllable canonical form realization is, using (10.12),

$$\mathbf{x}[n+1] = \begin{bmatrix} -1.5 & -2 & -3.5 & -2.5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u[n]$$

$$y[n] = [1.25 \quad -3 \quad -6.25 \quad 1.25]\mathbf{x}[n] + 1.5u[n]$$

and its observable canonical form realization is, using (10.18),

$$\mathbf{x}[n+1] = \begin{bmatrix} -1.5 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ -3.5 & 0 & 0 & 1 \\ -2.5 & 0 & 0 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 1.25 \\ -3 \\ -6.25 \\ 1.25 \end{bmatrix} u[n]$$

$$y[n] = [1 \quad 0 \quad 0 \quad 0]\mathbf{x}[n] + 1.5u[n]$$

We see that these realizations can be read out from the coefficients of the transfer function.

We mention that it is possible to find infinitely many other realizations for the transfer function in (10.11), but the canonical forms in (10.12) and (10.18) are most convenient to develop and use.

EXERCISE 10.3.1

Find two realizations for each of the following transfer functions:

(a) $H_1(z) = \dfrac{2z+3}{4z+10}$

(b) $H_2(z) = \dfrac{z^2+3z-4}{2z^3+3z+1}$

(c) $H_3(z) = \dfrac{2}{z^3}$

Answers

Same as those for Exercise 7.2.1.

The MATLAB function tf2ss, an acronym for transfer function to ss equation, carries out realizations. Its applications to CT and DT transfer functions are identical if DT transfer functions are expressed in the positive-power form. For the transfer function in (10.19), typing

```
n=[3 7 0 -2 10];d=[2 3 4 7 5];
[a,b,c,d]=tf2ss(n,d)
```

yields

$$a = \begin{matrix} -1.5000 & -2.0000 & -3.5000 & -2.5000 \\ 1.0000 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 \\ 0 & 0 & 1.0000 & 0 \end{matrix}$$

$$b = \begin{matrix} 1 \\ 0 \\ 0 \\ 0 \end{matrix}$$

$$c = 1.2500 \; -3.0000 \; -6.2500 \; 1.3500$$

$$d = 1.500$$

This is the controllable canonical form realization in Example 10.3.1. In using tf2ss, there is no need to normalize the leading coefficient and to carry out direct division. Thus its use is simple and straightforward.

10.3.1 Minimal Realizations

Consider a proper positive-power rational transfer function $H(z) = N(z)/D(z)$. If $N(z)$ and $D(z)$ are coprime (have no common factors), then the degree of $H(z)$ is defined as the degree of $D(z)$. For a degree N proper transfer function, if the dimension of a realization is N, then the realization is called a *minimal-dimensional realization* or, simply, a *minimal realization*.

Consider the transfer function in (10.19). Typing roots([3 7 0 -2 10]) in MATLAB yields the four roots of the numerator as $\{-1.78 \pm j0.47, 0.61 \pm j0.78\}$. Typing roots([2 3 4 7 5]) yields the four roots of the denominator as $\{0.3 \pm j1.39, -1.05 \pm j0.57\}$. We see that the numerator and denominator of (10.19) have no common roots; thus the transfer function in (10.19) has degree 4. Its two realizations in Example 10.3.1 have dimension 4; thus they are minimal realizations.

For a transfer function of degree N, it is not possible to find a realization of a dimension less than N. However, it is quite simple to find a realization of a dimension larger than N, as the next example illustrates.

EXAMPLE 10.3.2

Consider the transfer function

$$H(z) = \frac{Y(z)}{U(z)} = \frac{2}{z^2 + 0.3z + 0.8} \tag{10.20}$$

It has degree 2 and the following two-dimensional ss equation

$$\begin{aligned} \mathbf{x}[n+1] &= \begin{bmatrix} -0.3 & -0.8 \\ 1 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u[n] \\ y[n] &= [0 \quad 2]\mathbf{x}[n] \end{aligned} \tag{10.21}$$

is a minimal realization.

Now if we multiply the numerator and denominator of (10.20) by $(z-1)$ to yield

$$H(z) = \frac{2(z-1)}{(z^2 + 0.3z + 0.8)(z-1)} = \frac{2z - 2}{z^3 - 0.7z^2 + 0.5z - 0.8} \tag{10.22}$$

then we can obtain, using (10.12),

$$\begin{aligned} \mathbf{x}[n+1] &= \begin{bmatrix} 0.7 & -0.5 & 0.8 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u[n] \\ y[n] &= [0 \quad 2 \quad -2]\mathbf{x}[n] \end{aligned} \tag{10.23}$$

This is a nonminimal realization of (10.20) or (10.22).

Transfer functions are used to develop general properties of systems and to carry out design. State-space equations, as we will discuss in the next section, can be easily used to carry out computer computation. If we use a nonminimal realization, then it requires more computation and, consequently, incur more numerical errors. Thus there is no reason to use a nonminimal realization. When we are given a transfer function, if we first cancel out all common factors (if there are any) and then carry out realization, the resulting ss equation is automatically minimal.

10.4 MATLAB COMPUTATION

This section discusses computer computation of DT systems. We first use an example to discuss the reason of using ss equations. Consider the second-order difference equation

$$y[n+2] + 0.2y[n+1] + 0.82y[n] = 2.5u[n+2] - 0.87u[n+1] + 0.8336u[n] \quad (10.24)$$

Suppose we are interested in the output $y[n]$, for $0 \leq n \leq 100$, excited by the input $u[n] = 1/(n+1)$, for $n \geq 0$, and some initial conditions. The initial conditions are assumed to be $y[-1] = 2$, $y[-2] = -1$, $u[-1] = 0$, and $u[-2] = 0$. Because (10.24) involves only additions and multiplications, its solution can be obtained by direct substitution. First we write it as

$$y[n+2] = -0.2y[n+1] - 0.82y[n] + 2.5u[n+2] - 0.87u[n+1] + 0.8336u[n]$$

or

$$y[n] = -0.2y[n-1] - 0.82y[n-2] + 2.5u[n] - 0.87u[n-1] + 0.8336u[n-2]$$

Then we have

$$\begin{aligned}
n = 0: \quad y[0] &= -0.2y[-1] - 0.82y[-2] + 2.5u[0] - 0.87u[-1] + 0.8336u[-2] \\
&= -0.2 \cdot 2 - 0.82 \cdot (-1) + 2.5 \cdot 1 - 0.87 \cdot 0 + 0 = 2.92 \\
n = 1: \quad y[1] &= -0.2y[0] - 0.82y[-1] + 2.5u[1] - 0.87u[0] + 0.8336u[-1] \\
&= -0.2 \cdot 2.92 - 0.82 \cdot 2 + 2.5 \cdot 0.5 - 0.87 \cdot 1 + 0 = -1.844 \\
&\vdots
\end{aligned}$$

Proceeding forward, we can easily compute $y[n]$ recursively from $n = 0$ to 100. Although the preceding computation can be directly programmed, its programming is more complex than the one using an ss equation. Thus high-order difference equations are not directly programmed. Applying the z-transform to (10.24) and assuming zero initial conditions, we obtain

$$\begin{aligned}
H(z) = \frac{Y(z)}{U(z)} &= \frac{2.5z^2 - 0.87z + 0.8336}{z^2 + 0.2z + 0.82} \\
&= \frac{-1.37z - 1.2164}{z^2 + 0.2z + 0.82} + 2.5 \qquad (10.25)
\end{aligned}$$

Then we can obtain its controllable canonical form realization as

$$\begin{aligned}
\mathbf{x}[n+1] &= \mathbf{A}\mathbf{x}[n] + \mathbf{b}u[n] \\
y[n] &= \mathbf{c}\mathbf{x}[n] + du[n]
\end{aligned}$$

with

$$\begin{aligned}
\mathbf{A} &= \begin{bmatrix} -0.2 & -0.82 \\ 1 & 0 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
\mathbf{c} &= [-1.37 \quad -1.2164] \qquad d = 2.5
\end{aligned} \qquad (10.26)$$

Then its programming becomes

```
Given A, b, c, d, x(0), u(n), for n≥0,
Compute, from n=0 to N
y(n)=c*x(n)+d*u(n)
x(n+1)=A*x(n)+b*u(n)
```

This is simpler and more systematic than the programming using high-order difference equations. Thus DT ss equations are the choice of programming. Its actual programming depends on the language used such as FORTRAN, C, or C++ and is outside the scope of this text. We discuss in the following only the use of MATLAB functions.

In MATLAB, matrices are expressed row by row separated by semicolons. For example, the $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}$, and d in (10.26) are expressed in MATLAB as `a=[-0.2 -0.82;1 0]; b=[1;0]; c=[-1.37 -1.2164]; d=2.5`. Note that `b` has a semicolon because it has two rows, whereas `c` has no semicolon because it is a row vector. Associated with every DT system, there is a sampling period T. Let us select $T = 1$. To compute the output of the system excited by the initial state $x[0] = [2\ \ -4]'$ and the input $u[n] = 1/(n+1)$ for $n = 0 : 40$, we may use the MATLAB function `lsim`, an acronym for *linear simulation*. Let us type

```
a=[-0.2 -0.82;1 0];b=[1;0];c=[-1.37 -1.2164];d=2.5;T=1;
dog=ss(a,b,c,d,T);
n=0:40;t=n*T;
u=1./(n+1);x0=[2;-4];
[y,t]=lsim(dog,u,n,x0);
subplot(3,1,1)
stem(t,y)
subplot(3,1,2)
stairs(t,y)
subplot(3,1,3)
lsim(dog,u,n,x0)
```

The first line expresses $\{\mathbf{A}, \mathbf{b}, \mathbf{c}, d\}$ in MATLAB format. The second line defines the system. We call the system dog, which is defined using the state-space model, denoted by `ss`. It is important to have the fifth argument T inside the parentheses. *Without a T, it defines a CT system.* The third line denotes the number of samples to be computed and the corresponding time instants. The fourth line is the input and initial conditions. Note the use of dot slash (./). Most operations in MATLAB are defined for matrices; with a dot, an operation becomes element by element. We then use `lsim` to compute the output. The outputs are plotted in Figure 10.6(a) using the function `stem` and in Figure 10.6(b) using the function `stairs`. The actual output of the DT system is the one shown in Figure 10.6(a). The output shown in Figure 10.6(b) is actually the output of the DT system followed by a zero-order hold. A zero-order hold holds the value constant until

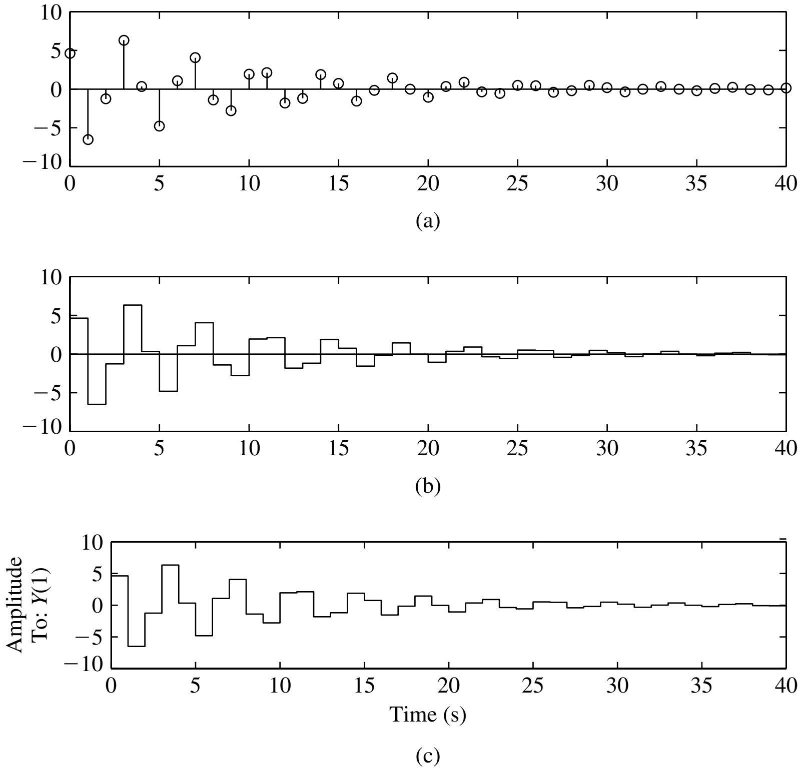


Figure 10.6 Output of (10.26): (**a**) Actual output. (**b**) and (**c**) Outputs of zero-order holds.

the arrival of the next value. If it is understood that the output of the DT system consists of only the values at the sampling instants, then Figure 10.6(b) is easier for viewing than Figure 10.6(a), especially if the sampling period is very small. Thus all outputs of DT systems in MATLAB are plotted using the function stairs instead of stem.

If the function lsim does not have the left-hand argument as shown in the last line of the preceding program, then MATLAB automatically plots the output as shown in Figure 10.6(c). We mention that the function lsim is developed for multi-output systems. Our system has only one output; thus Figure 10.6(c) has $Y(1)$ (first output) in the vertical ordinate of the plot.

Before proceeding, we mention that a DT system can also be defined using the transfer function model. However, if we use a transfer function, all initial conditions are automatically assumed to be zero. Therefore we cannot use the transfer function model to generate the plots in Figure 10.6. If we compute zero-state responses, there is no difference in using the state-space or transfer-function model. If we use the latter, MATLAB first transforms, using tf2ss, the transfer function into the controllable canonical form ss equation and then carries out computation.

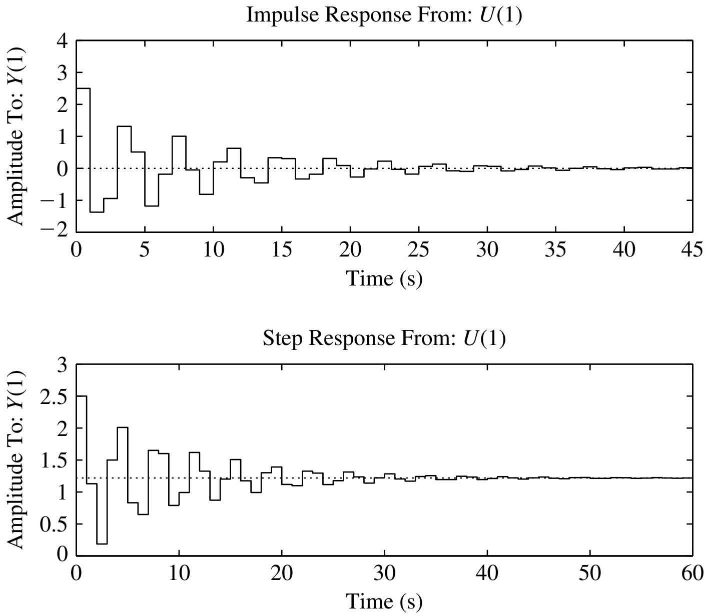


Figure 10.7 Impulse and step responses of (10.25).

To compute the impulse and step responses of (10.25) or (10.26), we can use the functions impulse and step. Typing

```
a=[1.8 1.105;-4 -2];b=[0.5;-0.8];c=[1.1 2.4];d=2.5;T=1;
sys=ss(a,b,c,d,T);
[or nu=[2.5 -0.87 0.8336];de=[1 0.2 0.82];sys=tf(nu,de,T);]
subplot(2,1,1)
impulse(sys)
subplot(2,1,2)
step(sys)
```

generates the impulse and step responses in Figure 10.7. Note that the program calls the DT system sys. In the program, we did not specify the number of samples to be computed. This is selected automatically by the functions. Unlike the CT case where the impulse can be used only if the direct transmission gain is zero, there is no such restriction in the DT case.

To conclude this section, we mention that MATLAB contains the functions dimpulse and dstep, which compute the impulse and step responses of DT systems. The first character “d” in dimpulse and dstep denotes discrete time. If we use impulse and step to compute the response of a DT system, then discrete time is embedded in defining the system by including the sampling period in the argument. We also mention that the MATLAB function filter can also be used to compute the response of DT systems. The inputs of filter, however, are the coefficients of the numerator and denominator of a negative-power transfer function. In conclusion, there are many MATLAB functions that can be used to compute responses of DT systems.

10.4.1 MATLAB Computation of Convolutions

This subsection discusses how the discrete convolution of two finite sequences is computed in MATLAB. Given two sequences of finite lengths. We call the shorter sequence $h[n]$ and call the longer one $u[n]$. Let

$$h[n] \qquad \text{for } n = 0 : N$$

and zero for $n < 0$ and $n > N$. Here we implicity assume $h[N] \neq 0$. We will consider $h[n]$ to be the impulse response of an FIR filter of length $N + 1$. Let

$$u[n] \qquad \text{for } n = 0 : P$$

and zero for $n < 0$ and $n > P$. It is a finite sequence of length $P + 1$.

To simplify the discussion, we assume $N = 5$. Then the convolution of $h[n]$ and $u[n]$ is

$$\begin{aligned} y[n] &= \sum_{m=-\infty}^{\infty} h[m]u[n-m] = \sum_{m=0}^{N=5} h[m]u[n-m] \\ &= h[0]u[n] + h[1]u[n-1] + h[2]u[n-2] + h[3]u[n-3] \\ &\quad + h[4]u[n-4] + h[5]u[n-5] \end{aligned} \tag{10.27}$$

This equation holds for all n. But we have $y[n]$ for $n < 0$ and $n > N + P$. Thus $y[n]$ is a sequence of length at most $N + P + 1$.

The equation in (10.27) is in fact a delayed-form difference equation of order $N = 5$. This is consistent with our discussion in Section 3.6.1 that there is no difference between convolutions and nonrecursive difference equations. Thus the result of the convolution of $h[n]$ and $u[n]$ equals the output of the nonrecursive difference equation. Applying the z-transform and assuming zero initial conditions, we obtain the transfer function of (10.27) as

$$\begin{aligned} H(z) &= \frac{Y(z)}{U(z)} = h[0] + h[1]z^{-1} + h[2]z^{-2} + h[3]z^{-4} + h[4]z^{-4} + h[5]z^{-5} \\ &= \frac{h[1]z^4 + h[2]z^3 + h[3]z^2 + h[4]z + h[5]}{z^5} + h[0] \end{aligned} \tag{10.28}$$

It is a transfer function of degree $N = 5$ and has the following realization, using (10.12):

$$\begin{aligned} \mathbf{x}[n+1] &= \mathbf{A}\mathbf{x}[n] + \mathbf{b}u[n] \\ y[n] &= \mathbf{c}\mathbf{x}[n] + h[0]u[n] \end{aligned}$$

with $\mathbf{x} = [x_1 \;\; x_2 \;\; x_3 \;\; x_4 \;\; x_5]'$ and

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{10.29}$$

$$\mathbf{c} = [h[1] \quad h[2] \quad h[3] \quad h[4] \quad h[5]] \qquad d = [h[0]]$$

The output of this ss equation excited by u will yield the convolution of h and u. The only problem is that the program in MATLAB generates an output with the same length as u, whereas the convolution has length $N+P+1$. This can be resolved by padding N trailing zeros to the end of u. Then the output will yield the convolution of h and u. Another problem is to select a sampling period. It can be selected as any positive number or, simply, as 1.

EXAMPLE 10.4.1

Compute the convolution of $h = [-1\ \ 0\ \ 1\ \ 2]$ and $u = [-2\ \ -1\ \ 0\ \ 1\ \ 2]$ with $N = 3$ and $P = 4$. We use h to form an ss equation of dimension 3. Because the convolution has length $3+4+1 = 8$, we pad 3 trailing zeros to u to form $u_e = [u\ \ 0\ \ 0\ \ 0]$. Typing

```
a=[0 0 0;1 0 0;0 1 0];b=[1;0;0];c=[0 1 2];d=-1;
ue=[-2 -1 0 1 2 0 0 0];
sys=ss(a,b,c,d,1);
y=lsim(sys,ue)
```

in MATLAB yields 2 1 -2 -6 -4 1 4 4. This is the result shown at the end of Section 3.3.1.

To conclude this subsection, we mention that FFT can also be used to compute convolutions. Recall that FFT is an efficient way of computing (5.35). Given x, fft(x) yields X_d in (5.35). Conversely, inverse FFT computes x from X_d; that is, ifft(Xd) yields x. For two sequences h and u of lengths $N+1$ and $P+1$, if we select $L \geq N+P+1$, then the MATLAB program ifft(fft(h,L).*fft(u,L)) will yield the convolution of h and u. See Reference 3. For example, typing

```
h=[-1 0 1 2];u=[-2 -1 0 1 2];
ifft(fft(h,8).*fft(u,8))
```

yields 2 1 -2 -6 -4 1 4 4. This is the same result as computed above.

If N and P are very large, FFT computation of convolutions requires a smaller number of operations than the one using ss equations. However, if we include other factors such as memory locations, numerical errors, and real-time computation, then FFT computation is not as desirable as the one using ss equations. See Reference 3. In any case, the MATLAB function conv is based on ss equations, not on FFT.

10.5 COMPLETE CHARACTERIZATION BY TRANSFER FUNCTIONS

An LTI lumped system can be described by a transfer function and an ss equation. We now compare these two descriptions. The transfer function describes only the relationship between the input and output. It does not reveal the internal structure of the system and is therefore called an *external description*. It is applicable only if the system is initially relaxed or all initial conditions are zero. The ss equation describes not only the relationship between the input and output but also the internal variables of the system and is therefore called an *internal* *description*. Furthermore, it is applicable no matter the initial conditions are zero or not. It is then natural to ask, Does the transfer function describe fully the system or is there any information of the system missing in the transfer function? We use DT basic block diagrams to discuss the issue.

Consider a DT basic block diagram that consists of N unit delay elements. If its transfer function has degree $N-1$ or less, then the basic block diagram is not completely described by its transfer function. If we assign the output of each unit delay element as a state variable and develop an N-dimensional ss equation, then the ss equation will describe fully the basic block diagram. However, the ss equation will not have the so-called controllability and observability property and is not used in design. In this situation, the basic block diagram has some redundancy in unit delay elements as far as the input and output are concerned and should not be used in practice. On the other hand, if

$$\begin{aligned}\text{Dimension of ss equation} &= \text{Degree of its transfer function}\\ &= \text{No. of unit delay elements}\end{aligned}$$

then the system has no redundancy and is a good system. Furthermore, the basic block diagram is completely characterized by its transfer function. Its ss equation also has the controllable and observable properties. In this case, there is no difference in using the transfer function or the ss equation to study the system. However, the former is more convenient to carry out analysis and design, and the latter is more convenient to carry out computer computation. In practice, we use transfer functions to carry out design. We find a transfer function to meet a given specification and then implement it using its minimal realization. Such an implementation will not have any redundancy. Thus all practical systems, unless inadvertently implemented, are completely characterized by their transfer functions.

If a DT system is completely characterized by its transfer function, then every zero-input response (response excited by some initial conditions), as in the CT case discussed in Section 7.6.1, can be generated as a zero-state response. Thus all properties of zero-input responses can be deduced from zero-state responses. In this case, we can disregard zero-input responses and study only zero-state responses using transfer functions.

PROBLEMS

10.1 Develop the direct-form basic block diagram for

$$y[n] + 0.5y[n-1] + 0.7y[n-2] = -u[n] + 2u[n-1] + 0.5u[n-2]$$

and then its direct-form II basic block diagram.

10.2 If we assign the outputs of the unit delay elements in Figure 10.3(b) as, from top, $x_4[n]$, $x_3[n]$, $x_2[n]$, $x_1[n]$, what is its ss equation?

10.3 Find a realization for each of the following transfer functions.

(a) $H_1(z) = \dfrac{6z+3}{2z+9}$

(b) $H_2(z) = \dfrac{z^2-1}{z^3+3z^2+2z}$

(c) $H_3(z) = \dfrac{1}{2z^4+3z^3+2z^2+5z+10}$

Are all realizations minimal? If not, find a corresponding minimal realization.

10.4 Develop basic block diagrams for the systems in (b) and (c) of Problem 10.3.

10.5 Consider the basic block diagram in Figure 10.8. Assign the output of each unit delay element as a state variable and then develop an ss equation to describe it. Compare your result with the one in Problem 7.4.

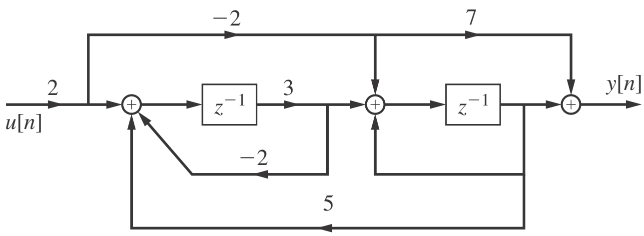


Figure 10.8

10.6 Use MATLAB to compute the step responses of a DT system with transfer function

$$H(z) = \frac{2z^3+z+1}{z^3-0.3z^2-0.23z+0.765}$$

and the sampling period $T = 0.5$. What is its steady-state response? Roughly how many seconds does it take to reach steady state?

10.7 Find the controllable-form realization of the transfer function in Problem 10.6 and then use the ss equation to compute its step response. Is the result the same as the one computed in Problem 10.6?

10.8 Write an ss equation to compute the convolution of [2 −3 −4 6] and ones(1,10). Is the result the same as the one computed using the MATLAB function `conv`?

ANSWERS TO SELECTED PROBLEMS

CHAPTER 1

1.5 $x_1(t) = r(-t+3) - r(-t+2) - q(-t) = q(t) - r(t-2) + r(t-3)$
$x_2(t) = 0.5[r(t) - r(t-2)] - [r(t-2) - r(t-3)] = 0.5[r(t) - r(t-2)]$
$\times [r(-t+3) - r(-t+2)]$

1.6 (a) $q(t) + q(t-2) - q(t-4)$
(b) $[r(t+3) - r(t+2)] - [r(t-2) - r(t-3)]$
(c) $-q(t+2) + 2q(t) - q(t-2)$
(d) $\sum_{k=-\infty}^{\infty} x(t-2k)$, where $x(t) = r(t) - r(t-2) - 2q(t-2)$ is the triangle in $[0, 2]$

1.8 (a) Neither even nor odd.
(b) Even.
(c) Odd.
(d) Neither even nor odd.

1.12 (a) $\delta(t) + \delta(t-2) - \delta(t-4)$
(b) $-\delta(t+2) + 2\delta(t) - \delta(t-2)$

1.14 (a) $\cos 3\pi = -1$
(b) 0
(c) $\cos(t+3)$
(d) 0
(e) $\cos(t+3)$

1.17 (a) $0.5r[-n+5] - 0.5r[-n+3] - q[-n-1]$
(b) $(r[n] - r[n-3])/3 - (r[n-3] - r[n-5])/2$

1.18 (a) $x_1[n] = \delta[n] + \delta[n-1] + \delta[n-2] + \delta[n-3] + 0.5\delta[n-4]$
(b) $x_2[n] = (1/3)\delta[n-1] + (2/3)\delta[n-2] + \delta[n-3] + 0.5\delta[n-4]$

1.21 $n = -5 : 5$
$y_1 = [0\ \ 0\ \ 0\ \ 1\ \ 2\ \ \underline{3}\ \ 4\ \ 5\ \ 5\ \ 5\ \ 5]$, where the underlined number indicates $n = 0$
$y_2[n] = 5$ for all n.

1.22 $x(t)$ is periodic with fundamental frequency 0.7 rad/s and fundamental period $2\pi/0.7$.

$x(t) = 2\sin(0\cdot t + \pi/2) + \sin(1.4t) + 4\sin(2.1t - \pi/2)$. It has frequencies 0, 1.4, 2.1 with corresponding magnitudes 2, 1, 4 and phases $\pi/2$, 0, $-\pi/2$.

$x(t) = 2\cos(0\cdot t) + \cos(1.4t - \pi/2) + 4\cos(2.1t + \pi)$. It has frequencies 0, 1.4, 2.1 with corresponding magnitudes 2, 1, 4 and phases 0, $-\pi/2$, π.

1.23 $x(t) = 2e^{j0}e^{j0\cdot t} + 0.5e^{-j\pi/2}e^{j1.4t} + 0.5e^{j\pi/2}e^{-j1.4t} + 2e^{j\pi}e^{j2.1t} + 2e^{j\pi}e^{-j2.1t}$

1.24 No. Yes.

1.25 It is periodic with fundamental frequency 0.7 rad/s and fundamental period $2\pi/0.7$.

$x(t) = 2.5e^{j\pi}e^{j0\cdot t} + 2.24e^{-j2.68}2^{j2.8t} + 2.24e^{j2.68}2^{-j2.8t} + 0.6e^{j4.9t} + 0.6e^{-j4.9t}$

1.26 $\sin 6.8\pi n$ is periodic with period $N = 5$ samples. $\sin 1.6\pi n$ is periodic with $N = 5$. $\cos 0.2n$, $\sin 4.9n$, and $-\sin 1.1n$ are not periodic. Their frequencies are not defined without specifying the sampling period.

1.27 Because $6.9\pi = 4.9\pi = 2.9\pi = 0.9\pi = -1.1\pi \pmod{2\pi}$, they all denote the same sequence. Its frequency is 0.9π rad/s.

1.28 For $T = 1$, the Nyquist frequency range (NFR) is $(-\pi, \pi] = (-3.14, 3.14]$. They have frequency 0.8π, 0.2, -1.38, -0.4π, and 1.1 rad/s, respectively.

1.29 The NFR for $T = 0.5$ is $(-2\pi, 2\pi]$. The frequency of $\sin 1.2\pi n = \sin 2.4\pi nT = \sin(-1.6\pi)nT$ is -1.6π rad/s.

The NFR for $T = 0.1$ is $(-10\pi, 10\pi]$. The frequency of $\sin 1.2\pi n = \sin 12\pi nT$ is -8π rad/s.

1.30 (a) 0
(b) π rad/s
(c) 2π rad/s

1.31 The NFR for $T = 1$ is $(-\pi, \pi]$. There are only N distinct $\cos\omega_k nT$, with frequencies

$$\omega_k = \frac{2\pi k}{N}\begin{cases} k = -N/2+1 : N/2 & \text{N even} \\ k = -(N-1)/2 : (N-1)/2 & \text{N odd} \end{cases}$$

1.32 The frequencies of the sampled sequence $\sin 4nT$, for $T = 0.3$, 0.6, 0.9, and 1.2, are 4, 4, -2.98, and -1.24, respectively. If $T < \pi/4 = 0.785$, then the frequency of $\sin 4nT$ is 4 rad/s.

1.33 The frequency of $\sin\pi nT$ is π for $T = 0.5, 0.9$, and $-\pi/3$ for $T = 1.5$. For $T = 1$, we have $\sin\pi nT = 0$, for all n, thus its frequency is 0. If $T < \pi/\pi = 1$, then the frequency of $\sin\pi t$ equals the frequency of its sampled sequence.

1.34 If $T = \pi/60$, then NFR $= (-60, 60]$ and $\sin 70nT = \sin(-50)nT = -\sin 50nT$. Thus we have

$$x(nT) = \sin 50nT + 2\sin 70nT = -\sin 50nT$$

and $\sin 70t$ cannot be detected from $x(nT)$.

If $T = \pi/45$, then NFR $= (-45, 45]$ and

$$x(nT) = -\sin 40nT - 2\sin 20nT$$

Thus $\sin 50t$ and $\sin 70t$ cannot be detected from $x(nT)$.

If $T = \pi/180$, then NFR $= (-180, 180]$ and

$$x(nT) = \sin 50nT + 2\sin 70nT$$

which yields $x(t)$ if nT is replaced by t.

1.35 If $T = \pi/40$, then NFR $= (-45, 45]$ and

$$x(nT) = \cos 40nT - 2\sin 20nT$$

If $T = \pi/60$, then NFR $= (-60, 60]$ and

$$x(nT) = \cos 50nT - 2\sin 50nT$$

If $T = \pi/180$, then NFR $= (-180, 180]$ and

$$x(nT) = \cos 50nT + 2\sin 70nT$$

If $T < \pi/70$, then all frequencies of $x(t)$ are retained in $x(nT)$.

1.36 For $T = 1$, NFR $= (-\pi, \pi]$ and $4.9 = -1.38$ (mod $2\pi = 6.28$). Thus we have

$$x(nT) = -2.5 + 2\sin 2.8nT - 4\cos 2.8nT + 1.2\cos 1.38nT$$

which does not contain all frequencies of $x(t)$.

1.37 For $T = 0.5$, NFR $= (-2\pi, 2\pi] = (-6.28, 6.28]$:

$$x(nT) = -2.5 + 2\sin 2.8nT - 4\cos 2.8nT + 1.2\cos 4.9nT$$

which contains all frequencies of $x(t)$.

CHAPTER 2

2.1 (a) Incorrect.
(b) Correct.
(c) Incorrect.

Note that initial conditions must also meet the additivity and homogeneity properties.

2.2 All are correct.

2.5 (a) Memoryless, nonlinear, time-invariant, causal.
(b) Memoryless, nonlinear, time-invariant, causal.
(c) With memory, nonlinear, time-invariant, causal.
(d) Memoryless, linear, time-varying, causal.
(e) With memory, linear, time-invariant, causal.
(f) With memory, linear, time-varying, causal.

2.6 Yes. Yes. No. Yes.

2.7 It is time-varying. Thus time shifting cannot be used. Because $u_3 = 2u_1 - u_2$, we have $y_3 = 2y_1 - y_2$. Although $u_4(t) = u_1(t-1)$ and $u_5(t) = u_1(t-3)$, we don't know what y_4 and y_5 are.

2.8 Because $u_2(t) = 2u_1(t) + 2u_1(t-1) - 0.5u_1(t-2)$, we have $y_2(t) = 2y_1(t) + 2y_1(t-1) - 0.5y_1(t-2)$.

2.9 y_1 and y_2 are plotted in Figure P2.9.
$y_3(t) = 0.5y_1(t) + 0.5y_1(t-1)$

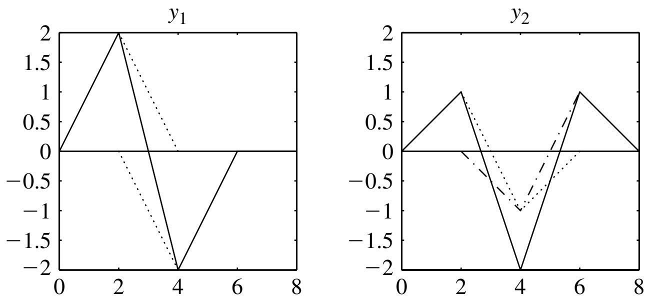


Figure P2.9

2.10 Because $y(t) = u(t)/R$, the input and output are linearly related.

2.11 $v_o = [10^5/(10^5 - 1)]v_i \approx v_i$

2.14 Using $i_- = -i_+ = 0$, we have $e_- = (R_1/2R_1)v_o = 0.5v_o$ and $i = (v - v_o)/R$. Using $e_- = e_+ = v$, we have $v_o = 2v$. Thus we have $i = -v/R$ or $v = -Ri$.

2.15 $y_1[n] = 2\delta[n] + 4\delta[n-1] + 6\delta[n-2] + 4\delta[n-3] + 2\delta[n-4]$
$y_2[n] = 2\delta[n] + 2\delta[n-2] + 2\delta[n-4]$

2.16 $q[n] = u[n] + u[n-3] + u[n-6] + \cdots$ and $\delta[n] = q[n] - q[n-1]$. Thus $y_q[n] = y[n] + y[n-3] + y[n-6] + \cdots$ and $h[n] = y_q[n] - y_q[n-1]$. They are plotted in Figure P2.16.

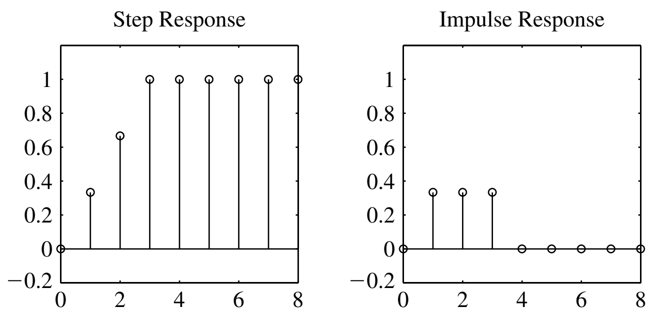


Figure P2.16

2.17 Figure 2.6(b) can be described by $y = au + b$. If we define $\bar{y} = y - b$, then it becomes $\bar{y} = au$ which is a linear equation.

CHAPTER 3

3.1 10889.05

3.2 The result is $[0\ \ 0\ \ \underline{1}\ \ 2\ \ 1\ \ -2\ \ 6\ \ 29\ \ 14\ \ 7\ \ -10\ \ 0\ \ 0]$, where the underlined number indicates $n = 0$.

3.3 $h(s)u(s) = s^8 + 2s^7 + s^6 - 2s^5 + 6s^4 + 29s^3 + 14s^2 + 7s - 10$

3.4 $n = 0:3$
$y = [2\ \ 3.44\ \ 4.87\ \ 6.45]$

3.5 $y[n+1] - 1.22y[n] = 2u[n+1]$

3.6 $y[n+1] - y[n] = u[n]$

3.7 Third order. But it can be reduced to second order as

$$2y[n+2] - y[n+1] + 3y[n] = u[n+2] + 2u[n]$$

3.8 $2y[n] - y[n-1] + 3y[n-2] = u[n] + 2u[n-2]$

3.9 Yes. Order 3.

3.10 $h[0] = 0$, $h[1] = 1$, $h[2] = 1$, and $h[n] = 0$, for $n \geq 3$. It is FIR with length 3. The equation can be reduced as

$$y[n] = u[n-1] + u[n-2]$$

It is a nonrecursive difference equation of order 2.

3.11 $h[0] = 0$, $h[1] = 1$, $h[2] = 1$, $h[3] = -4$, $h[4] = 8$, $h[5] = -16, \ldots$. It is IIR.

3.12 $h[n] = 1/20 = 0.05$ for $n = 0:19$
$y[n] = 0.05\,(u[n] + u[n-1] + u[n-2] + \cdots + u[n-19])$
$y[n] - y[n-1] = 0.05(u[n] - u[n-20])$
This recursive difference equation requires less computation.

3.13 $h(t) = \sum_{n=1}^{\infty} a^n \delta(t-n)$

3.14 $h(t) = \sum_{n=1}^{\infty} (-1)^{n-1} a^n \delta(t-n)$

3.15 They are plotted in Figure P3.15.

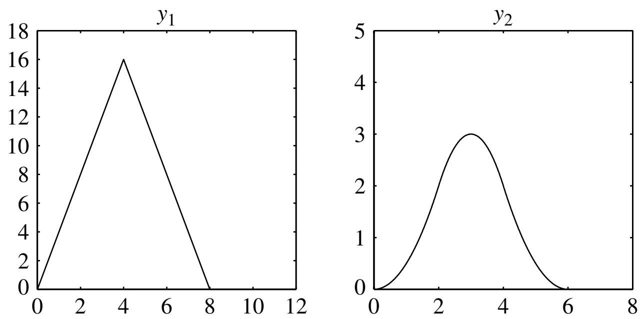


Figure P3.15

3.16 $y_1(t) = 2f_1(t-2)$
$y_2 = 2$

3.17 $y_2 = 2$ and $y_1(t)$ is plotted in Figure P3.17.

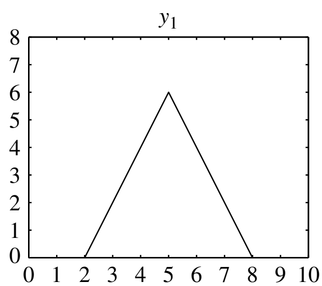


Figure P3.17

3.18 $4\ddot{y}(t) + \dot{y}(t) + 2y(t) = \dot{u}(t)$

3.19 $ml\ddot{\theta} + mg\sin\theta = u\cos\theta$

3.20 $v_o(t) = -RC\dot{v}_i(t)$

CHAPTER 4

4.1 Fundamental frequency 3 rad/s. Fundamental period $2\pi/3$.

$$\begin{aligned} x(t) = 3e^{j0} &+ 1.12e^{-j2.68}e^{j6t} + 1.12e^{j2.68}e^{-j6t} \\ &+ 1.57e^{-j\pi/2}e^{j9t} + 1.57e^{j\pi/2}e^{-j9t} \\ &+ 0.5e^{-j\pi}e^{j12t} + 0.5e^{j\pi}e^{-j12t} \end{aligned}$$

4.2 Fundamental period $\pi/2$. Fundamental frequency 4 rad/s.

$$y(t) = \sum_{m=-\infty}^{\infty} \frac{2}{\pi(1-4m^2)} e^{j4mt}$$

4.3 $\lim_{a\to 0}(1/Ta) = \infty$

4.5 Its total energy is infinity. Its average power is 16.92. 68%.

4.6 0.5, 0.25, 50%

4.7 0.5, 0.5, 100%

4.8 The SNR at the input is 10000. The SNR at the output is 0.0001.

4.9 All c_m with m odd are zero.

4.11 (a) Not defined
(b) $1/(j\omega + 2)$
(c) $4/(\omega^2 + 4)$
(d) $(10 + j\omega)/(12 + j\omega + \omega^2)$

4.14 $A\left[\frac{\sin a(\omega - \omega_c)}{\omega - \omega_c} + \frac{\sin a(\omega + \omega_c)}{\omega + \omega_c}\right]$

4.15 $0.7(j\omega + 10.1)/(-\omega^2 + j0.2\omega + 100.01)$

4.17 $275.6/2\pi$

4.18 $500/2\pi$

4.20 $X(\omega) = 2\pi[3\delta(\omega) + 1.12e^{-j2.68}\delta(\omega - 6) + 1.12e^{j2.68}\delta(\omega + 6) + 1.57e^{-j\pi/2}\delta(\omega - 9) + 1.57e^{j\pi/2}\delta(\omega + 9) + 0.5e^{-j\pi}\delta(\omega - 12) + 0.5e^{j\pi}\delta(\omega + 12)]$

4.21 $\frac{1}{j\omega + 1} - j3\pi\delta(\omega - 2) + j3\pi\delta(\omega + 2)$

CHAPTER 5

5.1 It does not exist. The sequence is not absolutely summable.

5.2 $-1 + 4\cos\omega$. Real-valued and periodic with period 2π.

5.3 $X_d(\omega) = e^{-j\omega}(-1 + 4\cos\omega)$. Its magnitude spectrum is the same as the one in Problem 5.2. Its phase spectrum differs from the one in Problem 5.2 by the linear phase $-\omega$.

5.4 $4j\sin 0.5\omega$. Periodic with period 4π.

5.5 $X_d(\omega) = e^{-j0.4\omega}[1 + 2\cos 0.2\omega + 2\cos 0.4\omega]$ or $e^{-j0.4\omega}\sin 0.5\omega/\sin 0.1\omega$

5.6 $H_d(\omega) = 1/(1 - 0.8e^{-j0.2\omega})$

5.7 $H_d(\omega) = 1/(1 + 0.8e^{-j0.2\omega})$

5.9 $X(\omega) = \pi\delta(\omega - 40) + \pi\delta(\omega + 40) - \pi\delta(\omega - 60) - \pi\delta(\omega + 60)$
$T_1 = \pi/100$, NFR $= (-100, 100]$, $X_d(\omega) = 100\delta(\omega - 40) + 100\delta(\omega + 40) - 100\delta(\omega - 60) - 100\delta(\omega + 60)$

5.10 $T_2 = \pi/50$, $x(nT_2) = 0$, $X_d(\omega) = 0$
$T_3 = \pi/30$, $x(nT_3) = \cos 20nT_3 - 1$, $X_d(\omega) = 30\delta(\omega - 20) + 30\delta(\omega + 20) - 60\delta(\omega)$

5.11 $X(\omega) = -j\pi\delta(\omega - 40) + j\pi\delta(\omega + 40) + \pi\delta(\omega - 60) + \pi\delta(\omega + 60)$
$T_1 = \pi/100$, $X_d(\omega) = -j100\delta(\omega - 40) + j100\delta(\omega + 40) + 100\delta(\omega - 60) + 100\delta(\omega + 60)$
$T_2 = \pi/50$, $X_d(\omega) = 70.7e^{-j\pi/4}\delta(\omega - 40) + 70.7e^{j\pi/4}\delta(\omega + 40)$
$T_3 = \pi/30$, $x(nT_3) = -\sin 20nT_3 + 1$, $X_d(\omega) = 30j\delta(\omega - 20) - 30j\delta(\omega + 20) + 60\delta(\omega)$

5.12 $X(\omega) = 1/(j\omega + 0.05)$. $X_d(\omega) = 1/(1 - e^{-0.05T}e^{-j\omega T})$

5.14 $T = 1$; $x = [2\ \ -1\ \ 2]$; fft(x) $= [3\ \ 1.5 + j2.6\ \ 1.5 - j2.6]$. They are located at $\omega = [0\ \ 2\pi/3\ \ 4\pi/3]$. The values of $X_d(\omega)$ in Problem 5.3 at these frequencies are the same as the output of fft.

5.15
```
x=[2 -1 2];T=1;N=8;D=2*pi/(N*T);X=fft(x,N);
mp=0:N/2;plot(mp*D,abs(X(mp+1)),mp*D,angle(X(mp+1)))
```

5.16 Replace $N = 8$ by $N = 1024$ in the program in Problem 5.15.

5.17 $T = 0.2$; $x = [1\ \ 1\ \ 1\ \ 1\ \ 1]$; fft(x) $= [5\ \ 0\ \ 0\ \ 0\ \ 0]$. They are located at $\omega = [0\ \ 2\pi\ \ 4\pi\ \ 6\pi\ \ 8\pi]$. The values of $X_d(\omega)$ in Problem 5.5 at these five frequencies are the same as the output of fft.

5.18
```
x=[1 1 1 1];T=0.2;N=8;D=2*pi/(N*T); Xd=fft(x,N);mp=0:N/2;
plot(mp*D,abs(Xd(mp+1)),mp*D,angle(Xd(mp+1)),':')
```

5.19 Replace $N = 8$ by $N = 1024$ in the program in Problem 5.18.

5.20
```
T=0.1;N=8192;D=2*pi/(N*T);n=0:N-1;t=n*T;
x=exp(-0.01*t).*sin(5*t);X=T*fft(x);
mp=0:N/2;plot(mp*D,abs(X(mp+1)))
```

CHAPTER 6

6.1 $(2s^2 + 2s + 12)/[(s + 2)(s^2 + 4)]$

6.2 (a) $(s^2 - s - 2)/(2s^2 + 4s + 10)$
(b) $(s^2 + 3s + 2)/(2s^3 + 4s^2 + 10s)$
(c) $(s^3 + 1)/(s^5 + 2)$

6.3 $v^{(4)}(t) + 3v^{(3)}(t) + 10v(t) = 2\ddot{r}(t) + 5\dot{r}(t) + 3r(t)$

6.4 $10/(s + 1)$

6.5 $(4s^2 + 1.9s + 9.04)/[(s^2 + 0.4s + 9.04)(s + 0.5)]$

6.6 $Y(s)/U(s) = [6s^2(3s + 2)]/[18s^3 + 72s^2 + 26s + 4]$
$I(s)/U(s) = [9s^2 + 6s + 1]/[18s^3 + 72s^2 + 26s + 4]$

6.8 (a) $1.5(s - 2)/(s + 2)^2$
(b) $0.5(s - 2)/[(s + 2 + j)(s + 2 - j)]$

Note that -1 is neither a pole nor a zero.

6.9 22 minutes 20 seconds

6.10 $h(t) = -2e^{-t} + (7/3)e^{-2t} + (2/3)e^{t}$, for $t \geq 0$
$y_q(t) = 2e^{-t} - (7/6)e^{-2t} + (2/3)e^{t} - 1.5$, for $t \geq 0$

6.11 $\omega_r = \sqrt{1 - \sigma^2}$

6.12 $h(t) = -0.4e^{-2t} + 0.2e^{-t}\sin 3t + 0.4e^{-t}\cos 3t$

6.13 $k_1 + k_2e^{-t} + k_3te^{-t} + k_4t^2e^{-t} + k_5e^{-3t}$

6.14 $k_1e^{-2t} + k_2te^{-2t} + k_3t^2e^{-2t} + k_4t^3e^{-2t} + k_5e^{-0.1t} + k_6e^{-t}\sin(3t + k_7) + 20$, for $t \geq 0$

6.15 $k_1e^{-t} + k_2\sin(10t + k_3) + k_4t\sin(10t + k_5)$. It grows unbounded. Theorem 6.4 does not hold because the system is not stable.

6.16 $k_1e^{-t} + k_2\sin(10t + k_3) + k_4\sin(2t + k_5)$. It consists of two sinusoids with frequency 10 and 2. Theorem 6.4 does not hold because the system is not stable.

6.19 The systems in Problems 6.10 and 6.15 are not stable. The rest are stable.

6.20 (a) Not stable.
(b) Not stable.
(c) Not stable.
(d) Stable.

6.23 $H(0) = 0.02e^{j\pi}$, $H(j5) = 0.07e^{j1.88}$, $H(j10) = 1.02e^{j0.23}$, $H(j20) = 0.07e^{-j1.4}$, $H(j100) = 0.01e^{-j\pi/2}$

6.24 $y_{ss}(t) = -0.04 + 0.07\sin(5t + 1.88) + 3.06\cos(10t + 0.23) - 0.02\sin(100t - 1.57) \approx 3.06\cos(10t + 0.23)$. 10 seconds. It is a bandpass filter.

6.25 The systems in Problems 6.12 and 6.13 both have time constant 1 and take 5 seconds to reach steady state. The system in Problem 6.14 has time constant 10 and takes 50 seconds to reach steady state. The time constants of the systems in Problems 6.10 and 6.15 are not defined.

6.26 Zero-input response $= 0.5e^{-0.4t}$. Zero-state response $= 5 - 5e^{-0.4t}$. Transient response $= -4.5e^{-0.4t}$. Steady-state response $= 5$.

6.27 Zero-input response $= 0$. Transient response $= -5e^{-0.4t}$.

6.28 (a) $H(s) = 10/(s + 10)$. $y_{ss}(t) = \sin 0.1t + 0.1\sin(100t - 1.57)$. Lowpass. Bandwidth = 10 rad/s.
(b) $H(s) = s/(s + 10)$. $y_{ss}(t) = 0.01\sin(0.1t + 1.57) + \sin(100t)$. Highpass. Bandwidth $= \infty$. Passband edge frequency = 10 rad/s.

6.29 $H(s) = 1/(RCs + 1)$. Time constant $= RC$.

6.31 $y(t) = 1 - e^{-t} - te^{-t} - 0.5t^2e^{-t}$. Transient response $= y_{tr}(t) = -e^{-t} - te^{-t} - 0.5t^2e^{-t}$. Time constant = 1. $|y_{tr}(0)| = 1$, $|y_{tr}(5)| = 0.125 > 0.01|y_{tr}(0)|$, and $|y_{tr}(9)| = 0.006 < 0.01|y_{tr}(0)|$.

6.33 $H(0) = k/a$ and $H(j\omega) \geq 0.707H(0)$ for all $|\omega| \leq a$

6.34 (a) $6/(\omega^2 + 9)$
(b) $5/(\omega^2 - j\omega + 6)$
(c) $-20/(-\omega^2 + 2j\omega + 101)$

6.35 The result is 0. The correct spectrum is $2\pi\delta(\omega)$.

CHAPTER 7

7.1 (a) Controllable form: $\dot{x} = -3x + u$, $y = -5x + 2u$
Observable form: $\dot{x} = -3x - 5u$, $y = x + 2u$
(b) Controllable form:

$$\dot{\mathbf{x}} = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u$$

$$y = [-3 \quad 0]\mathbf{x} + u$$

(c) Controllable form:

$$\dot{\mathbf{x}} = \begin{bmatrix} -1.5 & -1 & -2.5 & -5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}u$$

$$y = [0 \quad 0 \quad 0 \quad 0.5]\mathbf{x}$$

7.2 It has degree 2. Its minimal realization is

$$\dot{\mathbf{x}} = \begin{bmatrix} -2 & -1 \\ 1 & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u$$
$$y = [1 \quad -2]\mathbf{x}$$

7.4 Let x_1 (x_2) be the output of the left-hand (right-hand) side integrator. Then we have

$$\dot{\mathbf{x}} = \begin{bmatrix} -6 & 5 \\ 3 & 1 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 2 \\ -4 \end{bmatrix}u$$
$$y = [0 \quad 1]\mathbf{x} - 28u$$

7.5 $\dot{x}(t) = 2\cos t + 10\cos 100t$
$\int x(t)\,dt = -2\cos t - 0.001\cos 100t$

7.8 Typing

```
a=[-3 -2;1 0];b=[1;0];c=[0 2];d=0;a1=[-2 1 2;1 0 0;0 1 0];
b1=[1;0;0];c1=[0 2 -2];d1=0;step(a,b,c,d)
hold on,step(a1,b1,c1,d1)
```

yields the same results. Note the use of semicolons and commas.

7.9
```
a=[1.8 1.105;-4 -2];b=[0.5;-0.8];c=[1.1 2.4];d=2.5;
sys=ss(a,b,c,d);x0=[1;-2];t=0:0.01:40;u=sin(0.1*t.*t);
lsim(sys,u,t,x0).
```

7.10
```
a=[-3.125 -0.875;1 0];b=[1;0];c=[0 0.5];d=0;
sys=ss(a,b,c,d);t=0:0.01:18;u=[zeros(1,1801)];
subplot(2,1,1),impulse(sys)
subplot(2,1,2),lsim(sys,u,t,b)
```

7.11 1. $Z_2(s)$. No.

7.12 $1/[1 + Z_1(s) + Z_2(s)]$, $Z_2(s)/[1 + Z_1(s) + Z_2(s)]$. Yes.

7.13 $H(s) = 1$

7.15 (a) Let x_1 be the capacitor voltage and x_2 be the inductor current. Then we have

$$\dot{\mathbf{x}} = \begin{bmatrix} -1/6 & -1/3 \\ 2/3 & -2/3 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1/6 \\ 1/3 \end{bmatrix}u$$
$$y = [-1/3 \quad -2/3]\mathbf{x} + (1/3)u$$

(b) Let x_1 be the voltage across the 3-F capacitor and let x_2 be the voltage across the 1-F capacitor. Let x_3 be the current through the inductor. Then we have

$$\dot{\mathbf{x}} = \begin{bmatrix} -1/6 & -1/6 & 0 \\ -1/2 & -1/2 & -1 \\ 0 & 1/2 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1/3 \\ 1 \\ 0 \end{bmatrix} u$$

$$y = [0 \quad 1 \quad 0]\mathbf{x}$$

7.16 $\dot{x} = -(1/3)x + (1/3)u$
$y = x$

7.17 Let x_1 be the inductor current and x_2 be the capacitor voltage. Then we have

$$\dot{\mathbf{x}} = \begin{bmatrix} -1/3 & -1/3 \\ 1/2 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1/3 \\ 0 \end{bmatrix} u$$

$$y = [0 \quad 1]\mathbf{x}$$

7.21 The corner frequency is 0.5. The gain at low frequency is 1. Thus the transfer function is $1/(1 + s/0.5) = 1/(1 + 2s)$.

CHAPTER 8

8.1 $H(s) = 10^5/(s + 10^5)$, its operational frequency range is $[0, 10^5]$ rad/s.

8.2 $H(s) = 10^8/(11s + 553.3 - 10^7)$. It is unstable and its step response grows unbounded. It cannot be used as an inverting amplifier.

8.3 $H_i(s) = 11$. Its step response is 11, for $t \geq 0$.
$H(s) = (11 \times 10^7)/(11s + 10^7)$. Its time constant is 1.1 μs, and its step response reaches $H(0) = 11$ in 5.5 μs. $H(s)$ can be reduced as $H_r(s) = 11$ in the operational frequency range $[0, 0.91 \times 10^6]$.

8.4 $\dot{x} = -20x + u;\ y = -800x + 40u$. $R_2C = 2$ and $R_1C = 1/20$.

8.5 $0.1a$. Yes. 0.2. 1 s.

8.6 If $f = 2$, then its operational frequency range is $[4, \infty)$, which is smaller than $[1.25, \infty)$.

8.7 If $f = 0.2$, then its operational frequency range is $[0.4, \infty)$, which is larger than $[1.25, \infty)$.

8.8 It follows from $H_1(s)H_2(s) = H_2(s)H_1(s)$. If $H_i(s)$ are matrices, generally $H_1(s)H_2(s) \neq H_2(s)H_1(s)$.

8.9 $2t \cos 2t$. $\sin 2t + 2t \cos 2t$. They are different. Thus we cannot interchange the order of time-varying systems.

8.10 Because $e = -(10R/R)u - (10R/R_f)y = -A(u + \beta y) = -10u - 10\beta y$, we have $\beta = R/R_f$ or $R_f = R/\beta$.

8.11 $H_1(s)$ and $H_2(s)$ are stable. $H(s) = -10/(s - 19)$ is not stable. Not true.

8.12 $H_1(s)$ is not stable. $H(s) = 2/(s + 1)$ is stable. Not true.

8.13 $H_1(s)$ is not stable. $H(s) = (s^3 + 2s^2 - 2)/(s^4 + s^3 + 3s^2 + 2s + 1)$ is stable.

8.15 -9.9989. -9.9995. They differ by 0.006%.

8.17 Yes. No.

8.18 No. No.

8.20 $[(R_1 + R_2)(RCs + 1)^2]/[R_1(RCs)^2 + (2R_1 - R_2)RCs + R_1]$
$2R_1 = R_2$

8.22 $X(\omega) = \pi[\delta(\omega - 10) + \delta(\omega + 10)]$
$X_m(\omega) = 0.5\pi[\delta(\omega - 110) + \delta(\omega + 90) + \delta(\omega - 90) + \delta(\omega + 110)]$
$\bar{X}(\omega) = X_m(\omega) + 2\pi[\delta(\omega - 100) + \delta(\omega + 100)]$

8.24 Multiply $x_m(t)$ by $\cos 100t$ and then design a lowpass filter with gain 2 and passband edge frequency larger than 10 rad/s and stopband edge frequency smaller than 190.

8.26 $\omega_c > W$. Otherwise, shifted spectra overlap each other and the original spectrum cannot be recovered.

CHAPTER 9

9.1 $H(z) = z(4.7z - 4.36)/[(z - 1)(z - 0.8)(z + 0.7)]$

9.2 (a) $H_1(z) = (z^2 - z - 2)/(2z^2 + 4z + 10)$. Biproper.
(b) $H_2(z) = (z^{-1} + 3z^{-2} + 2z^{-3})/(2 + 4z^{-1} + 10z^{-2})$
$= (z^2 + 3z + 2)/(2z^3 + 4z^2 + 10z)$. Strictly proper.
(c) $H_3(z) = (z^{-2} + z^{-5})/(1 + z^{-5}) = (z^3 + 1)/(z^5 + 1)$. Strictly proper.

9.3 $v[n + 4] + 3v[n + 3] + 10v[n] = 2r[n + 2] + 5r[n + 1] + 3r[n]$
$v[n] + 3v[n - 1] + 10v[n - 4] = 2r[n - 2] + 5r[n - 3] + 3r[n - 4]$

9.4 $H(z) = (4.7z - 4.36)/(z^2 - 0.1z - 0.56)$

9.5 (a) Zero: -2; poles: $-0.5 \pm j0.5$
(b) Zeros: $3, -2$; poles: $0, -1, -1$.

9.6 $h[n] = -0.5(-1)^n + 0.5(0.8)^n$
$y_q[n] = -0.25(-1)^n - 2(0.8)^n + 2.25$

9.8 $y_q[n] = k_1 + k_2 n + k_3(0.78)^n \sin(0.88n + k_4)$

9.9 $y_q[n] = k_1 + k_2(-0.6)^n + k_3 n(-0.6)^n + k_4 n^2(-0.6)^n + k_5(0.5)^n + k_6(0.78)^n \sin(2.27n + k_7)$

9.10 (a) Not DT stable.
(b) Not DT stable.
(c) DT stable.

9.11 $z + 0.5$, $z + 2$, $z - 0.5$, $z - 2$.

9.13 No. $k_1(-1)^n + k_2 \sin(2n + k_3)$. It contains $k_1(-1)^n = k_1 \cos(\pi n)$, which has frequency π rad/s.

9.14 $y_{ss}[n] = 0.23$. 20 samples.

9.16 $y_{ss}[n] = 0 \times 2 + 0.055 \sin(0.1n + 1.56) + 8.4 \cos(3n + 0.57)$. 22 samples. Highpass.

9.18 Final value is infinity. $\lim_{z\to 1}(z - 1)X(z) = 0$. Because $x[n]$ does not approach a constant, the formula cannot be used.

CHAPTER 10

10.2

$$\mathbf{x}[n+1] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_5 & -a_4 & -a_3 & -a_2 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u[n]$$

$$y[n] = [\hat{b}_5 - a_5\hat{b}_1 \quad \hat{b}_4 - a_4\hat{b}_1 \quad \hat{b}_3 - a_3\hat{b}_1 \quad \hat{b}_2 - a_2\hat{b}_1]\mathbf{x}[n] + \hat{b}_1 u[n]$$

10.3 (a) $x[n+1] = -4.5x[n] + u[n]$
$y[n] = -12x[n] + 3u[n]$
(b) Controllable form:

$$\mathbf{x}[n+1] = \begin{bmatrix} -3 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u[n]$$

$$y[n] = [1 \quad 0 \quad -1]\mathbf{x}[n]$$

Minimal realization:

$$\mathbf{x}[n+1] = \begin{bmatrix} -2 & 0 \\ 1 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u[n]$$

$$y[n] = [1 \quad -1]\mathbf{x}[n]$$

(c) Controllable form:

$$\mathbf{x}[n+1] = \begin{bmatrix} -1.5 & -1 & -2.5 & -5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u[n]$$

$$y = [0 \quad 0 \quad 0 \quad 0.5]\mathbf{x}[n]$$

10.5 Let x_1 (x_2) be the output of the left-hand (right-hand) side unit delay element. Then we have

$$\mathbf{x}[n+1] = \begin{bmatrix} -6 & 5 \\ 3 & 1 \end{bmatrix} \mathbf{x}[n] + \begin{bmatrix} 2 \\ -4 \end{bmatrix} u[n]$$

$$y[n] = [0 \quad 1]\mathbf{x}[n] - 28u[n]$$

10.6
```
T=0.5;N=[2 0 1 1];D=[1 -0.3 -0.23 0.765];
s1=tf(N,D,T);step(s1)
```

10.7
```
T=0.5;N=[2 0 1 1];D=[1 -0.3 -0.23 0.765];
[a,b,c,d]=tf2ss(N,D);s2=ss(a,b,c,d,T);step(s2)
```

10.8
```
a=[0 0 0;1 0 0;0 1 0];b=[1;0;0];c=[-3 -4 6];d=2;T=1;
sys=ss(a,b,c,d,T);ue=[ones(1,10) 0 0 0];y=lsim(sys,ue)
```

REFERENCES

[1] C. B. Boyer, *A History of Mathematics,* New York: Wiley, 1968.
[2] C. T. Chen, *Digital Signal Processing: Spectral Computation and Filter Design,* New York: Oxford University Press, 2001.
[3] C. T. Chen, *Linear System Theory and Design,* 3rd ed., New York: Oxford University Press, 2000.
[4] C. T. Chen, *System and Signal Analysis,* 2nd ed., New York: Oxford University Press, 1994.
[5] C. T. Chen, *Analog and Digital Control System Design: Transfer-Function, State-Space, and Algebraic Methods,* New York: Oxford University Press, 1993.
[6] L. O. Chua, C. A. Desoer, and E. S. Kuo, *Linear and Nonlinear Circuits,* New York: McGraw-Hill, 1987.
[7] A. Dimarogonas, *Vibration for Engineers,* Upper Saddle River, NJ: Prentice-Hall, 1996.
[8] A. R. Hambley, *Electronics,* 2nd ed., Upper Saddle River, NJ: Prentice-Hall, 2000.
[9] S. Haykin, *Communication Systems,* 2nd ed., New York: John Wiley, 1983.
[10] W. H. Hayt, Jr., J. E. Kemmerly, and S. M. Durbin, *Engineering Circuit Analysis,* 6th ed., New York: McGraw-Hill, 2002.
[11] M. N. Horenstein, *Microelectronic Circuits and Devices,* 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1996.
[12] D. A. Johnson, J. R. Johnson, J. L. Hilburn, and P. D. Scott, *Electric Circuit Analysis,* 3rd ed., Upper Saddle River, NJ: Prentice-Hall, 1997.
[13] D. W. Kammler, *A First Course in Fourier Analysis,* Upper Saddle River, NJ: Prentice-Hall, 2000.
[14] T. Kailath, *Linear Systems,* Englewood Cliffs, NJ: Prentice-Hall, 1980.
[15] B. P. Lathi, *Modern Digital and Analog Communication Systems,* New York: Oxford University Press, 1998.
[16] C. D. McGillem and G. R. Cooper, *Continuous and Discrete Signal and System Analysis,* 3rd ed., New York: Oxford University Press, 1984.
[17] L. Meirovitch, *Fundamentals of Vibrations,* New York: McGraw-Hill, 2001.
[18] G. M. Miller, *Modern Electronic Communication,* 6th ed., Upper Saddle River, NJ: Prentice-Hall, 1999.
[19] A. V. Oppenheim, A. S. Willsky, with S. H. Nawab, *Signals and Systems,* 2nd ed., Upper Saddle River, NJ: Prentice-Hall, 1997.
[20] L. A. Pipes and L. R. Harville, *Applied Mathematics for Engineers and Physicists,* 3rd ed., New York: McGraw-Hill, 1970.
[21] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications,* Upper Saddle River, NJ: Prentice-Hall, 1996.
[22] M. S. Sarma, *Introduction to Electrical Engineering,* New York: Oxford University Press, 2001.

[23] A. S. Sedra and K. C. Smith, *Microelectronic Circuits,* 3rd ed., New York: Oxford University Press, 1991.

[24] D. H. Sheingold, *Analog–Digital Conversion Handbook,* 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 1986.

[25] T. Soderstrom and P. Stoica, *System Identification,* Englewood Cliffs, NJ: Prentice-Hall, 1989.

[26] S. Soloman, *Sensors Handbook,* New York: McGraw-Hill, 1999.

[27] K. Steiglitz, *A Digital Signal Processing Primer,* Reading, MA: Addison-Wesley, 1996.

[28] W. Tomasi, *Electronic Communications Systems,* 3rd ed., Upper Saddle River, NJ: Prentice-Hall, 1998.

[29] J. G. Truxal, *Introductory System Engineering,* New York: McGraw-Hill, 1972.

[30] J. Turner and M. Hill, *Instrumentation for Engineers and Scientists,* New York: Oxford University Press, 1999.

[31] R. Van De Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters,* Boston: Kluwer Academic Publishers, 1994.

[32] M. E. Van Valkenburg, *Analog Filter Design,* New York: Oxford University Press, 1982.

[33] A. H. Zemanian, *Distribution Theory and Transform Analysis,* New York: McGraw-Hill, 1965.

INDEX

Index Terms	Links		

Index Terms	Links		

Index Terms	Links		

Index Terms | Links

Index Terms	**Links**	

Index Terms	Links	

Index Terms	Links	

Index Terms	**Links**		

Index Terms	Links		

Index Terms	Links		

O

P

Index Terms	**Links**		

Q

R

Index Terms	**Links**		

Index Terms	**Links**	

S

Index Terms	Links		