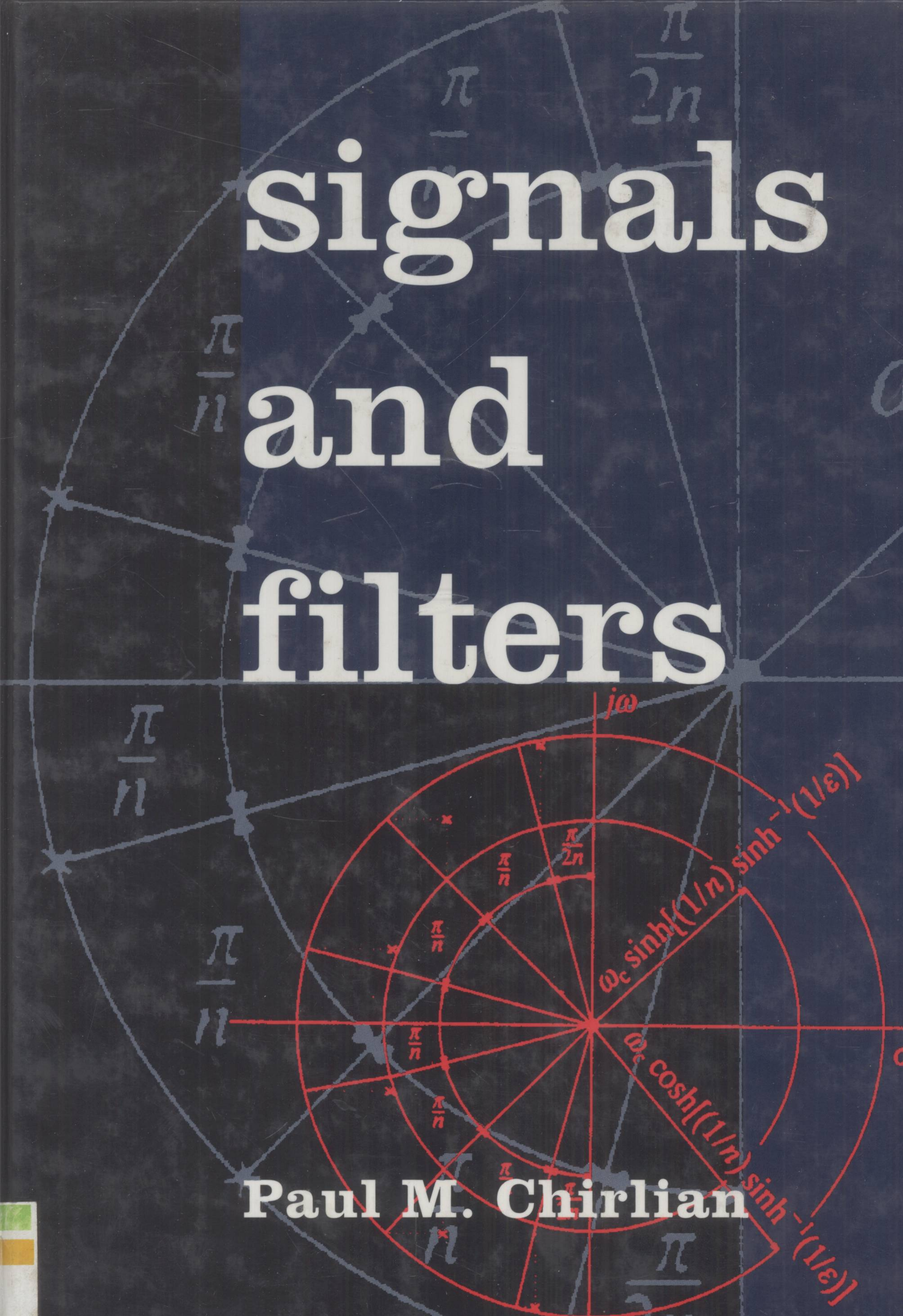
signals and filters



Paul M. Chirlian

Signals and Filters

Paul M. Chirlian

Professor of Electrical Engineering
University of New Orleans

VNR VAN NOSTRAND REINHOLD
New York

Library of Congress Catalog Card Number 93-12670
ISBN 0-442-01324-8

I(T)P Van Nostrand Reinhold is an International Thomson Publishing company.
ITP logo is a trademark under license.

Printed in the United States of America

Van Nostrand Reinhold
115 Fifth Avenue
New York, NY 10003

International Thomson Publishing GmbH
Königswinterer Str. 518
5300 Bonn 3
Germany

International Thomson Publishing
Berkshire House
168–173 High Holborn
London WC1V 7AA, England

International Thomson Publishing Asia
38 Kim Tian Rd., #0105
Kim Tian Plaza
Singapore 0316

Thomas Nelson Australia
102 Dodds Street
South Melbourne 3205
Victoria, Australia

International Thomson Publishing Japan
Kyowa Building, 3F
2-2-1 Hirakawacho
Chiyada-Ku, Tokyo 102
Japan

Nelson Canada
1120 Birchmount Road
Scarborough
Ontario M1K 5G4, Canada

16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Chirlian, Paul M.
Signals and filters / by Paul M. Chirlian.
p. cm.
Includes index.
ISBN 0-442-01324-8
1. Signal processing—Mathematics. 2. System analysis—Data processing. 3. Filters (Mathematics). I. Title.
TK5102.9.C48 1994
621.382—dc20 93-12670 CIP

To Barbara, Lisa, Peter, Jerry, Elizabeth, and Sandy

Preface

This is a textbook intended to be used in a two-course sequence on signals and systems that follows an introductory electrical engineering course. As such, it is assumed the student knows basic circuit concepts and has had a course in calculus and differential equations. This book provides depth of coverage in both continuous and discrete systems.

Systems courses should contain a substantial mathematical discussion of transforms and, in addition, there must be a meaningful discussion of practical systems. Rather than discuss a great many applications which, if considered in sufficient detail, would require more time than is available in a two-course sequence, analog and digital filtering is discussed in detail. It is important that the student be capable of both analyzing and designing, or synthesizing, linear systems. The chapters on the synthesis of signal processing systems reinforce both the basic discussions of analysis and provide the techniques needed to design the systems.

The book starts by introducing various properties of simple signals and systems. The fundamental ideas of linear, time invariant continuous-time and discrete-time systems are introduced in Chapters 2 and and 3. Continuous-time systems are considered in Chapters 4 through 6. Fourier and Laplace transforms are thoroughly covered in Chapters 4 and 5, respectively. The use of residues is discussed in conjunction with the evaluation of the inverse Laplace transform. Although a complete discussion of complex variable calculus is not included here, enough is covered so the student can understand how to apply residues to the evaluation of complex integrals. The two-sided Laplace transform is presented. General orthogonal polynomial expansion is discussed in Chapter 4.

Analog filtering is discussed in Chapter 6, both active and passive synthesis techniques. Butterworth, Chebyshev, and Thompson filters are discussed in detail. Frequency transformations are presented. The approximation problem is introduced and simple approximation is discussed.

Discrete-time systems are discussed in Chapters 7, 8, and 9. In addition to discussing the theoretical aspects of the Fourier transform, computations involving the fast Fourier transforms are presented. The z-transform is discussed in Chapter 8. Methods for obtaining the z-transform of a filter function from an analog filter function are presented. Fourier series procedures for the design of nonrecursive digital filters are developed, and windows are discussed in detail. Recursive filter synthesis, including wave digital filters, is developed. Linear phase filters are discussed. Multiplier coefficient sensitivity is considered in detail. Roundoff noise is considered and the implications of scaling are presented. Limit-cycle oscillations are discussed. Many problems are included at the end of each chapter.

Signals and Filters

Contents

Preface xi

Chapter 1 Introduction 1

1-1. Signals 2
1-2. Properties of Signals 3
1-3. Common Signals 7
1-4. Systems 13
1-5. Unit-step Response and Unit-impulse Response 19
1-6. Summary 21

Chapter 2 Linear, Time Invariant, Continuous-time Systems 25

2-1. Unit-impulse Response 25
2-2. Convolution 27
2-3. The Unit-step Response and General System Response 30
2-4. Singularity Functions 32
2-5. System Response to Complex Exponentials 34
2-6. Block Diagrams 36
2-7. Summary 37

Chapter 3 Linear, Time Invariant, Discrete-time Systems 41

3-1. Difference Equations – Unit-impulse Response 41
3-2. Convolution – The Unit-impulse Response 44
3-3. The Unit-step Response and General System Response 47
3-4. System Response to Complex Exponentials 49
3-5. Some Properties of the Convolution 50
3-6. Summary 54

Chapter 4 The Fourier Transform 58

4-1. More on the Unit-impulse 58
4-2. Fourier Transforms 61
4-3. Existence of the Fourier Transform 66
4-4. Some Common Fourier Transforms 69
4-5. Real and Imaginary Parts of the Fourier Transform 69
4-6. A Physical Interpretation of the Fourier Transform 70
4-7. Convolution – Transfer Functions 73
4-8. Transient and Sinusoidal Steady-state Responses 76
4-9. Some Useful Theorems 80
4-10. Some Special Forms of the Fourier Transform 87
4-11. Energy – Parseval's Theorem 88
4-12. Distortion – Band Limiting – Gibbs' Phenomenon 90
4-13. The Fourier Transform of Periodic Functions 95
4-14. The Fourier Series 99
4-15. Mean-square Error 102
4-16. Orthogonal Function Expansion 105
4-17. Summary 111

Chapter 5 The Laplace Transform 118

5-1. One-sided Laplace Transform – The Unilateral Laplace Transform 118
5-2. The Region of Existence of the Laplace Transform – The *s*-Plane – Poles and Zeros 122
5-3. Relation Between the Fourier Transform and the Laplace Transform 124
5-4. Some Common Laplace Transforms 127
5-5. Solution of Linear Integral-Differential Equations with Constant Coefficients 127
5-6. Transformed Networks 134
5-7. Convolution – Transfer Functions 136
5-8. Some Useful Theorems 140
5-9. Semiperiodic Functions 145
5-10. Contour Integration 147
5-11. Laplace Transform Inversion Integral – Jordan's Lemma 152
5-12. Two-sided Laplace Transform – The Bilateral Laplace Transform 158
5-13. Summary 160

Chapter 6 Analog Filters 166

6-1. A General Discussion of Filters 166
6-2. Design of Double Loaded Low-pass Networks 170
6-3. Active Filters 176
6-4. Butterworth Filters 183
6-5. Chebyshev Filters 186
6-6. Delay Circuits – Bessel Polynomials 190
6-7. Frequency Transformation 198
6-8. Approximation of Characteristics 206
6-9. Summary 210

Chapter 7 Discrete-time Systems – The Discrete Fourier Transform 216

7-1. The Discrete Fourier Transform 216
7-2. Properties of the Discrete Fourier Transform 223
7-3. The Fast Fourier Transform (FFT) – Time Decimation 229
7-4. The Fast Fourier Transform (FFT) – Frequency Decimation 236
7-5. The Sampling Theorem 239
7-6. Aliasing Errors 243
7-7. Analog Sampling 246
7-8. Convolution – Transfer Functions 253
7-9. Parseval's Theorem 259
7-10. Summary 261

Chapter 8 The z-Transform 265

8-1. The z-Transform 265
8-2. Inverse z-Transform 269
8-3. The z-Transform Inversion Integral 273
8-4. Solution of Difference Equations Using the z-Transform 275
8-5. Convolution – Transfer Functions 279
8-6. Sinusoidal Steady-state Response 284
8-7. z-Transform Theorems 286
8-8. Parseval's Theorem 289
8-9. Two-sided z-Transform – The Bilateral z-Transform 290
8-10. Summary 296

Chapter 9 Digital Filters 301

9-1. Procedure for Obtaining a Transfer Function in z-Transformed Form from a Known Laplace Transform – The Bilinear z-Transform 302
9-2. Impulse Invariant Method 309
9-3. Finite and Infinite Impulse Response 310
9-4. Design of Nonrecursive Digital Filters Using Fourier Series Methods 313
9-5. A Further Discussion of FIR Filter Approximation – Windowing 316
9-6. Basic Digital Filter Realizations 329
9-7. Transposed Forms 334
9-8. FIR Filters – Linear Phase Filters 336
9-9. Coefficient Quantization 342
9-10. Wave Digital Filters 348
9-11. Finite Data Word Length Effects 366
9-12. Fixed-point Arithmetic 370
9-13. Scaling 375
9-14. Limit Cycle Oscillations 383
9-15. Summary 386

Index 393

Chapter 1

Introduction

In this book we shall study *linear systems* and *signal processing*. In the most general sense a system is a collection of objects designed to perform some task. In this book we shall be concerned with systems that are used to perform operations on signals. In a somewhat simplified sense, a signal can be any mechanism that is used to transmit information. Examples of information transmitted by signals are: a telephone conversation, a television picture, a patient's heart rate supplied to a recovery room monitor, and the electromagnetic waves generated by a radar system that are used to determine an airplane's altitude. Signals are sent among the various parts of a digital computer and between computers.

Many signals consist of a varying electric voltage, current, or electromagnetic field. The signals discussed in the previous paragraph are often of this form. For example, an ordinary telephone produces a voltage that varies with the speaker's voice. The data transmitted between computers is often in the form of pulses of voltage. Signals can be transmitted using light waves where the intensity of the light is analogous to the magnitude of the voltage in an electrical signal. Signals need not be electrical in nature. For instance, ordinary sound signals are variations of air pressure.

Signal processing systems modify the signal in some way. For instance, a telephone system converts sound waves into electrical signals, transmits these signals to the receiver, and finally converts the electrical signals back into sound waves. The transmission of the signal is a complex operation that involves many types of signal processing. Let us illustrate this with a representative system. The original electrical signal is transmitted over wires to a central office. The signal might then be modified by a process called *modulation* to produce a high frequency signal with the information contained in the telephone conversation. This signal could be transmitted, using radio (electromagnetic) waves, possibly via satellite, to a receiving station where it would be converted back to its original electrical form. During the process of transmission, the signal could be corrupted by unwanted signals or noise. This interference would have to be removed, or at least substantially reduced. All of these operations are part of the topic of signal processing.

In this book we shall discuss many of the systems that are used to perform signal processing. Powerful mathematical techniques are used to analyze and design systems. We shall consider these mathematical techniques and then apply them to the design of filters that process signals.

1-1. SIGNALS

In this section we briefly consider the various types of signals to be discussed in this book. And as we go along, these discussions will be extended and expanded. The signals that we shall consider are real functions of time. For instance,

$$f(t) = u(t)\varepsilon^{-3t} \tag{1-1}$$

represents the signal illustrated in Fig. 1-1a. The function $u(t)$ is called the *unit-step* and it is shown in Fig. 1-1b. Note that the unit-step function is zero for $t < 0$ and is equal to 1 for $t > 0$. In general, the signals that we consider may be of different shapes, but we shall assume that each of them can be represented by some mathematical function.

The signal of Eq. (1-1) is represented by a function of the single variable t, representing time. Signals can be represented by functions of more than one variable. For instance, consider a black and white photograph. At any point, the intensity of the image can vary from black to white. In fact, the intensity could be represented by a numerical scale. For instance, a value of zero could represent a black point and a value of 10 could represent the whitest point in the picture. In such a scheme, the intensity could be expressed as a function of two variables representing the x and y coordinates of each point in the picture. Such a function could be expressed as $f(x, y)$. If the picture were received by a black and white television receiver, it would change with time. In this case, the intensity would be a function of three variables, wherein two of these variables represent position and the third represents time. This function could be expressed as $f(x, y, t)$.

Of course, functions can be expressed in terms of many different independent variables. For example, the production of wheat could

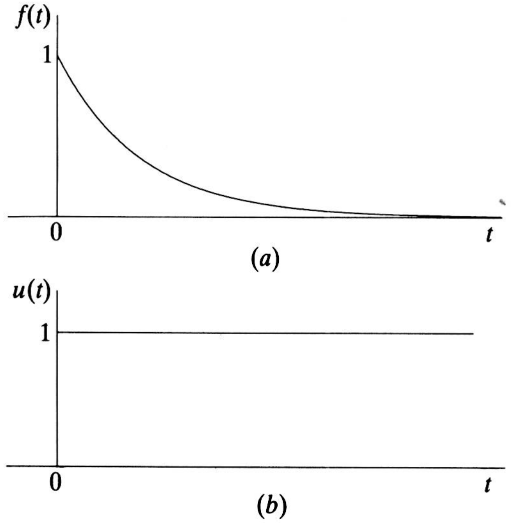


FIGURE 1-1. (a) A plot of the waveform $u(t)e^{-3t}$; (b) a plot of the unit-step function $u(t)$.

be expressed as a function of the average rainfall. However, this book deals with functions of time.

Signals are categorized as to whether they are *analog* or *digital*. In general, analog signals represent an operation in some direct way. For instance, if the signal of Fig. 1-1 represented the sound output of a musical instrument, then the sound level would be maximum at $t = 0$ and fall off exponentially. A digital signal, on the other hand, takes on one of only two values, *zero* (0) and *one* (1). The data is encoded according to some procedure. Binary numbers are often used to represent these values. A typical digital signal is shown in Fig. 1-2. Note that the signal level is either 0 or 1. (In a digital circuit these are often represented by two different voltage levels. These levels, however, do *not* have to be zero and one volt, respectively.) The presence of a pulse represents a one, while the absence of a pulse indicates a zero. Note that pulses in Fig. 1-2 are allowed to be present only in the time intervals,

$$0 \le t \le T \tag{1-2a}$$

$$2T \le t \le 3T \tag{1-2b}$$

$$\vdots$$

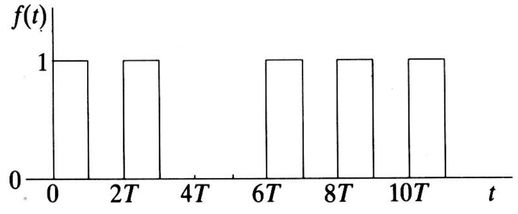


FIGURE 1-2. An example of an idealized digital signal.

The general expression for these time intervals is

$$2nT \le t \le (2n+1)T \quad n = 0, 1, 2, \ldots \tag{1-3}$$

Note that there is no signal allowed in the time intervals

$$(2n+1)T < t < (2n+2)T \quad n = 0, 1, 2 \tag{1-4}$$

Digital signals can be considered to transmit binary numbers. The signal of Fig. 1-2, for example, transmits the binary number 110111.

The blank intervals between the pulses as described by Eq. (1-4) do not transmit any data, either 0's or 1's. Thus, they can be considered to be wasted time. There are techniques that utilize this time. Using a procedure called *time multiplexing*, one or more additional signals can be transmitted in the blank spaces. In Fig. 1-2, the blank intervals between the pulses are of the same duration as the pulses themselves; this need not be the case. For example, the blank intervals could be longer than the pulses themselves. In this case, another signal could be transmitted during these "blank spaces". Naturally, the proper circuitry would be necessary to keep the two signals separate.

There are many techniques for transmitting information that is encoded in the form of binary data. In general, however, coded binary information is transmitted as a sequence of 0's and 1's.

In many cases the analog signal under consideration is continuous except for possibly a finite number of points at which it is discontinuous. Figure 1-1, for example, represents a signal that is continuous, except for a

FIGURE 1-3. An example of a complicated signal.

jump discontinuity at $t = 0$. On the other hand, many analog signals are essentially discontinuous in nature. We shall see that certain analog signals can be completely determined by specifying a sequence of values taken at specified *discrete* times. The sequence of samples is essentially discontinuous in nature. We shall illustrate such *sampled signals* in the next section and discuss them throughout the book.

Often, for instructional purposes, we discuss functions that have relatively simple expressions, such as that given by Eq. (1-1). Real signals are usually far more complex such as the sample shown in Fig. 1-3. When necessary we will work with complex signal forms. Most often, more insight can be gained by dealing with simpler signals.

1-2. PROPERTIES OF SIGNALS

Even and Odd Functions

There are two special types of functions called *even* functions and *odd* functions. An even function is characterized by the relation

$$f(t) = f(-t) \tag{1-5}$$

This means that if t is replaced by $-t$, *the value of the function is unchanged.* Figure 1-4 illustrates an even function. Note that the curve exhibits mirror symmetry about the $f(t)$ axis.

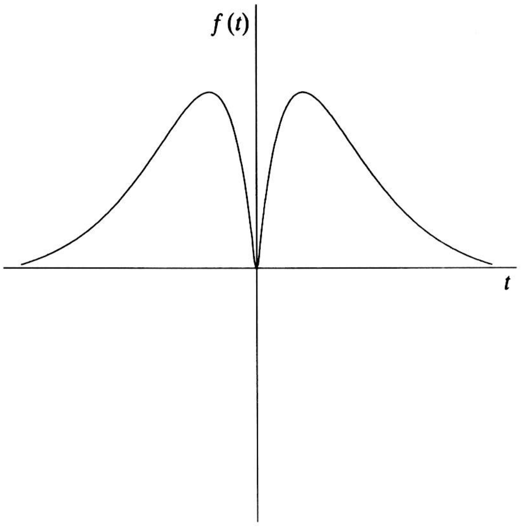


FIGURE 1-4. An even function.

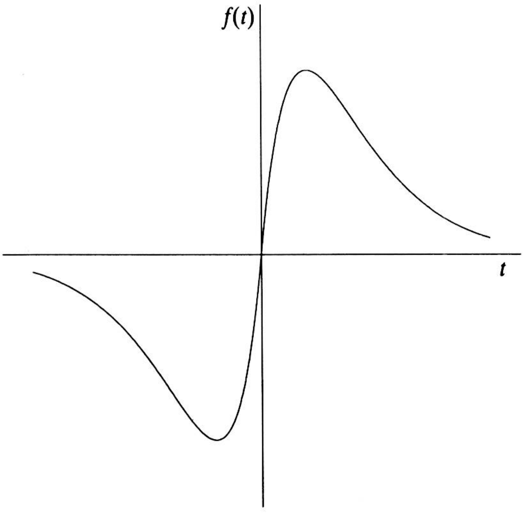


FIGURE 1-5. An odd function.

An odd function is characterized by the relation

$$f(t) = -f(-t) \tag{1-6}$$

This means that if t is replaced by $-t$, *the value of the function is multiplied by* -1. Figure 1-5 illustrates an odd function. If the function were multiplied by -1 for values of $t < 0$, the resulting plot would exhibit mirror symmetry about the $f(t)$ axis. That is, it would become an even function.

It is often helpful to write a function as the sum of an even and an odd function. Let us demonstrate that this can be done. Suppose that $f(t)$ is a function of time and that $f_e(t)$ and $f_o(t)$ are even and odd functions of time, respectively. Let us relate $f_e(t)$ and $f_o(t)$ to $f(t)$ in the following way:

$$f_e(t) = \frac{f(t) + f(-t)}{2} \tag{1-7}$$

and

$$f_o(t) = \frac{f(t) - f(-t)}{2} \tag{1-8}$$

It can be shown that $f_e(t)$ is an even function simply by replacing t by $-t$ and noting that $f(-[-t]) = f(t)$. When this substitution of $-t$ for t is made, the right side of Eq. (1-7) is unchanged. This demonstrates that $f_e(t)$ is indeed even. In a similar way, if t is replaced by $-t$ the right side of Eq. (1-8) is replaced by $[f(-t) - f(t)]/2$. This is the negative of the right side of Eq. (1-8). Therefore, it has been demonstrated that $f_o(t)$ is an odd function. If Eqs. (1-7) and (1-8) are added we obtain

$$f_e(t) + f_o(t) = f(t) \tag{1-9}$$

Thus, it has been demonstrated that any function can be expressed as a sum of even and odd parts, and that we can use Eqs. (1-7) and (1-8) to break a function into its even and odd parts.

Periodic Functions

A function that repeats itself every T seconds is called a *periodic* function. The mathematical definition of a periodic function is one that satisfies the relation

$$f(t) = f(t + T) \quad \text{for all } t \tag{1-10}$$

where T is a constant. When we deal with

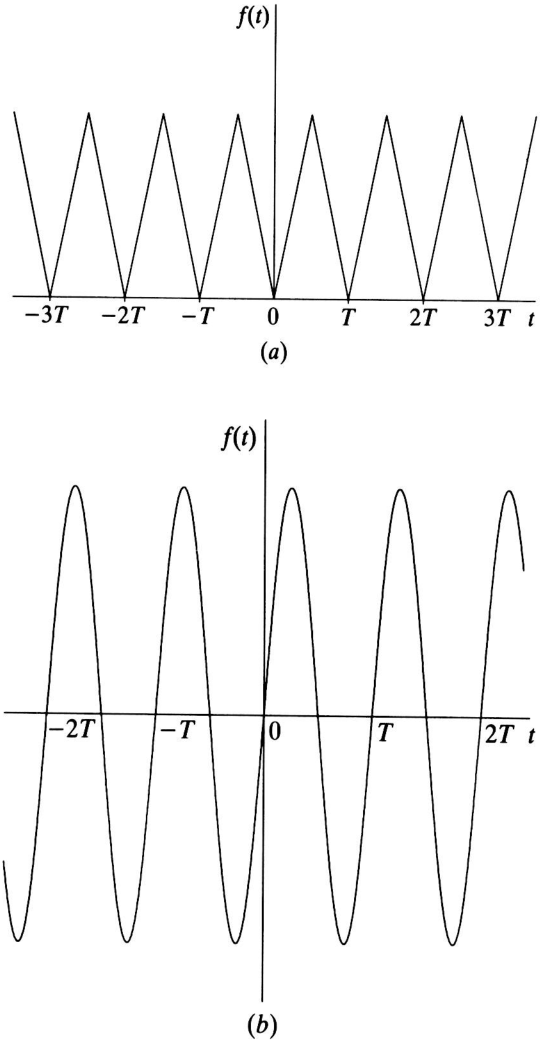


FIGURE 1-6. Periodic functions: (a) a triangular wave; (b) a sinusoid.

real functions with real arguments, T will be restricted to be a real constant. There is no loss in generality if T is required to be positive and we shall so restrict it. Two periodic functions are shown in Fig. 1-6. Figure 1-6a depicts a triangular wave while Fig. 1-6b illustrates a sine wave. In each case the period is T seconds. That is, the waveform repeats every T seconds.

The *frequency* of a waveform is the number of times per second that the waveform repeats. The frequency is equal to the reciprocal of the period. That is,

$$f = 1/T \tag{1-11}$$

The formula for the curve of Fig. 1-6b is

$$f(t) = \sin \omega t \tag{1-12}$$

where ω is called the *angular frequency* and is given by

$$\omega = 2\pi f \tag{1-13}$$

Discrete Signals

In Section 1-1 we mentioned that some analog signals could consist of a sequence of samples. Figure 1-7 illustrates a sampled sinusoidal signal. The sinusoid is drawn for illustrative purposes and is not a part of the actual signal. The heavy vertical lines capped by disks represent the samples. The vertical position of the disk indicates the numerical value of the sample. The actual sampled signal can take on different forms. For instance, short pulses each of whose height is equal to the height of the corresponding sample could be

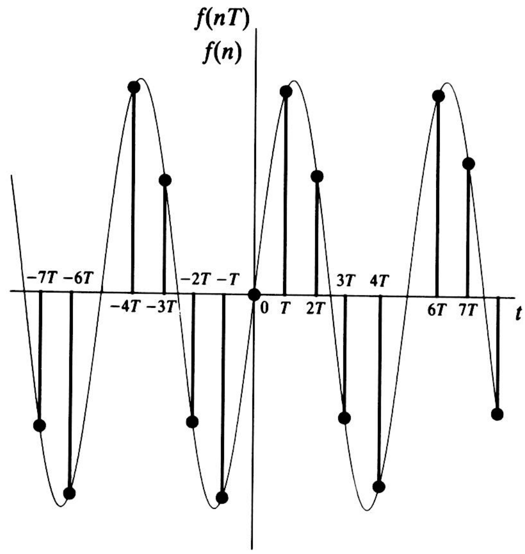


FIGURE 1-7. A sampled periodic signal. The large dots represent the values of the samples.

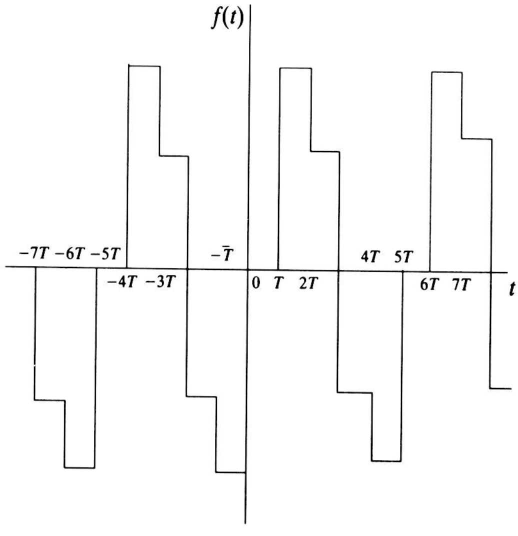


FIGURE 1-8. Another form of a sampled periodic signal; the waveform depends on the starting time of the sampling.

transmitted. In this case the actual signal would appear essentially as the diagram of Fig. 1-7 with the sinusoid and the disks omitted. Sampled signals are also called *discrete signals.*

Another form of discrete signal does not change value between sampling times. Instead the value of the signal remains constant at the last sampled value. For the sinusoid of Fig. 1-7, the discrete signal would be as shown in Fig. 1-8. These signals are not continuous; but they are analog signals. The signal level directly corresponds to a level in the sampled signal. Note that there are other forms of sampled analog signals. In one such sampling scheme, the width of the pulses corresponds to the level of the original signal.

Consider Fig. 1-7. The signal is sampled at times

$$t = 0, \pm T, \pm 2T, \ldots \qquad (1\text{-}14)$$

T is called the *sampling time.* The vertical axis is labeled $f(nT)$ to indicate the discrete nature of the operation. Common usage is somewhat sloppy concerning the notation used here. The T is often omitted causing the function to be written as $f(n)$. Note that, according to mathematical usage, if $f(t)$ is a function of the variable t, then $f(n)$ is the same function with t simply replaced by n. This is *not* what is meant by the commonly used notation of sampled signals where $f(t)$ and $f(n)$ are two different entities; $f(t)$ is the function that is defined for all time and $f(n)$ represents the function that is defined only at the sample times. This type of notation can be confusing. However, in most instances the usage clarifies the meaning. Note that, according to the notation of sampled signals, $f(n)$ and $f(nT)$ represent the same functions, but this violates common usage. When necessary, to avoid ambiguity, we shall include suitable explanatory material in the remainder of the book.

Note that we have labeled the axis of Fig. 1-8 as $f(t)$. This signal is defined for all values of time (except possibly at the discontinuities) and is expressed as a function of time. Note that

$$f(t) \neq -f(-t) \qquad (1\text{-}15)$$

even though the sinusoid from which the samples are derived is itself odd. Figure 1-8 is not odd. If, however, the start of the sampling time were shifted by $-T/2$ seconds, so that the sample interval containing the origin was symmetric about the origin, then the sampled signal would be odd as well. It should be noted that the form of sampled signals changes if the sampling times are changed. For example, the sampled signal of Fig. 1-8 would be odd if the start of sampling were shifted by $-T/2$ seconds.

The previous discussion of periodic functions can be applied to discrete signals as well. Note that the signal of Fig. 1-8 is periodic and that its period is the same as that of the original sinusoid. *This need not be the case.* If the sample interval were unrelated to the period of the sinusoid, then the sampled signal would not be periodic. Specifically, if a periodic signal is sampled and the ratio of the sampling time to the period is *not* a rational number (i.e., it cannot be expressed as the ratio of two integers), then the sampled signal will *not* be periodic. Note also that the period of the sampled signal can be different from the period of the signal being sampled. The comments that we make about periodicity apply to any type of sampled signal. For instance, the sampled signal of Fig. 1-7 is periodic. Of course if the signal is not periodic, the sampled signal will not, in general, be periodic.

We have considered discrete analog signals. If the values of the samples are converted into numbers and these numbers are transmitted, then a digital signal results.

1-3. COMMON SIGNALS

This section introduces several signals frequently used to test systems or as an aid in mathematical analysis.

Unit-step Function

The unit-step function was introduced in Section 1-1 and is illustrated in Fig. 1-1b. To review, the unit step is written as $u(t)$ and defined as

$$u(t) = 1 \quad t > 0 \tag{1-16a}$$

$$u(t) = 0 \quad t < 0 \tag{1-16b}$$

$$u(t) = \tfrac{1}{2} \quad t = 0 \tag{1-16c}$$

Previously we did not define the value of $u(t)$ at the point of its discontinuity, that is at $t = 0$. It is common to define this value as $\frac{1}{2}$. We shall see why it is reasonable to do this subsequently. For many applications, any finite value could be used for the value of $u(0)$ without changing the results of any calculations. The unit-step function is used to test systems and we shall discuss this use subsequently. The unit-step function is also used mathematically to establish signals that are 0 for $t < 0$.

Sinusoid

Another commonly encountered signal is the sinusoid. A simple sinusoidal signal is the sine wave, which we have illustrated in Fig. 1-6b. Another form of sinusoidal function is the cosine which is illustrated in Fig. 1-9a. The equation for this waveform is

$$f(t) = A \cos \omega t \tag{1-17}$$

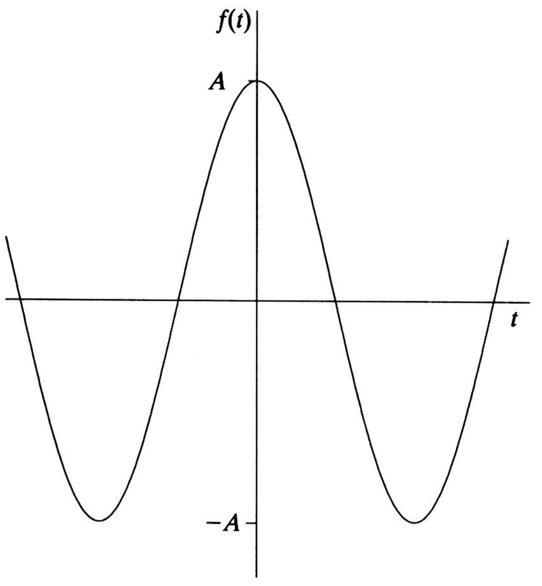
Note that A is the maximum value of $f(t)$. Figure 1-9b illustrates a signal which is a cosine wave for $t > 0$, but which is 0 for $t < 0$. The equation for this curve is

$$f(t) = Au(t) \cos \omega t \tag{1-18}$$

Notice how the unit-step function has been used to set $f(t)$ equal to 0 for values of $t < 0$. The sine and cosine waveforms are very similar and they shall be related later in this section. Subsequently, we will demonstrate that sines and cosines can be used to express many arbitrary functions of time.

Damped Sinusoid

The response of most practical systems is in the form of damped functions of time. Commonly, these are simple decaying



(*a*)

FIGURE 1-9. (a) The function $A \cos \omega t$. (*Continued*)