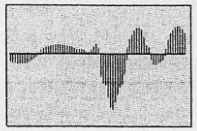

Discrete-Time Signals and Systems

2

2.0 INTRODUCTION

A signal can be defined as a function that conveys information, generally about the state or behavior of a physical system. Although signals can be represented in many ways, in all cases the information is contained in some pattern of variations. Signals are represented mathematically as functions of one or more independent variables. For example, a speech signal is represented mathematically as a function of time and a photographic image is represented as a brightness function of two spatial variables. It is a common convention, and one that will usually be followed in this book, to refer to the independent variable of the mathematical representation of a signal as time, although in specific examples the independent variable may in fact not represent time.

The independent variable in the mathematical representation of a signal may be either continuous or discrete. *Continuous-time signals* are defined along a continuum of times and thus are represented by a continuous independent variable. Continuous-time signals are often referred to as *analog signals*. *Discrete-time signals* are defined at discrete times and thus the independent variable has discrete values; i.e., discrete-time signals are represented as sequences of numbers. Signals such as speech or images may have either a continuous- or a discrete-variable representation, and if certain conditions hold, these representations are entirely equivalent. Besides the independent variables being either continuous or discrete, the signal amplitude may be either continuous or discrete. *Digital signals* are those for which both time and amplitude are discrete.

Signal processing systems may be classified along the same lines as signals. That is, continuous-time systems are systems for which both the input and the output are continuous-time signals, and discrete-time systems are those for which both the input and the output are discrete-time signals. Similarly, a digital system is one for which

both the input and the output are digital signals. Digital signal processing, then, deals with the transformation of signals that are discrete in both amplitude and time. The principal focus in this book is on discrete-time (rather than digital) signals and systems. The effects of signal amplitude quantization are considered in Sections 3.7, 6.7–6.10, and 9.8.

Discrete-time signals may arise by sampling a continuous-time signal or they may be generated directly by some discrete-time process. Whatever the origin of the discrete-time signals, discrete-time signal processing systems have many attractive features. They can be realized with great flexibility with a variety of technologies such as charge transport devices, surface acoustic wave devices, general-purpose digital computers, or high-speed microprocessors. Complete signal processing systems can be implemented using VLSI techniques. Discrete-time systems can be used to simulate analog systems or, more important, to realize signal transformations that cannot be implemented with continuous-time hardware. Thus, discrete-time representations of signals are often desirable when sophisticated and flexible signal processing is required.

In this chapter we consider the fundamental concepts of discrete-time signals and signal processing systems for one-dimensional signals. We emphasize the class of linear time-invariant discrete-time systems. Many of the properties and results that we derive in this and subsequent chapters will be similar to properties and results for linear time-invariant continuous-time systems as presented in a variety of texts (see, for example, Papoulis, 1980; Oppenheim and Willsky, 1983). In fact, it is possible to approach the discussion of discrete-time systems by treating sequences as analog signals that are impulse trains. This approach, if implemented carefully, can lead to correct results and has formed the basis for much of the classical discussion of sampled data systems (see, for example, Raggazzini and Franklin, 1958; Freeman, 1965; Phillips and Nagle, 1984). However, not all sequences arise from sampling a continuous-time signal, and many discrete-time systems are not simply approximations to corresponding analog systems. Furthermore, there are important and fundamental differences between discrete- and continuous-time systems. Therefore, rather than attempt to force results from continuous-time system theory into a discrete-time framework, we will derive parallel results starting within a framework and with notation suitable to discrete-time systems. Discrete-time signals will be related to continuous-time signals only when it is necessary and useful to do so.

2.1 DISCRETE-TIME SIGNALS: SEQUENCES

Discrete-time signals are represented mathematically as sequences of numbers. A sequence of numbers x, in which the nth number in the sequence is denoted $x[n]$,† is

† A sequence is simply a function whose domain is the set of integers. Note that we use [ ] to enclose the independent variable of such functions, and we use ( ) to enclose the independent variable of continuous-variable functions.

formally written as

$$x = \{x[n]\}, \qquad -\infty < n < \infty, \tag{2.1}$$

where n is an integer. In a practical setting, such sequences can often arise from periodic sampling of an analog signal. In this case, the numeric value of the nth number in the sequence is equal to the value of the analog signal $x_a(t)$ at time nT; i.e.,

$$x[n] = x_a(nT), \qquad -\infty < n < \infty. \tag{2.2}$$

The quantity T is called the *sampling period* and its reciprocal is the *sampling frequency*. Although sequences do not always arise from sampling analog waveforms, it is convenient to refer to $x[n]$ as the "nth sample" of the sequence. Also, although strictly speaking $x[n]$ denotes the nth number in the sequence, the notation of Eq. (2.1) is often unnecessarily cumbersome, and it is convenient and unambiguous to refer to "the sequence $x[n]$" when we mean the entire sequence, just as we referred to "the analog signal $x_a(t)$." Discrete-time signals (i.e., sequences) are often depicted graphically as shown in Fig. 2.1. Although the abscissa is drawn as a continuous line, it is important to recognize that $x[n]$ is defined only for integer values of n. It is not correct to think of $x[n]$ as being zero for n not an integer; $x[n]$ is simply undefined for noninteger values of n.

As an example, Fig. 2.2(a) shows a segment of a speech signal corresponding to acoustic pressure variation as a function of time, and Figure 2.2(b) shows a sequence of samples of the speech signal. Although the original speech signal is defined at all values of time t, the sequence contains information about the signal only at discrete time instants. From the sampling theorem, discussed in Chapter 3, the original signal can be reconstructed as accurately as desired from a corresponding sequence of samples if the samples are taken frequently enough.

2.1.1 Basic Sequences and Sequence Operations

In the analysis of discrete-time signal processing systems, sequences are manipulated in several basic ways. The product and sum of two sequences $x[n]$ and $y[n]$ are defined as the sample-by-sample product and sum, respectively. Multiplication of a sequence $x[n]$ by a number α is defined as multiplication of each sample value by α. A

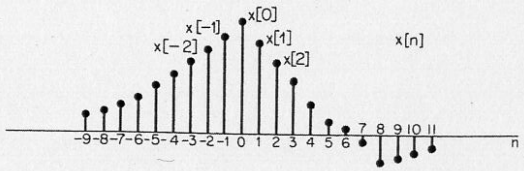

Figure 2.1 Graphical representation of a discrete-time signal.

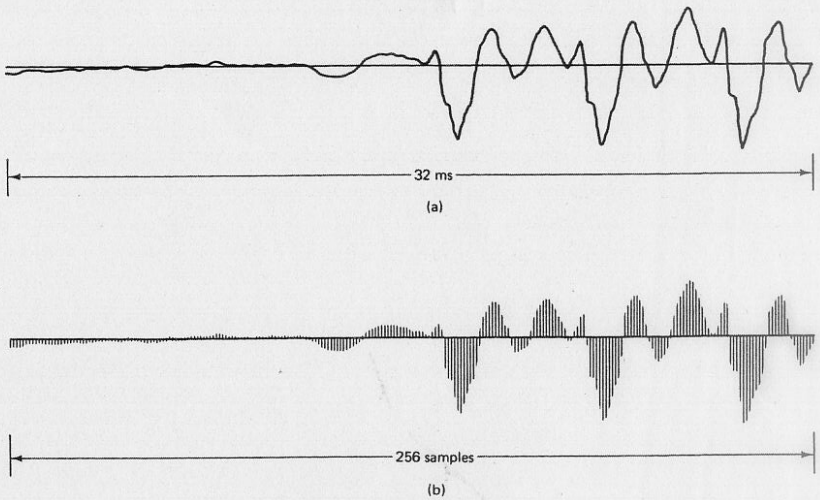

Figure 2.2 (a) Segment of a continuous-time speech signal. (b) Sequence of samples obtained from part (a) with $T = 125\ \mu s$.

sequence $y[n]$ is said to be a delayed or shifted version of a sequence $x[n]$ if $y[n]$ has values

$$y[n] = x[n - n_0], \tag{2.3}$$

where n_0 is an integer.

In discussing the theory of discrete-time signals and systems, several basic sequences are of particular importance. These sequences are shown in Fig. 2.3 and discussed below.

The *unit sample sequence* (Fig. 2.3a), denoted by $\delta[n]$, is defined as the sequence with values

$$\delta[n] = \begin{cases} 0, & n \neq 0, \\ 1, & n = 0. \end{cases} \tag{2.4}$$

As we will see, the unit-sample sequence plays the same role for discrete-time signals and systems that the unit impulse function (Dirac delta function) does for continuous-time signals and systems. For convenience, the unit sample sequence is often referred to as a *discrete-time impulse*, or simply as an *impulse*. It is important to note that a discrete-time impulse does not suffer from the mathematical complications of the continuous-time impulse. Its definition is simple and precise.

As we will see in the discussion of linear systems, one of the important aspects of the impulse sequence is that an arbitrary sequence can be represented as a sum of

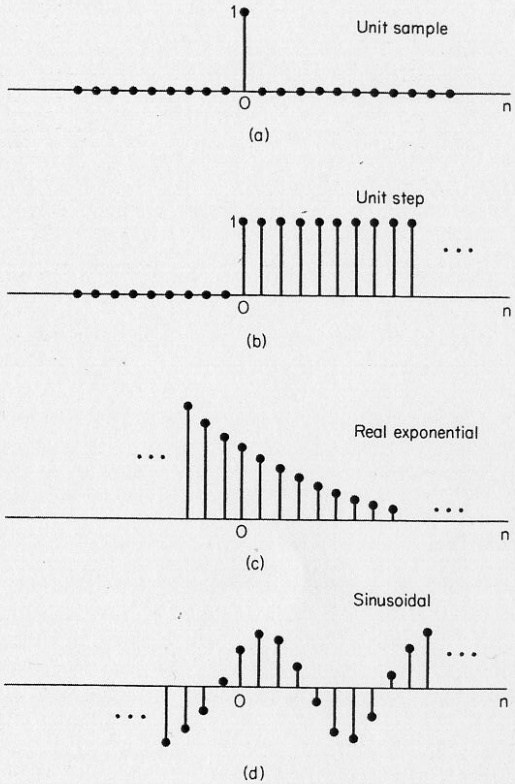

Figure 2.3 Some basic sequences. The sequences shown play important roles in the analysis and representation of discrete-time signals and systems.

scaled, delayed impulses. For example, the sequence $p[n]$ in Fig. 2.4 can be expressed as

$$p[n] = a_{-3}\delta[n+3] + a_1\delta[n-1] + a_2\delta[n-2] + a_7\delta[n-7]. \tag{2.5}$$

More generally, any sequence can be expressed as

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k]. \tag{2.6}$$

We will make specific use of Eq. (2.6) in discussing the representation of discrete-time linear systems.

The *unit step sequence* (Fig. 2.3b), denoted by $u[n]$, has values

$$u[n] = \begin{cases} 1, & n \geq 0, \\ 0, & n < 0. \end{cases} \tag{2.7}$$

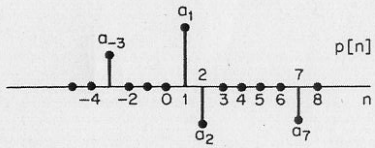

Figure 2.4 Example of a sequence to be represented as a sum of scaled and delayed impulses.

The unit step is related to the impulse by

$$u[n] = \sum_{k=-\infty}^{n} \delta[k], \tag{2.8}$$

i.e., the value of the unit step sequence at (time) index n is equal to the accumulated sum of the value at index n and all previous values of the impulse sequence. An alternative representation of the unit step in terms of the impulse is obtained by interpreting the unit step in Fig. 2.3(b) in terms of a sum of delayed impulses. This is expressed mathematically as

$$u[n] = \delta[n] + \delta[n-1] + \delta[n-2] + \cdots \tag{2.9a}$$

or

$$u[n] = \sum_{k=0}^{\infty} \delta[n-k]. \tag{2.9b}$$

Conversely, the impulse sequence can be expressed as the first backward difference of the unit step sequence, i.e.,

$$\delta[n] = u[n] - u[n-1]. \tag{2.10}$$

Exponential and *sinusoidal sequences* are extremely important in representing and analyzing linear time-invariant discrete-time systems. The general form for an exponential sequence is

$$x[n] = A\alpha^n. \tag{2.11}$$

If A and α are real numbers, then the sequence is real. If $0 < \alpha < 1$ and A is positive, then the sequence values are positive and decrease with increasing n, as in Fig. 2.3(c). For $-1 < \alpha < 0$, the sequence values alternate in sign but again decrease in magnitude with increasing n. If $|\alpha| > 1$, then the sequence grows in magnitude as n increases. A sinusoidal sequence has the general form

$$x[n] = A\cos(\omega_0 n + \phi) \qquad \text{for all } n, \tag{2.12}$$

with A real, and is illustrated in Fig. 2.3(d).

The exponential sequence $A\alpha^n$ with complex α has real and imaginary parts that are exponentially weighted sinusoids. Specifically, if $\alpha = |\alpha|e^{j\omega_0}$ and $A = |A|e^{j\phi}$, the sequence $A\alpha^n$ can be expressed in any of the following ways:

$$\begin{aligned} x[n] = A\alpha^n &= |A|e^{j\phi}|\alpha|^n e^{j\omega_0 n} \\ &= |A||\alpha|^n e^{j(\omega_0 n+\phi)} \\ &= |A||\alpha|^n \cos(\omega_0 n + \phi) + j|A||\alpha|^n \sin(\omega_0 n + \phi). \end{aligned} \tag{2.13}$$

The sequence oscillates with an exponentially growing envelope if $|\alpha| > 1$ or with an exponentially decaying envelope if $|\alpha| < 1$. (As a simple example consider the case $\omega_0 = \pi$.)

When $|\alpha| = 1$, the sequence is referred to as a *complex exponential sequence* and has the form

$$x[n] = |A|e^{j(\omega_0 n+\phi)} = |A|\cos(\omega_0 n + \phi) + j|A|\sin(\omega_0 n + \phi), \tag{2.14}$$

i.e., the real and imaginary parts of $e^{j\omega_0 n}$ vary sinusoidally with n. The fact that n is always an integer in Eq. (2.14) leads to some important differences between the properties of discrete-time and continuous-time complex exponential sequences and sinusoidal sequences. By analogy with the continuous-time case, the quantity ω_0 is called the *frequency* of the complex sinusoid or complex exponential and ϕ is called the *phase*. However, note that n is a dimensionless integer. Thus, the dimension of ω_0 must be radians. If we wish to maintain a closer analogy with the continuous-time case, we can specify the units of ω_0 to be radians/sample and the units of n to be samples.

A more interesting difference between continuous-time and discrete-time complex sinusoids is seen when we consider a frequency $(\omega_0 + 2\pi)$. In this case,

$$\begin{aligned} x[n] &= Ae^{j(\omega_0+2\pi)n} \\ &= Ae^{j\omega_0 n}e^{j2\pi n} = Ae^{j\omega_0 n}. \end{aligned} \tag{2.15}$$

Indeed, we can easily see that complex exponential sequences with frequencies $(\omega_0 + 2\pi r)$, where r is an integer, are indistinguishable from one another. An identical statement holds for sinusoidal sequences. Specifically, it is easily verified that

$$\begin{aligned} x[n] &= A\cos[(\omega_0 + 2\pi r)n + \phi] \\ &= A\cos(\omega_0 n + \phi). \end{aligned} \tag{2.16}$$

The implications of this property for sequences obtained by sampling sinusoids and other signals will be discussed in Chapter 3. For now, we simply conclude that when discussing complex exponential signals of the form $x[n] = Ae^{j\omega_0 n}$ or real sinusoidal signals of the form $x[n] = A\cos(\omega_0 n + \phi)$, we need only consider frequencies in a frequency interval of length 2π such as $-\pi < \omega_0 \le \pi$ or $0 \le \omega_0 < 2\pi$.

Another important difference between continuous-time and discrete-time complex exponentials and sinusoids concerns periodicity. In the continuous-time case, a

sinusoidal signal and a complex exponential signal are both periodic with period equal to 2π divided by frequency. In the discrete-time case, a periodic sequence is one for which

$$x[n] = x[n + N], \qquad \text{for all } n, \tag{2.17}$$

where the period N is necessarily an integer. If this condition for periodicity is tested for the discrete-time sinusoid, then

$$A\cos(\omega_0 n + \phi) = A\cos(\omega_0 n + \omega_0 N + \phi), \tag{2.18}$$

which requires

$$\omega_0 N = 2\pi k, \tag{2.19}$$

where k is an integer. A similar statement holds for the complex exponential sequence $Ce^{j\omega_0 n}$, i.e., periodicity with period N requires

$$e^{j\omega_0(n+N)} = e^{j\omega_0 n}, \tag{2.20}$$

which is true only for $\omega_0 N = 2\pi k$, as in Eq. (2.19). Consequently, complex exponential and sinusoidal sequences are not necessarily periodic with period $(2\pi/\omega_0)$ and, depending on the value of ω_0, may not be periodic at all. For example, for $\omega_0 = 3\pi/4$, the smallest value of N for which Eq. (2.19) can be satisfied with k an integer is $N = 8$ (corresponding to $k = 3$). For $\omega_0 = 1$, there are no integer values of N or k for which Eq. (2.19) is satisfied.

When we combine the condition of Eq. (2.19) with our previous observation that ω_0 and $(\omega_0 + 2\pi r)$ are indistinguishable frequencies, then it is clear that there are N distinguishable frequencies for which the corresponding sequences are periodic with period N. One set of frequencies is $\omega_k = 2\pi k/N$, $k = 0, 1, \ldots, N-1$. These properties of complex exponential and sinusoidal sequences are basic to both the theory and the design of computational algorithms for discrete-time Fourier analysis, and they will be discussed in more detail in Chapters 8 and 9.

Related to the preceding discussion is the fact that the interpretation of high and low frequencies is somewhat different for continuous-time and discrete-time sinusoidal and complex exponential signals. For a continuous-time sinusoidal signal $x(t) = A\cos(\Omega_0 t + \phi)$, as Ω_0 increases, $x(t)$ oscillates more and more rapidly. In contrast, for the discrete-time sinusoidal signal $x[n] = A\cos(\omega_0 n + \phi)$, as ω_0 increases from $\omega_0 = 0$ toward $\omega_0 = \pi$, $x[n]$ oscillates more and more rapidly. As ω_0 increases from $\omega_0 = \pi$ to $\omega_0 = 2\pi$, the oscillations become slower. This is illustrated in Fig. 2.5. In fact, because of the periodicity in ω_0 of sinusoidal and complex exponential sequences, $\omega_0 = 2\pi$ is indistinguishable from $\omega_0 = 0$ and, more generally, frequencies around $\omega_0 = 2\pi$ are indistinguishable from frequencies around $\omega_0 = 0$. As a consequence, for sinusoidal and complex exponential signals, values of ω_0 in the vicinity of $\omega_0 = 2\pi k$ for any integer value of k are typically referred to as low frequencies (relatively slow oscillations), while values of ω_0 in the vicinity of $\omega_0 = (\pi + 2\pi k)$ for any integer value of k are typically referred to as high frequencies (relatively rapid oscillations).

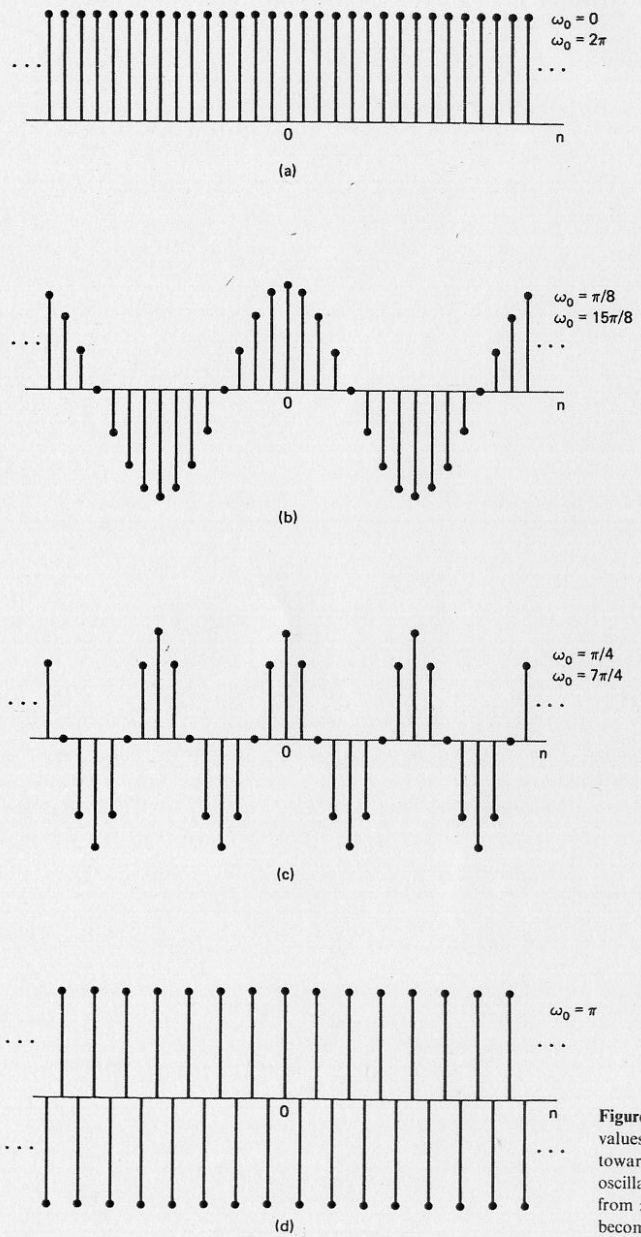

Figure 2.5 $\cos \omega_0 n$ for several different values of ω_0. As ω_0 increases from zero toward π (parts a–d), the sequence oscillates more rapidly. As ω_0 increases from π to 2π (parts d–a), the oscillations become slower.

2.2 DISCRETE-TIME SYSTEMS

A discrete-time system is defined mathematically as a transformation or operator that maps an input sequence with values $x[n]$ into an output sequence with values $y[n]$. This can be denoted as

$$y[n] = T\{x[n]\} \tag{2.21}$$

and is indicated pictorially in Fig. 2.6. Equation (2.21) represents a rule or formula for computing the output sequence values from the input sequence values. It should be emphasized that the value of the output sequence at each value of the index n may be a function of $x[n]$ at all values of n. The following examples illustrate some simple and useful systems.

Example 2.1 The Ideal Delay System

The ideal delay system is defined by the equation

$$y[n] = x[n - n_d], \qquad -\infty < n < \infty, \tag{2.22}$$

where n_d is a fixed positive integer called the delay of the system. In words, the ideal delay system simply shifts the input sequence to the right by n_d samples to form the output. If in Eq. (2.22) n_d is a fixed negative integer, then the system would shift the input to the left by n_d samples corresponding to a time advance.

In Example 2.1 only one sample of the input sequence is involved in determining a certain output sample. In the following example, this is not the case.

Example 2.2 Moving Average

The general moving average system is defined by the equation

$$\begin{aligned} y[n] &= \frac{1}{M_1 + M_2 + 1} \sum_{k=-M_1}^{M_2} x[n-k] \\ &= \frac{1}{M_1 + M_2 + 1} \{x[n + M_1] + x[n + M_1 - 1] + \cdots + x[n] \\ &\quad + x[n-1] + \cdots + x[n - M_2]\}. \end{aligned} \tag{2.23}$$

This system computes the nth sample of the output sequence as the average of $(M_1 + M_2 + 1)$ samples of the input sequence around the nth sample. Figure 2.7 shows an input sequence plotted as a function of a dummy index k and the samples involved in the computation of the output sample $y[n]$ for $n = 7$, $M_1 = 0$, and $M_2 = 5$. The output sample $y[7]$ is equal to $\frac{1}{6}$ times the sum of all the samples between the vertical dotted lines. To compute $y[8]$, both dotted lines would move one sample to the right.

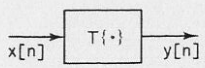

Figure 2.6 Representation of a discrete-time system, i.e., a transformation that maps an input sequence $x[n]$ into a unique output sequence $y[n]$.

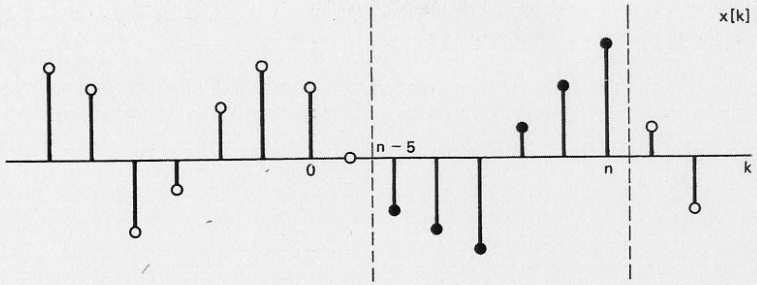

Figure 2.7 Sequence values involved in computing a causal moving average.

Classes of systems are defined by placing constraints on the properties of the transformation $T\{\cdot\}$. Doing so often leads to very general mathematical representations, as we will see. Of particular importance to consider are the system constraints and properties, discussed in Sections 2.2.1–2.2.5.

2.2.1 Memoryless Systems

A system is referred to as memoryless if the output $y[n]$ at every value of n depends only on the input $x[n]$ at the same value of n.

Example 2.3

An example of a memoryless system is one for which $x[n]$ and $y[n]$ are related by

$$y[n] = (x[n])^2 \quad \text{for each value of } n. \tag{2.24}$$

The system in Example 2.1 is not memoryless unless $n_d = 0$. In particular it is referred to as having "memory" whether n_d is positive (time delay) or negative (time advance). The system in Example 2.2 is not memoryless unless $M_1 = M_2 = 0$.

2.2.2 Linear Systems

The class of *linear systems* is defined by the principle of superposition. If $y_1[n]$ and $y_2[n]$ are the responses of a system when $x_1[n]$ and $x_2[n]$ are the respective inputs, then the system is linear if and only if

$$T\{x_1[n] + x_2[n]\} = T\{x_1[n]\} + T\{x_2[n]\} = y_1[n] + y_2[n] \tag{2.25a}$$

and

$$T\{ax[n]\} = aT\{x[n]\} = ay[n]. \tag{2.25b}$$

The first property is called the *additivity property* and the second is called the *homogeneity or scaling property*. These two properties can be combined into the principle of superposition, stated as

$$T\{ax_1[n] + bx_2[n]\} = aT\{x_1[n]\} + bT\{x_2[n]\} \tag{2.26}$$

for arbitrary constants a and b. This can be generalized to the superposition of many inputs. Specifically, if

$$x[n] = \sum_k a_k x_k[n], \tag{2.27a}$$

then

$$y[n] = \sum_k a_k y_k[n], \tag{2.27b}$$

where $y_k[n]$ is the system response to the input $x_k[n]$.

By using the definition of the principle of superposition, we can easily show that the systems of Examples 2.1 and 2.2 are linear systems. (See Problem 2.2.) An example of a nonlinear system is the system in Example 2.3.

Example 2.4

The system referred to as an accumulator and defined by the relation

$$y[n] = \sum_{k=-\infty}^{n} x[k] \tag{2.28}$$

is a linear system. The output at each sample instant n is equal to the accumulated sum of the value at n and all previous values of the input. Comparing Eq. (2.28) and Eq. (2.8), we note that if the input $x[n]$ is a unit impulse $\delta[n]$, then the output $y[n]$ is a unit step $u[n]$.

2.2.3 Time-Invariant Systems

A time-invariant system (equivalently often referred to as a shift-invariant system) is one for which a time shift or delay of the input sequence causes a corresponding shift in the output sequence. Specifically, suppose that a system transforms the input sequence with values $x[n]$ into the output sequence with values $y[n]$. The system is said to be time-invariant if for all n_0 the input sequence with values $x_1[n] = x[n - n_0]$ produces the output sequence with values $y_1[n] = y[n - n_0]$.

All of the systems in Examples 2.1–2.4 are time-invariant. The following example illustrates a system that is not time-invariant.

Example 2.5

The system defined by the relation

$$y[n] = x[Mn], \qquad -\infty < n < \infty, \tag{2.29}$$

with M a positive integer, is called a *compressor*. Specifically, it discards $(M - 1)$ samples out of M; i.e., it creates the output sequence by selecting every Mth sample. It should be clear simply from the description of its operation that this system is not time-invariant. We can see this mathematically by considering the response $y_1[n]$ to the input $x_1[n]$ defined as $x_1[n] = x[n - n_0]$. Then

$$y_1[n] = x_1[Mn] = x[Mn - n_0]. \tag{2.30}$$

But

$$y[n - n_0] = x[M(n - n_0)] \neq y_1[n]. \tag{2.31}$$

With $x_1[n]$ equal to $x[n]$ shifted by n_0, $y_1[n]$ is not $y[n]$ shifted by n_0, so the downsampler is not time-invariant, unless of course $M = 1$.

2.2.4 Causality

A system is causal if for every choice of n_0 the output sequence value at index $n = n_0$ depends only on the input sequence values for $n \le n_0$. This implies that if $x_1[n] = x_2[n]$ for $n \le n_0$, then $y_1[n] = y_2[n]$ for $n \le n_0$. That is, the system is *nonanticipative*. The system of Example 2.1 is causal for $n_d \ge 0$ and is noncausal for $n_d < 0$. The system of Example 2.2 is causal if $M_1 \le 0$ and $M_2 \ge -M_1$; otherwise it is noncausal. The system of Example 2.3 is causal, as is the accumulator of Example 2.4. However, the system of Example 2.5 is noncausal if $M > 1$ since $y[1] = x[M]$. Another noncausal system is given in the following example.

Example 2.6

The *forward difference system* is defined by the relation

$$y[n] = x[n+1] - x[n]. \tag{2.32}$$

Obviously $y[n]$ depends on $x[n+1]$; therefore the system is noncausal. However, the *backward difference system*, defined by

$$y[n] = x[n] - x[n-1], \tag{2.33}$$

is causal.

2.2.5 Stability

A system is stable in the bounded-input bounded-output (BIBO) sense if and only if every bounded input sequence produces a bounded output sequence. The input $x[n]$ is bounded if there exists a fixed positive finite value B_x such that

$$|x[n]| \le B_x < \infty \qquad \text{for all } n. \tag{2.34}$$

Stability requires that for every bounded input there exists a fixed positive finite value B_y such that

$$|y[n]| \le B_y < \infty \qquad \text{for all } n. \tag{2.35}$$

Examples 2.1, 2.2, 2.3, and 2.5 are stable systems. The accumulator of Example 2.4 is unstable. To demonstrate this we need only specify one bounded input for which the output is unbounded. Thus, consider $x[n] = u[n]$, which is clearly bounded. For this input to the accumulator, the output is

$$y[n] = \sum_{k=-\infty}^{n} u[k] = \begin{cases} 0, & n < 0, \\ (n+1), & n \ge 0. \end{cases} \tag{2.36}$$

Although this response is finite if n is finite, it is nevertheless unbounded, i.e., there is no fixed positive finite value B_y such that $(n+1) \le B_y < \infty$ for all n.

It is important to emphasize that the properties we have defined in this section are properties of *systems*, not of the inputs to the system. That is, we may be able to find inputs for which the properties hold, but the existence of the property for some inputs does not mean that the system has the property. For the system to have the property, it must hold for all inputs. For example, an unstable system may have some bounded inputs for which the output is bounded, but for the system to have the property of stability, it must be true that for *all* bounded inputs the output is bounded. If, as we were able to do with the accumulator system of Example 2.4, we can find just one input for which the system property does not hold, then we have shown that the system does *not* have that property.

2.3 LINEAR TIME-INVARIANT SYSTEMS

A particularly important class of systems consists of those that are linear and time-invariant. These two properties in combination lead to especially convenient representations for these systems. Most important, this class of systems has significant signal processing applications. The class of linear systems is defined by the principle of superposition in Eq. (2.26). If the linearity property is combined with the representation of a general sequence as a linear combination of delayed impulses as in Eq. (2.6), it follows that a linear system can be completely characterized by its impulse response. Specifically, let $h_k[n]$ be the response of the system to $\delta[n-k]$, an impulse occurring at $n = k$. Then from Eq. (2.6),

$$y[n] = T\left\{\sum_{k=-\infty}^{\infty} x[k]\delta[n-k]\right\}. \tag{2.37}$$

From the principle of superposition in Eq. (2.26), we can write

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]T\{\delta[n-k]\} = \sum_{k=-\infty}^{\infty} x[k]h_k[n]. \tag{2.38}$$

According to Eq. (2.38), the system response to any input can be expressed in terms of the response of the system to $\delta[n-k]$. If only linearity is imposed, $h_k[n]$ will depend on both n and k, in which case the computational usefulness of Eq. (2.38) is limited. We obtain a more useful result if we impose the additional constraint of time invariance.

The property of time invariance implies that if $h[n]$ is the response to $\delta[n]$, then the response to $\delta[n-k]$ is $h[n-k]$. With this additional constraint, Eq. (2.38) becomes

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]. \tag{2.39}$$

As a consequence of Eq. (2.39), a linear time-invariant system (which we will sometimes abbreviate as LTI) is completely characterized by its impulse response, $h[n]$, in the sense that, given $h[n]$, it is possible to use Eq. (2.39) to compute the output $y[n]$ due to *any* input $x[n]$.

Equation (2.39) is commonly called the *convolution sum.* If $y[n]$ is a sequence whose values are related to the values of two sequences $h[n]$ and $x[n]$ as in Eq. (2.39), we say that $y[n]$ is the convolution of $x[n]$ with $h[n]$ and represent this by the notation

$$y[n] = x[n] * h[n]. \tag{2.40}$$

The operation of discrete-time convolution takes two sequences $x[n]$ and $h[n]$ and produces a third sequence $y[n]$. Equation (2.39) expresses each sample of the output sequence in terms all of the samples of the input and the impulse response sequences.

The derivation of Eq. (2.39) suggests the interpretation that the input sample at $n = k$, represented as $x[k]\delta[n-k]$, is transformed by the system into an output sequence $x[k]h[n-k]$, $-\infty < n < \infty$ and that these sequences for each k are superimposed to form the overall output sequence. This interpretation is illustrated in Fig. 2.8, which shows an impulse response, a simple input sequence having three nonzero samples, the individual outputs due to each sample, and the composite output due to all the samples in the input sequences. Specifically, $x[n]$ can be decomposed as the sum of the three sequences $x[-2]\delta[n+2]$, $x[0]\delta[n]$, and $x[3]\delta[n-3]$ representing the three nonzero values in the sequence $x[n]$. The sequences $x[-2]h[n+2]$, $x[0]h[n]$, and $x[3]h[n-3]$ are the system responses to $x[-2]\delta[n+2]$, $x[0]\delta[n]$, and $x[3]\delta[n-3]$, respectively. The response to $x[n]$ is then the sum of these three individual responses.

Although the convolution-sum expression is analogous to the convolution integral of continuous-time linear system theory, the convolution sum should not be thought of as an approximation to the convolution integral. The convolution integral plays mainly a theoretical role in continuous-time linear system theory; we will see that the convolution sum, in addition to its theoretical importance, often serves as an explicit realization of a discrete-time linear system. Thus it is important to gain some insight into the properties of the convolution sum in actual calculations.

The preceding interpretation of Eq. (2.39) emphasizes that the convolution sum is a direct result of linearity and time invariance. However, a slightly different way of looking at Eq. (2.39) leads to a particularly useful computational interpretation. When viewed as a formula for computing a single value of the output sequence, Eq. (2.39) dictates that $y[n]$ (i.e., the nth value of the output) is obtained by multiplying the input sequence (expressed as a function of k) by the sequence whose values are $h[n-k]$, $-\infty < k < \infty$, and then, for any fixed value of n, summing all the values of the products $x[k]h[n-k]$, with k a counting index in the summation process. The operation of convolving two sequences thus involves doing the computation for all values of n, thus generating the complete output sequence $y[n]$, $-\infty < n < \infty$. The key to carrying out the computations of Eq. (2.39) to obtain $y[n]$ is understanding how to form the sequence $h[n-k]$, $-\infty < k < \infty$, for all values of n that are of interest. To this end, it is useful to note that

$$h[n-k] = h[-(k-n)]. \tag{2.41}$$

The interpretation of Eq. (2.41) is best done with an example.

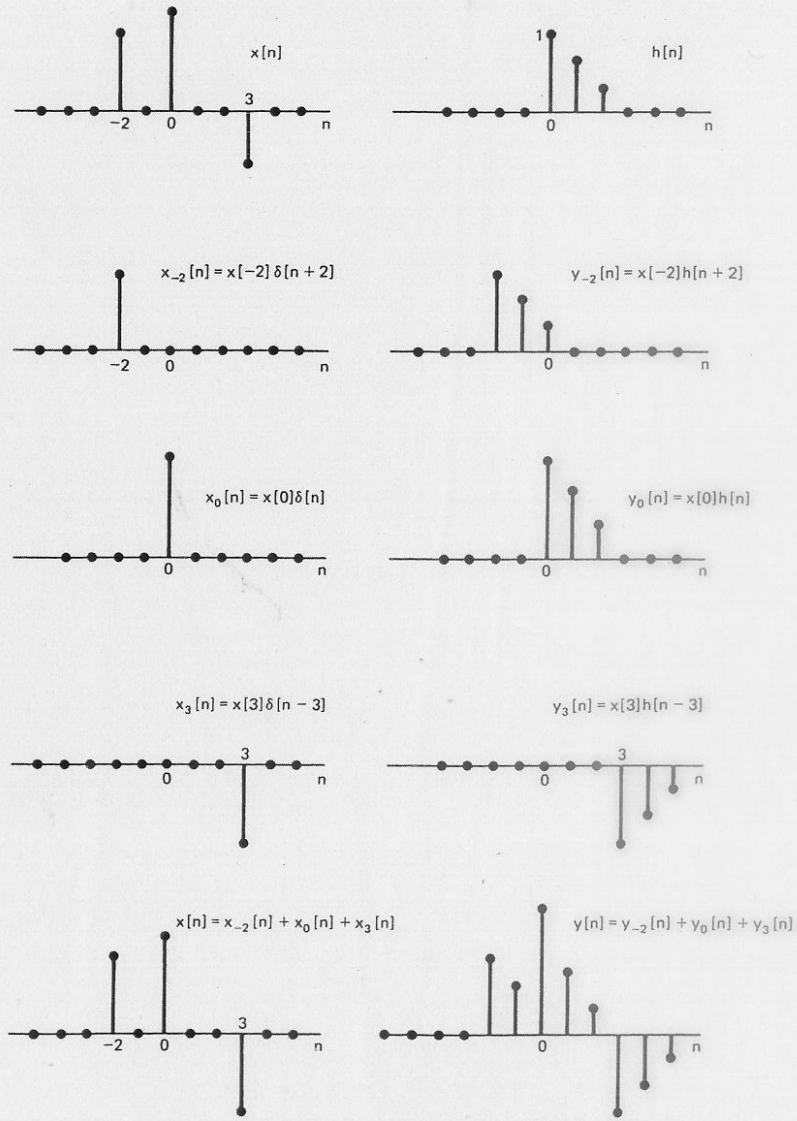

Figure 2.8 Representation of the output of a linear time-invariant system as the superposition of responses to individual samples of the input.

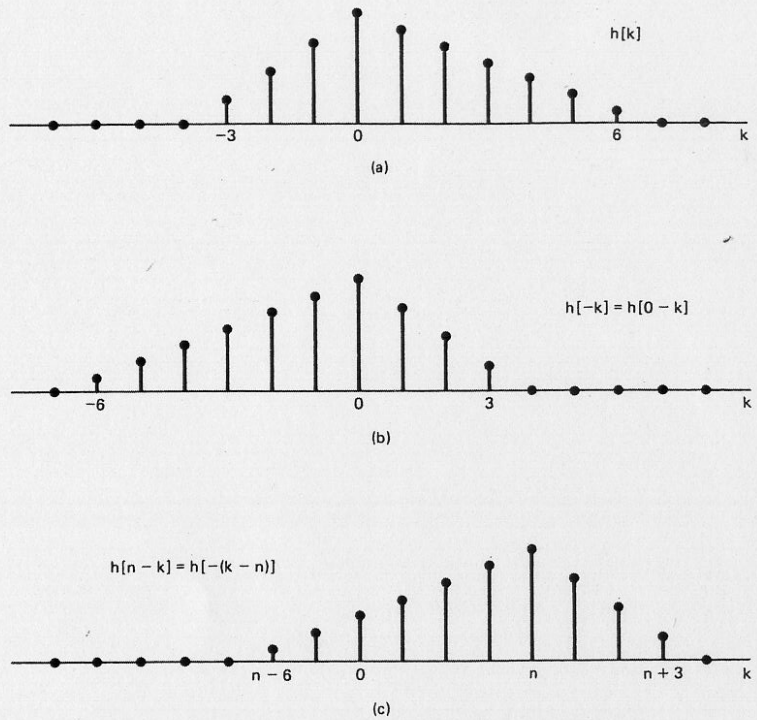

Figure 2.9 Forming the sequence $h[n-k]$. (a) The sequence $h[k]$ as a function of k. (b) The sequence $h[-k]$ as a function of k. (c) The sequence $h[n-k]$ as a function of k for $n = 4$.

Example 2.7

Suppose $h[k]$ is the sequence shown in Fig. 2.9(a). First consider $h[-k]$ plotted against k; $h[-k]$ is simply $h[k]$ reflected or "flipped" about $k = 0$, as shown in Fig. 2.9(b). Replacing k by $k - n$, where n is a fixed integer, leads to a shift of the origin of the sequence $h[-k]$ to $k = n$, as shown in Fig. 2.9(c) for the case $n = 4$.

Generalizing Example 2.7, it should be clear that in general the sequence $h[n-k]$, $-\infty < k < \infty$, is obtained by

1. reflecting $h[k]$ about the origin to obtain $h[-k]$;
2. shifting the origin of the reflected sequence to $k = n$.

To implement discrete-time convolution, the two sequences $x[k]$ and $h[n-k]$ are multiplied together for $-\infty < k < \infty$ and the products are summed to compute the output sample $y[n]$. To obtain another output sample, the origin of the sequence

$h[-k]$ is shifted to the new sample position and the process is repeated. This computational procedure applies whether the computations are carried out numerically on sampled data or analytically with sequences for which the sample values have simple formulas. The following example illustrates discrete-time convolution for the latter case.

Example 2.8

Consider a system with impulse response

$$h[n] = u[n] - u[n-N] = \begin{cases} 1, & 0 \le n \le N-1, \\ 0, & \text{otherwise.} \end{cases}$$

The input is

$$x[n] = a^n u[n].$$

To find the output at a particular index n, we must form the sums over all k of the product $x[k]h[n-k]$. In this case, we can find formulas for $y[n]$ for different sets of values of n. For example, Fig. 2.10(a) shows the sequences $x[k]$ and $h[n-k]$ plotted for n a negative integer. Clearly, all negative values of n give a similar picture; i.e., the nonzero portions of the sequences $x[k]$ and $h[n-k]$ do not overlap, so

$$y[n] = 0, \qquad n < 0.$$

Figure 2.10(b) illustrates the two sequences when $0 \le n$ and $n - N + 1 \le 0$. These two conditions can be combined into the single condition $0 \le n \le N-1$. By considering Fig. 2.10(b), we see that since

$$x[k]h[n-k] = a^k,$$

then

$$y[n] = \sum_{k=0}^{n} a^k \qquad \text{for } 0 \le n \le N-1. \tag{2.42}$$

The limits on the sum are determined directly from Fig. 2.10(b). Equation (2.42) shows that $y[n]$ is the sum of $n + 1$ terms of a geometric series where the ratio of terms is a. This sum can be expressed in closed form using the general formula

$$\sum_{k=N_1}^{N_2} \alpha^k = \frac{\alpha^{N_1} - \alpha^{N_2+1}}{1-\alpha}, \qquad N_2 \ge N_1. \tag{2.43}$$

Applying this to Eq. (2.42), we obtain

$$y[n] = \frac{1-a^{n+1}}{1-a}, \qquad 0 \le n \le N-1. \tag{2.44}$$

Finally, Fig. 2.10(c) shows the two sequences when $0 < n - N + 1$ or $N - 1 < n$. As before,

$$x[k]h[n-k] = a^k, \qquad n - N + 1 < k \le n,$$

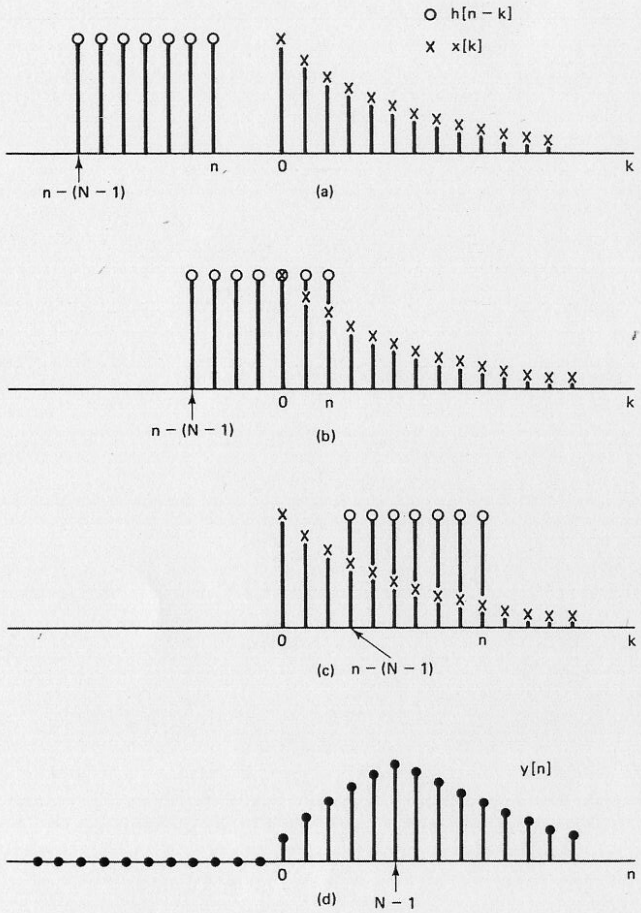

Figure 2.10 Sequences involved in computing a discrete convolution. (a)–(c) The sequences $x[k]$ and $h[n-k]$ as a function of k for different values of n. (Only nonzero samples are shown.) (d) Corresponding output sequence as a function of n.

but now the lower limit on the sum is $n - N + 1$, as seen in Fig. 2.10(c). Thus

$$y[n] = \sum_{k=n-N+1}^{n} a^k \qquad \text{for } N-1 < n. \tag{2.45}$$

Using Eq. (2.43) we obtain

$$y[n] = \frac{a^{n-N+1} - a^{n+1}}{1-a}$$

or

$$y[n] = a^{n-N+1}\left(\frac{1-a^N}{1-a}\right). \tag{2.46}$$

Thus, because of the piecewise-exponential nature of both the input and the unit sample response, we have been able to obtain the following closed-form expression for $y[n]$ as a function of the index n:

$$y[n] = \begin{cases} 0, & n < 0, \\ \dfrac{1-a^{n+1}}{1-a}, & 0 \le n \le N-1, \\ a^{n-N+1}\left(\dfrac{1-a^N}{1-a}\right), & N-1 < n. \end{cases} \tag{2.47}$$

This sequence is shown in Fig. 2.10(d).

Example 2.8 illustrates how the convolution sum can be computed analytically when the input and the impulse response are given by simple formulas. In such cases, the sums may have a compact form that may be derived using the formula for the sum of a geometric series or other "closed-form" formulas.† When no simple form is available, the convolution sum can still be evaluated numerically using the technique illustrated in Example 2.8 whenever the sums are finite, which will be the case if either the input sequence or the impulse response is of finite length, i.e., has a finite number of nonzero samples. We cannot overemphasize that neatly drawn and labeled sketches such as those in Fig. 2.10 are invaluable to ensure that the limits on the sum are correctly determined.

2.4 PROPERTIES OF LINEAR TIME-INVARIANT SYSTEMS

Since all linear time-invariant systems are described by the convolution sum of Eq. (2.39), the properties of this class of systems are defined by the properties of discrete-time convolution. Therefore the impulse response is a complete characterization of the properties of a specific linear time-invariant system.

Some general properties of the class of linear time-invariant systems can be found by considering properties of the convolution operation. For example, the convolution operation is commutative:

$$x[n] * h[n] = h[n] * x[n]. \tag{2.48}$$

This can be shown by applying a substitution of variables to Eq. (2.39). Specifically, with $m = n - k$,

$$y[n] = \sum_{m=-\infty}^{\infty} x[n-m]h[m], \tag{2.49}$$

† Such results are discussed, for example, in Grossman (1982). A very useful book of sums is Jolley (1961).

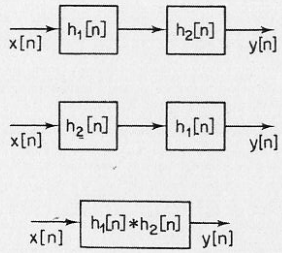

Figure 2.11 Three linear time-invariant systems with identical impulse responses.

so the roles of $x[n]$ and $h[n]$ in the summation are interchanged. That is, the order of the sequences in a convolution is unimportant, and hence the system output is the same if the roles of the input and impulse response are reversed. A linear time-invariant system with input $x[n]$ and impulse response $h[n]$ will have the same output as a linear time-invariant system with input $h[n]$ and impulse response $x[n]$. The convolution operation also distributes over addition, i.e.,

$$x[n] * (h_1[n] + h_2[n]) = x[n] * h_1[n] + x[n] * h_2[n].$$

This follows in a straightforward way from Eq. (2.39).

In a *cascade connection* of systems, the output of the first system is the input to the second, the output of the second is the input to the third, etc. The output of the last system is the overall output. Two linear time-invariant systems in cascade correspond to a linear time-invariant system with an impulse response that is the convolution of the impulse responses of the two systems. This is illustrated in Fig. 2.11. In the upper block diagram, the output of the first system will be $h_1[n]$ if $x[n] = \delta[n]$. Thus the output of the second system (and the overall impulse response of the system) will be

$$h[n] = h_1[n] * h_2[n]. \tag{2.50}$$

As a consequence of the commutative property of convolution, it follows that the impulse response of a cascade combination of linear time-invariant systems is independent of the order in which they are cascaded. This result is summarized in Fig. 2.11, where the three systems all have the same impulse response.

In a *parallel connection*, the systems have the same input and their outputs are summed to produce an overall output. It follows from the distributive property of convolution that two linear time-invariant systems in parallel are equivalent to a single system whose impulse response is the sum of the individual impulse responses, i.e.,

$$h[n] = h_1[n] + h_2[n]. \tag{2.51}$$

This is depicted in Fig. 2.12.

The constraints of linearity and time invariance define a class of systems with very special properties. Stability and causality represent additional properties, and it is often important to know whether a linear time-invariant system is stable and whether it is causal. Recall from Section 2.2.5 that a stable system is one for which

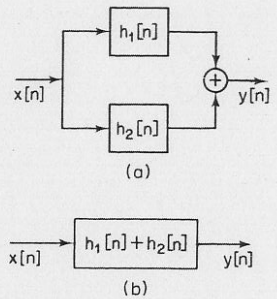

Figure 2.12 (a) Parallel combination of linear time-invariant systems. (b) An equivalent system.

every bounded input produces a bounded ouput. Linear time-invariant systems are stable if and only if the impulse response is absolutely summable, i.e., if

$$S = \sum_{k=-\infty}^{\infty} |h[k]| < \infty. \tag{2.52}$$

This can be shown as follows. From Eq. (2.49),

$$|y[n]| = \left| \sum_{k=-\infty}^{\infty} h[k]x[n-k] \right| \le \sum_{k=-\infty}^{\infty} |h[k]|\,|x[n-k]|. \tag{2.53}$$

If $x[n]$ is bounded so that

$$|x[n]| \le B_x,$$

then

$$|y[n]| \le B_x \sum_{k=-\infty}^{\infty} |h[k]|. \tag{2.54}$$

Thus $y[n]$ is bounded if Eq. (2.52) holds; i.e., Eq. (2.52) is a sufficient condition for stability. To show that it is also a necessary condition, we must show that if $S = \infty$, then a bounded input can be found that will cause an unbounded output. Such an input is the sequence with values

$$x[n] = \begin{cases} \dfrac{h^*[-n]}{|h[-n]|}, & h[n] \ne 0, \\ 0, & h[n] = 0, \end{cases} \tag{2.55}$$

where $h^*[n]$ is the complex conjugate of $h[n]$. The sequence $x[n]$ is clearly bounded by unity. However, the value of the output at $n = 0$ is

$$y[0] = \sum_{k=-\infty}^{\infty} x[-k]h[k] = \sum_{k=-\infty}^{\infty} \frac{|h[k]|^2}{|h[k]|} = S. \tag{2.56}$$

Thus if $S = \infty$, it is possible for a bounded input sequence to produce an unbounded output sequence.

The class of causal systems was defined in Section 2.2.4 as those systems for which the output $y[n_0]$ depends only on the input samples $x[n]$, for $n \leq n_0$. It follows from Eq. (2.39) or Eq. (2.49) that this definition implies the condition

$$h[n] = 0, \qquad n < 0, \tag{2.57}$$

for causality of linear time-invariant systems. (See Problem 2.12.) For this reason it is sometimes convenient to refer to a sequence that is zero for $n < 0$ as a *causal sequence*, meaning that it could be the impulse response of a causal system.

To illustrate how the properties of linear time-invariant systems are reflected in the impulse response, let us reconsider some of the systems defined in Examples 2.1–2.6. First note that only the systems of Examples 2.1, 2.2, 2.4, and 2.6 are linear and time-invariant. Although the impulse response of nonlinear or time-varying systems can be found, it is generally of limited interest since the convolution-sum formula and Eqs. (2.52) and (2.57), expressing stability and causality, do not apply to such systems.

First, let us find the impulse responses of the systems in Examples 2.1, 2.2, 2.4, and 2.6. We can do this by simply computing the response of each system to $\delta[n]$ using the defining relationship for the system. The resulting impulse responses are as follows:

Ideal Delay (Example 2.1)

$$h[n] = \delta[n - n_d], \qquad n_d \text{ a positive fixed integer} \tag{2.58}$$

Moving Average (Example 2.2)

$$\begin{aligned} h[n] &= \frac{1}{M_1 + M_2 + 1} \sum_{k=-M_1}^{M_2} \delta[n-k] \\ &= \begin{cases} \dfrac{1}{M_1 + M_2 + 1}, & -M_1 \leq n \leq M_2 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{2.59}$$

Accumulator (Example 2.4)

$$\begin{aligned} h[n] &= \sum_{k=-\infty}^{n} \delta[k] \\ &= \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases} \\ &= u[n] \end{aligned} \tag{2.60}$$

Forward Difference (Example 2.6)

$$h[n] = \delta[n+1] - \delta[n] \tag{2.61}$$

Backward Difference (Example 2.6)

$$h[n] = \delta[n] - \delta[n-1] \tag{2.62}$$

Next let us test the stability of each system by computing the sum

$$S = \sum_{n=-\infty}^{\infty} |h[n]|.$$

For the ideal delay, moving average, forward difference, and backward difference examples, it is clear that $S < \infty$ since the impulse response has only a finite number of nonzero samples. Such systems are called *finite-duration impulse response* (FIR) systems. Clearly FIR systems will always be stable as long as each of the impulse response values is finite in magnitude. The accumulator, however, is unstable because

$$S = \sum_{n=0}^{\infty} u[n] = \infty.$$

In Section 2.2.5 we also demonstrated the instability of the accumulator by giving an example of a bounded input (the unit step) for which the output is unbounded.

The impulse response of the accumulator is infinite in duration. This is an example of the class of systems referred to as *infinite-duration impulse response* (IIR) systems. An example of an IIR system that is stable is one having the impulse response $h[n] = a^n u[n]$ with $|a| < 1$. In this case

$$S = \sum_{n=0}^{\infty} |a|^n. \tag{2.63}$$

If $|a| < 1$, the formula for the sum of the terms of an infinite geometric series gives

$$S = \frac{1}{1 - |a|} < \infty. \tag{2.64}$$

If, on the other hand, $|a| \geq 1$, the sum is infinite and the system is unstable.

To test causality of the linear time-invariant systems in Examples 2.1, 2.2, 2.4, and 2.6, we can check to see if $h[n] = 0$ for $n < 0$. As discussed in Section 2.2.4, the ideal delay ($n_d \geq 0$ in Eq. 2.22) is causal. If $n_d < 0$, the system is noncausal. For the moving average, causality requires that $-M_1 \geq 0$ and $M_2 \geq 0$. The accumulator and backward difference systems are causal, and the forward difference system is noncausal.

The property of convolution can be used to describe and analyze interconnections of linear time-invariant systems. Consider the system of Fig. 2.13(a), which consists of a forward difference system cascaded with an ideal delay of one sample. According to the commutative property of convolution, the order in which systems are cascaded does not matter as long as they are linear and time-invariant. Therefore we obtain the same result when we compute the forward difference of a sequence and delay the result (Fig. 2.13a) as when we delay the sequence first and then compute the forward difference (Fig. 2.13b). Also, it follows from Eq. (2.50) that the

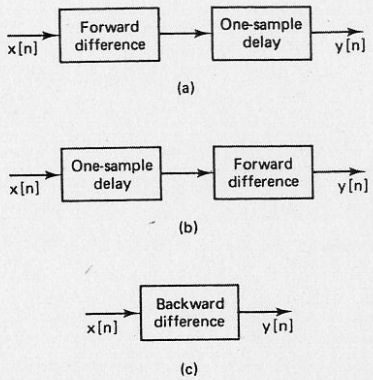

Figure 2.13 Equivalent systems found by using the commutative property of convolution.

overall impulse response of each cascade system is the convolution of the individual impulse responses. Therefore

$$\begin{aligned} h[n] &= (\delta[n+1] - \delta[n]) * \delta[n-1] \\ &= \delta[n-1] * (\delta[n+1] - \delta[n]) \\ &= \delta[n] - \delta[n-1]. \end{aligned} \tag{2.65}$$

Thus $h[n]$ is identical to the impulse response of the backward difference system, i.e., the cascaded systems of Figs. 2.13(a) and 2.13(b) can be replaced by a backward difference system, as shown in Fig. 2.13(c).

Note that the noncausal forward difference systems in Figs. 2.13(a) and (b) have been converted to causal systems by cascading them with a delay. In general, any noncausal FIR system can be made causal by cascading it with a sufficiently long delay.

Another example of cascaded systems introduces the concept of an *inverse system*. Consider the cascade of systems in Fig. 2.14. The impulse response of the cascade system is

$$\begin{aligned} h[n] &= u[n] * (\delta[n] - \delta[n-1]) \\ &= u[n] - u[n-1] \\ &= \delta[n]. \end{aligned} \tag{2.66}$$

That is, the cascade combination of an accumulator followed by a backward difference (or vice versa) yields a system whose overall impulse response is the impulse. Thus the output of the cascade system will always be equal to the input since

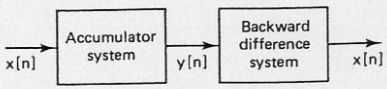

Figure 2.14 An accumulator in cascade with a backward difference. Since the backward difference is the inverse system for the accumulator, the cascade combination is equivalent to the identity system.

$x[n] * \delta[n] = x[n]$. In this case, the backward difference system compensates exactly for (or inverts) the effect of the accumulator, i.e., the backward difference system is the *inverse system* for the accumulator. From the commutative property of convolution, the accumulator is likewise the inverse system for the backward difference system. Also note that this example provides a system interpretation of Eqs. (2.8) and (2.10). In general, if a linear time-invariant system has impulse response $h[n]$, then its inverse system, if it exists, has impulse response $h_i[n]$ defined by the relation

$$h[n] * h_i[n] = h_i[n] * h[n] = \delta[n]. \tag{2.67}$$

Inverse systems are useful in many situations where it is necessary to compensate for the effects of a linear system. In general, it is difficult to solve Eq. (2.67) directly for $h_i[n]$ given $h[n]$. However, in Chapter 5 we will see that the Fourier transform and more generally the z-transform provide a straightforward method of finding an inverse system.

2.5 LINEAR CONSTANT-COEFFICIENT DIFFERENCE EQUATIONS

An important subclass of linear time-invariant systems consists of those systems for which the input $x[n]$ and the output $y[n]$ satisfy an Nth-order linear constant-coefficient difference equation of the form

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{n=0}^{M} b_n x[n-n]. \tag{2.68}$$

The properties discussed in Section 2.4 and some of the analysis techniques introduced there can be used to find difference equation representations for some of the linear time-invariant systems that we have defined.

Example 2.9

An example of the class of linear constant-coefficient difference equations is the accumulator (Example 2.4). To show that its input and output satisfy such a difference equation, consider Fig. 2.14 again. If $y[n]$ represents the output of the accumulator, then we have seen that the backward difference of $y[n]$ is equal to $x[n]$ since the backward difference system is the inverse of the accumulator; i.e.,

$$y[n] - y[n-1] = x[n]. \tag{2.69}$$

Thus we have shown that in addition to the defining relationship of Eq. (2.28), which is in the form of a convolution, the input and output also satisfy a linear constant-coefficient difference equation of the form of Eq. (2.68), where $N = 1$, $a_0 = 1$, $a_1 = -1$, $M = 0$, and $b_0 = 1$.

The difference equation in Eq. (2.69) as a representation of the accumulator can be better understood by rewriting it as

$$y[n] = y[n-1] + x[n]. \tag{2.70}$$

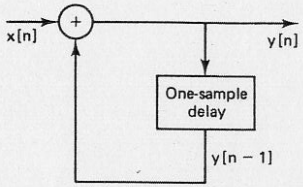

Figure 2.15 Block diagram of a recursive difference equation representing an accumulator.

According to Eq. (2.70), for each value of n we add the current input value $x[n]$ to the previously accumulated sum $y[n-1]$. This interpretation of the accumulator is represented in block diagram form in Fig. 2.15. Equation (2.70) and the block diagram in Fig. 2.15 are referred to as a *recursion representation* of the system since each value is computed using previously computed values. This general notion will be explored in more detail later in this section.

Example 2.10

Let us consider the moving average system of Example 2.2, with $M_1 = 0$ so that the system is causal. In this case, from Eq. (2.59) the impulse response is

$$h[n] = \frac{1}{(M_2+1)}(u[n] - u[n-M_2-1]), \tag{2.71}$$

from which it follows that

$$y[n] = \frac{1}{(M_2+1)}\sum_{k=0}^{M_2} x[n-k], \tag{2.72}$$

which is a special case of Eq. (2.68), with $N = 0$, $a_0 = 1$, and $b_k = 1/(M_2+1)$ for $0 \le k \le M_2$. The impulse response can also be expressed as

$$h[n] = \frac{1}{(M_2+1)}(\delta[n] - \delta[n-M_2-1]) * u[n], \tag{2.73}$$

which suggests that the causal moving average system can be represented as the cascade system of Fig. 2.16. We can obtain a difference equation for this block diagram by noting first that

$$x_1[n] = \frac{1}{(M_2+1)}(x[n] - x[n-M_2-1]). \tag{2.74}$$

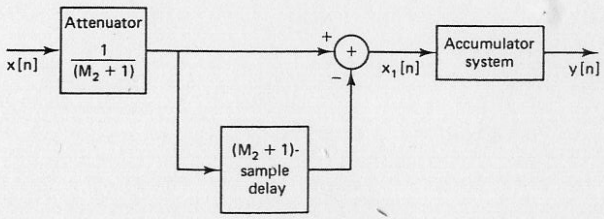

Figure 2.16 Block diagram of the recursive form of a moving average system.

From Eq. (2.69) of Example 2.9, the output of the accumulator satisfies the difference equation

$$y[n] - y[n-1] = x_1[n]$$

so that

$$y[n] - y[n-1] = \frac{1}{(M_2+1)}(x[n] - x[n-M_2-1]). \tag{2.75}$$

Again, we have a difference equation in the form of Eq. (2.68), but this time $N = 1$, $a_0 = 1$, $a_1 = -1$, and $b_0 = -b_{M_2+1} = 1/(M_2+1)$, and $b_k = 0$ otherwise.

Interestingly, there are several distinct forms for the representation of the moving average system in terms of difference equations. In Chapter 6 we will see that an unlimited number of distinct difference equations can be used to represent a given linear time-invariant input/output relation.

Just as in the case of linear constant-coefficient differential equations for continuous-time systems, without additional constraints or information a linear constant-coefficient difference equation for discrete-time systems does not provide a unique specification of the output for a given input. Specifically, suppose that for a given input $x_p[n]$ we have determined by some means one output sequence $y_p[n]$ so that an equation of the form of Eq. (2.68) is satisfied. Then the same equation with the same input is satisfied by any output of the form

$$y[n] = y_p[n] + y_h[n], \tag{2.76}$$

where $y_h[n]$ is any solution to Eq. (2.68) with $x[n] = 0$, i.e., to the equation

$$\sum_{k=0}^{N} a_k y_h[n-k] = 0. \tag{2.77}$$

Equation (2.77) is referred to as the *homogeneous equation* and $y_h[n]$ the *homogeneous solution*. The sequence $y_h[n]$ is in fact a member of a family of solutions of the form

$$y_h[n] = \sum_{m=1}^{N} A_m z_m^n. \tag{2.78}$$

Substituting Eq. (2.78) into Eq. (2.77) shows that the complex numbers z_m must be roots of the polynomial

$$\sum_{k=0}^{N} a_k z^{-k} = 0. \tag{2.79}$$

Equation (2.78) assumes that all N roots of the polynomial in Eq. (2.79) are distinct. The form of terms associated with multiple roots is slightly different, but there are always N undetermined coefficients. An example of the homogeneous solution with multiple roots is considered in Problem 2.13.

Since $y_h[n]$ has N undetermined coefficients, a set of N auxiliary conditions is required for unique specification of $y[n]$ for a given $x[n]$. These auxiliary conditions might consist of specifying fixed values of $y[n]$ at specific values of n, such as $y[-1]$,

$y[-2], \ldots, y[-N]$, and then solving a set of N linear equations for the N undetermined coefficients.

Alternatively, if the auxiliary conditions are a set of auxiliary values of $y[n]$, the other values of $y[n]$ can be generated by rewriting Eq. (2.68) as a recurrence formula, i.e., in the form

$$y[n] = -\sum_{k=1}^{N} \frac{a_k}{a_0} y[n-k] + \sum_{r=0}^{M} \frac{b_r}{a_0} x[n-r]. \tag{2.80}$$

If the input $x[n]$ together with a set of auxiliary values, say $y[-1], y[-2], \ldots, y[-N]$, is specified, then $y[0]$ can be determined from Eq. (2.80). With $y[0], y[-1], \ldots, y[-N+1]$ available, $y[1]$ can then be calculated, and so on. Using this procedure, $y[n]$ is said to be computed *recursively*; i.e., the output computation involves not only the input sequence but also previous values of the output sequence.

To generate values of $y[n]$ for $n < -N$ (again assuming that the values $y[-1], y[-2], \ldots, y[-N]$ are given as auxiliary conditions), we can rearrange Eq. (2.68) in the form

$$y[n-N] = -\sum_{k=0}^{N-1} \frac{a_k}{a_N} y[n-k] + \sum_{r=0}^{M} \frac{b_r}{a_N} x[n-r], \tag{2.81}$$

from which $y[-N-1], y[-N-2], \ldots$, can be computed recursively. The following example illustrates this recursive procedure.

Example 2.11

The difference equation satisfied by the input and output of a system is

$$y[n] = ay[n-1] + x[n]. \tag{2.82}$$

Consider the input $x[n] = K\delta[n]$, where K is an arbitrary number, and the auxiliary condition $y[-1] = c$. Beginning with this value, the output for $n > -1$ can be computed recursively as follows:

$$\begin{aligned}
y[0] &= ac + K, \\
y[1] &= ay[0] + 0 \\
&= a(ac + K) = a^2c + aK, \\
y[2] &= a(a^2c + aK) = a^3c + a^2K, \\
y[3] &= a(a^3c + a^2K) = a^4c + a^3K, \\
&\vdots
\end{aligned}$$

For this simple case, we can see that for $n \geq 0$, $y[n]$ is given by

$$y[n] = a^{n+1}c + a^nK \qquad \text{for } n \geq 0. \tag{2.83}$$

To determine the output for $n < 0$, we express the difference equation in the form

$$y[n-1] = a^{-1}(y[n] - x[n]), \tag{2.84a}$$

or

$$y[n] = a^{-1}(y[n+1] - x[n+1]). \tag{2.84b}$$

Using the auxiliary condition $y[-1] = c$, we can compute $y[n]$ for $n < -1$ as follows:

$$\begin{aligned}
y[-2] &= a^{-1}\{(y[-1] - x[-1])\} \\
&= a^{-1}c, \\
y[-3] &= a^{-1}a^{-1}c = a^{-2}c, \\
y[-4] &= a^{-1}a^{-2}c = a^{-3}c.
\end{aligned}$$

Clearly,

$$y[n] = a^{n+1}c \qquad \text{for } n \leq -1. \tag{2.85}$$

Thus, in summary, combining Eqs. (2.83) and (2.85) we obtain as the result of the recursive computation

$$y[n] = a^{n+1}c + Ka^nu[n] \qquad \text{for all } n. \tag{2.86}$$

Several important points are illustrated by the solution of Example 2.11. First note that we implemented the system by recursively computing the output in both positive and negative directions beginning with $n = -1$. Clearly this procedure is noncausal. Also note that when $K = 0$, the input is zero, but $y[n] = a^{n+1}c$. A linear system requires that the output be zero for all time when the input is zero for all time (see Problem 2.1). Consequently, this system is not linear. Furthermore, if the input were shifted by n_0 samples, i.e., $x_1[n] = K\delta[n - n_0]$, the output would be

$$y_1[n] = a^{n+1}c + Ka^{n-n_0}u[n - n_0], \tag{2.87}$$

and the system therefore is not time-invariant.

Our principal interest in this text is in systems that are linear and time-invariant, in which case the auxiliary conditions must be consistent with these additional requirements. In Chapter 4, when we discuss the solution of difference equations using the z-transform, we implicitly incorporate conditions of linearity and time invariance. As we will see in that discussion, even with the additional constraints of linearity and time invariance, the solution to the difference equation, and therefore the system, is not uniquely specified. In particular, there are in general both causal and noncausal linear time-invariant systems consistent with a given difference equation.

If a system is characterized by a linear constant-coefficient difference equation and is further specified to be linear, time-invariant, and causal, the solution is unique. In this case the auxiliary conditions are often stated as *initial-rest conditions*. In other words, the auxiliary information is that if the input $x[n]$ is zero for n less than some time n_0, then the output $y[n]$ is constrained to be zero for n less than n_0. This then provides sufficient initial conditions to obtain $y[n]$ for $n \geq n_0$ recursively using Eq. (2.80).

To summarize, for a system for which the input and output satisfy a linear constant-coefficient difference equation:

- The output for a given input is not uniquely specified. Auxiliary information or conditions are required.

- If the auxiliary information is in the form of N sequential values of the output, later values can be obtained by rearranging the difference equation as a recursive relation running forward in n, and prior values can be obtained by rearranging the difference equation as a recursive relation running backward in n.
- Linearity, time invariance, and causality of the system will depend on the auxiliary conditions. If an additional condition is that the system is initially at rest, then the system will be linear, time-invariant, and causal.

With the preceding discussion in mind, let us now reconsider Example 2.11, but with initial-rest conditions. With $x[n] = K\delta[n]$, $y[-1] = 0$ since $x[n] = 0$, $n < 0$. Consequently from Eq. (2.86),

$$y[n] = Ka^n u[n]. \tag{2.88}$$

If the input is instead $K\delta[n - n_0]$, again with initial-rest conditions, then the recursive solution is carried out using the initial condition $y[n] = 0$, $n < n_0$. Note that for $n_0 < 0$, initial rest implies that $y[-1] \neq 0$. That is, initial rest does not always mean $y[-1] = \cdots = y[-N] = 0$. It does mean that $y[n_0 - 1] = \cdots = y[n_0 - N] = 0$ if $x[n] = 0$ for $n < n_0$. Note also that the impulse response for the example is $h[n] = a^n u[n]$, i.e., it is zero for $n < 0$, consistent with the causality imposed by the assumption of initial rest.

The preceding discussion assumed that $N \geq 1$ in Eq. (2.68). If $N = 0$ in Eq. (2.68), no recursion is required to use the difference equation to compute the output and therefore no auxiliary conditions are required. That is,

$$y[n] = \sum_{k=0}^{M} \left(\frac{b_k}{a_0}\right) x[n-k]. \tag{2.89}$$

Equation (2.89) is in the form of a convolution, and by setting $x[n] = \delta[n]$ we see that the impulse response is

$$h[n] = \sum_{k=0}^{M} \left(\frac{b_k}{a_0}\right) \delta[n-k],$$

or

$$h[n] = \begin{cases} \left(\dfrac{b_k}{a_0}\right), & 0 \leq n \leq M, \\ 0, & \text{otherwise.} \end{cases} \tag{2.90}$$

The impulse response is obviously finite in duration. Indeed, the output of any FIR system can be computed nonrecursively using the difference equation of Eq. (2.89), where the coefficients are the values of the impulse sequence. The moving average system of Example 2.10 with $M_1 = 0$ is an example of a causal FIR system. An interesting feature of that system was that we also found a recursive equation for

the output. In Chapter 6 we will show that there are many possible ways of implementing a desired signal transformation using difference equations. Advantages of one method over another depend on practical considerations such as numerical accuracy, data storage, and the number of multiplications and additions required to compute each sample of the output.

2.6 FREQUENCY-DOMAIN REPRESENTATION OF DISCRETE-TIME SIGNALS AND SYSTEMS

In the previous sections we have introduced some of the fundamental concepts of the theory of discrete-time signals and systems. For linear time-invariant systems we saw that a representation of the input sequence as a weighted sum of delayed impulses leads to a representation of the output as a weighted sum of delayed responses. As with continuous-time signals, discrete-time signals may be represented in a number of different ways. For example, sinusoidal and complex exponential sequences play a particularly important role in representing discrete-time signals. This is because complex exponential sequences are eigenfunctions of linear time-invariant systems, and the response to a sinusoidal input is sinusoidal with the same frequency as the input and with amplitude and phase determined by the system. This fundamental property of linear time-invariant systems makes representations of signals in terms of sinusoids or complex exponentials (i.e., Fourier representations) very useful in linear system theory.

To demonstrate the eigenfunction property of complex exponentials for discrete-time systems, consider an input sequence $x[n] = e^{j\omega n}$ for $-\infty < n < \infty$, i.e., a complex exponential of radian frequency ω. From Eq. (2.49), the corresponding output of a linear time-invariant system with impulse response $h[n]$ is

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[k] e^{j\omega(n-k)} \\ &= e^{j\omega n}\left(\sum_{k=-\infty}^{\infty} h[k] e^{-j\omega k}\right). \end{aligned} \tag{2.91}$$

If we define

$$H(e^{j\omega}) = \sum_{k=-\infty}^{\infty} h[k] e^{-j\omega k}, \tag{2.92}$$

Eq. (2.91) becomes

$$y[n] = H(e^{j\omega}) e^{j\omega n}. \tag{2.93}$$

Consequently, $e^{j\omega n}$ is an eigenfunction of the system, and the associated eigenvalue is $H(e^{j\omega})$. From Eq. (2.93) we see that $H(e^{j\omega})$ describes the change in complex amplitude of a complex exponential as a function of the frequency ω. The eigenvalue $H(e^{j\omega})$

is called the *frequency response* of the system. In general, $H(e^{j\omega})$ is complex and can be expressed in terms of its real and imaginary parts as

$$H(e^{j\omega}) = H_R(e^{j\omega}) + jH_I(e^{j\omega}) \tag{2.94}$$

or in terms of magnitude and phase as

$$H(e^{j\omega}) = |H(e^{j\omega})|e^{j\measuredangle H(e^{j\omega})}. \tag{2.95}$$

Example 2.12

As a simple example of how we can find the frequency response of a linear time-invariant system, consider the ideal delay system defined by

$$y[n] = x[n - n_d], \tag{2.96}$$

where n_d is a fixed integer. If we consider $x[n] = e^{j\omega n}$ as input to this system, from Eq. (2.96) we have

$$y[n] = e^{j\omega(n-n_d)} = e^{-j\omega n_d}e^{j\omega n}.$$

Thus for any given value of ω, we obtain an output that is the input multiplied by a complex constant, the value of which depends on ω. The frequency response of the ideal delay is therefore

$$H(e^{j\omega}) = e^{-j\omega n_d}. \tag{2.97}$$

As an alternative method of obtaining the frequency response, recall that $h[n] = \delta[n - n_d]$ for the ideal delay system. Using Eq. (2.92) we obtain

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \delta[n - n_d]e^{-j\omega n} = e^{-j\omega n_d}.$$

Using the Euler relation, the real and imaginary parts of the frequency response are

$$H_R(e^{j\omega}) = \cos \omega n_d, \tag{2.98a}$$

$$H_I(e^{j\omega}) = -\sin \omega n_d. \tag{2.98b}$$

The magnitude and phase are

$$|H(e^{j\omega})| = 1, \tag{2.99a}$$

$$\measuredangle H(e^{j\omega}) = -\omega n_d. \tag{2.99b}$$

In Section 2.7 we will show that a broad class of signals can be represented as a linear combination of complex exponentials in the form

$$x[n] = \sum_k \alpha_k e^{j\omega_k n}. \tag{2.100}$$

From the principle of superposition, the corresponding output of a linear time-invariant system is

$$y[n] = \sum_k \alpha_k H(e^{j\omega_k})e^{j\omega_k n}. \tag{2.101}$$

Thus if we can find a representation of $x[n]$ as a superposition of complex exponential sequences, as in Eq. (2.100), then we can find the output using Eq. (2.101) if we know

the frequency response of the system. The following simple example illustrates this fundamental property of linear time-invariant systems.

Example 2.13

Since it is simple to express a sinusoid as a linear combination of complex exponentials, let us consider a sinusoidal input

$$x[n] = A\cos(\omega_0 n + \phi) = \frac{A}{2}e^{j\phi}e^{j\omega_0 n} + \frac{A}{2}e^{-j\phi}e^{-j\omega_0 n}. \tag{2.102}$$

From Eq. (2.93), the response to $x_1[n] = (A/2)e^{j\phi}e^{j\omega_0 n}$ is

$$y_1[n] = H(e^{j\omega_0})\frac{A}{2}e^{j\phi}e^{j\omega_0 n}. \tag{2.103a}$$

The response to $x_2[n] = (A/2)e^{-j\phi}e^{-j\omega_0 n}$ is

$$y_2[n] = H(e^{-j\omega_0})\frac{A}{2}e^{-j\phi}e^{-j\omega_0 n}. \tag{2.103b}$$

Thus the total response is

$$y[n] = \frac{A}{2}[H(e^{j\omega_0})e^{j\phi}e^{j\omega_0 n} + H(e^{-j\omega_0})e^{-j\phi}e^{-j\omega_0 n}]. \tag{2.104}$$

If $h[n]$ is real, it can be shown (see Problem 2.26) that $H(e^{-j\omega_0}) = H^*(e^{j\omega_0})$. Consequently

$$y[n] = A|H(e^{j\omega_0})|\cos(\omega_0 n + \phi + \theta), \tag{2.105}$$

where $\theta = \measuredangle H(e^{j\omega_0})$ is the phase of the system at frequency ω_0.

For the simple example of the ideal delay, $|H(e^{j\omega_0})| = 1$ and $\theta = -\omega_0 n_d$, as we determined in Example 2.12. Therefore

$$\begin{aligned} y[n] &= A\cos(\omega_0 n + \phi - \omega_0 n_d) \\ &= A\cos[\omega_0(n - n_d) + \phi], \end{aligned} \tag{2.106}$$

which is consistent with what we would obtain directly using the definition of the ideal delay system.

The concept of the frequency response of linear time-invariant systems is essentially the same for continuous-time and discrete-time systems. However, an important distinction arises because the frequency response of discrete-time linear time-invariant systems is *always* a periodic function of the frequency variable ω with period 2π. To show this, we substitute $\omega + 2\pi$ into Eq. (2.92) to obtain

$$H[e^{j(\omega+2\pi)}] = \sum_{n=-\infty}^{\infty} h[n]e^{-j(\omega+2\pi)n}. \tag{2.107}$$

Using the fact that $e^{\pm j2\pi n} = 1$ for n an integer, we have

$$e^{-j(\omega+2\pi)n} = e^{-j\omega n}e^{-j2\pi n} = e^{-j\omega n}.$$

Therefore we see that

$$H(e^{j(\omega+2\pi)}) = H(e^{j\omega}) \tag{2.108}$$

and, more generally,

$$H(e^{j(\omega+2\pi r)}) = H(e^{j\omega}) \qquad \text{for } r \text{ an integer,} \tag{2.109}$$

i.e., $H(e^{j\omega})$ is periodic with period 2π. Note that this is obviously true for the ideal delay system since $e^{-j(\omega+2\pi)n_d} = e^{-j\omega n_d}$, when n_d is an integer.

The reason for this periodicity is related directly to our earlier observation that the sequence

$$\{e^{j\omega n}\}, \qquad -\infty < n < \infty,$$

is indistinguishable from the sequence

$$\{e^{j(\omega+2\pi)n}\}, \qquad -\infty < n < \infty.$$

Since these two sequences have identical values for all n, the system must respond identically to both input sequences. This requires that Eq. (2.108) hold.

Since $H(e^{j\omega})$ is periodic with period 2π, and since frequencies ω and $\omega + 2\pi$ are indistinguishable, it follows that we need only specify $H(e^{j\omega})$ over an interval of length 2π, e.g., $0 \le \omega < 2\pi$ or $-\pi < \omega \le \pi$. The inherent periodicity defines the frequency response everywhere outside the chosen interval. For simplicity and for consistency with the continuous-time case, it is generally convenient to specify $H(e^{j\omega})$ over the interval $-\pi < \omega \le \pi$. With respect to this interval, the low frequencies are frequencies close to zero, while the "high frequencies" are frequencies close to $\pm\pi$. Recalling that frequencies differing by an integer multiple of 2π are indistinguishable, we might generalize the preceding statement as follows: The "low frequencies" are those that are close to an even multiple of π, while the "high frequencies" are those that are close to an odd multiple of π, consistent with our earlier discussion in Section 2.1.

Example 2.14

An important class of linear time-invariant systems includes those systems for which the frequency response is unity over a certain range of frequencies and is zero at the remaining frequencies. These correspond to *ideal frequency-selective filters*. In keeping with the preceding discussion, the frequency response of an ideal lowpass filter is shown in Figure 2.17(a). Because of the inherent periodicity of the discrete-time frequency response, it has the appearance of a multiband filter since frequencies around $\omega = 2\pi$ are indistinguishable from frequencies around $\omega = 0$. In effect, however, the frequency response passes only low frequencies and rejects high frequencies. Since the frequency response is completely specified by its behavior over the interval $-\pi < \omega < \pi$, the ideal lowpass filter frequency response is more typically shown only in the interval $-\pi < \omega < \pi$ as in Fig. 2.17(b). The frequency response for ideal highpass, bandstop, and bandpass filters are shown in Figures 2.18(a), (b), and (c), respectively.

Example 2.15

The impulse response of the moving average system of Example 2.2 is

$$h[n] = \begin{cases} \dfrac{1}{M_1 + M_2 + 1}, & -M_1 \le n \le M_2, \\ 0, & \text{otherwise.} \end{cases}$$

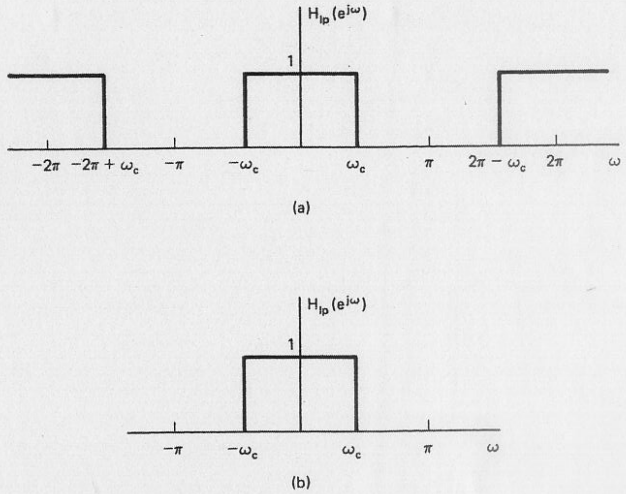

Figure 2.17 Ideal lowpass filter showng (a) periodicity of the frequency response, (b) one period of the periodic frequency response.

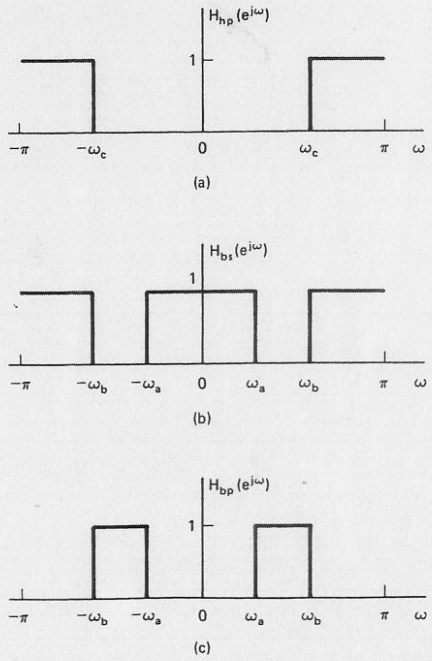

Figure 2.18 Ideal frequency-selective filters. (a) Highpass filter. (b) Bandstop filter. (c) Bandpass filter. In each case, the frequency response is periodic with period 2π. Only one period is shown.

Therefore the frequency response is

$$H(e^{j\omega}) = \frac{1}{M_1 + M_2 + 1} \sum_{n=-M_1}^{M_2} e^{-j\omega n}. \tag{2.110}$$

Equation (2.110) can be expressed in closed form by using Eq. (2.43) so that

$$\begin{aligned} H(e^{j\omega}) &= \frac{1}{M_1 + M_2 + 1} \frac{e^{j\omega M_1} - e^{-j\omega(M_2+1)}}{1 - e^{-j\omega}} \\ &= \frac{1}{M_1 + M_2 + 1} e^{-j\omega(M_2 - M_1 + 1)/2} \frac{e^{j\omega(M_1+M_2+1)/2} - e^{-j\omega(M_1+M_2+1)/2}}{1 - e^{-j\omega}} \\ &= \frac{1}{M_1 + M_2 + 1} e^{-j\omega(M_2 - M_1)/2} \frac{e^{j\omega(M_1+M_2+1)/2} - e^{-j\omega(M_1+M_2+1)/2}}{e^{j\omega/2} - e^{-j\omega/2}} \\ &= \frac{1}{M_1 + M_2 + 1} e^{-j\omega(M_2 - M_1)/2} \frac{\sin[\omega(M_1 + M_2 + 1)/2]}{\sin(\omega/2)}. \end{aligned} \tag{2.111}$$

The magnitude and phase of $H(e^{j\omega})$ are plotted in Fig. 2.19 for $M_1 = 0$, $M_2 = 4$. Note that $H(e^{j\omega})$ is periodic as required of the frequency response of a discrete-time system. Also note that $|H(e^{j\omega})|$ falls off at "high frequencies" and $\measuredangle H(e^{j\omega})$, i.e., the phase of $H(e^{j\omega})$, varies linearly with ω. This attenuation of the high frequencies suggests that the system will smooth out rapid variations in the input sequence; i.e., it is a rough approximation to a lowpass filter. This is consistent with what we would intuitively expect about the behavior of the moving average system.

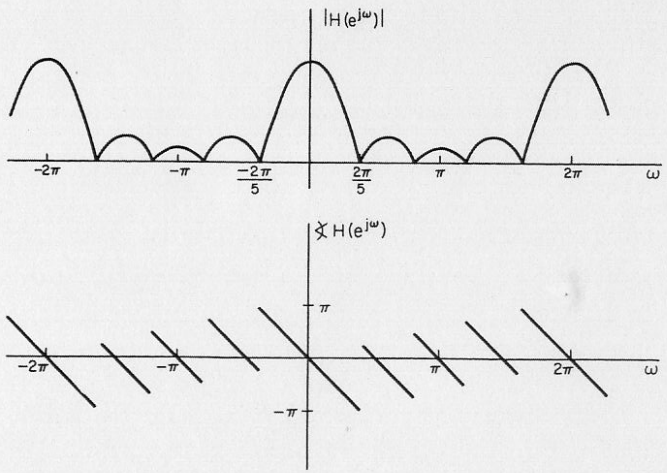

Figure 2.19 (a) Magnitude and (b) phase of the frequency response of the moving average system for the case $M_1 = 0$ and $M_2 = 4$.

One of the advantages of the frequency-response representation of a linear time-invariant system is that interpretations of system behavior such as the one we made in Example 2.15 often follow easily. We will elaborate on this point in considerably more detail in Chapter 5. At this point, however, let us return to the question of how we may find representations of the form of Eq. (2.100) for an arbitrary input sequence.

2.7 REPRESENTATION OF SEQUENCES BY FOURIER TRANSFORMS

Many sequences can be represented by a Fourier integral of the form

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n}\, d\omega, \tag{2.112}$$

where $X(e^{j\omega})$ is given by

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}. \tag{2.113}$$

Equations (2.112) and (2.113) together form a *Fourier representation* for the sequence. Equation (2.112), the *inverse Fourier transform*, is a *synthesis* formula. That is, it represents $x[n]$ as a superposition of infinitesimally small complex sinusoids of the form

$$\frac{1}{2\pi} X(e^{j\omega}) e^{j\omega n}\, d\omega,$$

with ω ranging over an interval of length 2π and with $X(e^{j\omega})$ determining the relative amount of each complex sinusoidal component. Although in writing Eq. (2.112) we have chosen the range of values for ω between $-\pi$ and $+\pi$, any interval of length 2π can be used. Equation (2.113), the *Fourier transform*,† is an expression for computing $X(e^{j\omega})$ from the sequence $x[n]$, i.e., for *analyzing* the sequence $x[n]$ to determine how much of each frequency component is required to synthesize $x[n]$ using Eq. (2.112).

In general, the Fourier transform is a complex-valued function of ω. As with the frequency response, we sometimes express $X(e^{j\omega})$ in rectangular form as

$$X(e^{j\omega}) = X_R(e^{j\omega}) + jX_I(e^{j\omega}) \tag{2.114a}$$

or in polar form as

$$X(e^{j\omega}) = |X(e^{j\omega})| e^{j\measuredangle X(e^{j\omega})}. \tag{2.114b}$$

The quantities $|X(e^{j\omega})|$ and $\measuredangle X(e^{j\omega})$ are called the *magnitude* and *phase*, respectively, of the Fourier transform. The Fourier transform is also sometimes referred to as the

† Sometimes we will refer to Eq. (2.113) more explicitly as the discrete-time Fourier transform, particularly when it is important to distinguish it from the continuous-time Fourier transform.

Fourier spectrum or simply the *spectrum*. Also, the terminology *magnitude spectrum* or *amplitude spectrum* is sometimes used to refer to $|X(e^{j\omega})|$ and the angle or phase $\measuredangle X(e^{j\omega})$ is sometimes called the *phase spectrum*.

The phase $\measuredangle X(e^{j\omega})$ is not uniquely specified by Eq. (2.114b) since any integer multiple of 2π may be added to $\measuredangle X(e^{j\omega})$ at any value of ω without affecting the result of the complex exponentiation. When we specifically want to refer to the principal value, i.e., $\measuredangle X(e^{j\omega})$ restricted to the range of values between $-\pi$ and $+\pi$, we will denote this as $\mathrm{ARG}[X(e^{j\omega})]$. In Chapter 5 we will also find it useful to refer to a phase function that is a continuous function of ω for $0 < \omega < \pi$, and this will be denoted $\arg[X(e^{j\omega})]$.

By comparing Eqs. (2.92) and (2.113) we can see that the frequency response of a linear time-invariant system is simply the Fourier transform of the impulse response and that therefore the impulse response can be obtained from the frequency response by applying the inverse Fourier transform integral; i.e.,

$$h[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(e^{j\omega})e^{j\omega n}\,d\omega. \tag{2.115}$$

As discussed previously, the frequency response is a periodic function. Likewise, the Fourier transform is periodic with period 2π. Indeed, Eq. (2.113) is of the form of a Fourier series for the continuous-variable periodic function $X(e^{j\omega})$, and Eq. (2.112), which expresses the sequence values $x[n]$ in terms of the periodic function $X(e^{j\omega})$, is of the form of the integral that would be used to obtain the coefficients in the Fourier series. Our use of Eqs. (2.112) and (2.113) focuses on the representation of the sequence $x[n]$. Nevertheless it is useful to be aware of the equivalence between the Fourier series representation of continuous-variable periodic functions and the Fourier transform representation of discrete-time signals, since all the familiar properties of Fourier series can be applied, with appropriate interpretation of variables, to the Fourier transform representation of a sequence.

We have not yet shown explicitly that Eqs. (2.112) and (2.113) are inverses of each other or considered the question of how broad a class of signals can be represented in the form of Eq. (2.112). To demonstrate that Eq. (2.112) is the inverse of Eq. (2.113), we can find $X(e^{j\omega})$ using Eq. (2.113) and then substitute the result into Eq. (2.112). Specifically, consider

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{m=-\infty}^{\infty} x[m]e^{-j\omega m}\right)e^{j\omega n}\,d\omega = \hat{x}[n], \tag{2.116}$$

where tentatively we have used $\hat{x}[n]$ to denote the result of the Fourier synthesis. We wish to show that $\hat{x}[n] = x[n]$ if $X(e^{j\omega})$ can be found using Eq. (2.113). Note that the "dummy index" of summation has been changed to m to distinguish it from n, the variable index in Eq. (2.112). If the infinite sum converges uniformly† for all ω, then

† Uniform convergence means that for every value of ω,

$$\lim_{M\to\infty} |X(e^{j\omega}) - X_M(e^{j\omega})| = 0.$$

we can interchange the order of integration and summation to obtain

$$\hat{x}[n] = \sum_{m=-\infty}^{\infty} x[m]\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\omega(n-m)}\,d\omega\right). \tag{2.117}$$

Evaluating the integral within the parentheses gives

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j\omega(n-m)}\,d\omega &= \frac{\sin \pi(n-m)}{\pi(n-m)}\\ &= \begin{cases}1, & m = n,\\ 0, & m \neq n,\end{cases}\\ &= \delta[n-m].\end{aligned}$$

Thus

$$\hat{x}[n] = \sum_{m=-\infty}^{\infty} x[m]\delta[n-m] = x[n],$$

which is what we set out to show.

The question of determining the class of signals that can be represented by Eq. (2.112) is equivalent to considering the convergence of the infinite sum in Eq. (2.113). That is, we are concerned with the conditions that must be satisfied by the terms in the sum in Eq. (2.113) such that

$$|X(e^{j\omega})| < \infty \qquad \text{for all } \omega,$$

where $X(e^{j\omega})$ is the limit as $M \to \infty$ of the finite sum

$$X_M(e^{j\omega}) = \sum_{n=-M}^{M} x[n]e^{-j\omega n}. \tag{2.118}$$

A sufficient condition for convergence can be found as follows:

$$\begin{aligned}|X(e^{j\omega})| &= \left|\sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}\right| < \infty\\ &\le \sum_{n=-\infty}^{\infty} |x[n]||e^{-j\omega n}| < \infty\\ &\le \sum_{n=-\infty}^{\infty} |x[n]| < \infty.\end{aligned}$$

Thus, if $x[n]$ is *absolutely summable*, then $X(e^{j\omega})$ exists. Furthermore, in this case the series can be shown to converge uniformly to a continuous function of ω.

Since a stable sequence is, by definition, absolutely summable, all stable sequences have Fourier transforms. It also follows, then, that any stable *system* will have a finite and continuous frequency response.

Absolute summability is a sufficient condition for existence of a Fourier transform representation. In Examples 2.12 and 2.15, we computed the Fourier

transforms of the sequences $\delta[n - n_d]$ and $[1/(M_1 + M_2 + 1)](u[n + M_1] - u[n - M_2 - 1])$. These sequences are absolutely summable since they are finite in length. Clearly, any finite-length sequence is absolutely summable and thus will have a Fourier transform representation. In the context of linear time-invariant systems, any FIR system will be stable and therefore will have a finite, continuous frequency response. When a sequence has infinite length, we must be concerned about convergence of the infinite sum. The following example illustrates this case.

Example 2.16

Let $x[n] = a^n u[n]$. The Fourier transform of this sequence is

$$X(e^{j\omega}) = \sum_{n=0}^{\infty} a^n e^{-j\omega n} = \sum_{n=0}^{\infty} (ae^{-j\omega})^n$$

$$= \frac{1}{1 - ae^{-j\omega}} \qquad \text{if } |ae^{-j\omega}| < 1 \quad \text{or} \quad |a| < 1.$$

Clearly the condition $|a| < 1$ is the condition for absolute summability of $x[n]$; i.e.,

$$\sum_{n=0}^{\infty} |a|^n = \frac{1}{1 - |a|} < \infty \qquad \text{if } |a| < 1. \tag{2.119}$$

Absolute summability is a *sufficient* condition for existence of a Fourier transform representation, and it also guarantees uniform convergence. Some sequences are not absolutely summable but are square summable, i.e.,

$$\sum_{n=-\infty}^{\infty} |x[n]|^2 < \infty. \tag{2.120}$$

Such sequences can be represented by a Fourier transform if we are willing to relax the condition of uniform convergence of the infinite sum defining $X(e^{j\omega})$. Specifically, in this case we have mean-square convergence, i.e., with

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \tag{2.121a}$$

and

$$X_M(e^{j\omega}) = \sum_{n=-M}^{M} x[n]e^{-j\omega n}, \tag{2.121b}$$

then

$$\lim_{M\to\infty} \int_{-\pi}^{\pi} |X(e^{j\omega}) - X_M(e^{j\omega})|^2 \, d\omega \to 0. \tag{2.122}$$

In other words, the error $|X(e^{j\omega}) - X_M(e^{j\omega})|$ may not approach zero at each value of ω as $M \to \infty$, but the total "energy" in the error does. Example 2.17 illustrates this case.

Example 2.17

Let us determine the impulse response of the ideal lowpass filter discussed in Example 2.14. The frequency response is

$$H_{lp}(e^{j\omega}) = \begin{cases} 1, & |\omega| < \omega_c, \\ 0, & \omega_c < |\omega| \le \pi, \end{cases} \tag{2.123}$$

with periodicity 2π also understood. The impulse response $h_{lp}[n]$ can be found using the Fourier transform synthesis equation (2.112):

$$\begin{aligned} h_{lp}[n] &= \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} e^{j\omega n} \, d\omega \\ &= \frac{\sin \omega_c n}{\pi n}, \qquad -\infty < n < \infty. \end{aligned} \tag{2.124}$$

We note that since $h_{lp}[n]$ is nonzero for $n < 0$, the ideal lowpass filter is noncausal. Also, $h_{lp}[n]$ is *not* absolutely summable. The sequence values approach zero as $n \to \infty$, but only as $1/n$. This is because $X(e^{j\omega})$ is discontinuous at $\omega = \omega_c$. Since $x[n]$ is not absolutely summable,

$$\sum_{n=-\infty}^{\infty} \frac{\sin \omega_c n}{\pi n} e^{-j\omega n}$$

does not converge uniformly for all values of ω. To obtain an intuitive feeling for this, let us consider $H_M(e^{j\omega})$ as the sum of a finite number of terms:

$$H_M(e^{j\omega}) = \sum_{n=-M}^{M} \frac{\sin \omega_c n}{\pi n} e^{-j\omega n}. \tag{2.125}$$

We can show that $H_M(e^{j\omega})$ can be expressed as

$$H_M(e^{j\omega}) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \frac{\sin[(2M+1)(\omega - \theta)/2]}{\sin[(\omega - \theta)/2]} \, d\theta.$$

The function $H_M(e^{j\omega})$ is evaluated in Fig. 2.20 for several values of M. Note that as M increases, the oscillatory behavior at $\omega = \omega_c$ (often referred to as the Gibbs phenomenon) is more rapid, but the size of the ripples does not decrease. In fact, it can be shown that as $M \to \infty$ the maximum amplitude of the oscillations does not go to zero but the oscillations converge in location toward the point $\omega = \omega_c$. Thus, the infinite sum does not converge uniformly to the discontinuous function $H_{lp}(e^{j\omega})$ of Eq. (2.123). However, $h_{lp}[n]$ as given in Eq. (2.124) is square summable, and correspondingly $H_M(e^{j\omega})$ converges in the mean-square sense to $X(e^{j\omega})$, i.e.,

$$\lim_{M\to\infty} \int_{-\pi}^{\pi} |H(e^{j\omega}) - H_M(e^{j\omega})|^2 \, d\omega \to 0.$$

Although the error between $\lim_{M\to\infty} H_M(e^{j\omega})$ and $H_{lp}(e^{j\omega})$ might seem unimportant because the two functions differ only at $\omega = \omega_c$, we will see in Chapter 7 that the behavior of finite sums has important implications in the design of discrete-time systems for filtering.

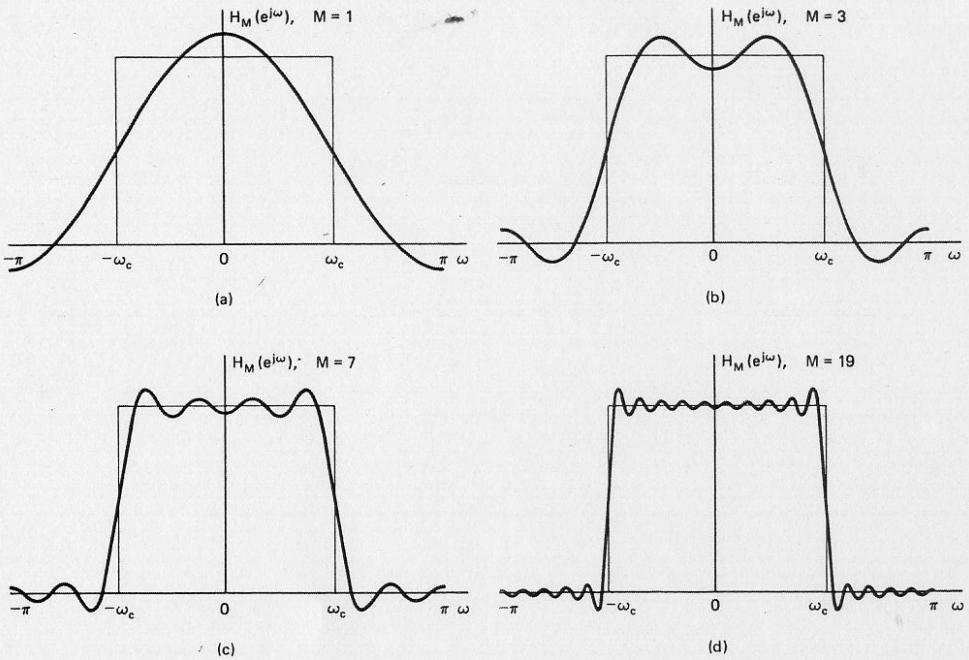

Figure 2.20 Convergence of the Fourier transform. The oscillatory behavior at $\omega = \omega_c$ is often called the Gibbs phenomenon.

It is sometimes useful to have a Fourier transform representation for certain sequences that are neither absolutely summable nor square summable. We illustrate several of these in the following examples.

Example 2.18

Consider the sequence $x[n] = 1$ for all n. This sequence is neither absolutely nor square summable, and Eq. (2.113) does not converge in either the uniform or mean-square sense for this case. However, it is possible and useful to define the Fourier transform of this sequence to be the periodic impulse train

$$X(e^{j\omega}) = \sum_{r=-\infty}^{\infty} 2\pi\delta(\omega + 2\pi r). \tag{2.126}$$

The impulses in this case are functions of a continuous variable and therefore are of "infinite height, zero width, and unit area," consistent with the fact that Eq. (2.113) does not converge. The use of Eq. (2.126) as a Fourier representation of the sequence $x[n] = 1$ is justified principally because formal substitution of Eq. (2.126) into Eq. (2.112) leads to the correct result. Example 2.19 represents a generalization of this example.

Example 2.19

Consider a sequence $x[n]$ whose Fourier transform is the periodic impulse train

$$X(e^{j\omega}) = \sum_{r=-\infty}^{\infty} 2\pi\delta(\omega - \omega_0 + 2\pi r). \tag{2.127}$$

Substituting $X(e^{j\omega})$ into the inverse Fourier transform integral of Eq. (2.112) and noting that the integration involves only one period of $X(e^{j\omega})$, we have

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} 2\pi\delta(\omega - \omega_0)e^{j\omega n}\,d\omega. \tag{2.128}$$

Using the definition of the impulse function, it follows that

$$x[n] = e^{j\omega_0 n} \qquad \text{for any } n.$$

For $\omega_0 = 0$ this reduces to the sequence considered in Example 2.18.

Clearly, $x[n]$ in Example 2.19 is not absolutely summable, nor is it square summable, and $|X(e^{j\omega})|$ is not finite for all ω. Thus the mathematical statement

$$\sum_{n=-\infty}^{\infty} e^{j\omega_0 n}e^{-j\omega n} = \sum_{r=-\infty}^{\infty} 2\pi\delta(\omega - \omega_0 + 2\pi r) \tag{2.129}$$

must be interpreted in a special way. Such an interpretation is provided by the theory of generalized functions (Lighthill, 1958). Using that theory, the concept of a Fourier transform representation can be extended rigorously to the class of sequences that can be expressed as a sum of discrete frequency components such as

$$x[n] = \sum_k a_k e^{j\omega_k n}, \qquad -\infty < n < \infty. \tag{2.130}$$

From the result of Example 2.19, it follows that

$$X(e^{j\omega}) = \sum_{r=-\infty}^{\infty}\sum_k 2\pi a_k\delta(\omega - \omega_k + 2\pi r) \tag{2.131}$$

is a consistent Fourier transform representation of $x[n]$ in Eq. (2.130).

Another sequence that is neither absolutely nor square summable is the unit step sequence $u[n]$. Although it is not completely straightforward to show, this sequence can be represented by the following Fourier transform:

$$U(e^{j\omega}) = \frac{1}{1 - e^{-j\omega}} + \sum_{r=-\infty}^{\infty} \pi\delta(\omega + 2\pi r). \tag{2.132}$$

2.8 SYMMETRY PROPERTIES OF THE FOURIER TRANSFORM

In using Fourier transforms, it is useful to have a detailed knowledge of the way that properties of the sequence manifest themselves in the Fourier transform and vice versa. In this section and Section 2.9 we discuss and summarize a number of such properties.

Symmetry properties of the Fourier transform are often very useful. The following discussion presents these properties, and the proofs are considered in Problems 2.29 and 2.30. Before presenting the properties, we begin with some definitions.

A *conjugate-symmetric sequence* $x_e[n]$ is defined as a sequence for which $x_e[n] = x_e^*[-n]$, and a *conjugate-antisymmetric sequence* $x_o[n]$ is defined as a sequence for which $x_o[n] = -x_o^*[-n]$, where * denotes complex conjugation. Any sequence $x[n]$ can be expressed as a sum of a conjugate-symmetric and conjugate-antisymmetric sequence. Specifically,

$$x[n] = x_e[n] + x_o[n], \tag{2.133a}$$

where

$$x_e[n] = \tfrac{1}{2}(x[n] + x^*[-n]) \tag{2.133b}$$

and

$$x_o[n] = \tfrac{1}{2}(x[n] - x^*[-n]). \tag{2.133c}$$

A real sequence that is conjugate-symmetric such that $x_e[n] = x_e[-n]$ is called an *even sequence*, and a real sequence that is conjugate-antisymmetric such that $x_o[n] = -x_o[-n]$ is called an *odd sequence*.

A Fourier transform $X(e^{j\omega})$ can be decomposed into a sum of conjugate-symmetric and conjugate-antisymmetric functions as

$$X(e^{j\omega}) = X_e(e^{j\omega}) + X_o(e^{j\omega}), \tag{2.134a}$$

where

$$X_e(e^{j\omega}) = \tfrac{1}{2}[X(e^{j\omega}) + X^*(e^{-j\omega})] \tag{2.134b}$$

and

$$X_o(e^{j\omega}) = \tfrac{1}{2}[X(e^{j\omega}) - X^*(e^{-j\omega})]. \tag{2.134c}$$

Clearly, $X_e(e^{j\omega})$ is conjugate-symmetric and $X_o(e^{j\omega})$ is conjugate-antisymmetric; i.e.,

$$X_e(e^{j\omega}) = X_e^*(e^{-j\omega}) \tag{2.135a}$$

and

$$X_o(e^{j\omega}) = -X_o^*(e^{-j\omega}). \tag{2.135b}$$

TABLE 2.1 SYMMETRY PROPERTIES OF THE FOURIER TRANSFORM

Sequence $x[n]$	Fourier Transform $X(e^{j\omega})$
1. $x^*[n]$	$X^*(e^{-j\omega})$
2. $x^*[-n]$	$X^*(e^{j\omega})$
3. $\mathscr{Re}\{x[n]\}$	$X_e(e^{j\omega})$ (conjugate-symmetric part of $X(e^{j\omega})$)
4. $j\,\mathscr{Im}\{x[n]\}$	$X_o(e^{j\omega})$ (conjugate-antisymmetric part of $X(e^{j\omega})$)
5. $x_e[n]$ (conjugate-symmetric part of $x[n]$)	$X_R(e^{j\omega})$
6. $x_o[n]$ (conjugate-antisymmetric part of $x[n]$)	$jX_I(e^{j\omega})$
The following properties apply only when $x[n]$ is real.	
7. Any real $x[n]$	$X(e^{j\omega}) = X^*(e^{-j\omega})$ (Fourier transform is conjugate-symmetric)
8. Any real $x[n]$	$X_R(e^{j\omega}) = X_R(e^{-j\omega})$ (real part is even)
9. Any real $x[n]$	$X_I(e^{j\omega}) = -X_I(e^{-j\omega})$ (imaginary part is odd)
10. Any real $x[n]$	$\lvert X(e^{j\omega})\rvert = \lvert X(e^{-j\omega})\rvert$ (magnitude is even)
11. Any real $x[n]$	$\sphericalangle X(e^{j\omega}) = -\sphericalangle X(e^{-j\omega})$ (phase is odd)
12. $x_e[n]$ (even part of $x[n]$)	$X_R(e^{j\omega})$
13. $x_0[n]$ (odd part of $x[n]$)	$jX_I(e^{j\omega})$

As with sequences, if a real function of a continuous variable is conjugate-symmetric, it is referred to as an *even function*, and a real conjugate-antisymmetric function of a continuous variable is referred to as an *odd function*.

The symmetry properties of the Fourier transform are summarized in Table 2.1. The first six properties apply for a general complex sequence $x[n]$ with Fourier transform $X(e^{j\omega})$. Properties 1 and 2 are considered in Problem 2.29. Property 3 follows from properties 1 and 2 together with the fact that the Fourier transform of the sum of two sequences is the sum of their Fourier transforms. Specifically, the Fourier transform of $\mathscr{Re}\{x[n]\} = \frac{1}{2}(x[n] + x^*[n])$ is the conjugate-symmetric part of $X(e^{j\omega})$, or $X_e(e^{j\omega})$. Similarly, $j\mathscr{Im}\{x[n]\} = \frac{1}{2}(x[n] - x^*[n])$ or, equivalently, $j\mathscr{Im}\{x[n]\}$ has a Fourier transform that is the conjugate-antisymmetric component $X_o(e^{j\omega})$ corresponding to property 4. By considering the Fourier transform of $x_e[n]$ and $x_o[n]$, the conjugate-symmetric and conjugate-antisymmetric components of $x[n]$, properties 5 and 6 follow.

If $x[n]$ is a real sequence, these symmetry properties become particularly straightforward and useful. Specifically, for a real sequence the Fourier transform is conjugate-symmetric; i.e., $X(e^{j\omega}) = X^*(e^{-j\omega})$ (property 7). Expressing $X(e^{j\omega})$ in terms of its real and imaginary parts as

$$X(e^{j\omega}) = X_R(e^{j\omega}) + jX_I(e^{j\omega}), \tag{2.136}$$

properties 8 and 9 then follow, specifically,

$$X_R(e^{j\omega}) = X_R(e^{-j\omega}) \tag{2.137a}$$

and

$$X_I(e^{j\omega}) = -X_I(e^{-j\omega}). \tag{2.137b}$$

In other words, the real part of the Fourier transform is an even function, and the imaginary part is an odd function if the sequence is real. In a similar manner, by expressing $X(e^{j\omega})$ in polar form,

$$X(e^{j\omega}) = |X(e^{j\omega})|e^{j\measuredangle X(e^{j\omega})}, \tag{2.138}$$

it follows that for a real sequence $x[n]$, the magnitude of the Fourier transform $|X(e^{j\omega})|$ is an even function of ω and the phase, $\measuredangle X(e^{j\omega})$, can be chosen to be an odd function of ω (properties 10 and 11). Also for a real sequence, the even part of $x[n]$ transforms to $X_R(e^{j\omega})$ and the odd part of $x[n]$ transforms to $jX_I(e^{j\omega})$ (properties 12 and 13).

Example 2.20

To illustrate a few of these properties, let us return to the sequence of Example 2.16, where we showed that the Fourier transform of the real sequence $x[n] = a^n u[n]$ is

$$X(e^{j\omega}) = \frac{1}{1 - ae^{-j\omega}} \quad \text{if } |a| < 1. \tag{2.139}$$

Then

$$X^*(e^{-j\omega}) = \frac{1}{1 - ae^{-j\omega}} = X(e^{j\omega}), \qquad \text{(property 7)}$$

$$X_R(e^{j\omega}) = \frac{1 - a\cos\omega}{1 + a^2 - 2a\cos\omega} = X_R(e^{-j\omega}), \qquad \text{(property 8)}$$

$$X_I(e^{j\omega}) = \frac{a\sin\omega}{1 + a^2 - 2a\cos\omega} = -X_I(e^{-j\omega}), \qquad \text{(property 9)}$$

$$|X(e^{j\omega})| = \frac{1}{(1 + a^2 - 2a\cos\omega)^{1/2}} = |X(e^{-j\omega})|, \qquad \text{(property 10)}$$

$$\measuredangle X(e^{j\omega}) = \tan^{-1}\left(\frac{a\sin\omega}{1 - a\cos\omega}\right) = -\measuredangle X(e^{-j\omega}). \qquad \text{(property 11)}$$

These functions are plotted in Fig. 2.21 for $a > 0$, specifically $a = 0.9$ (solid curve) and $a = 0.5$ (dashed curve). In Problem 2.27 we consider the corresponding plots for $a < 0$.

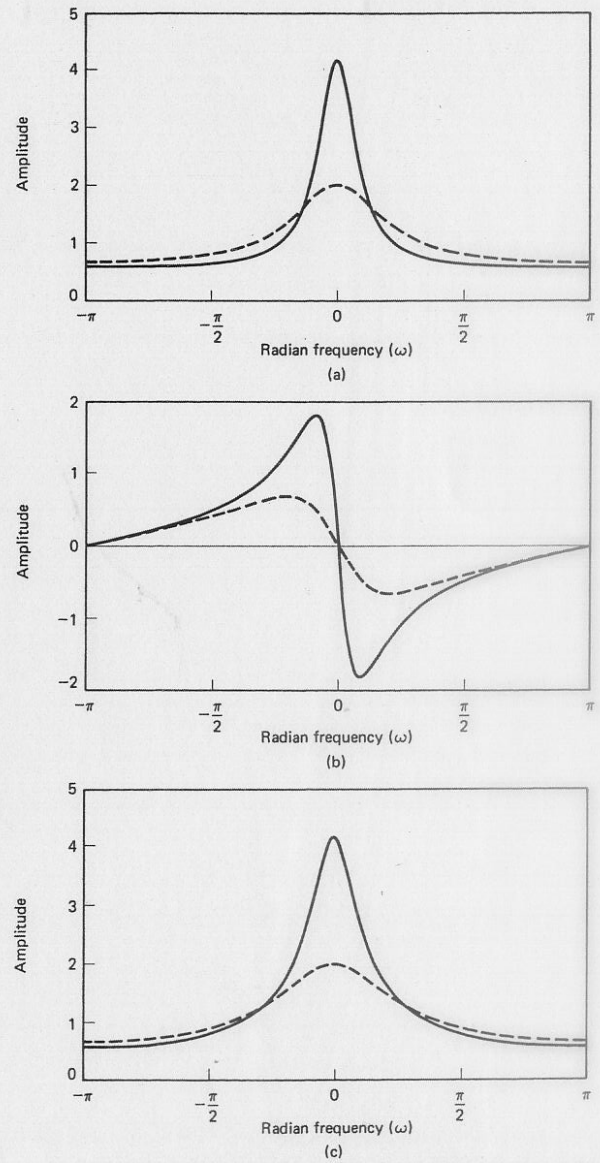

Figure 2.21 Frequency response for a system with impulse response $h[n] = a^n u[n]$. (a) Real part. (b) Imaginary part. (c) Magnitude. $a > 0$; $a = 0.9$ (solid curve) and $a = 0.5$ (dashed curve).

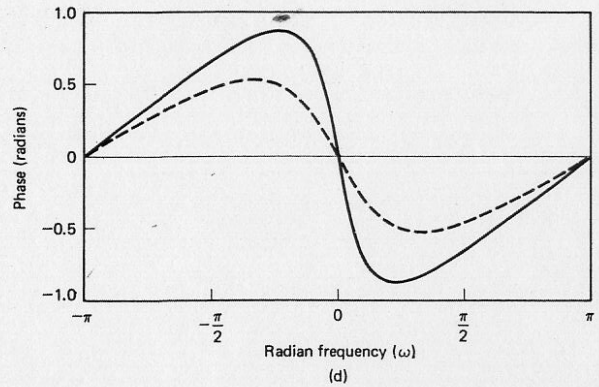

Figure 22.1 (*continued*) (d) Phase.

2.9 FOURIER TRANSFORM THEOREMS

In addition to the symmetry properties, a variety of theorems (presented in Sections 2.9.1–2.9.7) relate operations on the sequence to operations on the Fourier transform. We will see that these theorems are quite similar in most cases to corresponding theorems for continuous-time signals and their Fourier transforms. To facilitate the statement of these theorems, we introduce the following operator notation:

$$X(e^{j\omega}) = \mathscr{F}\{x[n]\},$$
$$x[n] = \mathscr{F}^{-1}\{X(e^{j\omega})\},$$
$$x[n] \overset{\mathscr{F}}{\longleftrightarrow} X(e^{j\omega}).$$

That is, $\mathscr{F}$ denotes the operation of "taking the Fourier transform of $x[n]$," and $\mathscr{F}^{-1}$ is the inverse of that operation. Most of the theorems will be stated without proof. The proofs, which are left as exercises (Problem 2.31), generally involve only simple manipulations of variables of summation or integration. The theorems in this section are summarized in Table 2.2.

2.9.1 Linearity of the Fourier Transform

If

$$x_1[n] \overset{\mathscr{F}}{\longleftrightarrow} X_1(e^{j\omega})$$

and

$$x_2[n] \overset{\mathscr{F}}{\longleftrightarrow} X_2(e^{j\omega}),$$

TABLE 2.2 FOURIER TRANSFORM THEOREMS

Sequence $x[n]$ $y[n]$	Fourier Transform $X(e^{j\omega})$ $Y(e^{j\omega})$
1. $ax[n] + by[n]$	$aX(e^{j\omega}) + bY(e^{j\omega})$
2. $x[n - n_d]$ (n_d an integer)	$e^{-j\omega n_d}X(e^{j\omega})$
3. $e^{j\omega_0 n}x[n]$	$X(e^{j(\omega-\omega_0)})$
4. $x[-n]$	$X(e^{-j\omega})$ if $x[n]$ real $X^*(e^{j\omega})$
5. $nx[n]$	$j\dfrac{dX(e^{j\omega})}{d\omega}$
6. $x[n] * y[n]$	$X(e^{j\omega})Y(e^{j\omega})$
7. $x[n]y[n]$	$\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} X(e^{j\theta})Y(e^{j(\omega-\theta)})d\theta$
Parseval's Theorem	
8. $\displaystyle\sum_{n=-\infty}^{\infty} \lvert x[n]\rvert^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \lvert X(e^{j\omega})\rvert^2\, d\omega$	
9. $\displaystyle\sum_{n=-\infty}^{\infty} x[n]y^*[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})Y^*(e^{j\omega})\, d\omega$	

then

$$ax_1[n] + bx_2[n] \overset{\mathscr{F}}{\longleftrightarrow} aX_1(e^{j\omega}) + bX_2(e^{j\omega}). \tag{2.140}$$

2.9.2 Time Shifting and Frequency Shifting

If

$$x[n] \overset{\mathscr{F}}{\longleftrightarrow} X(e^{j\omega}),$$

then for the time-shifted sequence,

$$x[n - n_d] \overset{\mathscr{F}}{\longleftrightarrow} e^{-j\omega n_d}X(e^{j\omega}) \tag{2.141}$$

and for the frequency-shifted Fourier transform,

$$e^{j\omega_0 n}x[n] \overset{\mathscr{F}}{\longleftrightarrow} X(e^{j(\omega-\omega_0)}). \tag{2.142}$$

2.9.3. Time Reversal

If

$$x[n] \overset{\mathscr{F}}{\longleftrightarrow} X(e^{j\omega}),$$

then if the sequence is time reversed,

$$x[-n] \overset{\mathscr{F}}{\longleftrightarrow} X(e^{-j\omega}). \tag{2.143}$$

If $x[n]$ is real, this theorem reduces to

$$x[-n] \overset{\mathscr{F}}{\longleftrightarrow} X^*(e^{j\omega}). \tag{2.144}$$

2.9.4 Differentiation in Frequency

If

$$x[n] \overset{\mathscr{F}}{\longleftrightarrow} X(e^{j\omega}),$$

then

$$nx[n] \overset{\mathscr{F}}{\longleftrightarrow} j\frac{dX(e^{j\omega})}{d\omega}. \tag{2.145}$$

2.9.5 Parseval's Theorem

If

$$x[n] \overset{\mathscr{F}}{\longleftrightarrow} X(e^{j\omega}),$$

then

$$E = \sum_{n=-\infty}^{\infty} |x[n]|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |X(e^{j\omega})|^2\, d\omega. \tag{2.146}$$

The function $|X(e^{j\omega})|^2$ is called the *energy density spectrum* since it determines how the energy is distributed in frequency. Necessarily, the energy density spectrum is defined only for finite-energy signals. A more general form of Parseval's theorem is shown in Problem 2.34.

2.9.6 The Convolution Theorem

If

$$x[n] \overset{\mathscr{F}}{\longleftrightarrow} X(e^{j\omega})$$

and

$$h[n] \overset{\mathscr{F}}{\longleftrightarrow} H(e^{j\omega})$$

and if

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k] = x[n] * h[n], \tag{2.147}$$

then

$$Y(e^{j\omega}) = X(e^{j\omega})H(e^{j\omega}). \tag{2.148}$$

Thus convolution of sequences implies multiplication of the corresponding Fourier transforms. Note that the time shifting property is a special case of the convolution property since

$$\delta[n-n_d] \overset{\mathscr{F}}{\longleftrightarrow} e^{-j\omega n_d} \tag{2.149}$$

and if $h[n] = \delta[n-n_d]$, then $y[n] = x[n] * \delta[n-n_d]$. Therefore,

$$H(e^{j\omega}) = e^{-j\omega n_d} \quad \text{and} \quad Y(e^{j\omega}) = e^{-j\omega n_d}X(e^{j\omega}).$$

A formal derivation of the convolution theorem is easily achieved by applying the definition of the Fourier transform to $y[n]$ as expressed in Eq. (2.147). This theorem can also be interpreted as a direct consequence of the eigenfunction property of complex exponentials for linear time-invariant systems. Recall that $H(e^{j\omega})$ is the frequency response of the linear time-invariant system whose impulse response is $h[n]$. Also recall that if

$$x[n] = e^{j\omega n},$$

then

$$y[n] = H(e^{j\omega})e^{j\omega n}.$$

That is, complex exponentials are *eigenfunctions* of linear time-invariant systems where $H(e^{j\omega})$, the Fourier transform of $h[n]$, is the eigenvalue. From the definition of integration, the Fourier transform synthesis equation corresponds to the representation of a sequence $x[n]$ as a superposition of complex exponentials of infinitesimal size, i.e.,

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})e^{j\omega n}\, d\omega = \lim_{\Delta\omega\to 0}\frac{1}{2\pi}\sum_k X(e^{jk\Delta\omega})e^{jk\Delta\omega n}\,\Delta\omega.$$

By the eigenfunction property of linear systems and by the principle of superposition, the corresponding output will be

$$y[n] = \lim_{\Delta\omega\to 0}\frac{1}{2\pi}\sum_k H(e^{jk\Delta\omega})X(e^{jk\Delta\omega})e^{jk\Delta\omega n}\,\Delta\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(e^{j\omega})X(e^{j\omega})e^{j\omega n}\, d\omega.$$

Thus we conclude that

$$Y(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega}),$$

as in Eq. (2.148).

2.9.7 The Modulation or Windowing Theorem

If

$$x[n] \overset{\mathscr{F}}{\longleftrightarrow} X(e^{j\omega})$$

and

$$w[n] \overset{\mathscr{F}}{\longleftrightarrow} W(e^{j\omega})$$

and if

$$y[n] = x[n]w[n], \tag{2.150}$$

then

$$Y(e^{j\omega}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\theta})W(e^{j(\omega-\theta)})\,d\theta. \tag{2.151}$$

Equation (2.151) is a periodic convolution, i.e., it is a convolution of two periodic functions with the limits of integration extending over only one period. The duality inherent in most Fourier transform theorems is evident when we compare the convolution and modulation theorems. However, in contrast to the continuous-time case, where this duality is complete, in the discrete-time case fundamental differences arise because the Fourier transform is a sum while the inverse transform is an integral with a periodic integrand. While for continuous time we can state that convolution in the time domain is represented by multiplication in the frequency domain and vice versa, in discrete-time this statement must be modified somewhat. Specifically, discrete-time convolution of sequences (the convolution sum) is equivalent to multiplication of corresponding periodic Fourier transforms, and multiplication of sequences is equivalent to *periodic* convolution of corresponding Fourier transforms.

The theorems of this section and a number of fundamental Fourier transform pairs are summarized in Tables 2.2 and 2.3, respectively. One of the ways that knowledge of Fourier transform theorems and properties is useful is in determining Fourier transforms or inverse transforms. Often by using the theorems and known transform pairs it is possible to represent a sequence in terms of operations on other sequences for which the transform is known, thereby simplifying an otherwise difficult or tedious problem. Examples 2.21–2.24 illustrate this.

Example 2.21

Let $x[n] = a^n u[n-5]$. This sequence can be expressed as

$$x[n] = a^5 a^{n-5} u[n-5]. \tag{2.152}$$

Therefore, using Theorems 1 and 2 of Table 2.2 and the transform pair 4 of Table 2.3, we obtain

$$X(e^{j\omega}) = \frac{a^5 e^{-j\omega 5}}{1 - ae^{-j\omega}}. \tag{2.153}$$

Example 2.22

Suppose that

$$X(e^{j\omega}) = \frac{1}{(1 - ae^{-j\omega})(1 - be^{-j\omega})}. \tag{2.154}$$

Direct substitution of $X(e^{j\omega})$ into Eq. (2.112) leads to an integral that is difficult to evaluate by ordinary real integration techniques. However, using the technique of partial

TABLE 2.3 FOURIER TRANSFORM PAIRS

Sequence	Fourier Transform
1. $\delta[n]$	1
2. $\delta[n-n_0]$	$e^{-j\omega n_0}$
3. $1 \quad (-\infty < n < \infty)$	$\sum_{k=-\infty}^{\infty} 2\pi\delta(\omega + 2\pi k)$
4. $a^n u[n] \quad (\lvert a\rvert < 1)$	$\dfrac{1}{1 - ae^{-j\omega}}$
5. $u[n]$	$\dfrac{1}{1 - e^{-j\omega}} + \sum_{k=-\infty}^{\infty} \pi\delta(\omega + 2\pi k)$
6. $(n+1)a^n u[n] \quad (\lvert a\rvert < 1)$	$\dfrac{1}{(1 - ae^{-j\omega})^2}$
7. $\dfrac{r^n \sin \omega_p(n+1)}{\sin \omega_p} u[n] \quad (\lvert r\rvert < 1)$	$\dfrac{1}{1 - 2r\cos\omega_p e^{-j\omega} + r^2 e^{-j2\omega}}$
8. $\dfrac{\sin \omega_c n}{\pi n}$	$X(e^{j\omega}) = \begin{cases} 1, & \lvert\omega\rvert < \omega_c, \\ 0, & \omega_c < \lvert\omega\rvert \le \pi \end{cases}$
9. $x[n] = \begin{cases} 1, & 0 \le n \le M \\ 0, & \text{otherwise} \end{cases}$	$\dfrac{\sin[\omega(M+1)/2]}{\sin(\omega/2)} e^{-j\omega M/2}$
10. $e^{j\omega_0 n}$	$\sum_{k=-\infty}^{\infty} 2\pi\delta(\omega - \omega_0 + 2\pi k)$
11. $\cos(\omega_0 n + \phi)$	$\pi \sum_{k=-\infty}^{\infty} [e^{j\phi}\delta(\omega - \omega_0 + 2\pi k) + e^{-j\phi}\delta(\omega + \omega_0 + 2\pi k)]$

fraction expansion, which we discuss in detail in Chapter 4, $X(e^{j\omega})$ can be expanded into the form

$$X(e^{j\omega}) = \frac{a/(a-b)}{1 - ae^{-j\omega}} - \frac{b/(a-b)}{1 - be^{-j\omega}}. \tag{2.155}$$

Using Theorem 1 of Table 2.2 and the transform pair 4 of Table 2.3, it follows that

$$x[n] = \left(\frac{a}{a-b}\right)a^n u[n] - \left(\frac{b}{a-b}\right)b^n u[n]. \tag{2.156}$$

Example 2.23

Consider the Fourier transform

$$X(e^{j\omega}) = \begin{cases} e^{-j\omega n_d}, & \omega_c < \lvert\omega\rvert < \pi, \\ 0, & \lvert\omega\rvert < \omega_c, \end{cases} \tag{2.157}$$

with period 2π understood. This Fourier transform can be expressed as

$$X(e^{j\omega}) = e^{-j\omega n_d}(1 - R(e^{j\omega})),$$

where $R(e^{j\omega})$ is periodic with period 2π and

$$R(e^{j\omega}) = \begin{cases} 1, & |\omega| < \omega_c, \\ 0, & \omega_c < |\omega| < \pi. \end{cases}$$

Using the result of Example 2.17 to obtain the inverse transform of $R(e^{j\omega})$, together with properties 1 and 2 of Table 2.2, we have for $x[n]$

$$\begin{aligned} x[n] &= \delta[n - n_d] - r[n - n_d] \\ &= \delta[n - n_d] - \frac{\sin \omega_c(n - n_d)}{\pi(n - n_d)}. \end{aligned}$$

Example 2.24

Let us determine the impulse response for a stable linear time-invariant system for which the input $x[n]$ and output $y[n]$ satisfy the linear constant-coefficient difference equation

$$y[n] - \tfrac{1}{2}y[n-1] = x[n] - \tfrac{1}{4}x[n-1]. \tag{2.158}$$

In Chapter 5 we will see that the z-transform is more useful than the Fourier transform for dealing with difference equations. However, this example offers a hint of the utility of transform methods for analysis of linear systems. To find the impulse response, we set $x[n] = \delta[n]$; with $h[n]$ denoting the impulse response, Eq. (2.158) becomes

$$h[n] - \tfrac{1}{2}h[n-1] = \delta[n] - \tfrac{1}{4}\delta[n-1]. \tag{2.159}$$

Applying the Fourier transform to both sides of Eq. (2.159) and using properties 1 and 2 of Table 2.2, we obtain

$$H(e^{j\omega}) - \tfrac{1}{2}e^{-j\omega}H(e^{j\omega}) = 1 - \tfrac{1}{4}e^{-j\omega} \tag{2.160}$$

or

$$H(e^{j\omega}) = \frac{1 - \frac{1}{4}e^{-j\omega}}{1 - \frac{1}{2}e^{-j\omega}}. \tag{2.161}$$

To obtain $h[n]$, we want to determine the inverse Fourier transform of $H(e^{j\omega})$. Toward this end, we rewrite Eq. (2.161) as

$$H(e^{j\omega}) = \frac{1}{1 - \frac{1}{2}e^{-j\omega}} - \frac{\frac{1}{4}e^{-j\omega}}{1 - \frac{1}{2}e^{-j\omega}}. \tag{2.162}$$

From transform 4 of Table 2.3,

$$(\tfrac{1}{2})^n u[n] \overset{\mathscr{F}}{\longleftrightarrow} \frac{1}{1 - \frac{1}{2}e^{-j\omega}}.$$

Combining this transform with property 3 of Table 2.2,

$$-(\tfrac{1}{4})(\tfrac{1}{2})^{n-1}u[n-1] \overset{\mathscr{F}}{\longleftrightarrow} -\frac{\frac{1}{4}e^{-j\omega}}{1 - \frac{1}{2}e^{-j\omega}}. \tag{2.163}$$

Based on property 1 of Table 2.2, then,

$$h[n] = (\tfrac{1}{2})^n u[n] - (\tfrac{1}{4})(\tfrac{1}{2})^{n-1}u[n-1]. \tag{2.164}$$

2.10 DISCRETE-TIME RANDOM SIGNALS

The preceding sections have focused on mathematical representations of discrete-time signals and systems and the insights that derive from such mathematical representations. We have seen that discrete-time signals and systems have both a time-domain and a frequency-domain representation, each with an important place in the theory and design of discrete-time signal processing systems. Until now we have assumed that the signals are deterministic, i.e., that each value of a sequence is uniquely determined by a mathematical expression, a table of data, or a rule of some type.

In many situations, the processes that generate signals are so complex as to make precise description of a signal extremely difficult or undesirable, if not impossible. In such cases, modeling the signal as a stochastic process is analytically useful. As an example we will see in Chapter 6 that many of the effects encountered in implementing digital signal processing algorithms with finite register length can be represented by additive noise, i.e., a stochastic sequence. Many mechanical systems generate acoustic or vibratory signals that can be processed to diagnose potential failure; again, signals of this type are often best modeled in terms of stochastic signals. Speech signals to be processed for automatic recognition or bandwidth compression and music to be processed for quality enhancement are two more of many examples.

A stochastic signal is considered to be a member of an ensemble of discrete-time signals that is characterized by a set of probability density functions. In other words, for a specific signal at a particular time, the amplitude of the signal sample at that time is assumed to have been determined by an underlying scheme of probabilities.

The key to the mathematical representation of such signals lies in their description in terms of averages. While stochastic signals are not absolutely summable or square summable and consequently do not directly have Fourier transforms, many (but not all) of the properties of such signals can be summarized in terms of the *autocorrelation* or *autocovariance sequence*, for which the Fourier transform often exists. As we will discuss in this section, the Fourier transform of the autocovariance sequence has a useful interpretation in terms of the frequency distribution of the power in the signal. The use of the autocovariance sequence and its transform has another important advantage: the effect of processing stochastic signals with a discrete-time linear system can be conveniently described in terms of the effect of the system on the autocovariance sequence.

In the following discussion, we assume that the reader is familiar with the basic concepts of stochastic processes such as averages, correlation and covariance functions, and power spectrum. A brief review and summary of notation and concepts is provided in Appendix A. A more detailed presentation of the theory of random signals can be found in a variety of excellent texts, such as Davenport (1970) and Papoulis (1984). Our primary objective in this section is to present a specific set of results pertaining to representation of stochastic signals that will be useful in subsequent chapters. Therefore, we focus on wide-sense stationary random signals and their representation in the context of processing with linear time-invariant

systems. Although for simplicity we assume that $x[n]$ and $h[n]$ are real-valued, the results are easily generalized to the complex case.

Consider a stable linear time-invariant system with impulse response $h[n]$. Let $x[n]$ be a real-valued sequence that is a sample sequence of a wide-sense stationary discrete-time random process. Then the output of the linear system is a sample function of an output random process related to the input process by the linear transformation

$$y[n] = \sum_{k=-\infty}^{\infty} h[n-k]x[k] = \sum_{k=-\infty}^{\infty} h[k]x[n-k].$$

As we have shown, since the system is stable, $y[n]$ will be bounded if $x[n]$ is bounded. We will see below that if the input is stationary,† then so is the output. The input signal may be characterized by its mean m_x and its autocorrelation function $\phi_{xx}[m]$, or we may also have additional information about first- or even second-order probability distributions. In characterizing the output random process $y[n]$ we desire similar information. For many applications, it is sufficient to characterize both the input and output in terms of simple averages, such as the mean, variance, and autocorrelation. Therefore, we will derive input-output relationships for these quantities.

The means of the input and output processes are

$$m_{x_n} = \mathscr{E}\{\mathbf{x}_n\}, \qquad m_{y_n} = \mathscr{E}\{\mathbf{y}_n\}, \tag{2.165}$$

where $\mathscr{E}\{\cdot\}$ denotes the expected value. For notational convenience we will be less careful about distinguishing between the random variables $\mathbf{x}_n$ and $\mathbf{y}_n$ and their values $x[n]$ and $y[n]$, so that Eqs. (2.165) will alternatively be written

$$m_x[n] = \mathscr{E}\{x[n]\}, \qquad m_y[n] = \mathscr{E}\{y[n]\}. \tag{2.166}$$

If $x[n]$ is stationary, then $m_x[n]$ is independent of n and will be written as m_x, with similar notation for $m_y[n]$ if $y[n]$ is stationary.

The mean of the output process is

$$m_y[n] = \mathscr{E}\{y[n]\} = \sum_{k=-\infty}^{\infty} h[k]\mathscr{E}\{x[n-k]\},$$

where we have used the fact that the expected value of a sum is the sum of the expected values. Since the input is stationary, $m_x[n-k] = m_x$ and consequently

$$m_y[n] = m_x \sum_{k=-\infty}^{\infty} h[k]. \tag{2.167}$$

From Eq. (2.167) we see that the mean of the output is also constant. An equivalent expression to Eq. (2.167) in terms of the frequency response is

$$m_y = H(e^{j0})m_x. \tag{2.168}$$

† In the remainder of the text, we will use the term *stationary* to mean "wide-sense stationary."

Assuming temporarily that the output is nonstationary, the autocorrelation function of the output process for a real input is

$$\begin{aligned}\phi_{yy}[n, n+m] &= \mathscr{E}\{y[n]y[n+m]\} \\ &= \mathscr{E}\left\{\sum_{k=-\infty}^{\infty}\sum_{r=-\infty}^{\infty} h[k]h[r]x[n-k]x[n+m-r]\right\} \\ &= \sum_{k=-\infty}^{\infty} h[k] \sum_{r=-\infty}^{\infty} h[r]\mathscr{E}\{x[n-k]x[n+m-r]\}.\end{aligned}$$

Since $x[n]$ is assumed to be stationary, $\mathscr{E}\{x[n-k]x[n+m-r]\}$ depends only on the time difference $m+k-r$. Therefore,

$$\phi_{yy}[n, n+m] = \sum_{k=-\infty}^{\infty} h[k] \sum_{r=-\infty}^{\infty} h[r]\phi_{xx}[m+k-r] = \phi_{yy}[m]. \tag{2.169}$$

That is, the output autocorrelation sequence also depends only on the time difference m. Thus for a linear time-invariant system that is excited by a wide-sense stationary input, the output is also wide-sense stationary.

By making the substitution $\ell = r - k$, Eq. (2.169) can be expressed as

$$\begin{aligned}\phi_{yy}[m] &= \sum_{\ell=-\infty}^{\infty} \phi_{xx}[m-\ell] \sum_{k=-\infty}^{\infty} h[k]h[\ell+k] \\ &= \sum_{\ell=-\infty}^{\infty} \phi_{xx}[m-\ell]c[\ell],\end{aligned} \tag{2.170}$$

where we have defined

$$c[\ell] = \sum_{k=-\infty}^{\infty} h[k]h[\ell+k]. \tag{2.171}$$

A sequence of the form of $c[\ell]$ is often called an *aperiodic autocorrelation sequence* or simply the *autocorrelation sequence of* $h[n]$. It should be emphasized that $c[\ell]$ is the autocorrelation of an aperiodic, i.e., finite-energy, sequence and should not be confused with the autocorrelation of an infinite-energy sequence. Indeed, it can be seen that $c[\ell]$ is simply the discrete convolution of $h[n]$ with $h[-n]$. Equation (2.170), then, can be interpreted to mean that the autocorrelation of the output of a linear system is the convolution of the input autocorrelation with the aperiodic autocorrelation of the system impulse response.

Equation (2.170) suggests that Fourier transforms may be useful in characterizing the response of a linear time-invariant system to a stochastic input. Assume for convenience that $m_x = 0$; i.e., the autocorrelation and autocovariance sequences are identical. Then, with $\Phi_{xx}(e^{j\omega})$, $\Phi_{yy}(e^{j\omega})$, and $C(e^{j\omega})$ denoting the Fourier transforms of $\phi_{xx}[m]$, $\phi_{yy}[m]$, and $c[\ell]$, respectively, from Eq. (2.170)

$$\Phi_{yy}(e^{j\omega}) = C(e^{j\omega})\Phi_{xx}(e^{j\omega}). \tag{2.172}$$

Also, from Eq. (2.171)

$$C(e^{j\omega}) = H(e^{j\omega})H^*(e^{j\omega})$$
$$= |H(e^{j\omega})|^2,$$

so

$$\Phi_{yy}(e^{j\omega}) = |H(e^{j\omega})|^2\Phi_{xx}(e^{j\omega}). \tag{2.173}$$

Equation (2.173) provides the motivation for the term *power density spectrum*. Specifically,

$$\mathscr{E}\{y^2[n]\} = \phi_{yy}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{yy}(e^{j\omega})\,d\omega \tag{2.174}$$
$$= \text{total average power in output.}$$

Substituting Eq. (2.173) into Eq. (2.174), we have

$$\mathscr{E}\{y^2[n]\} = \phi_{yy}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi}|H(e^{j\omega})|^2\Phi_{xx}(e^{j\omega})\,d\omega. \tag{2.175}$$

Suppose that $H(e^{j\omega})$ is an ideal bandpass filter, as shown in Fig. 2.18(c). We recall that $\phi_{xx}[m]$ is an even sequence, so

$$\Phi_{xx}(e^{j\omega}) = \Phi_{xx}(e^{-j\omega}).$$

Likewise, $|H(e^{j\omega})|^2$ is an even function of ω. Therefore, we can write

$$\phi_{yy}[0] = \text{average power in output}$$
$$= \frac{1}{\pi}\int_{\omega_a}^{\omega_b}\Phi_{xx}(e^{j\omega})\,d\omega + \frac{1}{\pi}\int_{-\omega_b}^{-\omega_a}\Phi_{xx}(e^{j\omega})\,d\omega. \tag{2.176}$$

Thus the area under $\Phi_{xx}(e^{j\omega})$ for $\omega_a \le |\omega| \le \omega_b$ can be taken to represent the mean-square value of the input in that frequency band. We observe that the output power must remain nonnegative, so

$$\lim_{(\omega_b-\omega_a)\to 0}\phi_{yy}[0] \ge 0.$$

This result, together with Eq. (2.176), implies that

$$\Phi_{xx}(\omega) \ge 0. \tag{2.177}$$

Thus we note that the power density function of a real signal is real, even, and nonnegative.

Another important result concerns the cross-correlation between the input and output of a linear time-invariant system:

$$\phi_{xy}[m] = \mathscr{E}\{x[n]y[n+m]\}$$
$$= \mathscr{E}\{x[n]\sum_{k=-\infty}^{\infty}h[k]x[n+m-k]\}$$
$$= \sum_{k=-\infty}^{\infty}h[k]\phi_{xx}[m-k]. \tag{2.178}$$

In this case we note that the cross-correlation between input and output is the convolution of the impulse response with the input autocorrelation sequence.

The Fourier transform of Eq. (2.178) is

$$\Phi_{xy}(e^{j\omega}) = H(e^{j\omega})\Phi_{xx}(e^{j\omega}). \tag{2.179}$$

This result has a useful application when the input is white noise; i.e., $\phi_{xx}[m] = \sigma_x^2\delta[m]$. Substituting into Eq. (2.176), we note that

$$\phi_{xy}[m] = \sigma_x^2 h[m], \tag{2.180}$$

i.e., for a zero-mean white-noise input, the cross-correlation between input and output of a linear system is proportional to the impulse response of the system. Similarly, the power spectrum of a white-noise input is

$$\Phi_{xx}(e^{j\omega}) = \sigma_x^2, \qquad -\pi \le \omega \le \pi. \tag{2.181}$$

Thus from Eq. (2.179)

$$\Phi_{xy}(e^{j\omega}) = \sigma_x^2 H(e^{j\omega}), \tag{2.182}$$

i.e., the cross power spectrum is in this case proportional to the frequency response of the system. Equations (2.180) and (2.182) may serve as the basis for estimating the impulse response or frequency response of a linear time-invariant system if it is possible to observe the output of the system in response to a white-noise input.

2.11 SUMMARY

In this chapter we have considered a number of basic definitions relating to discrete-time signals and systems. We considered the definition of a set of basic sequences, the definition and representation of linear time-invariant systems in terms of the convolution sum, and some implications of stability and causality. The class of systems for which the input and output satisfy a linear constant-coefficient difference equation with initial rest conditions was shown to be an important subclass of linear time-invariant systems. The recursive solution of such difference equations was discussed and the classes of FIR and IIR systems defined.

An important means for the analysis and representation of linear time-invariant systems lies in their frequency-domain representation. The response of a system to a complex exponential input was considered, leading to the definition of the frequency response. The relation between impulse response and frequency response was then interpreted as a Fourier transform pair.

We called attention to many properties of Fourier transform representations and discussed a variety of useful Fourier transform pairs. Tables 2.1 and 2.2 summarize the properties and theorems and Table 2.3 contains some useful Fourier transform pairs.

The chapter concludes with an introduction to discrete-time random signals. These basic ideas and results will be developed further and used in later chapters.

Although the material in this chapter was presented without direct reference to continuous-time signals, an important class of discrete-time signal processing problems arises from sampling such signals. Consequently, in Chapter 3 we consider the relationship between continuous-time signals and sequences obtained by periodic sampling.

PROBLEMS

2.1. Consider an arbitrary linear system with input $x[n]$ and output $y[n]$. Show that if $x[n] = 0$ for all n, then $y[n]$ must also be zero for all n.

2.2. Using the definition of linearity (Eqs. 2.25), show that the ideal delay system (Example 2.1) and the moving average system (Example 2.2) are both linear systems.

2.3. For each of the following systems, determine whether the system is (1) stable, (2) causal, (3) linear, (4) time-invariant, and (5) memoryless.

(a) $T(x[n]) = g[n]x[n]$

(b) $T(x[n]) = \sum_{k=n_0}^{n} x[k]$

(c) $T(x[n]) = \sum_{k=n-n_0}^{n+n_0} x[k]$

(d) $T(x[n]) = x[x - n_0]$

(e) $T(x[n]) = e^{x[n]}$

(f) $T(x[n]) = ax[n] + b$

(g) $T(x[n]) = x[-n]$

(h) $T(x[n]) = x[n] + 3u[n+1]$

2.4. The system T in Fig. P2.4 is known to be *time-invariant*. When the inputs to the system are $x_1[n]$, $x_2[n]$, and $x_3[n]$, the responses of the system are $y_1[n]$, $y_2[n]$, and $y_3[n]$ as shown.

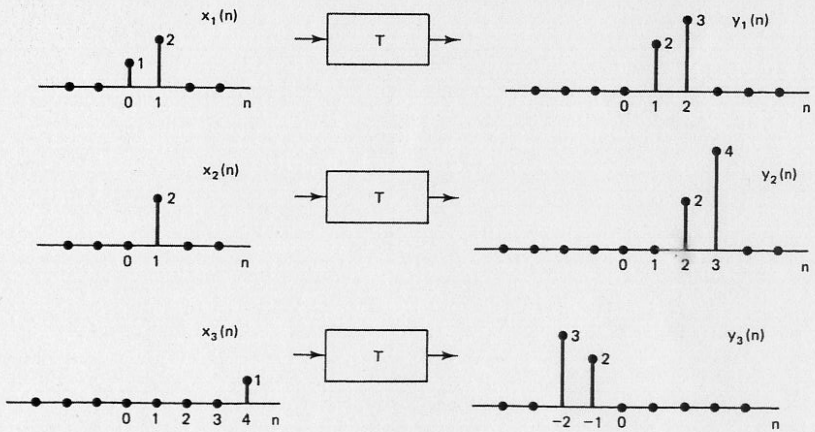

Figure P2.4

(a) Determine whether the system T could be linear.

(b) If the input $x[n]$ to the system T is $\delta[n]$, what is the system response $y[n]$?

(c) Determine all possible inputs $x[n]$ for which the response of the system T can be determined from the given information alone.

2.5. The system L in Fig. P2.5 is known to be *linear*. Shown are three output signals, $y_1[n]$, $y_2[n]$, and $y_3[n]$, in response to the input signals $x_1[n]$, $x_2[n]$, and $x_3[n]$, respectively.

(a) Determine whether the system L could be time-invariant.

(b) If the input $x[n]$ to the system L is $\delta[n]$, what is the system response $y[n]$?

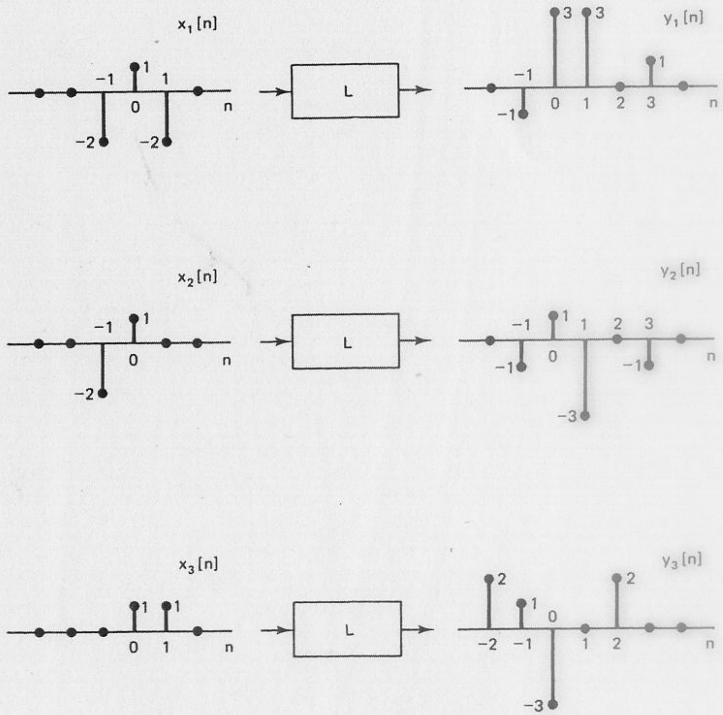

Figure P2.5

2.6. For each of the pairs of sequences in Fig. P2.6, use discrete convolution to find the response to the input $x[n]$ of the linear time-invariant system with impulse response $h[n]$.