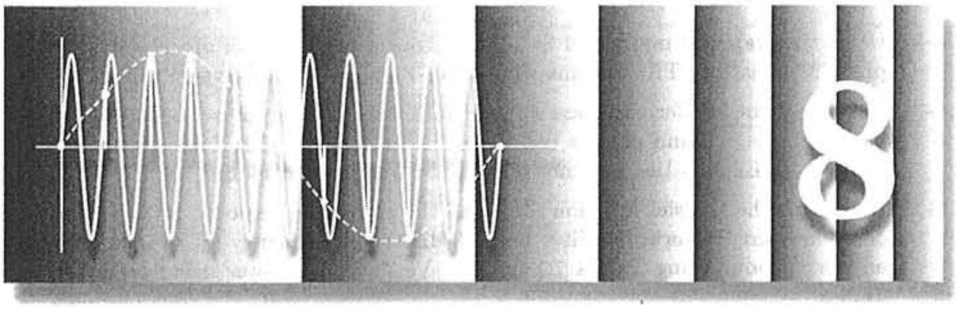

Discrete-Time Signals and Systems

In this chapter we introduce the basic concepts of discrete-time signals and systems.

8.1 Introduction

Signals specified over a continuous range of t are **continuous-time signals**, denoted by the symbols $f(t), y(t)$, etc. Systems whose inputs and outputs are continuous-time signals are **continuous-time systems**. In contrast, signals defined only at discrete instants of time are **discrete-time signals**. Systems whose inputs and outputs are discrete-time signals are called **discrete-time systems**. A digital computer is a familiar example of this type of system. We consider here uniformly spaced discrete instants such as $\ldots, -2T, -T, 0, T, 2T, 3T, \ldots, kT, \ldots$. Discrete-time signals can therefore be specified as $f(kT)$, $y(kT)$, and so on (k, integer). We further simplify this notation to $f[k], y[k]$, etc., where it is understood that $f[k] = f(kT)$ and that k is an integer. A typical discrete-time signal, depicted in Fig. 8.1, is therefore a sequence of numbers. This signal may be denoted by $f(kT)$ and viewed as a function of time t where signal values are specified at $t = kT$. It may also be denoted by $f[k]$ and viewed as a function of k (k, integer). For instance, a continuous-time exponential $f(t) = e^{-t}$, when sampled every $T = 0.1$ second, results in a discrete-time signal $f(kT)$ given by

$$f(kT) = e^{-kT} = e^{-0.1k}$$

Clearly, this signal is a function of k and may be expressed as $f[k]$. We can plot this signal as a function of t or as a function of k (k, integer). The representation $f[k]$ is more convenient and will be followed throughout this book. A discrete-time signal therefore may be viewed as a sequence of numbers, and a discrete-time system may be seen as processing a sequence of numbers $f[k]$ and yielding as output another sequence of numbers $y[k]$.

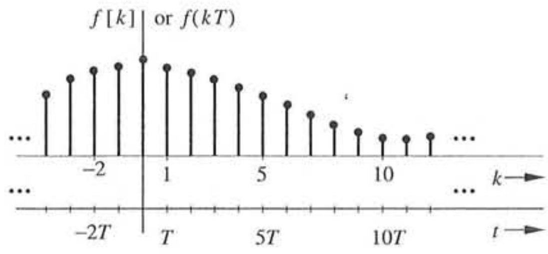

Fig. 8.1 A discrete-time signal.

Discrete-time signals arise naturally in situations which are inherently discrete-time, such as population studies, amortization problems, national income models, and radar tracking. They may also arise as a result of sampling continuous-time signals in sampled data systems, digital filtering, and so on. Digital filtering is a particularly interesting application in which continuous-time signals are processed by discrete-time systems, using appropriate interfaces at the input and output, as illustrated in Fig. 8.2. A continuous-time signal $f(t)$ is first sampled to convert it into a discrete-time signal $f[k]$, which is then processed by a discrete-time system to yield the output $y[k]$. A continuous-time signal $y(t)$ is finally constructed from $y[k]$. We shall use the notations C/D and D/C for continuous-to-discrete-time and discrete-to-continuous-time conversion. Using the interfaces in this manner, we can process a continuous-time signal with an appropriate discrete-time system. As we shall see later in our discussion, discrete-time systems have several advantages over continuous-time systems. For this reason, there is an accelerating trend toward processing continuous-time signals with discrete-time systems.

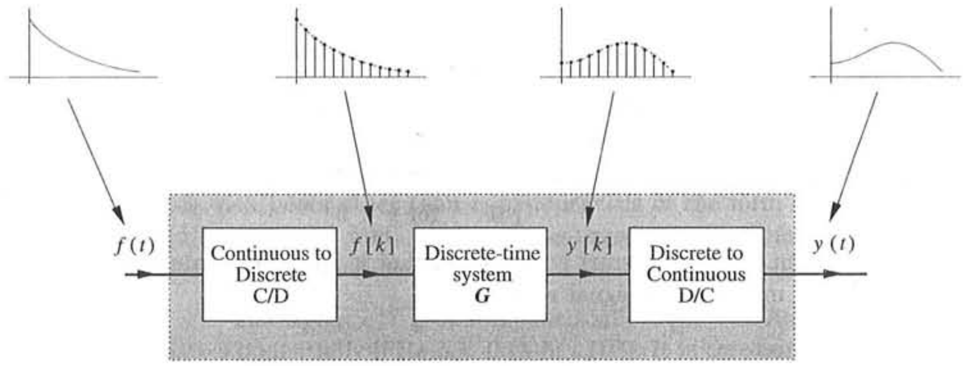

Fig. 8.2 Processing a continuous-time signal by a discrete-time system.

8.2 Some Useful Discrete-Time Signal Models

We now discuss some important discrete-time signal models which are encountered frequently in the study of discrete-time signals and systems.

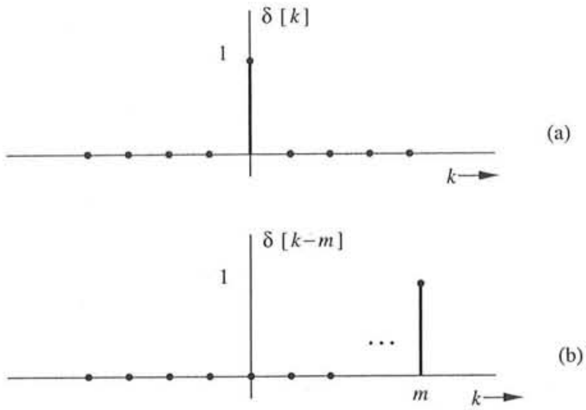

Fig. 8.3 Discrete-time impulse function.

1. Discrete-Time Impulse Function $\delta[k]$

The discrete-time counterpart of the continuous-time impulse function $\delta(t)$ is $\delta[k]$, defined by

$$\delta[k] = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases} \tag{8.1}$$

This function, also called the unit impulse sequence, is shown in Fig. 8.3a. The time-shifted impulse sequence $\delta[k-m]$ is depicted in Fig. 8.3b. Unlike its continuous-time counterpart $\delta(t)$, this is a very simple function without any mystery.

Later, we shall express an arbitrary input $f[k]$ in terms of impulse components. The (zero-state) system response to input $f[k]$ can then be obtained as the sum of system responses to impulse components of $f[k]$.

2. Discrete-Time Unit Step Function $u[k]$

The discrete-time counterpart of the unit step function $u(t)$ is $u[k]$ (Fig. 8.4), defined by

$$u[k] = \begin{cases} 1 & \text{for } k \geq 0 \\ 0 & \text{for } k < 0 \end{cases} \tag{8.2}$$

If we want a signal to start at $k = 0$ (so that it has a zero value for all $k < 0$), we need only multiply the signal with $u[k]$.

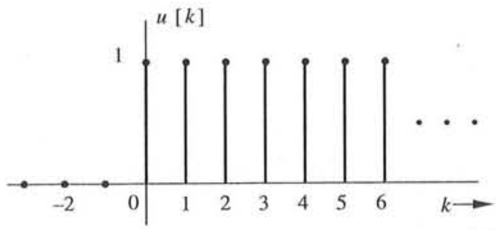

Fig. 8.4 A discrete-time unit step function $u[k]$.

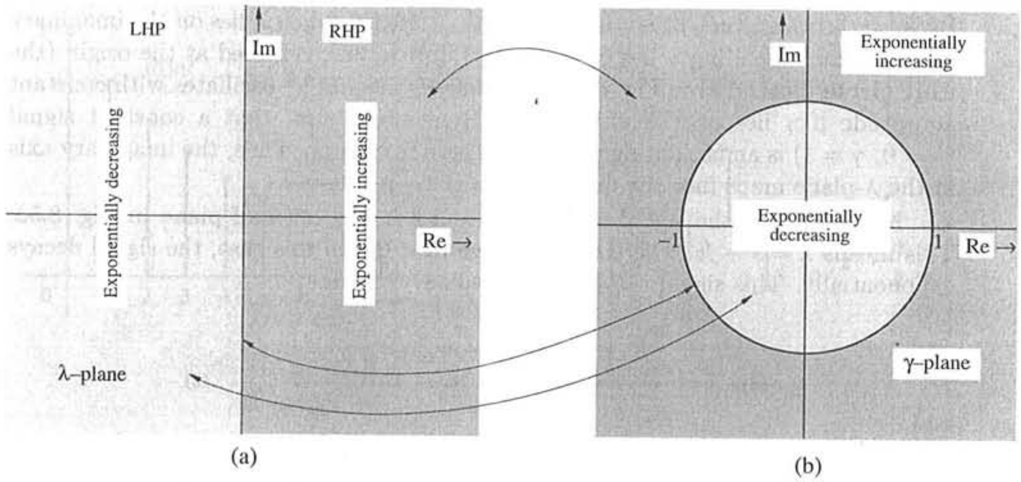

Fig. 8.5 The λ-plane, the γ-plane and their mapping.

3. Discrete-Time Exponential γ^k

A continuous-time exponential $e^{\lambda t}$ can be expressed in an alternate form as

$$e^{\lambda t} = \gamma^t \qquad (\gamma = e^\lambda \text{ or } \lambda = \ln \gamma) \tag{8.3a}$$

For example, $e^{-0.3t} = (0.7408)^t$ because $e^{-0.3} = 0.7408$. Conversely, $4^t = e^{1.386t}$ because $\ln 4 = 1.386$, that is, $e^{1.386} = 4$. In the study of continuous-time signals and systems we prefer the form $e^{\lambda t}$ rather than γ^t. The discrete-time exponential can also be expressed in two forms as

$$e^{\lambda k} = \gamma^k \qquad (\gamma = e^\lambda \text{ or } \lambda = \ln \gamma) \tag{8.3b}$$

For example, $e^{3k} = (e^3)^k = (20.086)^k$. Similarly, $5^k = e^{1.609k}$ because $5 = e^{1.609}$. In the study of discrete-time signals and systems, unlike the continuous-time case, the form γ^k proves more convenient than the form $e^{\lambda k}$. Because of unfamiliarity with exponentials with bases other than e, exponentials of the form γ^k may seem inconvenient and confusing at first. The reader is urged to plot some exponentials to acquire a sense of these functions.

Nature of γ^k: The signal $e^{\lambda k}$ grows exponentially with k if Re $\lambda > 0$ (λ in RHP), and decays exponentially if Re $\lambda < 0$ (λ in LHP). It is constant or oscillates with constant amplitude if Re $\lambda = 0$ (λ on the imaginary axis). Clearly, the location of λ in the complex plane indicates whether the signal $e^{\lambda k}$ grows exponentially, decays exponentially, or oscillates with constant frequency (Fig. 8.5a). A constant signal ($\lambda = 0$) is also an oscillation with zero frequency. We now find a similar criterion for determining the nature of γ^k from the location of γ in the complex plane.

Figure 8.5a shows a complex plane (λ-plane). Consider a signal $e^{j\Omega k}$. In this case, $\lambda = j\Omega$ lies on the imaginary axis (Fig. 8.5a), and therefore is a constant-amplitude oscillating signal. This signal $e^{j\Omega k}$ can be expressed as γ^k, where $\gamma = e^{j\Omega}$.

Because the magnitude of $e^{j\Omega}$ is unity, $|\gamma| = 1$. Hence, when λ lies on the imaginary axis, the corresponding γ lies on a circle of unit radius, centered at the origin (the **unit circle** illustrated in Fig. 8.5b). Therefore, a signal γ^k oscillates with constant amplitude if γ lies on the unit circle. Remember, also, that a constant signal ($\lambda = 0$, $\gamma = 1$) is an oscillating signal with zero frequency. Thus, the imaginary axis in the λ-plane maps into the unit circle in the γ-plane.

Next consider the signal $e^{\lambda k}$, where λ lies in the left-half plane in Fig. 8.5a. This means $\lambda = a + jb$, where a is negative ($a < 0$). In this case, the signal decays exponentially. This signal can be expressed as γ^k, where

$$\gamma = e^{\lambda} = e^{a+jb} = e^a\, e^{jb}$$

and

$$|\gamma| = |e^a|\,|e^{jb}| = e^a \qquad \text{because } |e^{jb}| = 1$$

Also, a is negative ($a < 0$). Hence, $|\gamma| = e^a < 1$. This result means that the corresponding γ lies inside the unit circle. Therefore, a signal γ^k decays exponentially if γ lies within the unit circle (Fig. 8.5b). If, in the above case we had selected a to be positive, (λ in the right-half plane), then $|\gamma| > 1$, and γ lies outside the unit circle. Therefore, a signal γ^k grows exponentially if γ lies outside the unit circle (Fig. 8.5b).

To summarize, the imaginary axis in the λ-plane maps into the unit circle in the γ-plane. The left-half plane in the λ-plane maps into the inside of the unit circle and the right-half of the λ-plane maps into the outside of the unit circle in the γ-plane, as depicted in Fig. 8.5. This fact means that the signal γ^k grows exponentially with k if γ is outside the unit circle ($|\gamma| > 1$), and decays exponentially if γ is inside the unit circle ($|\gamma| < 1$). The signal is constant or oscillates with constant amplitude if γ is on the unit circle ($|\gamma| = 1$).

Observe that

$$\gamma^{-k} = \left(\frac{1}{\gamma}\right)^k \tag{8.4}$$

Figures 8.6a and 8.6b show plots of $(0.8)^k$, and $(-0.8)^k$, respectively. Figures 8.6c and 8.6d show plots of $(0.5)^k$, and $(1.1)^k$, respectively. These plots verify our earlier conclusions about the location of γ and the nature of signal growth. Observe that a signal $(-\gamma)^k$ alternates sign successively (is positive for even values of k and negative for odd values of k, as depicted in Fig. 8.6b). Also, the exponential $(0.5)^k$ decays faster than $(0.8)^k$. The exponential $(0.5)^k$ can also be expressed as 2^{-k} because $(0.5)^{-1} = 2$ [see Eq. (8.4)].

△ **Exercise E8.1**

Sketch signals (a) $(1)^k$ (b) $(-1)^k$ (c) $(0.5)^k$ (d) $(-0.5)^k$ (e) $(0.5)^{-k}$ (f) 2^{-k} (g) $(-2)^k$. Express these exponentials as γ^k, and plot γ in the complex plane for each case. Verify that γ^k decays exponentially with k if γ lies inside the unit circle, and that γ^k grows with k if γ is outside the unit circle. If γ is on the unit circle, γ^k is constant or oscillates with a constant amplitude.

Hint: $(1)^k = 1$ for all k. However, $(-1)^k = 1$ for even values of k and is -1 for odd values of k. Therefore, $(-1)^k$ switches back and forth from 1 to -1 (oscillates with a constant amplitude). Note also that Eq. (8.4) yields $(0.5)^{-k} = 2^k$ ▽

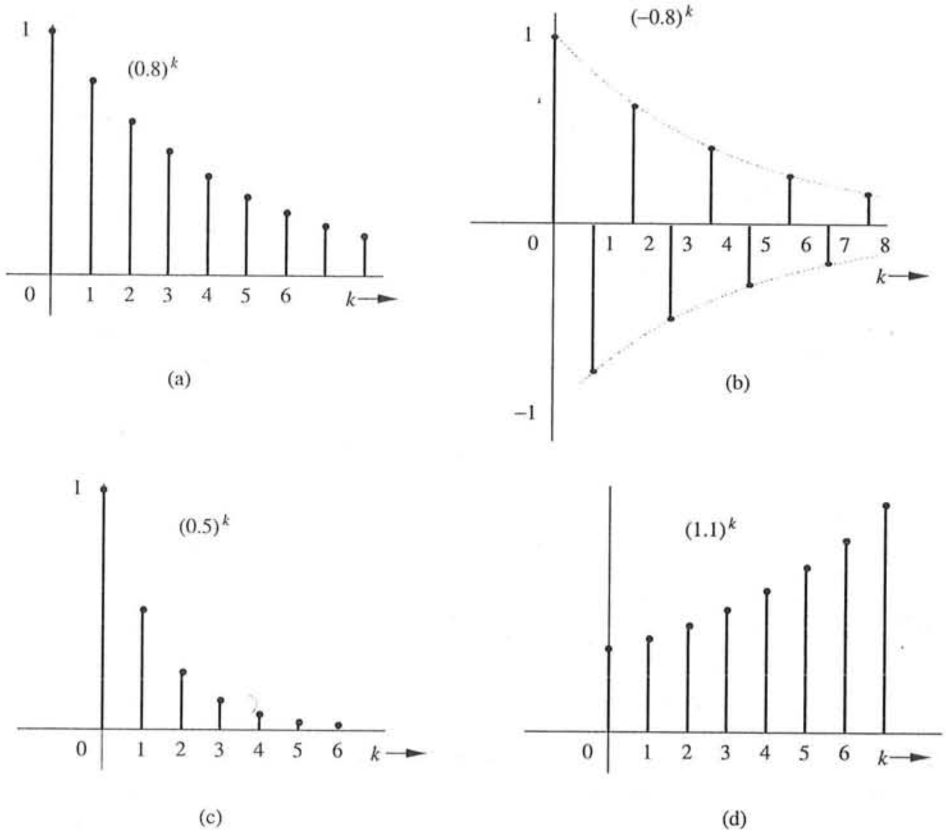

Fig. 8.6 discrete-time exponentials γ^k.

△ **Exercise E8.2**

(a) Show that (i) $(0.25)^{-k} = 4^k$ (ii) $4^{-k} = (0.25)^k$ (iii) $e^{2t} = (7.389)^t$ (iv) $e^{-2t} = (0.1353)^t = (7.389)^{-t}$ (v) $e^{3k} = (20.086)^k$ (vi) $e^{-1.5k} = (0.2231)^k = (4.4817)^{-k}$

(b) Show that (i) $2^k = e^{0.693k}$ (ii) $(0.5)^k = e^{-0.693k}$ (iii) $(0.8)^{-k} = e^{0.2231k}$ ▽

⊙ **Computer Example C8.1**

Sketch the discrete-time signals (a) $(-0.5)^k$ (b) $(2)^{-k}$ (c) $(-2)^k$

```
(a) k=0:5; k=k'; fk1=(-0.5).^k; stem(k,fk)
(b) k=0:5; k=k'; fk=2.^(-k); stem(k,fk)
(c) k=0:5; k=k';fk=(-2).^k; stem(k,fk3)
```
⊙

4. Discrete-Time Exponential $e^{j\Omega k}$

A general discrete-time exponential $e^{j\Omega k}$ (also called **phasor**) is a complex valued function of k and therefore its graphical description requires two plots (real part and imaginary part or magnitude and angle). To avoid two plots, we shall plot the values of $e^{j\Omega k}$ in the complex plane for various values of k, as illustrated in Fig. 8.7. The function $f[k] = e^{j\Omega k}$ takes on values e^{j0}, $e^{j\Omega}$, $e^{j2\Omega}$, $e^{j3\Omega}$, ... at $k = 0, 1, 2, 3, \ldots$, respectively. For the sake of simplicity we shall ignore the negative values of k for the time being. Note that

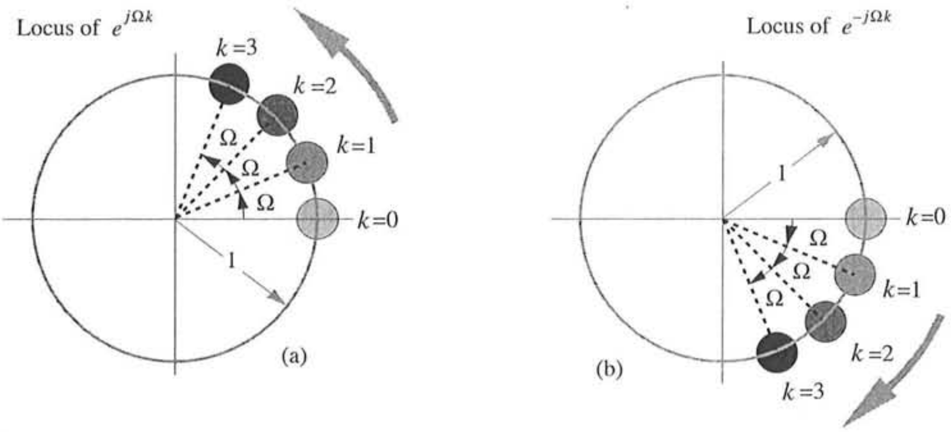

Fig. 8.7 Locus of (a) $e^{j\Omega k}$ (b) $e^{-j\Omega k}$.

$$e^{j\Omega k} = re^{j\theta}, \qquad r = 1, \text{ and } \theta = k\Omega$$

This fact shows that the magnitude and angle of $e^{j\Omega k}$ are 1 and $k\Omega$, respectively. Therefore, the points e^{j0}, $e^{j\Omega}$, $e^{j2\Omega}$, $e^{j3\Omega}$, ..., $e^{jk\Omega}$, ... lie on a circle of unit radius (unit circle) at angles 0, Ω, 2Ω, $3\Omega, \ldots, k\Omega, \ldots$ respectively, as shown in Fig. 8.7a. For each unit increase in k, the function $f[k] = e^{j\Omega k}$ moves along the unit circle counterclockwise by an angle Ω. Therefore, the locus of $e^{j\Omega k}$ may be viewed as a phasor rotating counterclockwise at a uniform speed of Ω radians per unit sample interval. The exponential $e^{-j\Omega k}$, on the other hand, takes on values $e^{j0} = 1$, $e^{-j\Omega}$, $e^{-j2\Omega k}$, $e^{-j3\Omega}, \ldots$ at $k = 0, 1, 2, 3, \ldots$, as depicted in Fig. 8.7b. Therefore, $e^{-j\Omega k}$ may be viewed as a phasor rotating clockwise at a uniform speed of Ω radians per unit sample interval.

Using Euler's formula, we can express an exponential $e^{j\Omega k}$ in terms of sinusoids of the form $\cos(\Omega k + \theta)$, and vice versa

$$e^{j\Omega k} = (\cos \Omega k + j \sin \Omega k) \tag{8.5a}$$

$$e^{-j\Omega k} = (\cos \Omega k - j \sin \Omega k) \tag{8.5b}$$

These equations show that **the frequency of both $e^{j\Omega k}$ and $e^{-j\Omega k}$ is** Ω (radians/sample). Therefore, the frequency of $e^{j\Omega k}$ is $|\Omega|$. Because of Eqs. (8.5), exponentials and sinusoids have similar properties and peculiarities. The discrete-time sinusoids will be considered next.

5. Discrete-Time Sinusoid $\cos(\Omega k + \theta)$

A general discrete-time sinusoid can be expressed as $C \cos(\Omega k + \theta)$, where C is the *amplitude*, Ω is the *frequency* (in radians per sample), and θ is the *phase* (in radians). Figure 8.8 shows a discrete-time sinusoid $\cos(\frac{\pi}{12}k + \frac{\pi}{4})$.

Here we make one basic observation. Because $\cos(-x) = \cos(x)$,

$$\cos(-\Omega k + \theta) = \cos(\Omega k - \theta) \tag{8.6}$$

This shows that both $\cos(\Omega k + \theta)$ and $\cos(-\Omega k + \theta)$ have the same frequency (Ω). Therefore, **the frequency of** $\cos(\Omega k + \theta)$ **is** $|\Omega|$.

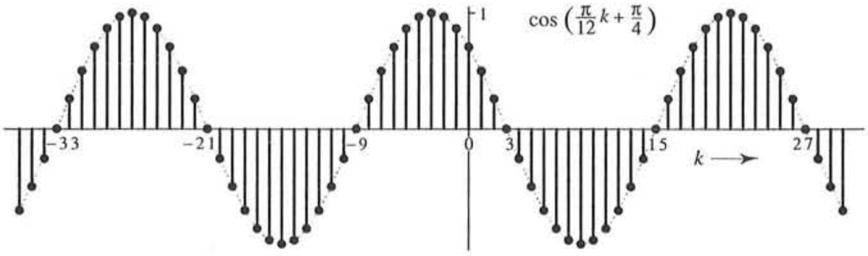

Fig. 8.8 A discrete-time sinusoid $\cos(\frac{\pi}{12}k+\frac{\pi}{4})$.

⊙ **Computer Example C8.2**

Sketch the discrete-time sinusoid $\cos\left(\frac{\pi}{12}k+\frac{\pi}{4}\right)$

```
k=-36:30; k=k';
fk=cos(k*pi/12+pi/4);
stem(k,fk)
```
⊙

Sampled Continuous-Time Sinusoid Yields a Discrete-Time Sinusoid

A continuous-time sinusoid $\cos \omega t$ sampled every T seconds yields a discrete-time sequence whose kth element (at $t = kT$) is $\cos \omega kT$. Thus, the sampled signal $f[k]$ is given by

$$\begin{aligned} f[k] &= \cos \omega kT \\ &= \cos \Omega k \qquad \text{where } \Omega = \omega T \end{aligned} \tag{8.7}$$

Clearly, a continuous-time sinusoid $\cos \omega t$ sampled every T seconds yields a discrete-time sinusoid $\cos \Omega k$, where $\Omega = \omega T$. Superficially, it may appear that a discrete-time sinusoid is a continuous-time sinusoid's cousin in a striped suit. As we shall see, however, some of the properties of discrete-time sinusoids are very different from those of continuous-time sinusoids. In the continuous-time case, the period of a sinusoid can take on any value; integral, fractional, or even irrational. The discrete-time signal, in contrast, is specified only at integral values of k. Therefore, the period must be an integer (in terms of k) or an integral multiple of T (in terms of variable t).

Some Peculiarities of Discrete-Time Sinusoids

There are two unexpected properties of discrete-time sinusoids which distinguish them from their continuous-time relatives.

1. A continuous-time sinusoid is always periodic regardless of the value of its frequency ω. But a discrete-time sinusoid $\cos \Omega k$ is periodic only if Ω is 2π times some rational number ($\frac{\Omega}{2\pi}$ is a rational number).
2. A continuous-time sinusoid $\cos \omega t$ has a unique waveform for each value of ω. In contrast, a sinusoid $\cos \Omega k$ does not have a unique waveform for each value of Ω. In fact, discrete-time sinusoids with frequencies separated by multiples of 2π are identical. Thus, a sinusoid $\cos \Omega k = \cos(\Omega+2\pi)k = \cos(\Omega+4\pi)k = \cdots$.

We now examine each of these peculiarities.

1 Not All Discrete-Time Sinusoids Are Periodic

A discrete-time signal $f[k]$ is said to be N_0-periodic if

$$f[k] = f[k+N_0] \tag{8.8}$$

for some positive integer N_0. The smallest value of N_0 that satisfies Eq. (8.8) is the **period** of $f[k]$. Figure 8.9 shows an example of a periodic signal of period 6. Observe that each period contains 6 samples (or values). If we consider the first cycle to start at $k = 0$, the last sample (or value) in this cycle is at $k = N_0 - 1 = 5$ (not at $k = N_0 = 6$). Note also that, by definition, a periodic signal must begin at $k = -\infty$ (everlasting signal) for the reasons discussed in Sec. 1.2-4.

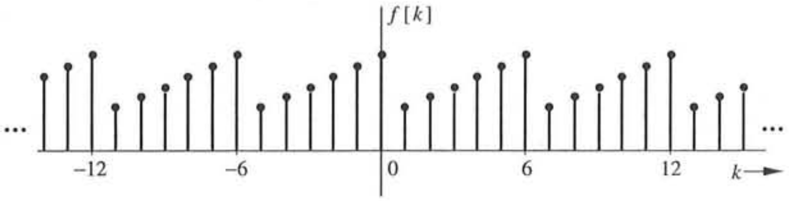

Fig. 8.9 Discrete-time periodic signal.

If a signal $\cos \Omega k$ is N_0-periodic, then

$$\begin{aligned}\cos \Omega k &= \cos \Omega(k + N_0) \\ &= \cos(\Omega k + \Omega N_0)\end{aligned}$$

This result is possible only if ΩN_0 is an integral multiple of 2π; that is,

$$\Omega N_0 = 2\pi m \qquad m \text{ integer}$$

or

$$\frac{\Omega}{2\pi} = \frac{m}{N_0} \tag{8.9a}$$

Because both m and N_0 are integers, Eq. (8.9a) implies that the sinusoid $\cos \Omega k$ is periodic only if $\frac{\Omega}{2\pi}$ is a rational number. In this case the period N_0 is given by [Eq. (8.9a)]

$$N_0 = m\left(\frac{2\pi}{\Omega}\right) \tag{8.9b}$$

To compute N_0, we must choose the smallest value of m that will make $m(\frac{2\pi}{\Omega})$ an integer. For example, if $\Omega = \frac{4\pi}{17}$, then the smallest value of m that will make $m\frac{2\pi}{\Omega} = m\frac{17}{2}$ an integer is 2. Therefore

$$N_0 = m\frac{2\pi}{\Omega} = 2\frac{17}{2} = 17$$

Using a similar argument, we can show that this discussion also applies to a discrete-time exponential $e^{j\Omega k}$. Thus, a discrete-time exponential $e^{j\Omega k}$ is periodic only if $\frac{\Omega}{2\pi}$ is a rational number.†

Physical Explanation of the Periodicity Relationship

Qualitatively, this result can be explained by recognizing that a discrete-time sinusoid $\cos \Omega k$ can be obtained by sampling a continuous-time sinusoid $\cos \Omega t$ at unit time interval $T = 1$; that is, $\cos \Omega t$ sampled at $t = 0, 1, 2, 3, \ldots$. This fact

†We can also demonstrate this point by observing that if $e^{j\Omega k}$ is N_0-periodic, then

$$e^{j\Omega k} = e^{j\Omega(k+N_0)} = e^{j\Omega k}e^{j\Omega N_0}$$

This result is possible only if $\Omega N_0 = 2\pi m$ (m, an integer). This conclusion leads to Eq. (8.9b).

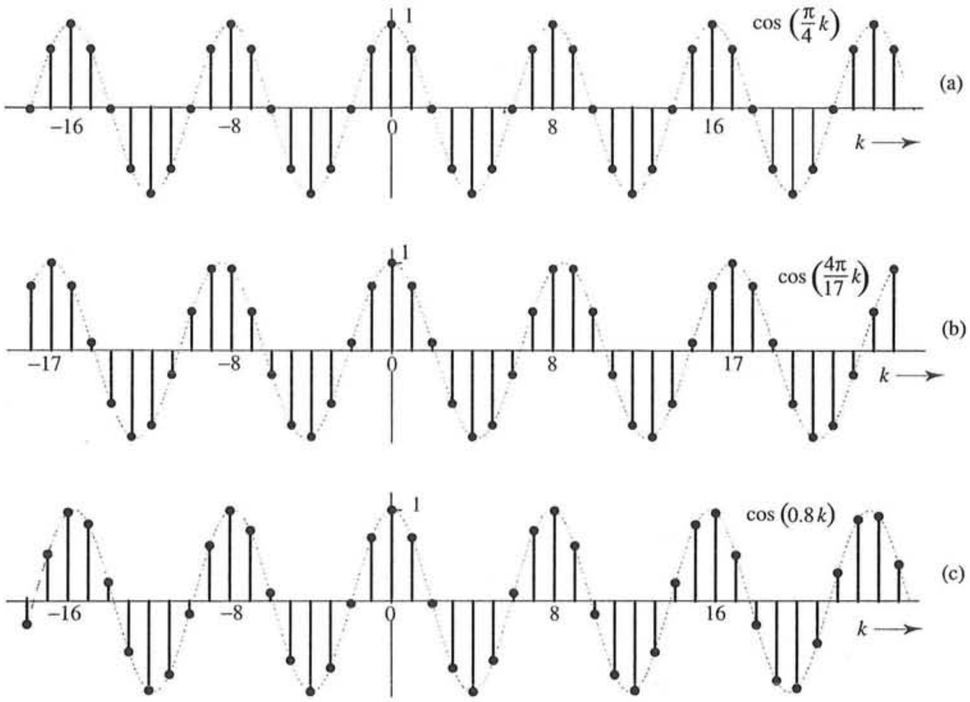

Fig. 8.10 Physical explanation of the periodicity relationship.

means $\cos \Omega t$ is the envelope of $\cos \Omega k$. Since the period of $\cos \Omega t$ is $2\pi/\Omega$, there are $2\pi/\Omega$ number of samples (elements) of $\cos \Omega k$ in one cycle of its envelope. This number may or may not be an integer.

Figure 8.10 shows three sinusoids $\cos\left(\frac{\pi}{4}k\right)$, $\cos\left(\frac{4\pi}{17}k\right)$, and $\cos(0.8k)$. Figure 8.10a shows $\cos\left(\frac{\pi}{4}k\right)$, for which there are exactly 8 samples in each cycle of its envelope $\left(\left(\frac{2\pi}{\Omega} = 8\right)\right.$. Thus, $\cos\left(\frac{\pi}{4}k\right)$ repeats every cycle of its envelope. Clearly, $\cos(4k/\pi)$ is periodic with period 8. On the other hand, Fig. 8.10b, which shows $\cos\left(\frac{4\pi}{17}k\right)$, has an average of $\frac{2\pi}{\Omega} = 8.5$ samples (not an integral number) in one cycle of its envelope. Therefore, the second cycle of the envelope will not be identical to the first cycle. But there are 17 samples (an integral number) in 2 cycles of its envelope. Hence, the pattern becomes repetitive every 2 cycles of its envelope. Therefore, $\cos\left(\frac{4\pi}{17}k\right)$ is also repetitive but its period is 17 samples (two cycles of its envelope). This observation indicates that a signal $\cos \Omega k$ is periodic only if we can fit an integral number (N_0) of samples in m integral number of cycles of its envelope so that the pattern becomes repetitive every m cycles of its envelope. Because the period of the envelope is $\frac{2\pi}{\Omega}$, we conclude that

$$N_0 = m\left(\frac{2\pi}{\Omega}\right)$$

which is precisely the condition of periodicity in Eq. (8.9b). If $\frac{\Omega}{2\pi}$ is irrational, it is impossible to fit an integral number (N_0) of samples in an integral number (m) of cycles of its envelope, and the pattern can never become repetitive. For instance, the sinusoid $\cos(0.8k)$ in Figure 8.10c has an average of 2.5π samples (an irrational number) per envelope cycle, and the pattern can never be made repetitive over any integral number (m) of cycles of its envelope; so $\cos(0.8k)$ is not periodic.

△ **Exercise E8.3**

State with reasons if the following sinusoids are periodic. If periodic, find the period.
(i) $\cos\left(\frac{3\pi}{7}k\right)$ **(ii)** $\cos\left(\frac{10}{7}k\right)$ **(iii)** $\cos\left(\sqrt{\pi}k\right)$

Ans: **(i)** Periodic: period $N_0 = 14$. **(ii)** and **(iii)** Aperiodic: $\Omega/2\pi$ irrational. ▽

⊙ **Computer Example C8.3**

Sketch and verify if $\cos\left(\frac{3\pi}{7}k\right)$ is periodic.

According to Eq. (8.9b), the smallest value of m that will make $N_0 = m\left(\frac{2\pi}{\Omega}\right) = m\left(\frac{14}{3}\right)$ an integer is 3. Therefore, $N_0 = 14$. This result means $\cos\left(\frac{3\pi}{7}k\right)$ is periodic and its period is 14 samples in three cycles of its envelop. This assertion can be verified by the following MATLAB commands:

```
t=-5*pi:pi/100:5*pi; t=t';
ft=cos(3*pi*t/7);
plot(t,ft,':'), hold on
k=-15:15; k=k';
fk=cos(k*3*pi/7);
stem(k,fk), hold off
```
⊙

2 Nonuniqueness of Discrete-Time Sinusoid Waveforms

A continuous-time sinusoid $\cos \omega t$ has a unique waveform for every value of ω in the range 0 to ∞. Increasing ω results in a sinusoid of ever increasing frequency. Such is not the case for the discrete-time sinusoid $\cos \Omega k$ because

$$\cos(\Omega \pm 2\pi m)k = \cos(\Omega k \pm 2\pi m k)$$

Now, if m is an integer, mk is also an integer, and the above equation reduces to

$$\cos(\Omega \pm 2\pi m)k = \cos \Omega k \qquad m \text{ integer} \tag{8.10}$$

This result shows that a discrete-time sinusoid of frequency Ω is indistinguishable from a sinusoid of frequency Ω plus or minus an integral multiple of 2π. This statement certainly does not apply to continuous-time sinusoids.

This result means that discrete-time sinusoids of frequencies separated by integral multiples of 2π are identical. The most dramatic consequence of this fact is that a discrete-time sinusoid $\cos(\Omega k + \theta)$ has a unique waveform only for the values of Ω over a range of 2π. We may select this range to be 0 to 2π, or π to 3π, or even $-\pi$ to π. The important thing is that the range must be of width 2π. A sinusoid of any frequency outside this interval is identical to a sinusoid of frequency within this range of width 2π. We shall select this range $-\pi$ to π and call it **the fundamental range of frequencies**. Thus, a sinusoid of any frequency Ω is identical to some sinusoid of frequency Ω_f in the fundamental range $-\pi$ to π. Consider, for example, sinusoids of frequencies $\Omega = 8.7\pi$ and 9.6π. We can add or subtract any integral multiple of 2π from these frequencies and the sinusoids will still remain unchanged. To reduce these frequencies to the fundamental range ($-\pi$ to π), we need to subtract $4 \times 2\pi = 8\pi$ from 8.7π and subtract $5 \times 2\pi = 10\pi$ from 9.6π, to yield frequencies 0.7π and -0.4π, respectively. Thus

$$\begin{aligned}\cos(8.7\pi k + \theta) &= \cos(0.7\pi k + \theta)\\ \cos(9.6\pi k + \theta) &= \cos(-0.4\pi k + \theta)\end{aligned} \tag{8.11}$$

This result shows that a sinusoid $\cos(\Omega k + \theta)$ *can always be expressed as* $\cos(\Omega_f k + \theta)$, *where* $-\pi \le \Omega_f < \pi$ *(the fundamental frequency range).* The reader should get used to the fact that the range of discrete-time frequencies is only 2π. We may select this range to be from $-\pi$ to π or from 0 to 2π, or any other interval of width 2π. It is most convenient to use the range from $-\pi$ to π. At times, however, we shall find it convenient to use the range from 0 to 2π. Thus, in the discrete-time world, frequencies can be considered to lie only in the fundamental frequency range (from $-\pi$ to π, for instance). Sinusoids of frequencies outside the fundamental frequencies do exist technically. But physically, they cannot be distinguished from the sinusoids of frequencies within the fundamental range. Thus, a discrete-time sinusoid of any frequency, no matter how high, is identical to a sinusoid of some frequency within the fundamental range ($-\pi$ to π).

The above results, derived for discrete-time sinusoids, are also applicable to discrete-time exponentials of the form $e^{j\Omega k}$. For example

$$e^{j(\Omega \pm 2\pi m)k} = e^{j\Omega k} e^{\pm j2\pi mk} = e^{j\Omega k} \qquad m, \text{ integer} \tag{8.12}$$

Here we have used the fact that $e^{\pm j2\pi n} = 1$ for all integral values of n. This result means that discrete-time exponentials of frequencies separated by integral multiples of 2π are identical.

Further Reduction in the Frequency Range of Distinguishable Discrete-Time Sinusoids

We shall now show that the range of frequencies that can be distinguished can be further reduced from $(-\pi, \pi)$ to $(0, \pi)$. According to Eq. (8.6), $\cos(-\Omega k + \theta) = \cos(\Omega k - \theta)$. In other words, the frequencies in the range (0 to $-\pi$) can be expressed as frequencies in the range (0 to π) with opposite phase. For example, the second sinusoid in Eq. (8.11) can be expressed as

$$\cos(9.6\pi k + \theta) = \cos(-0.4\pi k + \theta) = \cos(0.4\pi k - \theta) \tag{8.13}$$

This result shows that a sinusoid of any frequency Ω can always be expressed as a sinusoid of a frequency $|\Omega_f|$, where $|\Omega_f|$ lies in the range 0 to π. Note, however, a possible sign change in the phases of the two sinusoids. In other words, a discrete-time sinusoid of any frequency, no matter how high, is identical in every respect to a sinusoid within the fundamental frequency range, such as $-\pi$ to π. In contrast, a discrete-time sinusoid of any frequency, no matter how high, can be expressed, with a possible sign change in phase, as a sinusoid of frequency in the range $(0, \pi)$; that is, within **half the fundamental frequency range**.

A systematic procedure to reduce the frequency of a sinusoid $\cos(\Omega k + \theta)$ is to express Ω as†

$$\Omega = \Omega_f + 2\pi m \qquad |\Omega_f| \le \pi, \quad \text{and } m \text{ an integer} \tag{8.14}$$

This procedure is always possible. The reduced frequency of the sinusoid $\cos(\Omega k + \theta)$ is then $|\Omega_f|$.

---

†Equation (8.14) can also be expressed as $\Omega_f = \Omega|_{\text{modulo } 2\pi}$

■ **Example 8.1**

Consider sinusoids of frequencies Ω equal to (a) 0.5π (b) 1.6π (c) 2.5π (d) 5.6π (e) 34.116. Each of these sinusoids is equivalent to a sinusoid of some frequency $|\Omega_f|$ in the range 0 to π. We shall now determine these frequencies. This goal is readily accomplished by expressing the frequency Ω as in Eq. (8.14).

(a) The frequency 0.5π is in the range (0 to π) so that it cannot be reduced further.

(b) The frequency $1.6\pi = 2\pi - 0.4\pi$, and $\Omega_f = -0.4\pi$. Therefore, a sinusoid of frequency 1.6π can be expressed as a sinusoid of frequency $|\Omega_f| = 0.4\pi$.

(c) $2.5\pi = 2\pi + 0.5\pi$, and $\Omega_f = 0.5\pi$. Therefore, a sinusoid of frequency 2.5π can be expressed as a sinusoid of frequency $|\Omega_f| = 0.5\pi$.

(d) $5.6\pi = 3(2\pi) - 0.4\pi$, and $\Omega_f = -0.4\pi$. Therefore, a sinusoid of frequency 5.6π can be expressed as a sinusoid of frequency $|\Omega_f| = 0.4\pi$.

(e) $34.116 = 5(2\pi) + 2.7$, and $\Omega_f = 2.7$. Therefore, a sinusoid of frequency 34.116 can be expressed as a sinusoid of frequency $|\Omega_f| = 2.7$. ■

The fundamental range frequencies can be determined by using a simple graphical artifice as follows: mark all the frequencies on a tape using a linear scale, starting with zero frequency. Now wind this tape continuously around the two poles, one at $|\Omega_f| = 0$ and the other at $|\Omega_f| = \pi$, as illustrated in Fig. 8.11. The reduced value of any frequency marked on the tape is its projection on the horizontal ($|\Omega_f|$) axis. For instance, the reduced frequency corresponding to $\Omega = 1.6\pi$ is 0.4π (the projection of 1.6π on the horizontal Ω_f axis). Similarly, frequencies 2.5π, 5.6π, and 34.116 correspond to frequencies 0.5π, 0.4π, and 2.7 on the $|\Omega_f|$ axis.

△ **Exercise E8.4**

Show that the sinusoids of frequencies $\Omega =$ (a) 2π (b) 3π (c) 5π (d) 3.2π (e) 22.1327 (f) $\pi+2$ can be expressed as sinusoids of frequencies (a) 0 (b) π (c) π (d) 0.8π (e) 3 (f) $\pi - 2$, respectively. ▽.

△ **Exercise E8.5**

Show that a discrete-time sinusoid of frequency $\pi + x$ can be expressed as a sinusoid with frequency $\pi - x$ $(0 \le x \le \pi)$. This fact shows that a sinusoid with frequency above π by amount x has the frequency identical to a sinusoid of frequency below π by the same amount x, and the maximum rate of oscillation occurs at $\Omega = \pi$. As Ω increases beyond π, the rate of oscillation actually decreases. ▽.

⊙ **Computer Example C8.4**

In the fundamental range of frequencies from $-\pi$ to π find a sinusoid that is indistinguishable from the sinusoid $\cos\left(\frac{3\pi}{7}k\right)$. Verify by plotting these two sinusoids that they are indeed identical.

The sinusoid $\cos\left(\frac{3\pi}{7}k\right)$ is identical to the sinusoid $\cos\left(\frac{3\pi}{7} - 2\pi\right)k = \cos\left(-\frac{11\pi}{7}k\right) = \cos\left(\frac{11\pi}{7}k\right)$. We may verify that these two sinusoids are identical.

```
k=-15:15; k=k';
fk1=cos(3*pi*k/7);
fk2=cos(11*pi*k/7);
stem(k,fk1,'x'),hold on,
stem(k,fk2),hold off
```
⊙

Physical Explanation of Nonuniqueness of Discrete-Time Sinusoids

Nonuniqueness of discrete-time sinusoids is easy to prove mathematically. But why does it happen physically? We now give here two different physical explanations of this intriguing phenomenon.

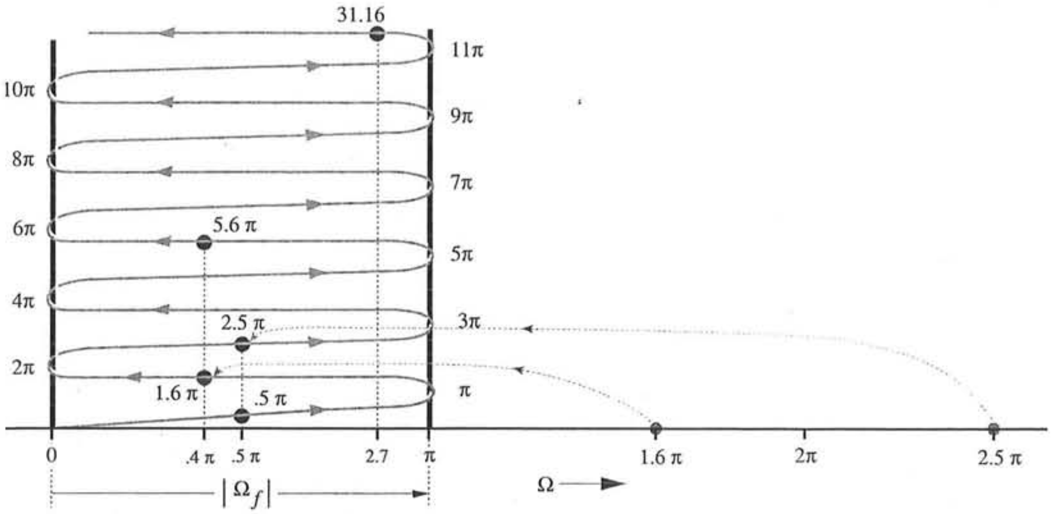

Fig. 8.11 A graphical artifice to determine the reduced frequency of a discrete-time sinusoid.

The First Explanation

Recall that sampling a continuous-time sinusoid cos Ωt at unit time intervals ($T = 1$) generates a discrete-time sinusoid cos Ωk. Thus, by sampling at unit intervals, we generate a discrete-time sinusoid of frequency Ω (rad/sample) from a continuous-time sinusoid of frequency Ω (rad/s). Superficially, it appears that since a continuous-time sinusoid waveform is unique for each value of Ω, the resulting discrete-time sinusoid must also have a unique waveform for each Ω. Recall, however, that there is a unit time interval between samples. If a continuous-time sinusoid executes several cycles during unit time (between successive samples), it will not be visible in its samples. The sinusoid may just as well not have executed those cycles. Another low frequency continuous-time sinusoid could also give the same samples. Figure 8.12 shows how the samples of two very different continuous-time sinusoids of different frequencies generate identical discrete-time sinusoid. This illustration explains why two discrete-time sinusoids whose frequencies Ω are nominally different have the same waveform.

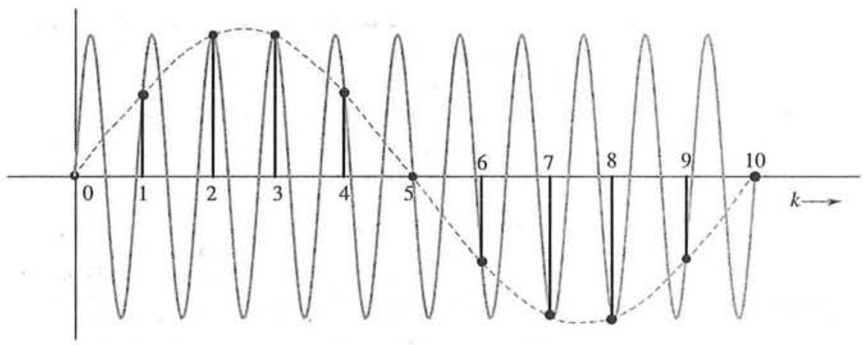

Fig. 8.12 Physical explanation of nonuniqueness of Discrete-time sinusoid waveforms.

Human Eye is a Lowpass Filter

Figure 8.12 also brings out one interesting fact; that a human eye is a lowpass filter. Both the continuous-time sinusoids in Fig. 8.12 have the same set of samples. Yet, when we see the samples, we interpret them as the samples of the lower frequency sinusoid. The eye does not see (or cannot reconstruct) the wiggles of the higher frequency sinusoid between samples because the eye is basically a lowpass filter.

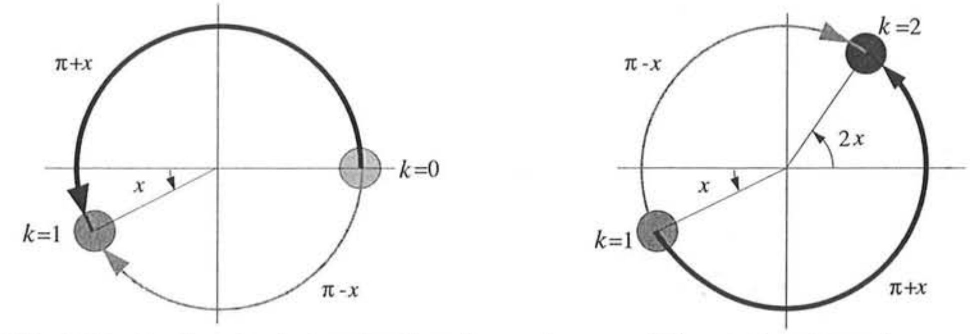

Fig. 8.13 Another physical explanation of nonuniqueness of discrete-time sinusoid waveforms.

The Second Explanation

Here we shall present a quantitative argument using a discrete-time exponential rather than a discrete-time sinusoid. As explained earlier, a discrete-time exponential $e^{j\Omega k}$ can be viewed as a phasor rotating counterclockwise at a uniform angular velocity of Ω rad/sample, as shown in Fig. 8.7a. A similar argument shows that the exponential $e^{-j\Omega k}$ is a phasor rotating clockwise at a uniform angular velocity of Ω radians per sample, as depicted in Fig. 8.7b. The angular velocity of both these rotating phasors is Ω rad. Therefore, as the frequency Ω increases, the angular velocity also increases. This, however, is true only for values of Ω in the range 0 to π. Something very interesting happens when the frequency Ω increases beyond π. Let $\Omega = \pi + x$ where $x < \pi$. Figure 8.13a shows the phasor progressing from $k = 0$ to $k = 1$, and Fig. 8.13b shows the same phasor progressing from $k = 1$ to $k = 2$. Because the phasor rotates at a speed of $\Omega = \pi + x$ radians/sample, the phasor angles at $k = 0$, 1, and 2 are 0, $\pi + x$ and $2\pi + 2x = 2x$, respectively. In both the figures, the phasor is progressing counterclockwise at a velocity of $(\pi + x)$ rad/sample. But we may also interpret this motion as the phasor moving clockwise (shown in gray) at a lower speed of $(\pi - x)$ rad/sample. Either of these interpretations describes the phasor motion correctly. If this motion could be seen by a human eye, which is a lowpass filter, it will automatically interpret the speed as $\pi - x$, the lower of the two speeds. This is the stroboscopic effect observed in movies, where at certain speeds, carriage wheels appear to move backwards.†

†A stroboscope is a source of light that flashes periodically on an object, thus generating a sampled image of that object. When a stroboscope flashes on a rotating object, such as a wheel, the wheel appears to rotate at a certain speed. Now increase the actual speed of rotation (while maintaining the same flashing rate). If the speed is increased beyond some critical value, the wheels appear to rotate backwards because of the lowpass filtering effect described above in the text. As we continue to increase the speed further, the backward rotation appears to slow down continuously to zero speed (where the wheels appear stationary), and reverse the direction again. This effect is often observed in movies in scenes with running carriages. A movie reel consists of a sequence of photographs shot at discrete instants, and is basically a sampled signal.

Thus, in a signal $e^{j\Omega k}$, the frequency $\Omega = \pi + x$ appears as frequency $\pi - x$. Therefore, as Ω increases beyond π, the actual frequency decreases, until at $\Omega = 2\pi$ ($x = \pi$), the actual frequency is zero ($\pi - x = 0$). As we increase Ω beyond 2π, the same cycle of events repeats. For instance, $\Omega' = 2.5\pi$ is the same as $\Omega = 0.5\pi$.

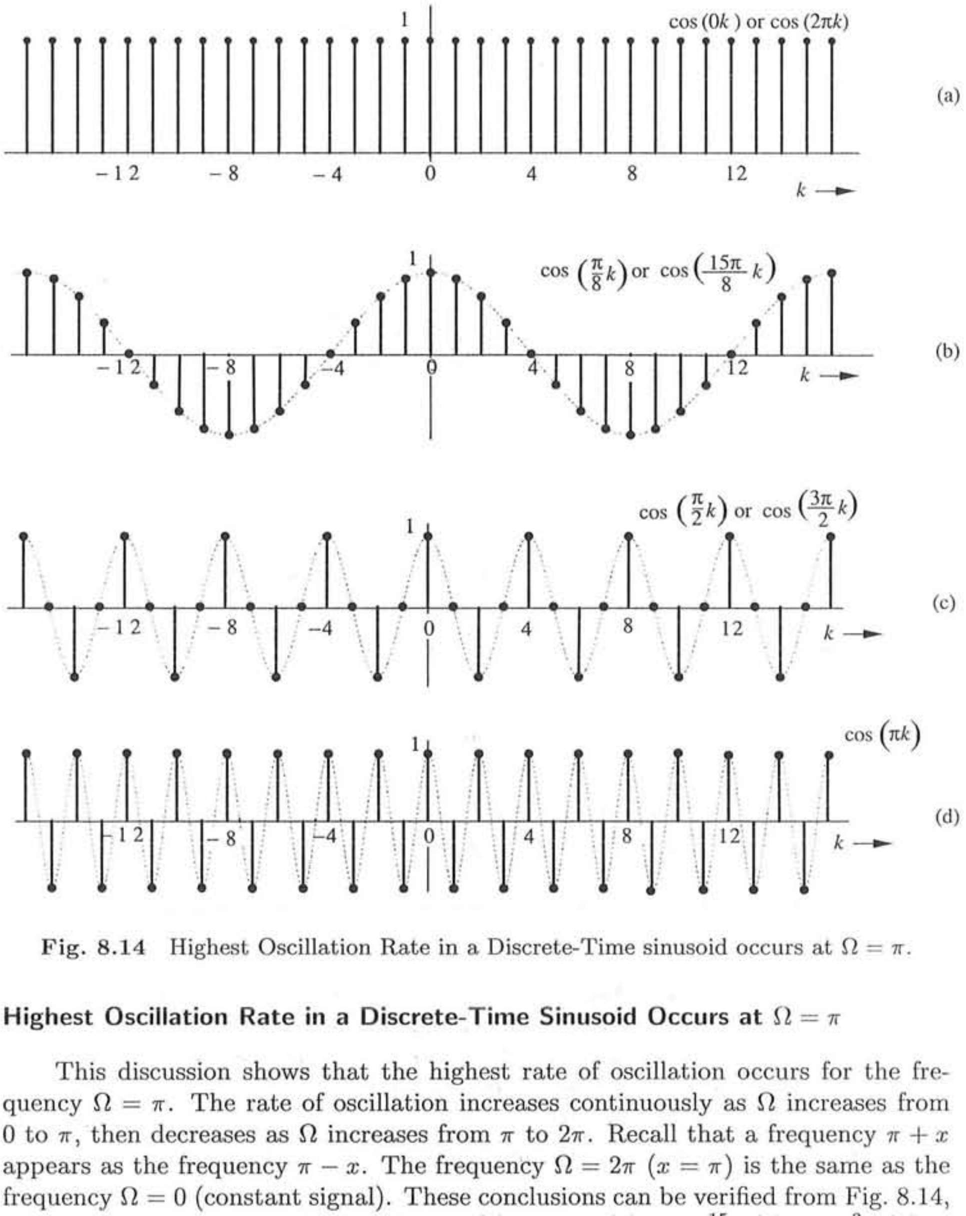

Fig. 8.14 Highest Oscillation Rate in a Discrete-Time sinusoid occurs at $\Omega = \pi$.

Highest Oscillation Rate in a Discrete-Time Sinusoid Occurs at $\Omega = \pi$

This discussion shows that the highest rate of oscillation occurs for the frequency $\Omega = \pi$. The rate of oscillation increases continuously as Ω increases from 0 to π, then decreases as Ω increases from π to 2π. Recall that a frequency $\pi + x$ appears as the frequency $\pi - x$. The frequency $\Omega = 2\pi$ ($x = \pi$) is the same as the frequency $\Omega = 0$ (constant signal). These conclusions can be verified from Fig. 8.14, which shows sinusoids of frequencies $\Omega =$ (a) 0 or 2π (b) $\frac{\pi}{8}$ or $\frac{15\pi}{8}$ (c) $\frac{\pi}{2}$ or $\frac{3\pi}{2}$ (d) π.

6. Exponentially Varying Discrete-Time Sinusoid $\gamma^k \cos(\Omega k + \theta)$

This is a sinusoid $\cos(\Omega k + \theta)$ with an exponentially varying amplitude γ^k. It is obtained by multiplying the sinusoid $\cos(\Omega k + \theta)$ by an exponential γ^k. Figure 8.15

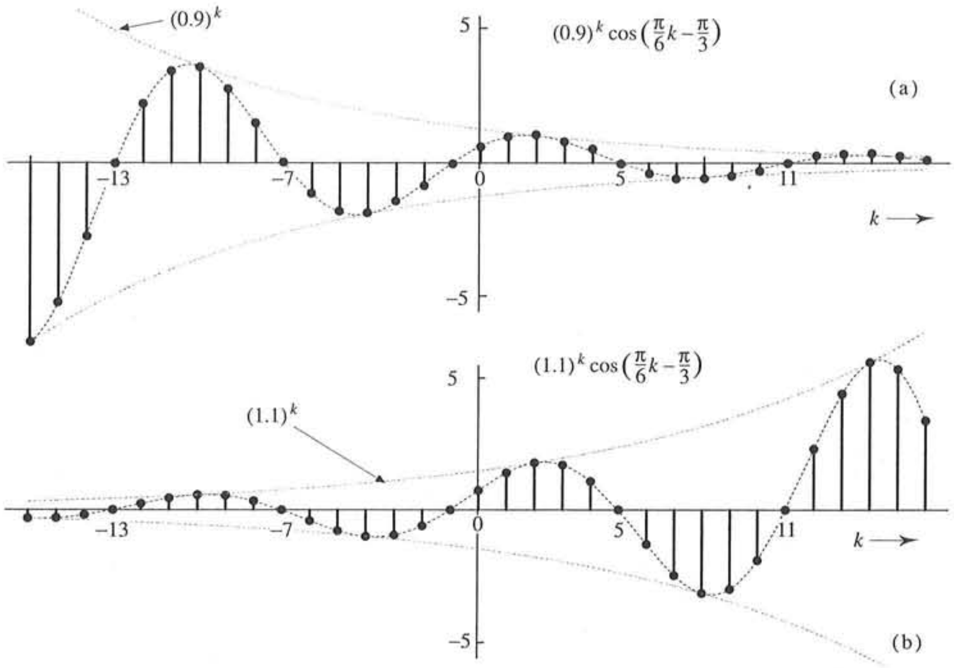

Fig. 8.15 Examples of exponentially varying discrete-time sinusoids.

shows signals $(0.9)^k \cos\left(\frac{\pi}{6}k - \frac{\pi}{3}\right)$, and $(1.1)^k \cos\left(\frac{\pi}{6}k - \frac{\pi}{3}\right)$. Observe that if $|\gamma| < 1$, the amplitude decays, and if $|\gamma| > 1$, the amplitude grows exponentially.

8.2-1 Size of a Discrete-Time Signal

Arguing along the lines similar to those used in continuous-time signals, the size of a discrete-time signal $f[k]$ will be measured by its energy E_f defined by

$$E_f = \sum_{k=-\infty}^{\infty} |f[k]|^2 \tag{8.15}$$

This definition is valid for real or complex $f[k]$. For this measure to be meaningful, the energy of a signal must be finite. A necessary condition for the energy to be finite is that the signal amplitude must $\to 0$ as $|k| \to \infty$. Otherwise the sum in Eq. (8.15) will not converge. If E_f is finite, the signal is called an **energy signal**.

In some cases, for instance, when the amplitude of $f[k]$ does not $\to 0$ as $|k| \to \infty$, then the signal energy is infinite, and a more meaningful measure of the signal in such a case would be the time average of the energy (if it exists), which is the signal power P_f defined by

$$P_f = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{-N}^{N} |f[k]|^2 \tag{8.16}$$

For periodic signals, the time averaging need be performed only over one period in view of the periodic repetition of the signal. If P_f is finite and nonzero, the signal is called a **power signal**. As in the continuous-time case, a discrete-time signal can either be an energy signal or a power signal, but cannot be both at the same time. Some signals are neither energy nor power signals.

△ **Exercise E8.7**

(a) Show that the signal $a^k u[k]$ is an energy signal of energy $\frac{1}{1-|a|^2}$ if $|a| < 1$. It is a power signal of power $P_f = 0.5$ if $|a| = 1$. It is neither an energy signal nor a power signal if $|a| > 1$. ▽

8.3 Sampling Continuous-Time Sinusoid and Aliasing

On the surface, the fact that discrete-time sinusoids of frequencies differing by $2\pi m$ are identical may appear innocuous. But in reality it creates a serious problem for processing continuous-time signals by digital filters. A continuous-time sinusoid $f(t) = \cos \omega t$ sampled every T seconds ($t = kT$) results in a discrete-time sinusoid $f[k] = \cos \omega kT$. Thus, the sampled signal $f[k]$ is given by

$$\begin{aligned} f[k] &= \cos \omega kT \\ &= \cos \Omega k \qquad \text{where } \Omega = \omega T \end{aligned}$$

Recall that the discrete-time sinusoids $\cos \Omega k$ have unique waveforms only for the values of frequencies in the range $\Omega \leq \pi$ or $\omega T \leq \pi$ (fundamental frequency range). We know that a sinusoid of frequency $\Omega > \pi$ appears as a sinusoid of a lower frequency $\Omega \leq \pi$. For a sampled continuous-time sinusoid, this fact means that samples of a sinusoid of frequency $\omega > \pi/T$ appear as samples of a sinusoid of lower frequency $\omega \leq \pi/T$. The mechanism of how the samples of continuous-time sinusoids of two (or more) different frequencies can generate the same discrete-time signal is shown in Fig. 8.12. *This phenomenon is known as* **aliasing** *because, through sampling, two entirely different analog sinusoids take on the same "discrete-time" identity.*

Aliasing causes ambiguity in digital signal processing, which makes it impossible to determine the true frequency of the sampled signal. Therefore, aliasing is highly undesirable and should be avoided. To avoid aliasing, the frequencies of the continuous-time sinusoids to be processed should be kept within the range $\omega T \leq \pi$ or $\omega \leq \pi/T$. Under this condition, the question of ambiguity or aliasing does not arise because any continuous-time sinusoid of frequency in this range has a unique waveform when it is sampled. Therefore, if ω_h is the highest frequency to be processed, then, to avoid aliasing,

$$\omega_h \leq \frac{\pi}{T} \tag{8.17a}$$

If $\mathcal{F}_h$ is the highest frequency in Hertz, $\mathcal{F}_h = \omega_h/2\pi$, and, according to Eq. (8.17a),

$$\mathcal{F}_h \leq \frac{1}{2T} \tag{8.17b}$$

or

$$T \leq \frac{1}{2\mathcal{F}_h} \tag{8.17c}$$

This result shows that discrete-time signal processing places the limit on the highest frequency $\mathcal{F}_h$ that can be processed for a given value of the sampling interval T

according to Eq. (8.17b). But we can process a signal of any frequency (without aliasing) by choosing a sufficiently low value of T according to Eq. (8.17c). The sampling rate or sampling frequency $\mathcal{F}_s$ is the reciprocal of the sampling interval T, and, according to Eq. (8.17c),†

$$\mathcal{F}_s = \frac{1}{T} \geq 2\mathcal{F}_h \tag{8.18}$$

This result, which is a special case of the **sampling theorem** (proved in Chapter 5), states that to process a continuous-time sinusoid by a discrete-time system, the sampling rate must not be less than twice the frequency (in Hz) of the sinusoid. In short, **a sampled sinusoid must have a minimum of two samples per cycle**. For a sampling rate below this minimum value, the output signal will be aliased, which means the signal will be mistaken for a sinusoid of lower frequency.

Equation 8.18 indicates that $\mathcal{F}_h$, the highest frequency that can be processed, is half the sampling frequency $\mathcal{F}_s$. This means the range of frequencies that can be processed without aliasing is from 0 to $\mathcal{F}_s/2$

$$0 \leq \mathcal{F} \leq \frac{\mathcal{F}_s}{2} \tag{8.19}$$

Frequencies greater than $\mathcal{F}_s/2$ (half the sampling frequency) will be aliased and appear as frequencies lower than $\mathcal{F}_s/2$. The aliasing appears as a folding back of frequencies about $\mathcal{F}_s/2$. Hence, this frequency is also known as the folding frequency. The details of this folding are explained more fully in Fig. 5.6.

The folding process is multilayered, as depicted in Fig. 8.11. The spectrum first folds back at the folding frequency, and then again folds forward at the origin, then back again at the folding frequency, and so on. We can find the aliased frequency (the reduced frequency) using an equation similar to Eq. (8.14) applicable to sampled continuous-time sinusoids.

We saw that a continuous-time sinusoid of frequency ω appears as a discrete-time sinusoid of frequency $\Omega = \omega T$. Hence, if ω_f is the reduced (aliased) frequency corresponding to a sinusoid of frequency ω, then, according to Eq. (8.14)

$$\omega T = \omega_f T + 2\pi m \qquad |\omega_f| T \leq \pi, \quad \text{and } m \text{ an integer} \tag{8.20}$$

When we express the radian frequencies in Hertz ($\omega = 2\pi\mathcal{F}$, etc.), and use the fact that the sampling frequency $\mathcal{F}_s = \frac{1}{T}$, Eq. (8.20) becomes

$$\mathcal{F} = \mathcal{F}_f + m\mathcal{F}_s \qquad |\mathcal{F}_f| \leq \frac{\mathcal{F}_s}{2}, \quad \text{and } m \text{ an integer} \tag{8.21}$$

Thus, if a continuous-time sinusoid of frequency $\mathcal{F}$ Hz is sampled at a rate of $\mathcal{F}_s$ Hz (samples/second), the resulting samples would appear as if they had come from a continuous-time sinusoid of a lower (aliased) frequency $|\mathcal{F}_f|$. For instance, if a continuous-time sinusoid of frequency 10 kHz were sampled at a rate of 3 kHz (3000 samples/second), the resulting samples will appear as if they had come from a continuous-time sinusoid of frequency 1 kHz because $10,000 = 1,000 + 3(3000)$. Note, however, if the frequency of a sinusoid is less than the folding frequency $\mathcal{F}_s/2$ (half the sampling frequency), there is no aliasing. Thus, the condition for the absence of aliasing is that the frequency of a sinusoid must be less than half the sampling frequency (the folding frequency).

†In some special cases, where the signal spectrum contains an impulse at $\mathcal{F}_h$, the sampling rate $\mathcal{F}_s$ must be greater than $2\mathcal{F}_h$ (see footnote on p.321)

Example 8.2

Determine the maximum sampling interval T that can be used in a discrete-time oscillator which generates a sinusoid of 50 kHz.

Here the highest frequency $\mathcal{F}_h = 50$ kHz. Therefore, according to Eq. (8.17c)

$$T \leq \frac{1}{2\mathcal{F}_h} = 10\,\mu\text{s}$$

The sampling interval must not be greater than $10\,\mu$s. The minimum sampling frequency is $\mathcal{F}_s = \frac{1}{T} = 100\,\text{kHz}$. If we use $T = 10\,\mu$s, the oscillator output will exhibit two samples per cycle. If we require the oscillator output to have 20 samples per cycle, then we must use $T = 1\,\mu$s (sampling frequency $\mathcal{F}_s = 1$ MHz). ■

Example 8.3

A discrete-time amplifier uses a sampling interval $T = 25\,\mu$s. What is the highest frequency of a signal that can be processed with this amplifier without aliasing?

According to Eq. (8.17b)

$$\mathcal{F}_h = \frac{1}{2T} = 20 \text{ kHz} \quad ■$$

Example 8.4

A sampler with sampling interval $T = 0.001$ second (1 ms.) samples continuous-time sinusoids of the following frequencies: (a) 400 Hz (b) 1 kHz (c) 1.4 kHz (d) 1.6kHz (e) 3.522 kHz. Determine the aliased frequencies of the resulting sampled signals.

The sampling frequency is $\mathcal{F}_s = 1/T = 1,000$. The folding frequency $\mathcal{F}_s/2 = 500$. Hence, sinusoids below 500 Hz will not be aliased and sinusoids of frequency above 500 Hz will be aliased. Using Eq. (8.21), we find:

(a) 400 Hz is less than 500 Hz (the folding frequency, which is half the sampling frequency $\mathcal{F}_s$). Hence, there is no aliasing.

(b) $1000 = 0 + 1000$ so that $\mathcal{F}_f = 0$ and the aliased frequency ($|\mathcal{F}_f|$) is zero. The sampled signal appears as samples of a dc signal.

(c)$1400 = 400 + 1000$ so that $\mathcal{F}_f = 400$ and the aliased frequency ($|\mathcal{F}_f|$) is 400 Hz. The sampled signal appears as samples of a signal of frequency 400 Hz.

(d) $1600 = -400 + 2(1000)$ so that $\mathcal{F}_f = -400$ and the aliased frequency ($|\mathcal{F}_f|$) is 400 Hz. The sampled signal appears as samples of a signal of frequency 400 Hz.

(e) $3522 = -478 + 4(1000)$ so that $\mathcal{F}_f = -478$ and the aliased frequency ($|\mathcal{F}_f|$) is 478 Hz. The sampled signal appears as samples of a signal of frequency 478 Hz.

Graphically, we can solve this problem using the artifice in Fig. 8.11. The folding frequency is 500 Hz instead of π. In case (a), the frequency 400 Hz is below the folding frequency 500 Hz. Hence, the samples of this sinusoid will not be aliased. For case (b), the frequency 1000 Hz, when folded back at 500 Hz terminates at the origin $\mathcal{F} = 0$. Hence, the aliased frequency is 0. For case (c), the frequency 1400 Hz folds back at 500 Hz, then folds forward at 0, and terminates at 400 Hz. Similarly, for case (d), the frequency 1600 Hz folds back at 500, then folds forward at 0, and folds back again at 500 Hz to terminate at 400 Hz, and so on. ■

8.4 Useful Signal Operations

Signal operations discussed for continuous-time systems also apply to discrete-time systems with some modification in time scaling. Since the independent variable in our signal description is time, the operations are called time shifting, time inversion (or time reversal), and time scaling. However, this discussion is valid for functions having independent variables other than time (e.g., frequency or distance).

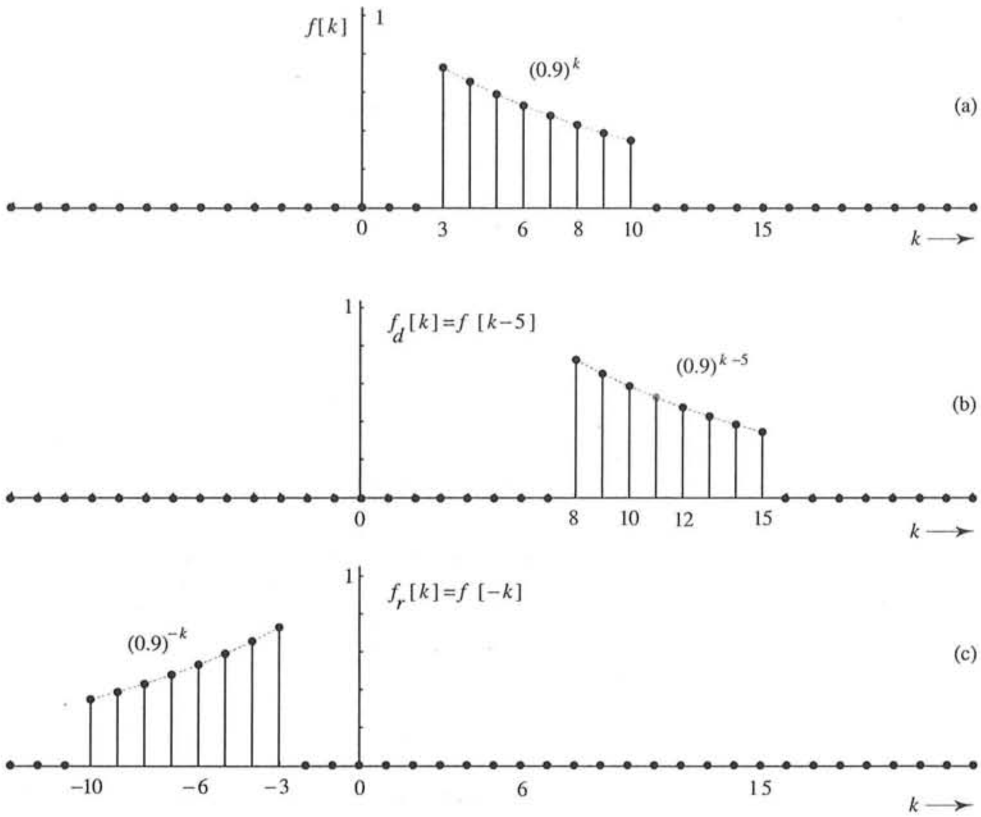

Fig. 8.16 Time-shifting and time inversion of a signal.

8.4-1 Time Shifting

Following the argument used for continuous-time signals, we can show that to time shift a signal $f[k]$ by m units, we replace k with $k-m$. Thus, $f[k-m]$ represents $f[k]$ time shifted by m units. If m is positive, the shift is to the right (delay). If m is negative, the shift is to the left (advance). Thus, $f[k-2]$ is $f[k]$ delayed (right-shifted) by 2 units, and $f[k+2]$ is $f[k]$ advanced (left-shifted) by 2 units. The signal $f_d[k]$ in Fig. 8.16b, being the signal in Fig. 8.16a delayed by 5 units, is the same as $f[k]$ with k replaced by $k-5$. Now, $f[k] = (0.9)^k$ for $3 \le k \le 10$. Therefore, $f_d[k] = (0.9)^{k-5}$ for $3 \le k-5 \le 10$ or $8 \le k \le 15$, as illustrated in Fig. 8.16b.

8.4-2 Time Inversion (or Reversal)

Following the argument used for continuous-time signals, we can show that to time invert a signal $f[k]$, we replace k with $-k$. This operation rotates the signal about the vertical axis. Figure 8.16c shows $f_r[k]$, which is the time-inverted signal $f[k]$ in Fig. 8.16a. The expression for $f_r[k]$ is the same as that for $f[k]$ with k replaced by $-k$. Because $f[k] = (0.9)^k$ for $3 \le k \le 10$, $f_r[k] = (0.9)^{-k}$ for $3 \le -k \le 10$; that is, $-3 \ge k \ge -10$, as shown in Fig. 8.16c.

8.4-3 Time Scaling

Following the argument used for continuous-time signals, we can show that to time scale a signal $f[k]$ by a factor a, we replace k with ak. However, because the discrete-time argument k can take only integral values, certain restrictions and changes in the procedure are necessary.

Time Compression: Decimation or Downsampling

Consider a signal

$$f_c[k] = f[2k] \tag{8.22}$$

The signal $f_c[k]$ is the signal $f[k]$ compressed by a factor 2. Observe that $f_c[0] = f[0]$, $f_c[1] = f[2]$, $f_c[2] = f[4]$, and so on. This fact shows that $f_c[k]$ is made up of even numbered samples of $f[k]$. The odd numbered samples of $f[k]$ are missing (Fig. 8.17b).† This operation loses part of the data, and that is why such time compression is called **decimation** or **downsampling**. In the continuous-time case, time compression merely speeds up the signal without loss of data. In general, $f[mk]$ (m integer) consists of only every mth sample of $f[k]$.

Time Expansion

Consider a signal

$$f_e[k] = f\left[\frac{k}{2}\right] \tag{8.23}$$

The signal $f_e[k]$ is the signal $f[k]$ expanded by a factor 2. According to Eq. (8.23), $f_e[0] = f[0]$, $f_e[1] = f[1/2]$, $f_e[2] = f[1]$, $f_e[3] = f[3/2]$, $f_e[4] = f[2]$, $f_e[5] = f[5/2]$, $f_e[6] = f[3]$, and so on. Now, $f[k]$ is defined only for integral values of k, and is zero (or undefined) for all fractional values of k. Therefore, for $f_e[k]$, its odd numbered samples $f_e[1]$, $f_e[3]$, $f_e[5]$, ... are all zero (or undefined), as depicted in Fig. 8.17c. In general, a function $f_e[k] = f[k/m]$ (m integer) is defined for $k = 0, \pm m, \pm 2m, \pm 3m, \ldots$, and is zero (or undefined) for all remaining values of k.

Interpolation

In the time-expanded signal in fig. 8.17c, the missing odd numbered samples can be reconstructed from the nonzero valued samples using some suitable interpolation formula. Figure 8.17d shows such an interpolated function $f_i[k]$, where the missing samples are constructed using an ideal lowpass filter interpolation formula (5.10b). In practice, we may use a realizable interpolation, such as a linear interpolation, where $f_i[1]$ is taken as the mean of $f_i[0]$ and $f_i[2]$. Similarly, $f_i[3]$ is taken as the mean of $f_i[2]$ and $f_i[4]$, and so on. This process of time expansion and inserting the missing samples using an interpolation is called **interpolation** or **upsampling**. In this operation, we increase the number of samples.

△ **Exercise E8.6**

Show that for a linearly interpolated function $f_i[k] = f[k/2]$, the odd numbered samples interpolated values are $f_i[k] = \frac{1}{2}\left\{f[\frac{k-1}{2}] + f[\frac{k+1}{2}]\right\}$. ▽

---

†Odd numbered samples of $f[k]$ can be retained (and even numbered samples omitted) by using the transform

$$f_c[k] = f[2k+1]$$

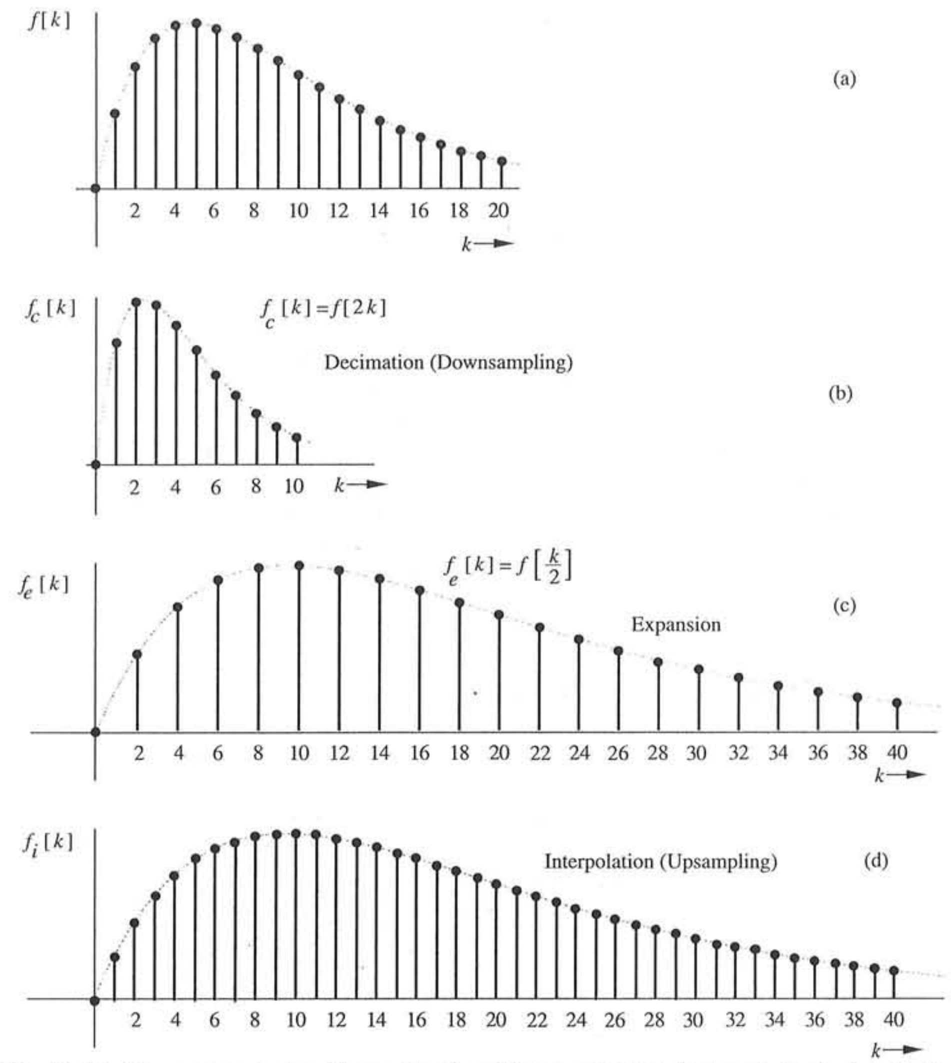

Fig. 8.17 Time compression (decimation) and time expansion (interpolation) of a signal.

8.5 Examples of Discrete-Time Systems

We shall give here three examples of discrete-time systems. In the first two examples, the signals are inherently discrete-time. In the third example, a continuous-time signal is processed by a discrete-time system, as illustrated in Fig. 8.2, by discretizing the signal through sampling.

Example 8.5

A person makes a deposit (the input) in a bank regularly at an interval of T (say, 1 month). The bank pays a certain interest on the account balance during the period T and mails out a periodic statement of the account balance (the output) to the depositor. Find the equation relating the output $y[k]$ (the balance) to the input $f[k]$ (the deposit).

In this case, the signals are inherently discrete-time. Let

$$
\begin{aligned}
f[k] &= \text{the deposit made at the } k\text{th discrete instant} \\
y[k] &= \text{the account balance at the } k\text{th instant computed} \\
&\quad\ \text{immediately after the } k\text{th deposit } f[k] \text{ is received} \\
r &= \text{interest per dollar per period } T
\end{aligned}
$$

The balance $y[k]$ is the sum of (i) the previous balance $y[k-1]$, (ii) the interest on $y[k-1]$ during the period T, and (iii) the deposit $f[k]$

$$
\begin{aligned}
y[k] &= y[k-1] + ry[k-1] + f[k] \\
&= (1+r)y[k-1] + f[k]
\end{aligned}
\tag{8.24}
$$

or

$$y[k] - ay[k-1] = f[k] \qquad a = 1 + r \tag{8.25a}$$

In this example the deposit $f[k]$ is the input (cause) and the balance $y[k]$ is the output (effect).

We can express Eq. (8.25a) in an alternate form. The choice of index k in Eq. (8.25a) is completely arbitrary, so we can substitute $k+1$ for k to obtain

$$y[k+1] - ay[k] = f[k+1] \tag{8.25b}$$

We also could have obtained Eq. (8.25b) directly by realizing that $y[k+1]$, the balance at the $(k+1)$st instant, is the sum of $y[k]$ plus $ry[k]$ (the interest on $y[k]$) plus the deposit (input) $f[k+1]$ at the $(k+1)$st instant.

For a hardware realization of such a system, we rewrite Eq. (8.25a) as

$$y[k] = ay[k-1] + f[k] \tag{8.25c}$$

Figure 8.18 shows the hardware realization of this equation using a single time delay of T units.† To understand this realization, assume that $y[k]$ is available. Delaying it by T, we generate $y[k-1]$. Next, we generate $y[k]$ from $f[k]$ and $y[k-1]$ according to Eq. (8.25c).

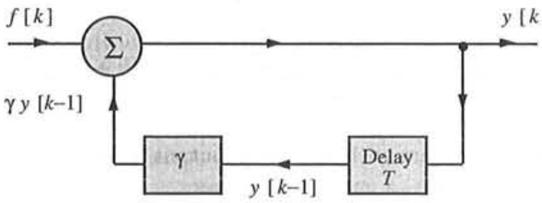

Fig. 8.18 Realization of the savings account system.

A withdrawal is a negative deposit. Therefore, this formulation can handle deposits as well as withdrawals. It also applies to a loan payment problem with the initial value $y[0] = -M$, where M is the amount of the loan. A loan is an initial deposit with a negative value. Alternately, we may treat a loan of M dollars taken at $k = 0$ as an input of $-M$ at $k = 0$ [see Prob. 9.4-9]. ■

†The time delay in Fig. 8.18 need not be T. The use of any other value will result in a time-scaled output.

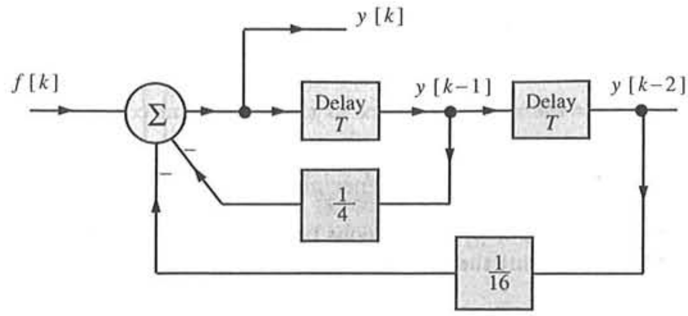

Fig. 8.19 Realization of a second-order discrete-time system in Example 8.6.

Example 8.6

In the kth semester, $f[k]$ number of students enroll in a course requiring a certain textbook. The publisher sells $y[k]$ new copies of the book in the kth semester. On the average, one quarter of students with books in saleable condition resell their books at the end of the semester, and the book life is three semesters. Write the equation relating $y[k]$, the new books sold by the publisher, to $f[k]$, the number of students enrolled in the kth semester, assuming that every student buys a book.

In the kth semester, the total books $f[k]$ sold to students must be equal to $y[k]$ (new books from the publisher) plus used books from students enrolled in the two previous semesters (because the book life is only three semesters). There are $y[k-1]$ new books sold in the $(k-1)$st semester, and one quarter of these books; that is, $\frac{1}{4}y[k-1]$ will be resold in the kth semester. Also, $y[k-2]$ new books are sold in the $(k-2)$nd semester, and one quarter of these; that is, $\frac{1}{4}y[k-2]$ will be resold in the $(k-1)$st semester. Again a quarter of these; that is, $\frac{1}{16}y[k-2]$ will be resold in the kth semester. Therefore, $f[k]$ must be equal to the sum of $y[k]$, $\frac{1}{4}y[k-1]$, and $\frac{1}{16}y[k-2]$.

$$y[k] + \tfrac{1}{4}y[k-1] + \tfrac{1}{16}y[k-2] = f[k] \tag{8.26a}$$

Equation (8.26a) can also be expressed in an alternative form by realizing that this equation is valid for any value of k. Therefore, replacing k by $k+2$, we obtain

$$y[k+2] + \tfrac{1}{4}y[k+1] + \tfrac{1}{16}y[k] = f[k+2] \tag{8.26b}$$

This is the alternative form of Eq. (8.26a).

For a realization of a system with this input-output equation, we rewrite Eq. (8.26a) as

$$y[k] = -\tfrac{1}{4}y[k-1] - \tfrac{1}{16}y[k-2] + f[k] \tag{8.26c}$$

Figure 8.19 shows a hardware realization of Eq. (8.26c) using two time delays (here the time delay T is a semester). To understand this realization, assume that $y[k]$ is available. Then, by delaying it successively, we generate $y[k-1]$ and $y[k-2]$. Next we generate $y[k]$ from $f[k]$, $y[k-1]$, and $y[k-2]$ according to Eq. (8.26c). ■

Equations (8.25) and (8.26) are examples of difference equations; the former is a first-order and the latter is a second-order difference equation. Difference equations also arise in numerical solution of differential equations.

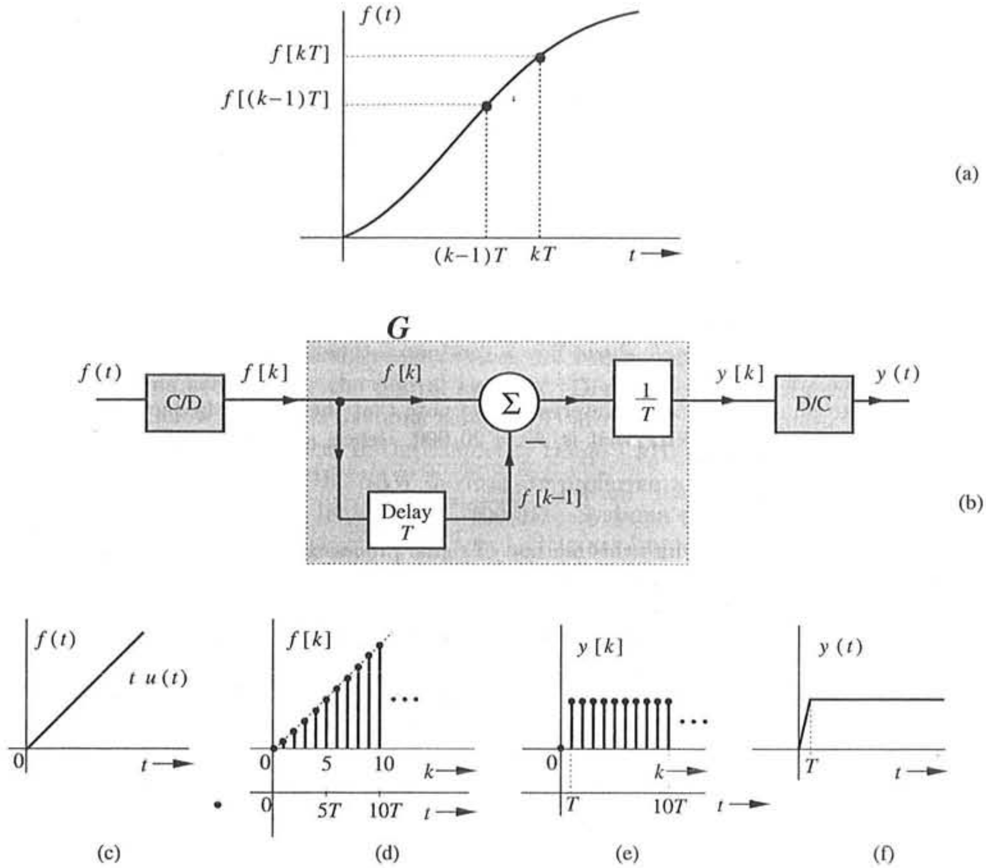

Fig. 8.20 Digital differentiator and its realization.

Example 8.7: Digital Differentiator

Design a discrete-time system, like the one in Fig. 8.2, to differentiate continuous-time signals. Determine the sampling interval if this differentiator is used in an audio system where the input signal bandwidth is below 20 kHz.

In this case, the output $y(t)$ is required to be the derivative of the input $f(t)$. The discrete-time processor (system) G processes the samples of $f(t)$ to produce the discrete-time output $y[k]$. Let $f[k]$ and $y[k]$ represent the samples T seconds apart of the signals $f(t)$ and $y(t)$, respectively; that is,

$$f[k] = f(kT) \qquad \text{and} \qquad y[k] = y(kT) \tag{8.27}$$

The signals $f[k]$ and $y[k]$ are the input and the output for the discrete-time system G. Now, we require that

$$y(t) = \frac{df}{dt} \tag{8.28}$$

Therefore, at $t = kT$ (see Fig. 8.20a)

$$\begin{aligned} y(kT) &= \left.\frac{df}{dt}\right|_{t=kT} \\ &= \lim_{T\to 0} \frac{1}{T}\left[f(kT) - f[(k-1)T]\right] \end{aligned}$$

By using the notation in Eq. (8.27), the above equation can be expressed as

$$y[k] = \lim_{T\to 0} \frac{1}{T}\{f[k] - f[k-1]\}$$

This is the input-output relationship for G required to achieve our objective. In practice, the sampling interval T cannot be zero. Assuming T to be sufficiently small, the above equation can be expressed as

$$y[k] \simeq \frac{1}{T}\{f[k] - f[k-1]\} \tag{8.29}$$

The approximation improves as T approaches 0. A discrete-time processor G to realize Eq. (8.29) is shown inside the shaded box in Fig. 8.20b. The system in Fig. 8.20b acts as a differentiator. This example shows how a continuous-time signal can be processed by a discrete-time system.

To determine the sampling interval T, we note that the highest frequency that will appear at the input is 20 kHz; that is, $\mathcal{F}_h = 20,000$. Hence, according to Eq. (8.17c)

$$T \leq \frac{1}{40,000} = 25\ \mu\text{s}$$

To gain some insight into this method of signal processing, let us consider the differentiator in Fig. 8.20b with a ramp input $f(t) = t$, depicted in Fig. 8.20c. If the system were to act as a differentiator, then the output $y(t)$ of the system should be the unit step function $u(t)$. Let us investigate how the system performs this particular operation and how well it achieves the objective.

The samples of the input $f(t) = t$ at the interval of T seconds act as the input to the discrete-time system G. These samples, denoted by a compact notation $f[k]$, are, therefore,

$$\begin{aligned} f[k] &= f(t)|_{t=kT} = t|_{t=kT} \qquad t \geq 0 \\ &= kT \qquad k \geq 0 \end{aligned}$$

Figure 8.20d shows the sampled signal $f[k]$. This signal acts as an input to the discrete-time system G. Figure 8.20b shows that the operation of G consists of subtracting a sample from the previous (delayed) sample and then multiplying the difference with $1/T$. From Fig. 8.20d, it is clear that the difference between the successive samples is a constant $kT - (k-1)T = T$ for all samples, except for the sample at $k = 0$ (because there is no previous sample at $k = 0$). The output of G is $1/T$ times the difference T, which is unity for all values of k, except $k = 0$, where it is zero. Therefore, the output $y[k]$ of G consists of samples of unit values for $k \geq 1$, as illustrated in Fig. 8.20e. The D/C (discrete-time to continuous-time) converter converts these samples into a continuous-time signal $y(t)$, as shown in Fig. 8.20f. Ideally, the output should have been $y(t) = u(t)$. This deviation from the ideal is caused by the fact that we have used a nonzero sampling interval T. As T approaches zero, the output $y(t)$ approaches the desired output $u(t)$. ∎

△ **Exercise E8.8**

Design a discrete-time system, such as in Fig. 8.2, to integrate continuous-time signals.

Hint: If $f(t)$ and $y(t)$ are the input and the output of an integrator, then $\frac{dy}{dt} = f(t)$. Approximation (similar to that in Example 8.7) of this equation at $t = kT$ yields $y[k] - y[k-1] = Tf[k]$. Show a realization of this system. ▽

Practical Realization of Discrete-Time Systems

These examples show that the basic elements required in the realization of discrete-time systems are time delays, scalar multipliers, and adders (summers).

We show in Chapter 11 that this is generally true of discrete-time systems. The discrete-time systems can be realized in two ways:

1. By using digital computers which readily perform the operations of adding, multiplying, and delaying. Minicomputers and microprocessors are well suited for this purpose, especially for signals with frequencies below 100 kHz.
2. By using special-purpose time-delay devices that have been developed in the last two decades. These include monolithic MOS charge-transfer devices (CTD) such as charge-coupled devices (CCD) and bucket brigade devices (BBD), which are implemented on silicon substrate as integrated circuit elements. In addition, there are surface acoustic wave (SAW) devices built on piezoelectric substrates. Systems using these devices are less expensive but are not as reliable or as accurate as the digital systems. Digital systems are preferable for signals below 100 kHz. Systems using CTD are suitable and competitive with those using SAW devices in the frequency range 1 kHz to 20 MHz. At frequencies higher than 20 MHz, SAW devices are preferred and are the only realistic choice for frequencies higher than 50 MHz. Systems using SAW devices with frequencies in the range of 10 MHz to 1 GHz are implemented routinely.[1]

There is a basic difference between continuous-time systems and analog systems. The same is true of discrete-time and digital systems. This is fully explained in Secs. 1.7-6 and 1.7-7.† For historical reasons, digital computers (rather than time-delay elements, such as CCD or SAW devices) were used in the realization of early discrete-time systems. Because of this fact, the terms *digital filters* and *discrete-time systems* are used synonymously in the literature. This distinction is irrelevant in the analysis of discrete-time systems. For this reason, in this book, the term *digital filters* implies *discrete-time systems*, and *analog filters* means *continuous-time systems*. Moreover, the terms C/D (continuous-to-discrete-time) and D/C will be used interchangeably with terms A/D (analog-to-digital) and D/A, respectively.

Advantages of Digital Signal Processing

1. Digital filters have a greater degree of precision and stability. They can be perfectly duplicated without having to worry about component value tolerances as in analog case.
2. Digital filters are more flexible. Their characteristics can be easily altered simply by changing the program.
3. A greater variety of filters can be realized by digital systems.
4. Very low frequency filters, if realized by continuous-time systems, require prohibitively bulky components. Such is not the case with digital filters.
5. Digital signals can be stored easily on magnetic tapes or disks without deterioration of signal quality.
6. More sophisticated signal processing algorithms can be used to process digital signals.
7. Digital filters can be time shared, and therefore can serve a number of inputs simultaneously.

†The terms *discrete-time* and *continuous-time* qualify the nature of a signal along the time axis (horizontal axis). The terms *analog* and *digital*, in contrast, qualify the nature of the signal amplitude (vertical axis).

8. Using integrated circuit technology, they can be fabricated in small packages requiring low power consumption.

   Some more advantages of using digital signals are listed in Sec. 5.1-3.

8.6 Summary

Signals specified only at discrete instants such as $t = 0, T, 2T, 3T, \ldots, kT$ are discrete-time signals. Basically, it is a sequence of numbers. Such a signal may be viewed as a function of time t, where the signal is defined or specified only at $t = kT$ with k any positive or negative integer. The signal therefore may be denoted as $f(kT)$. Alternately, such a signal may be viewed as a function of k, where k is any positive or negative integer. The latter approach results in a more compact notation such as $f[k]$, which is convenient and easier to manipulate. A system whose inputs and outputs are discrete-time signals is a discrete-time system.

In the study of continuous-time systems, exponentials with the natural base; that is, exponentials of the form $e^{\lambda t}$, where λ is complex in general, are more natural and convenient. In contrast, in the study of discrete-time systems, exponentials with a general base; that is, exponentials of the form γ^k, where γ is complex in general, are more convenient. One form of exponential can be readily converted to the other form by noting that $e^{\lambda k} = \gamma^k$, where $\gamma = e^\lambda$, or $\lambda = \ln \gamma$, and λ as well as γ are complex in general. The exponential γ^k grows exponentially with k if $|\gamma| > 1$ (γ outside the unit circle), and decays exponentially if $|\gamma| < 1$ (γ within the unit circle). If $|\gamma| = 1$; that is, if γ lies on the unit circle, the exponential is either a constant or oscillates with a constant amplitude.

Discrete-time sinusoids have two properties not shared by their continuous-time cousins. First, a discrete-time sinusoid $\cos \Omega k$ is periodic only if $\Omega/2\pi$ is a rational number. Second, discrete-time sinusoids whose frequencies Ω differ by an integral multiple of 2π are identical. Consequently, a discrete-time sinusoid of any frequency Ω is identical to some discrete-time sinusoid whose frequency lies in the interval $-\pi$ to π (called the fundamental frequency range). Further, because $\cos(-\Omega k + \theta) = \cos(\Omega k - \theta)$, a sinusoid of a frequency in the range from $-\pi$ to 0 can be expressed as a sinusoid of frequency in the range 0 to π. Thus, a discrete-time sinusoid of any frequency can be expressed as a sinusoid of frequency in the range 0 to π. Thus, in practice, a discrete-time sinusoid frequency is at most π. The highest rate of oscillation in a discrete-time sinusoid occurs when its frequency is π. In a given time, a sinusoid of frequency other than π will have a fewer number of cycles (or oscillations) than the sinusoid of frequency π. This peculiarity of nonuniqueness of waveforms in discrete-time sinusoids of different frequencies has a far reaching consequences in signal processing by discrete-time systems.

One useful measure of the size of a discrete-time signal is its energy defined by the sum $\sum_k |f[k]|^2$, if it is finite. If the signal energy is infinite, the proper measure is its power, if it exists. The signal power is the time average of its energy (averaged over the entire time interval from $k = -\infty$ to ∞). For periodic signals, the time averaging need be performed only over one period in view of the periodic repetition of the signal. Signal power is also equal to the mean squared value of the signal (averaged over the entire time interval from $k = -\infty$ to ∞).

Sampling a continuous-time sinusoid $\cos(\omega t + \theta)$ at uniform intervals of T seconds results in a discrete-time sinusoid $\cos(\Omega k + \theta)$, where $\Omega = \omega T$. A continuous time sinusoid of frequency $\mathcal{F}$ Hz must be sampled at a rate no less than $2\mathcal{F}$ Hz. Otherwise, the resulting sinusoid is aliased; that is, it appears as a sampled version of a sinusoid of lower frequency.

Discrete-time signals classification is identical to that of continuous-time signals, discussed in chapter 1.

A signal $f[k]$ delayed by m time units (right-shifted) is given by $f[k-m]$. On the other hand, $f[k]$ advanced (left-shifted) by m time units is given by $f[k+m]$. A signal $f[k]$, when time inverted, is given by $f[-k]$. These operations are the same as those for the continuous-time case. The case of time scaling, however, is somewhat different because of the discrete nature of variable k. Unlike the continuous-time case, where time compression results in the same data at a higher speed, time compression in the discrete-time case eliminates part of the data. Consequently, this operation is called *decimation* or *downsampling*. Time expansion operation of discrete-time signals results in time expanding the signal, thus creating zero-valued samples in between. We can reconstruct the zero-valued samples using interpolation from the nonzero samples. The interpolation, thus, creates additional samples in between using the interpolation process. For this reason, this operation is called *interpolation* or *upsampling*.

Discrete-time systems may be used to process discrete-time signals, or to process continuous-time signals using appropriate interfaces at the input and output. At the input, the continuous-time input signal is converted into a discrete-time signal through sampling. The resulting discrete-time signal is now processed by the discrete-time system yielding a discrete-time output. The output interface now converts the discrete-time output into a continuous-time output. Discrete-time systems are characterized by difference equations.

Discrete-time systems can be realized by using scalar multipliers, summers, and time delays. These operations can be readily performed by digital computers. Time delays also can be obtained from charge coupled devices (CCD), bucket brigade devices (BBD), and surface acoustic wave devices (SAW). Several advantages of discrete-time systems over continuous-time systems are discussed in Sec. 8.5. Because of these advantages, discrete-time systems are replacing continuous-time systems in several applications.

References

1. Milstein, L. B., and P.K. Das, "Surface Acoustic wave Devices," IEEE Communication Society Magazine, vol. 17, No. 5, pp. 25-33, September 1979.

Problems

8.2-1 The following signals are in the form $e^{\lambda k}$. Express them in the form γ^k: **(a)** $e^{-0.5k}$ **(b)** $e^{0.5k}$ **(c)** $e^{-j\pi k}$ **(d)** $e^{j\pi k}$. In each case show the locations of λ and γ in the complex plane. Verify that an exponential is growing if γ lies outside the unit circle (or if λ lies in the RHP), is decaying if γ lies within the unit circle (or if λ lies in the