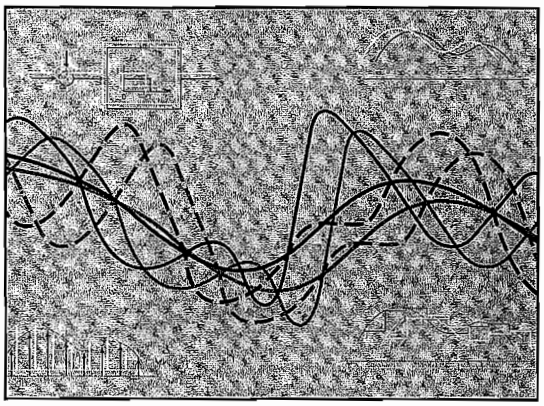
7

SAMPLING



7.0 INTRODUCTION

Under certain conditions, a continuous-time signal can be completely represented by and recoverable from knowledge of its values, or *samples,* at points equally spaced in time. This somewhat surprising property follows from a basic result that is referred to as the *sampling theorem.* This theorem is extremely important and useful. It is exploited, for example, in moving pictures, which consist of a sequence of individual frames, each of which represents an instantaneous view (i.e., a sample in time) of a continuously changing scene. When these samples are viewed in sequence at a sufficiently fast rate, we perceive an accurate representation of the original continuously moving scene. As another example, printed pictures typically consist of a very fine grid of points, each corresponding to a sample of the spatially continuous scene represented in the picture. If the samples are sufficiently close together, the picture appears to be spatially continuous, although under a magnifying glass its representation in terms of samples becomes evident.

Much of the importance of the sampling theorem also lies in its role as a bridge between continuous-time signals and discrete-time signals. As we will see in this chapter, the fact that under certain conditions a continuous-time signal can be completely recovered from a sequence of its samples provides a mechanism for representing a continuous-time signal by a discrete-time signal. In many contexts, processing discrete-time signals is more flexible and is often preferable to processing continuous-time signals. This is due in large part to the dramatic development of digital technology over the past few decades, resulting in the availability of inexpensive, lightweight, programmable, and easily reproducible discrete-time systems. The concept of sampling, then, suggests an extremely attractive and widely employed method for using discrete-time system technology to implement continuous-time systems and process continuous-time signals: We exploit sampling to

convert a continuous-time signal to a discrete-time signal, process the discrete-time signal using a discrete-time system, and then convert back to continuous time.

In the following discussion, we introduce and develop the concept of sampling and the process of reconstructing a continuous-time signal from its samples. In this discussion, we both identify the conditions under which a continuous-time signal can be exactly reconstructed from its samples and examine the consequences when these conditions are not satisfied. Following this, we explore the processing of continuous-time signals that have been converted to discrete-time signals through sampling. Finally, we examine the sampling of discrete-time signals and the related concepts of decimation and interpolation.

ESENTATION OF A CONTINUOUS-TIME SIGNAL BY ITS SAMPLES: SAMPLING THEOREM

In general, in the absence of any additional conditions or information, we would not expect that a signal could be uniquely specified by a sequence of equally spaced samples. For example, in Figure 7.1 we illustrate three different continuous-time signals, all of which have identical values at integer multiples of T; that is,

$$x_1(kT) = x_2(kT) = x_3(kT).$$

Clearly, an infinite number of signals can generate a given set of samples. As we will see, however, if a signal is band limited—i.e., if its Fourier transform is zero outside a finite band of frequencies—and if the samples are taken sufficiently close together in relation to the highest frequency present in the signal, then the samples *uniquely* specify the signal, and we can reconstruct it perfectly. This result, known as the *sampling theorem,* is of profound importance in the practical application of the methods of signal and system analysis.

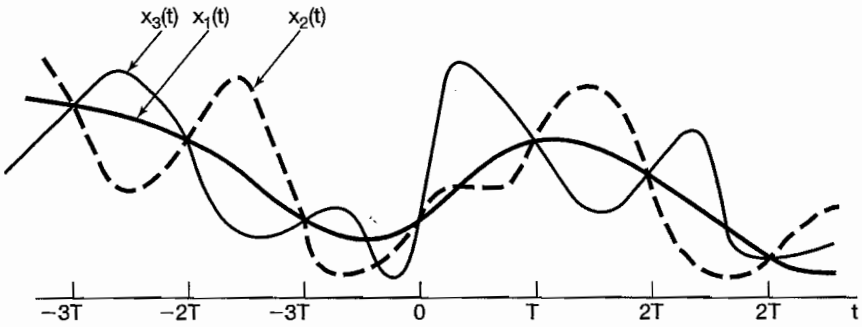


Figure 7.1 Three continuous-time signals with identical values at integer multiples of T.

7.1.1 Impulse-Train Sampling

In order to develop the sampling theorem, we need a convenient way in which to represent the sampling of a continuous-time signal at regular intervals. A useful way to do this is through the use of a periodic impulse train multiplied by the continuous-time signal $x(t)$ that we wish to sample. This mechanism, known as *impulse-train sampling,* is depicted in Figure 7.2. The periodic impulse train $p(t)$ is referred to as the *sampling function,* the period T as the *sampling period,* and the fundamental frequency of $p(t)$, $\omega_s = 2\pi/T$, as the *sampling frequency.* In the time domain,

$$x_p(t) = x(t)p(t), \tag{7.1}$$

where

$$p(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT). \tag{7.2}$$

Because of the sampling property of the unit impulse discussed in Section 1.4.2, we know that multiplying $x(t)$ by a unit impulse samples the value of the signal at the point at which the impulse is located; i.e., $x(t)\delta(t-t_0) = x(t_0)\delta(t-t_0)$. Applying this to eq. (7.1), we see, as illustrated in Figure 7.2, that $x_p(t)$ is an impulse train with the amplitudes of

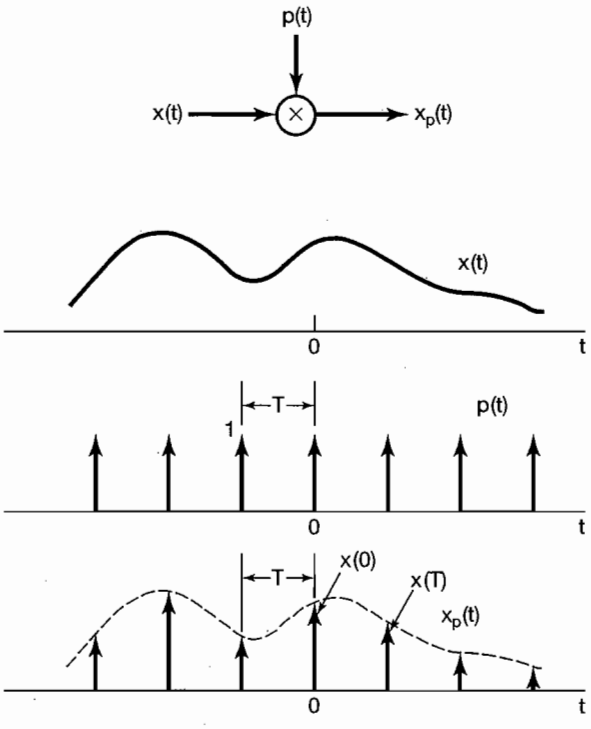


Figure 7.2 Impulse-train sampling.

the impulses equal to the samples of $x(t)$ at intervals spaced by T; that is,

$$x_p(t) = \sum_{n=-\infty}^{+\infty} x(nT)\delta(t - nT). \tag{7.3}$$

From the multiplication property (Section 4.5), we know that

$$X_p(j\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\theta)P(j(\omega - \theta))d\theta. \tag{7.4}$$

and from Example 4.8,

$$P(j\omega) = \frac{2\pi}{T}\sum_{k=-\infty}^{+\infty} \delta(\omega - k\omega_s). \tag{7.5}$$

Since convolution with an impulse simply shifts a signal [i.e., $X(j\omega) * \delta(\omega - \omega_0) = X(j(\omega - \omega_0))$], it follows that

$$X_p(j\omega) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X(j(\omega - k\omega_s)). \tag{7.6}$$

That is, $X_p(j\omega)$ is a periodic function of ω consisting of a superposition of shifted replicas of $X(j\omega)$, scaled by $1/T$, as illustrated in Figure 7.3. In Figure 7.3(c), $\omega_M < (\omega_s - \omega_M)$, or equivalently, $\omega_s > 2\omega_M$, and thus there is no overlap between the shifted replicas of $X(j\omega)$, whereas in Figure 7.3(d), with $\omega_s < 2\omega_M$, there is overlap. For the case illustrated in Figure 7.3(c), $X(j\omega)$ is faithfully reproduced at integer multiples of the sampling frequency. Consequently, if $\omega_s > 2\omega_M$, $x(t)$ can be recovered exactly from $x_p(t)$ by means of

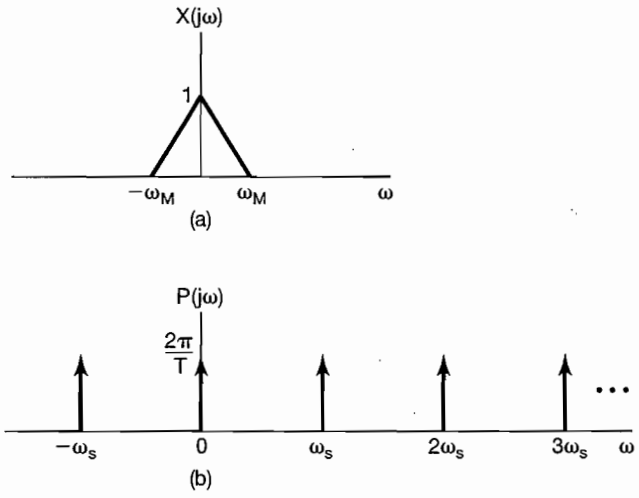


Figure 7.3 Effect in the frequency domain of sampling in the time domain: (a) spectrum of original signal; (b) spectrum of sampling function;

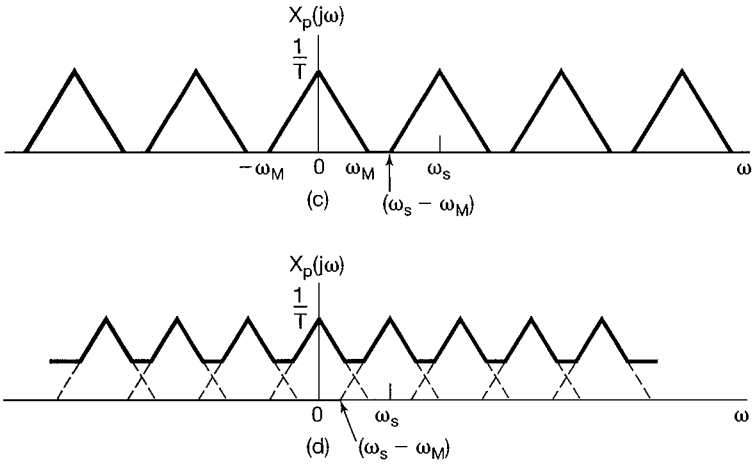


Figure 7.3 *Continued* (c) spectrum of sampled signal with $\omega_s > 2\omega_M$; (d) spectrum of sampled signal with $\omega_s < 2\omega_M$.

a lowpass filter with gain T and a cutoff frequency greater than ω_M and less than $\omega_s - \omega_M$, as indicated in Figure 7.4. This basic result, referred to as the *sampling theorem,* can be stated as follows: [1]

Sampling Theorem:
Let $x(t)$ be a band-limited signal with $X(j\omega) = 0$ for $|\omega| > \omega_M$. Then $x(t)$ is uniquely determined by its samples $x(nT)$, $n = 0, \pm 1, \pm 2, \ldots$, if

$$\omega_s > 2\omega_M,$$

where

$$\omega_s = \frac{2\pi}{T}.$$

Given these samples, we can reconstruct $x(t)$ by generating a periodic impulse train in which successive impulses have amplitudes that are successive sample values. This impulse train is then processed through an ideal lowpass filter with gain T and cutoff frequency greater than ω_M and less than $\omega_s - \omega_M$. The resulting output signal will exactly equal $x(t)$.

[1]The important and elegant sampling theorem was available for many years in a variety of forms in the mathematics literature. See, for example, J. M. Whittaker, "Interpolatory Function Theory," (New York: Stecher-Hafner Service Agency, 1964), chap. 4. It did not appear explicitly in the literature of communication theory until the publication in 1949 of the classic paper by Shannon entitled "Communication in the Presence of Noise" (*Proceedings of the IRE,* January 1949, pp. 10–21). However, H. Nyquist in 1928 and D. Gabor in 1946 had pointed out, based on the use of the Fourier Series, that $2TW$ numbers are sufficient to represent a function of duration T and highest frequency W. [H. Nyquist, "Certain Topics in Telegraph Transmission Theory," *AIEE Transactions,* 1928, p. 617; D. Gabor, "Theory of Communication,"*Journal of IEE* 93, no. 26 (1946), p. 429.]

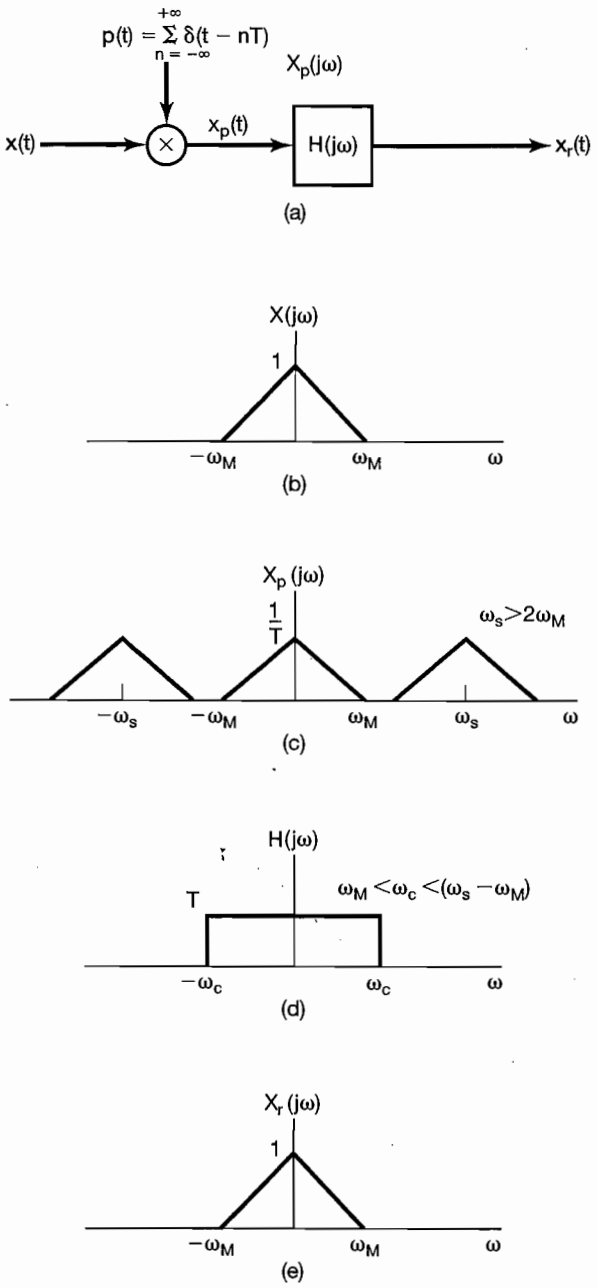


Figure 7.4 Exact recovery of a continuous-time signal from its samples using an ideal lowpass filter: (a) system for sampling and reconstruction; (b) representative spectrum for $x(t)$; (c) corresponding spectrum for $x_p(t)$; (d) ideal lowpass filter to recover $X(j\omega)$ from $X_p(j\omega)$; (e) spectrum of $x_r(t)$.

The frequency $2\omega_M$, which, under the sampling theorem, must be exceeded by the sampling frequency, is commonly referred to as the *Nyquist rate.*[2]

As discussed in Chapter 6, ideal filters are generally not used in practice for a variety of reasons. In any practical application, the ideal lowpass filter in Figure 7.4 would be

[2]The frequency ω_M corresponding to one-half the Nyquist rate is often referred to as the *Nyquist frequency.*

replaced by a nonideal filter $H(j\omega)$ that approximated the desired frequency characteristic accurately enough for the problem of interest (i.e., $H(j\omega) \simeq 1$ for $|\omega| < \omega_M$, and $H(j\omega) \simeq 0$ for $|\omega| > \omega_S - \omega_M$). Obviously, any such approximation in the lowpass filtering stage will lead to some discrepancy between $x(t)$ and $x_r(t)$ in Figure 7.4 or, equivalently, between $X(j\omega)$ and $X_r(j\omega)$. The particular choice of nonideal filter is then dictated by the acceptable level of distortion for the application under consideration. For convenience and to emphasize basic principles such as the sampling theorem, we will regularly assume the availability and use of ideal filters throughout this and the next chapter, with the understanding that in practice such a filter must be replaced by a nonideal filter designed to approximate the ideal characteristics accurately enough for the problem at hand.

7.1.2 Sampling with a Zero-Order Hold

The sampling theorem, which is most easily explained in terms of impulse-train sampling, establishes the fact that a band-limited signal is uniquely represented by its samples. In practice, however, narrow, large-amplitude pulses, which approximate impulses, are also relatively difficult to generate and transmit, and it is often more convenient to generate the sampled signal in a form referred to as a *zero-order hold.* Such a system samples $x(t)$ at a given instant and holds that value until the next instant at which a sample is taken, as illustrated in Figure 7.5. The reconstruction of $x(t)$ from the output of a zero-order hold can again be carried out by lowpass filtering. However, in this case, the required filter no longer has constant gain in the passband. To develop the required filter characteristic, we first note that the output $x_0(t)$ of the zero-order hold can in principle be generated by impulse-train sampling followed by an LTI system with a rectangular impulse response, as depicted in Figure 7.6. To reconstruct $x(t)$ from $x_0(t)$, we consider processing $x_0(t)$ with an LTI system with impulse response $h_r(t)$ and frequency response $H_r(j\omega)$. The cascade of this system with the system of Figure 7.6 is shown in Figure 7.7, where we wish to specify $H_r(j\omega)$ so that $r(t) = x(t)$. Comparing the system in Figure 7.7 with that in Figure 7.4, we see that $r(t) = x(t)$ if the cascade combination of $h_0(t)$ and $h_r(t)$ is the ideal lowpass filter $H(j\omega)$ used in Figure 7.4. Since, from Example 4.4 and the time-shifting property in Section 4.3.2,

$$H_0(j\omega) = e^{-j\omega T/2}\left[\frac{2\sin(\omega T/2)}{\omega}\right], \tag{7.7}$$

this requires that

$$H_r(j\omega) = \frac{e^{j\omega T/2}H(j\omega)}{\dfrac{2\sin(\omega T/2)}{\omega}}. \tag{7.8}$$

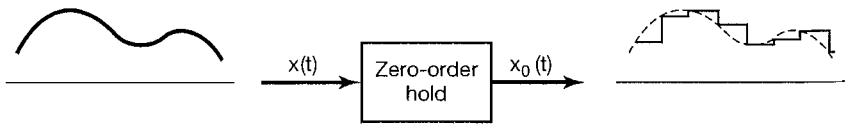


Figure 7.5 Sampling utilizing a zero-order hold.

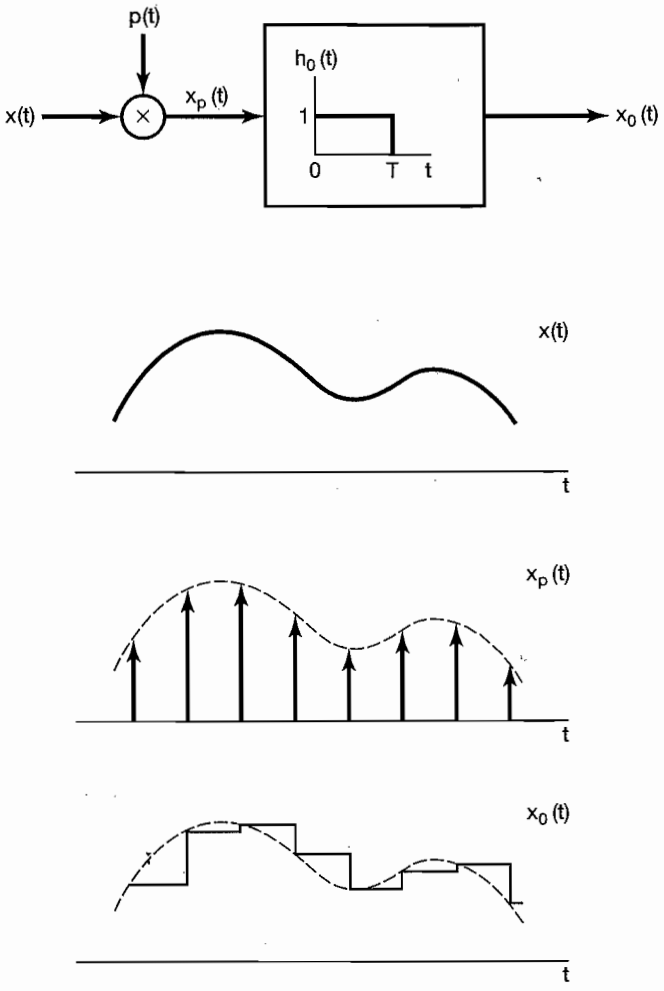


Figure 7.6 Zero-order hold as impulse-train sampling followed by an LTI system with a rectangular impulse response.

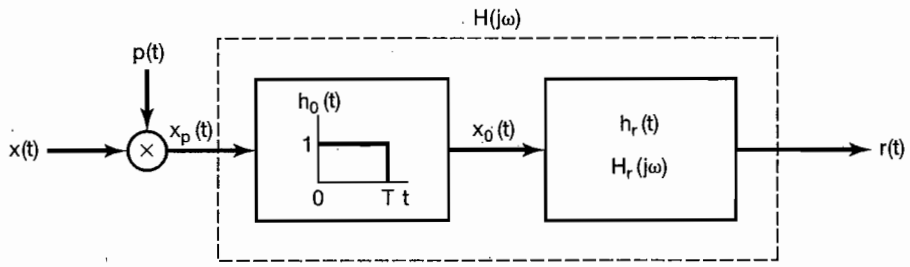


Figure 7.7 Cascade of the representation of a zero-order hold (Figure 7.6) with a reconstruction filter.

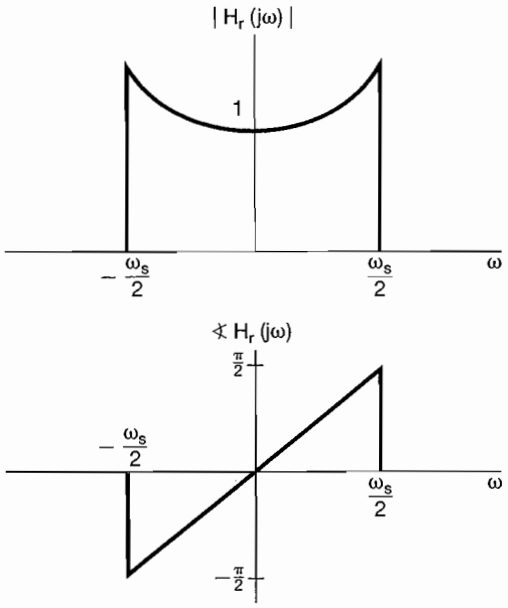


Figure 7.8 Magnitude and phase for the reconstruction filter for a zero-order hold.

For example, with the cutoff frequency of $H(j\omega)$ equal to $\omega_s/2$, the ideal magnitude and phase for the reconstruction filter following a zero-order hold is that shown in Figure 7.8.

Once again, in practice the frequency response in eq. (7.8) cannot be exactly realized, and thus an adequate approximation to it must be designed. In fact, in many situations, the output of the zero-order hold is considered an adequate approximation to the original signal by itself, without any additional lowpass filtering, and in essence represents a possible, although admittedly very coarse, interpolation between the sample values. Alternatively, in some applications, we may wish to perform some smoother interpolation between sample values. In the next section, we explore in more detail the general concept of interpreting the reconstruction of a signal from its samples as a process of interpolation.

7.2 RECONSTRUCTION OF A SIGNAL FROM ITS SAMPLES USING INTERPOLATION

Interpolation, that is, the fitting of a continuous signal to a set of sample values, is a commonly used procedure for reconstructing a function, either approximately or exactly, from samples. One simple interpolation procedure is the zero-order hold discussed in Section 7.1. Another useful form of interpolation is *linear interpolation,* whereby adjacent sample points are connected by a straight line, as illustrated in Figure 7.9. In more

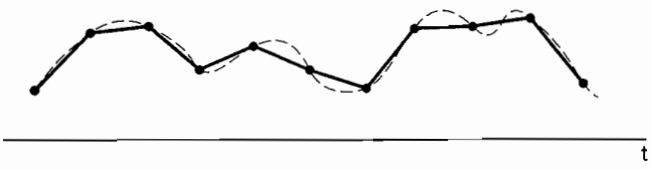


Figure 7.9 Linear interpolation between sample points. The dashed curve represents the original signal and the solid curve the linear interpolation.

complicated interpolation formulas, sample points may be connected by higher order polynomials or other mathematical functions.

As we have seen in Section 7.1, for a band-limited signal, if the sampling instants are sufficiently close, then the signal can be reconstructed exactly; i.e., through the use of a lowpass filter, exact interpolation can be carried out between the sample points. The interpretation of the reconstruction of $x(t)$ as a process of interpolation becomes evident when we consider the effect in the time domain of the lowpass filter in Figure 7.4. In particular, the output is

$$x_r(t) = x_p(t) * h(t)$$

or, with $x_p(t)$ given by eq. (7.3),

$$x_r(t) = \sum_{n=-\infty}^{+\infty} x(nT)h(t - nT). \tag{7.9}$$

Equation (7.9) describes how to fit a continuous curve between the sample points $x(nT)$ and consequently represents an interpolation formula. For the ideal lowpass filter $H(j\omega)$ in Figure 7.4,

$$h(t) = \frac{\omega_c T \sin(\omega_c t)}{\pi \omega_c t}, \tag{7.10}$$

so that

$$x_r(t) = \sum_{n=-\infty}^{+\infty} x(nT)\frac{\omega_c T}{\pi}\frac{\sin(\omega_c(t - nT))}{\omega_c(t - nT)}. \tag{7.11}$$

The reconstruction according to eq. (7.11) with $\omega_c = \omega_s/2$ is illustrated in Figure 7.10. Figure 7.10(a) represents the original band-limited signal $x(t)$, and Figure 7.10(b) represents $x_p(t)$, the impulse train of samples. In Figure 7.10(c), the superposition of the individual terms in eq. (7.11) is illustrated.

Interpolation using the impulse response of an ideal lowpass filter as in eq. (7.11) is commonly referred to as *band-limited interpolation,* since it implements exact reconstruction if $x(t)$ is band limited and the sampling frequency satisfies the conditions of the sampling theorem. As we have indicated, in many cases it is preferable to use a less accurate, but simpler, filter or, equivalently, a simpler interpolating function than the function in eq. (7.10). For example, the zero-order hold can be viewed as a form of interpolation between sample values in which the interpolating function $h(t)$ is the impulse response $h_0(t)$ depicted in Figure 7.6. In that sense, with $x_0(t)$ in the figure corresponding to the approximation to $x(t)$, the system $h_0(t)$ represents an approximation to the ideal lowpass filter required for the exact interpolation. Figure 7.11 shows the magnitude of the transfer function of the zero-order-hold interpolating filter, superimposed on the desired transfer function of the exact interpolating filter.

Both from Figure 7.11 and from Figure 7.6, we see that the zero-order hold is a very rough approximation, although in some cases it is sufficient. For example, if additional

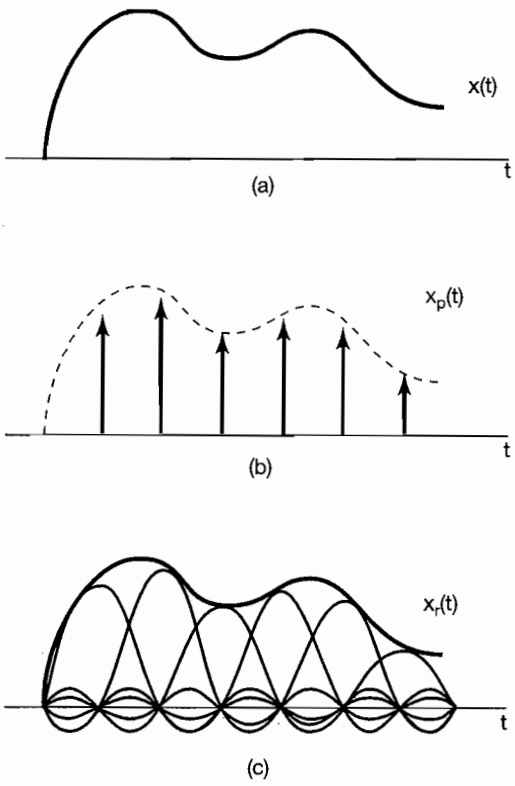


Figure 7.10 Ideal band-limited interpolation using the sinc function: (a) band-limited signal $x(t)$; (b) impulse train of samples of $x(t)$; (c) ideal band-limited interpolation in which the impulse train is replaced by a superposition of sinc functions [eq. (7.11)].

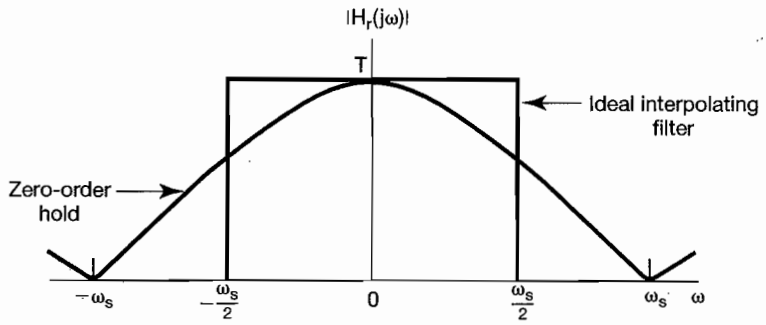


Figure 7.11 Transfer function for the zero-order hold and for the ideal interpolating filter.

lowpass filtering is naturally applied in a given application, it will tend to improve the overall interpolation. This is illustrated in the case of pictures in Figure 7.12. Figure 7.12(a) shows pictures with impulse sampling (i.e., sampling with spatially narrow pulses). Figure 7.12(b) is the result of applying a two-dimensional zero-order hold to Figure 7.12(a), with a resulting mosaic effect. However, the human visual system inherently imposes lowpass filtering, and consequently, when viewed at a distance, the discontinuities in the mosaic are smoothed. For example, in Figure 7.12(c) a

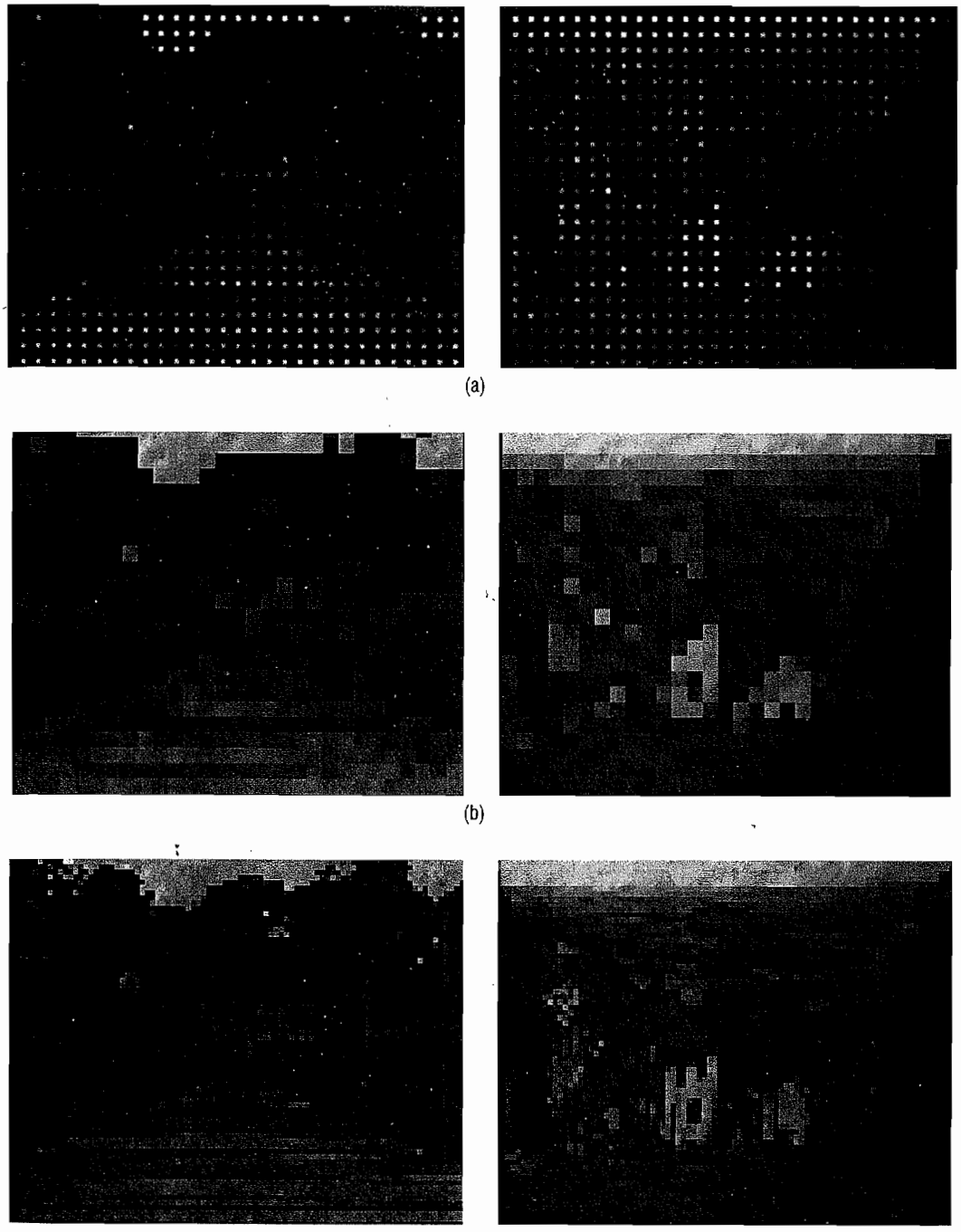


Figure 7.12 (a) The original pictures of Figures 6.2(a) and (g) with impulse sampling; (b) zero-order hold applied to the pictures in (a). The visual system naturally introduces lowpass filtering with a cutoff frequency that decreases with distance. Thus, when viewed at a distance, the discontinuities in the mosaic in Figure 7.12(b) are smoothed; (c) result of applying a zero-order hold after impulse sampling with one-fourth the horizontal and vertical spacing used in (a) and (b).

zero-order hold is again used, but here the sample spacing in each direction is one-fourth that in Figure 7.12(a). With normal viewing, considerable lowpass filtering is naturally applied, although the mosaic effect is still evident.

If the crude interpolation provided by the zero-order hold is insufficient, we can use a variety of smoother interpolation strategies, some of which are known collectively as *higher order holds.* In particular, the zero-order hold produces an output signal, as in Figure 7.5, that is discontinuous. In contrast, linear interpolation, as illustrated in Figure 7.9, yields reconstructions that are continuous, although with discontinous derivatives due to the changes in slope at the sample points. Linear interpolation, which is sometimes referred to as a first-order hold, can also be viewed as interpolation in the form of Figure 7.4 and eq. (7.9) with $h(t)$ triangular, as illustrated in Figure 7.13. The associated transfer function is also shown in the figure and is

$$H(j\omega) = \frac{1}{T}\left[\frac{\sin(\omega T/2)}{\omega/2}\right]^2. \tag{7.12}$$

The transfer function of the first-order hold is shown superimposed on the transfer function for the ideal interpolating filter. Figure 7.14 corresponds to the same pictures as those in Figure 7.12(b), but with a first-order hold applied to the sampled picture. In an analogous fashion, we can define second- and higher order holds that produce reconstructions with a higher degree of smoothness. For example, the output of a second-order hold provides an interpolation of the sample values that is continuous and has a continuous first derivative and discontinuous second derivative.

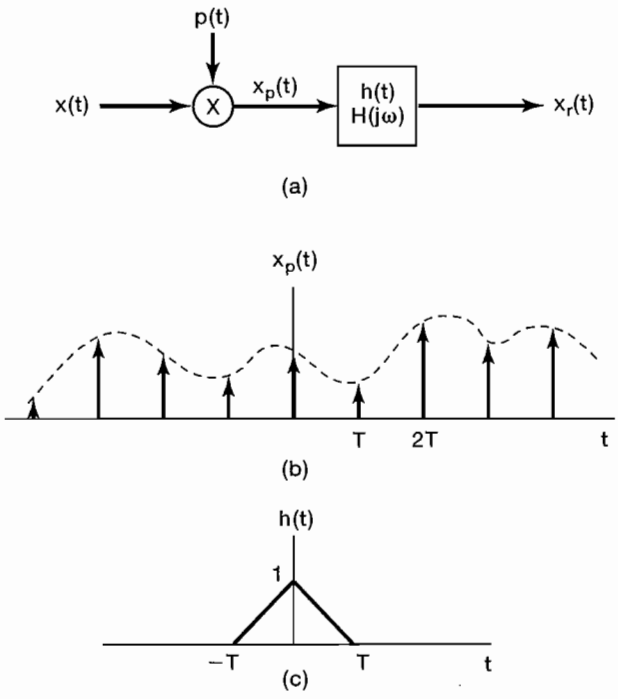


Figure 7.13 Linear interpolation (first-order hold) as impulse-train sampling followed by convolution with a triangular impulse response: (a) system for sampling and reconstruction; (b) impulse train of samples; (c) impulse response representing a first-order hold;

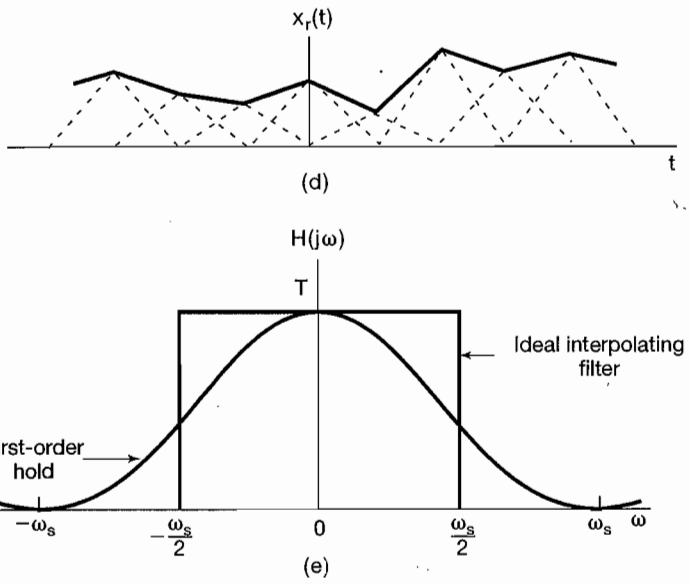


Figure 7.13 *Continued* (d) first-order hold applied to the sampled signal; (e) comparison of transfer function of ideal interpolating filter and first-order hold.

.3 THE EFFECT OF UNDERSAMPLING: ALIASING

In previous sections in this chapter, it was assumed that the sampling frequency was sufficiently high that the conditions of the sampling theorem were met. As illustrated in Figure 7.3, with $\omega_s > 2\omega_M$, the spectrum of the sampled signal consists of scaled replications of the spectrum of $x(t)$, and this forms the basis for the sampling theorem. When

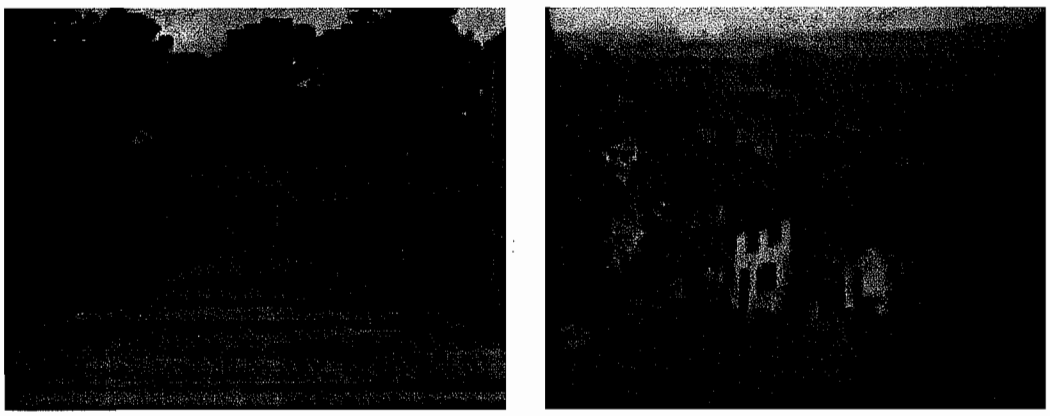


Figure 7.14 Result of applying a first-order hold rather than a zero-order hold after impulse sampling with one-third the horizontal and vertical spacing used in Figures 7.12(a) and (b).

$\omega_s < 2\omega_M$, $X(j\omega)$, the spectrum of $x(t)$, is no longer replicated in $X_p(j\omega)$ and thus is no longer recoverable by lowpass filtering. This effect, in which the individual terms in eq. (7.6) overlap, is referred to as *aliasing,* and in this section we explore its effect and consequences.

Clearly, if the system of Figure 7.4 is applied to a signal with $\omega_s < 2\omega_M$, the reconstructed signal $x_r(t)$ will no longer be equal to $x(t)$. However, as explored in Problem 7.25, the original signal and the signal $x_r(t)$ that is reconstructed using bandlimited interpolation will always be equal at the sampling instants; that is, for any choice of ω_s,

$$x_r(nT) = x(nT), \qquad n = 0, \pm 1, \pm 2, \ldots. \tag{7.13}$$

Some insight into the relationship between $x(t)$ and $x_r(t)$ when $\omega_s < 2\omega_M$ is provided by considering in more detail the comparatively simple case of a sinusoidal signal. Thus, let

$$x(t) = \cos \omega_0 t, \tag{7.14}$$

with Fourier transform $X(j\omega)$ as indicated in Figure 7.15(a). In this figure, we have graphically distinguished the impulse at ω_0 from that at $-\omega_0$ for convenience. Let us consider $X_p(j\omega)$, the spectrum of the sampled signal, and focus in particular on the effect of a change in the frequency ω_0 with the sampling frequency ω_s fixed. In Figures 7.15(b)–(e), we illustrate $X_p(j\omega)$ for several values of ω_0. Also indicated by a dashed line is the passband of the lowpass filter of Figure 7.4 with $\omega_c = \omega_s/2$. Note that no aliasing occurs in (b) and (c), since $\omega_0 < \omega_s/2$, whereas aliasing does occur in (d) and (e). For each of the four cases, the lowpass filtered output $x_r(t)$ is given as follows:

(a) $\omega_0 = \dfrac{\omega_s}{6}$; $\qquad x_r(t) = \cos \omega_0 t = x(t)$

(b) $\omega_0 = \dfrac{2\omega_s}{6}$; $\qquad x_r(t) = \cos \omega_0 t = x(t)$

(c) $\omega_0 = \dfrac{4\omega_s}{6}$; $\qquad x_r(t) = \cos(\omega_s - \omega_0)t \neq x(t)$

(d) $\omega_0 = \dfrac{5\omega_s}{6}$; $\qquad x_r(t) = \cos(\omega_s - \omega_0)t \neq x(t)$.

When aliasing occurs, the original frequency ω_0 takes on the identity of a lower frequency, $\omega_s - \omega_0$. For $\omega_s/2 < \omega_0 < \omega_s$, as ω_0 increases relative to ω_s, the output frequency $\omega_s - \omega_0$ decreases. When $\omega_s = \omega_0$, for example, the reconstructed signal is a constant. This is consistent with the fact that, when sampling once per cycle, the samples are all equal and would be identical to those obtained by sampling a constant signal ($\omega_0 = 0$). In Figure 7.16, we have depicted, for each of the four cases in Figure 7.15, the signal $x(t)$, its samples, and the reconstructed signal $x_r(t)$. From the figure, we can see how the lowpass

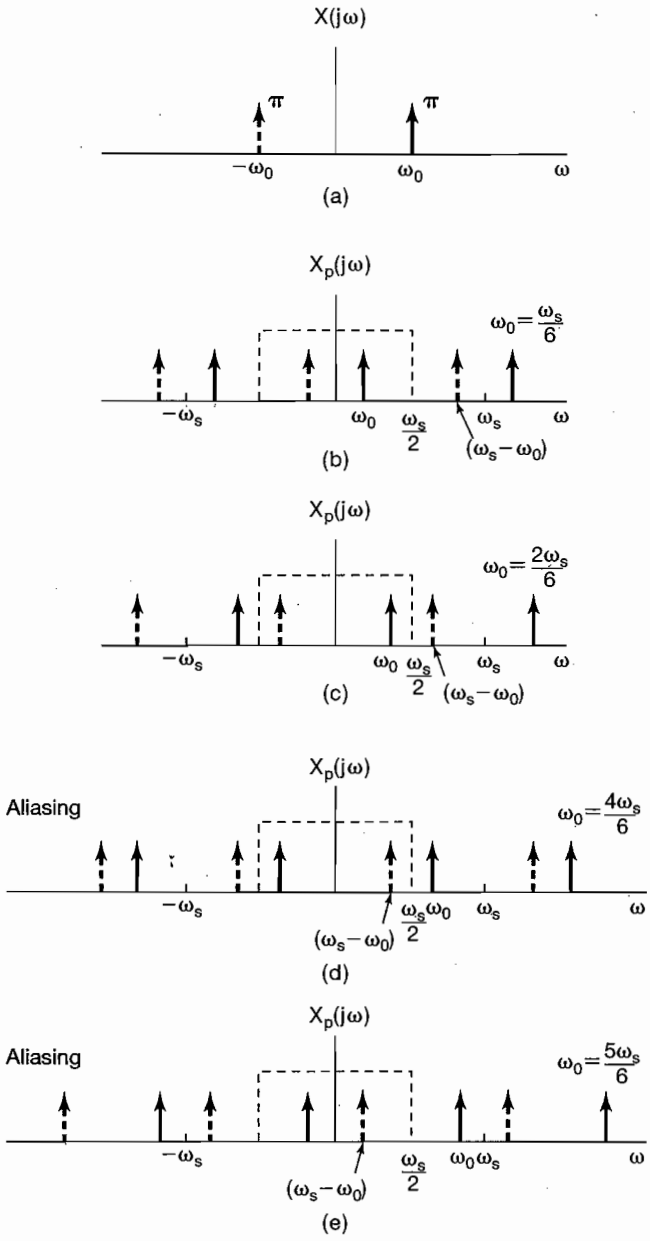


Figure 7.15 Effect in the frequency domain of oversampling and undersampling: (a) spectrum of original sinusoidal signal; (b), (c) spectrum of sampled signal with $\omega_s > 2\omega_0$; (d), (e) spectrum of sampled signal with $\omega_s < 2\omega_0$. As we increase ω_0 in moving from (b) through (d), the impulses drawn with solid lines move to the right, while the impulses drawn with dashed lines move to the left. In (d) and (e), these impulses have moved sufficiently that there is a change in the ones falling within the passband of the ideal lowpass filter.

filter interpolates between the samples, in particular always fitting a sinusoid of frequency less than $\omega_s/2$ to the samples of $x(t)$.

As a variation on the preceding examples, consider the signal

$$x(t) = \cos(\omega_0 t + \phi). \tag{7.15}$$

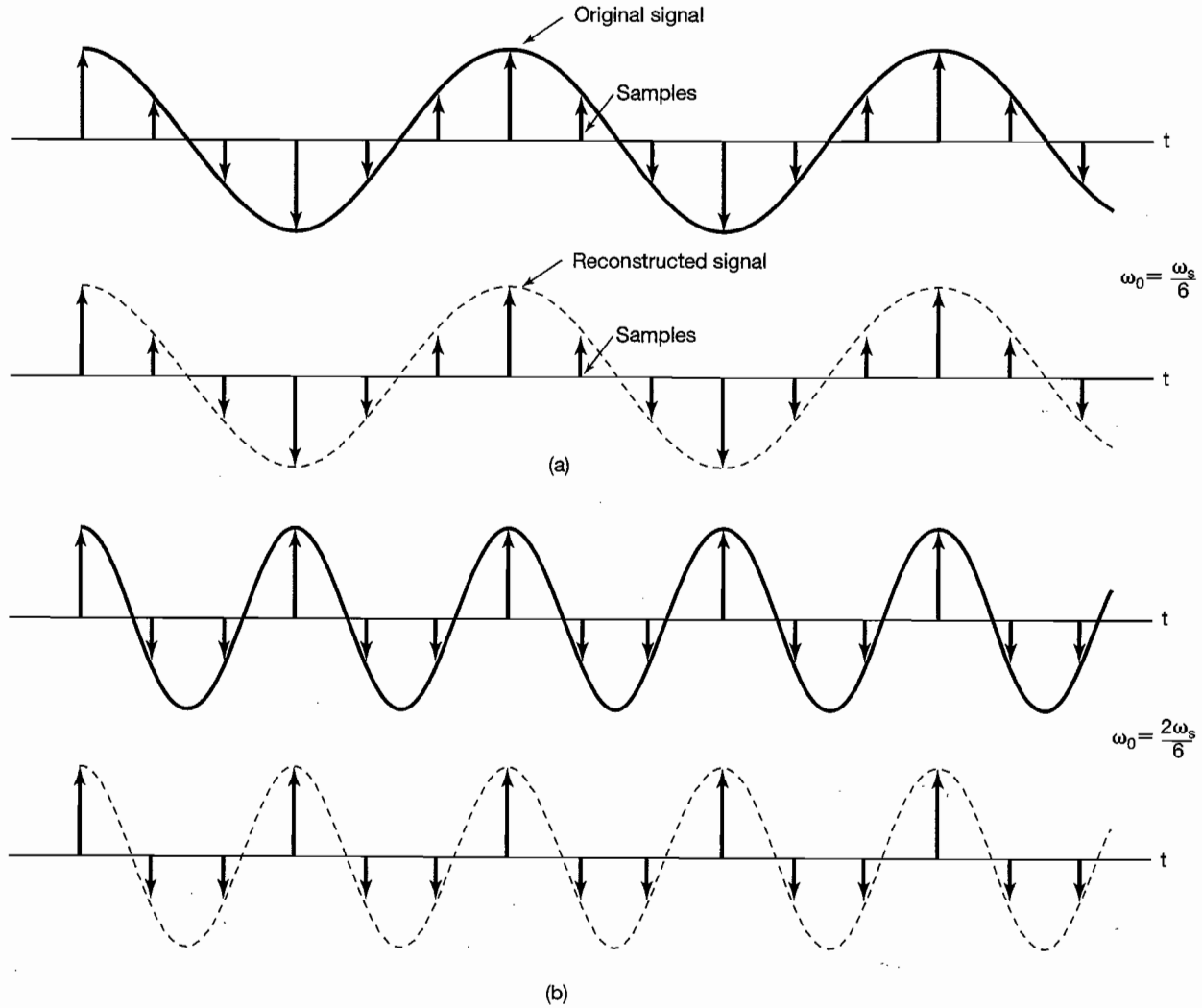


Figure 7.16 Effect of aliasing on a sinusoidal signal. For each of four values of ω_0, the original sinusoidal signal (solid curve), its samples, and the reconstructed signal (dashed curve) are illustrated: (a) $\omega_0 = \omega_s/6$; (b) $\omega_0 = 2\omega_s/6$; (c) $\omega_0 = 4\omega_s/6$; (d) $\omega_0 = 5\omega_s/6$. In (a) and (b) no aliasing occurs, whereas in (c) and (d) there is aliasing.

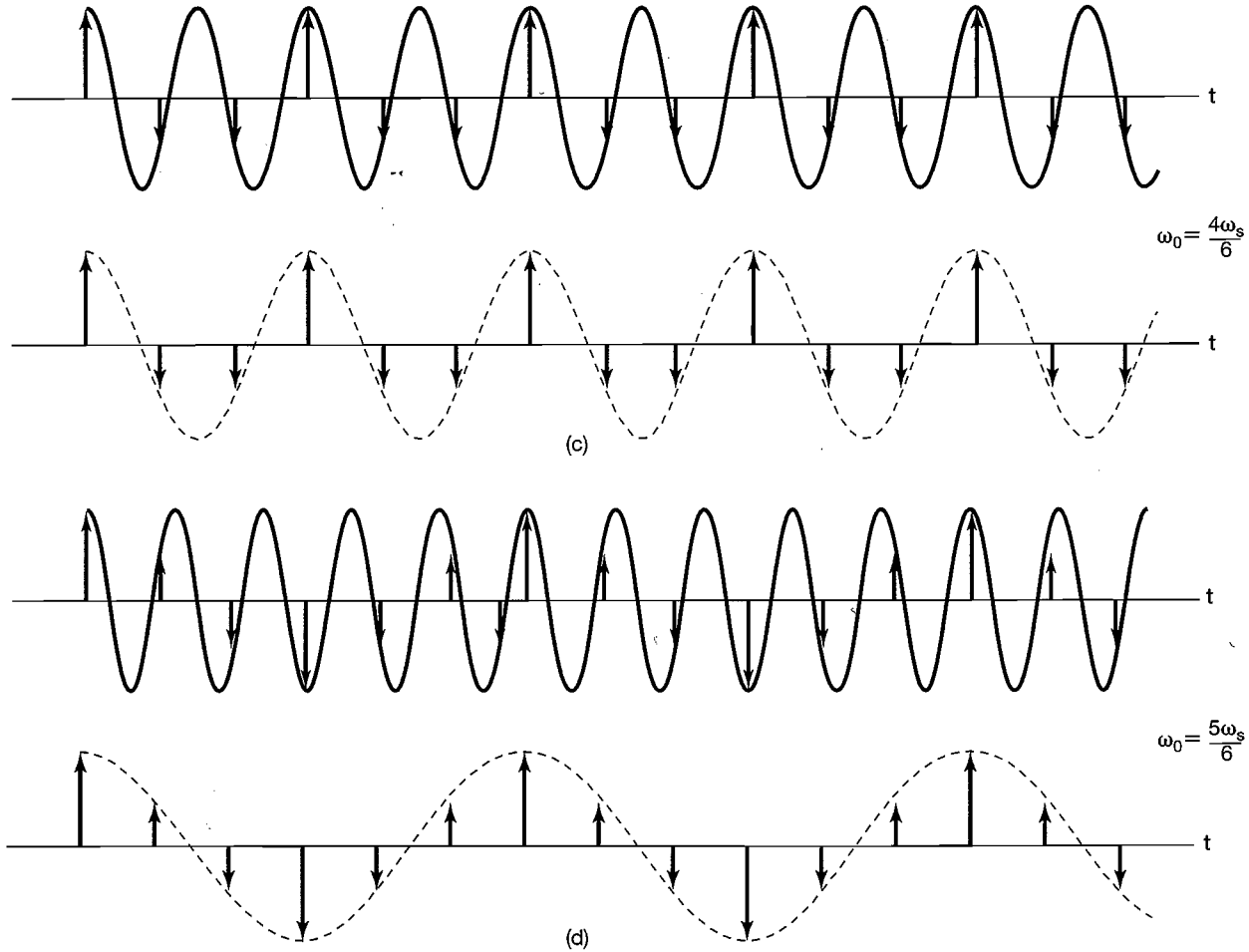


Figure 7.16 *Continued*

In this case, the Fourier transform of $x(t)$ is essentially the same as Figure 7.15(a), except that the impulse indicated with a solid line now has amplitude $\pi e^{j\phi}$, while the impulse indicated with a dashed line has amplitude with the opposite phase, namely, $\pi e^{-j\phi}$. If we now consider the same set of choices for ω_0 as in Figure 7.15, the resulting spectra for the sampled versions of $\cos(\omega_0 t + \phi)$ are exactly as in the figure, with all solid impulses having amplitude $\pi e^{j\phi}$ and all dashed ones having amplitude $\pi e^{-j\phi}$. Again, in cases (b) and (c) the condition of the sampling theorem is met, so that $x_r(t) = \cos(\omega_0 t + \phi) = x(t)$, while in (d) and (e) we again have aliasing. However, we now see that there has been a reversal in the solid and dashed impulses appearing in the passband of the lowpass filter. As a result, we find that in these cases, $x_r(t) = \cos[(\omega_s - \omega_0)t - \phi]$, where we have a change in the sign of the phase ϕ, i.e., a *phase reversal.*

It is important to note that the sampling theorem explicity requires that the sampling frequency be *greater than* twice the highest frequency in the signal, rather than greater than or equal to twice the highest frequency. The next example illustrates that sampling a sinusoidal signal at exactly twice its frequency (i.e., exactly two samples per cycle) is not sufficient.

Example 7.1

Consider the sinusoidal signal

$$x(t) = \cos\left(\frac{\omega_s}{2}t + \phi\right),$$

and suppose that this signal is sampled, using impulse sampling, at exactly twice the frequency of the sinusoid, i.e., at sampling frequency ω_s. As shown in Problem 7.39, if this impulse-sampled signal is applied as the input to an ideal lowpass filter with cutoff frequency $\omega_s/2$, the resulting output is

$$x_r(t) = (\cos\phi)\cos\left(\frac{\omega_s}{2}t\right).$$

As a consequence, we see that perfect reconstruction of $x(t)$ occurs only in the case in which the phase ϕ is zero (or an integer multiple of 2π). Otherwise, the signal $x_r(t)$ does not equal $x(t)$.

As an extreme example, consider the case in which $\phi = -\pi/2$, so that

$$x(t) = \sin\left(\frac{\omega_s}{2}t\right).$$

This signal is sketched in Figure 7.17. We observe that the values of the signal at integer multiples of the sampling period $2\pi/\omega_s$ are zero. Consequently, sampling at this rate produces a signal that is identically zero, and when this zero input is applied to the ideal lowpass filter, the resulting output $x_r(t)$ is also identically zero.

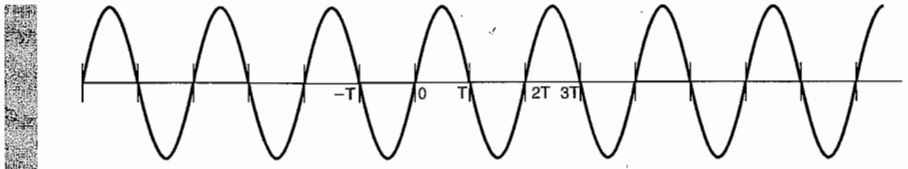


Figure 7.17 Sinusoidal signal for Example 7.1.

The effect of undersampling, whereby higher frequencies are reflected into lower frequencies, is the principle on which the stroboscopic effect is based. Consider, for example, the situation depicted in Figure 7.18, in which we have a disc rotating at a constant rate with a single radial line marked on the disc. The flashing strobe acts as a sampling system, since it illuminates the disc for extremely brief time intervals at a periodic rate. When the strobe frequency is much higher than the rotational speed of the disc, the speed of rotation of the disc is perceived correctly. When the strobe frequency becomes less than twice the rotational frequency of the disc, the rotation appears to be at a lower frequency than is actually the case. Furthermore, because of phase reversal, the disc will appear to be rotating in the wrong direction! Roughly speaking, if we track the position of a fixed line on the disc at successive samples, then when $\omega_0 < \omega_s < 2\omega_0$, so that we sample somewhat more frequently than once per revolution, samples of the disc will show the fixed line in positions that are successively displaced in a counterclockwise direction, opposite to the clockwise rotation of the disc itself. At one flash per revolution, corresponding to $\omega_s = \omega_0$, the radial line appears stationary (i.e., the rotational frequency of the disc and its harmonics have been aliased to zero frequency). A similar effect is commonly observed in Western movies,

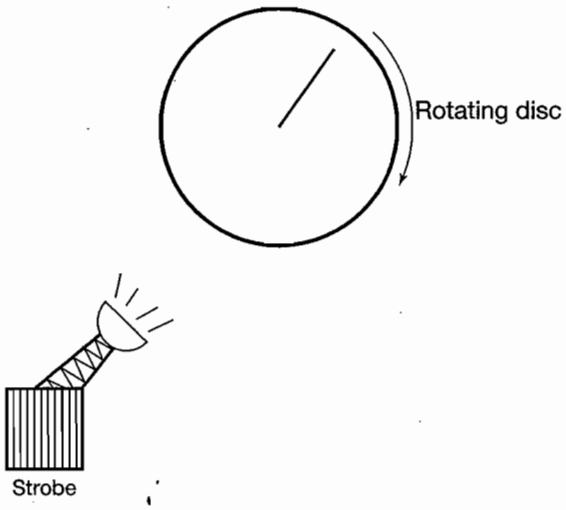


Figure 7.18 Strobe effect.

where the wheels of a stagecoach appear to be rotating more slowly than would be consistent with the coach's forward motion, and sometimes in the wrong direction. In this case, the sampling process corresponds to the fact that moving pictures are a sequence of individual frames with a rate (usually between 18 and 24 frames per second) corresponding to the sampling frequency.

The preceding discussion suggests interpreting the stroboscopic effect as an example of a useful application of aliasing due to undersampling. Another practical application of aliasing arises in a measuring instrument referred to as a *sampling oscilloscope*. This instrument is intended for observing very high-frequency waveforms and exploits the principles of sampling to alias these frequencies into ones that are more easily displayed. The sampling oscilloscope is explored in more detail in Problem 7.38.

7.4 DISCRETE-TIME PROCESSING OF CONTINUOUS-TIME SIGNALS

In many applications, there is a significant advantage offered in processing a continuous-time signal by first converting it to a discrete-time signal and, after discrete-time processing, converting back to a continuous-time signal. The discrete-time signal processing can be implemented with a general- or special-purpose computer, with microprocessors, or with any of the variety of devices that are specifically oriented toward discrete-time signal processing.

In broad terms, this approach to continuous-time signal processing can be viewed as the cascade of three operations, as indicated in Figure 7.19, where $x_c(t)$ and $y_c(t)$ are continuous-time signals and $x_d[n]$ and $y_d[n]$ are the discrete-time signals corresponding to $x_c(t)$ and $y_c(t)$. The overall system is, of course, a continuous-time system in the sense that its input and output are both continuous-time signals. The theoretical basis for converting a continuous-time signal to a discrete-time signal and reconstructing a continuous-time signal from its discrete-time representation lies in the sampling theorem, as discussed in Section 7.1. Through the process of periodic sampling with the sampling frequency consistent with the conditions of the sampling theorem, the continuous-time signal $x_c(t)$ is exactly represented by a sequence of instantaneous sample values $x_c(nT)$; that is, the discrete-time sequence $x_d[n]$ is related to $x_c(t)$ by

$$x_d[n] = x_c(nT). \tag{7.16}$$

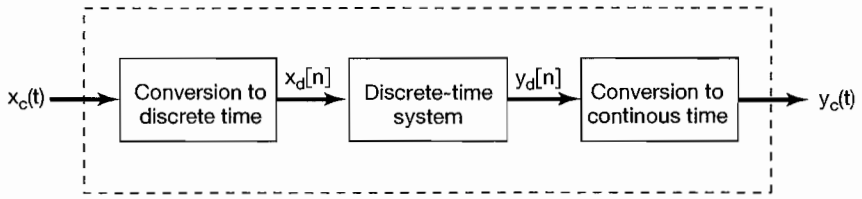


Figure 7.19 Discrete-time processing of continuous-time signals.

The transformation of $x_c(t)$ to $x_d[n]$ corresponding to the first system in Figure 7.19 will be referred to as *continuous-to-discrete-time conversion* and will be abbreviated C/D. The reverse operation corresponding to the third system in Figure 7.19 will be abbreviated D/C, representing *discrete-time to continuous-time conversion*. The D/C operation performs an interpolation between the sample values provided to it as input. That is, the D/C operation produces a continuous-time signal $y_c(t)$ which is related to the discrete-time signal $y_d[n]$ by

$$y_d[n] = y_c(nT).$$

This notation is made explicit in Figure 7.20. In systems such as digital computers and digital systems for which the discrete-time signal is represented in digital form, the device commonly used to implement the C/D conversion is referred to as an *analog-to-digital* (A-to-D) converter, and the device used to implement the D/C conversion is referred to as a *digital-to-analog* (D-to-A) converter.

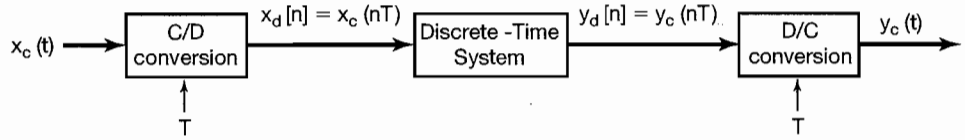


Figure 7.20 Notation for continuous-to-discrete-time conversion and discrete-to-continuous-time conversion. T represents the sampling period.

To understand further the relationship between the continuous-time signal $x_c(t)$ and its discrete-time representation $x_d[n]$, it is helpful to represent C/D as a process of periodic sampling followed by a mapping of the impulse train to a sequence. These two steps are illustrated in Figure 7.21. In the first step, representing the sampling process, the impulse train $x_p(t)$ corresponds to a sequence of impulses with amplitudes corresponding to the samples of $x_c(t)$ and with a time spacing equal to the sampling period T. In the conversion from the impulse train to the discrete-time sequence, we obtain $x_d[n]$, corresponding to the same sequence of samples of $x_c(t)$, but with unity spacing in terms of the new independent variable n. Thus, in effect, the conversion from the impulse train sequence of samples to the discrete-time sequence of samples can be thought of as a normalization in time. This normalization in converting $x_p(t)$ to $x_d[n]$ is evident in Figures 7.21(b) and (c), in which $x_p(t)$ and $x_d[n]$ are respectively illustrated for sampling rates of $T = T_1$ and $T = 2T_1$.

It is also instructive to examine the processing stages in Figure 7.19 in the frequency domain. Since we will be dealing with Fourier transforms in both continuous and discrete time, in this section only we distinguish the continuous-time and discrete-time frequency variables by using ω in continuous time and Ω in discrete time. For example, the continuous-time Fourier transforms of $x_c(t)$ and $y_c(t)$ are $X_c(j\omega)$ and $Y_c(j\omega)$, respectively, while the discrete-time Fourier transforms of $x_d[n]$ and $y_d[n]$ are $X_d(e^{j\Omega})$ and $Y_d(e^{j\Omega})$, respectively.

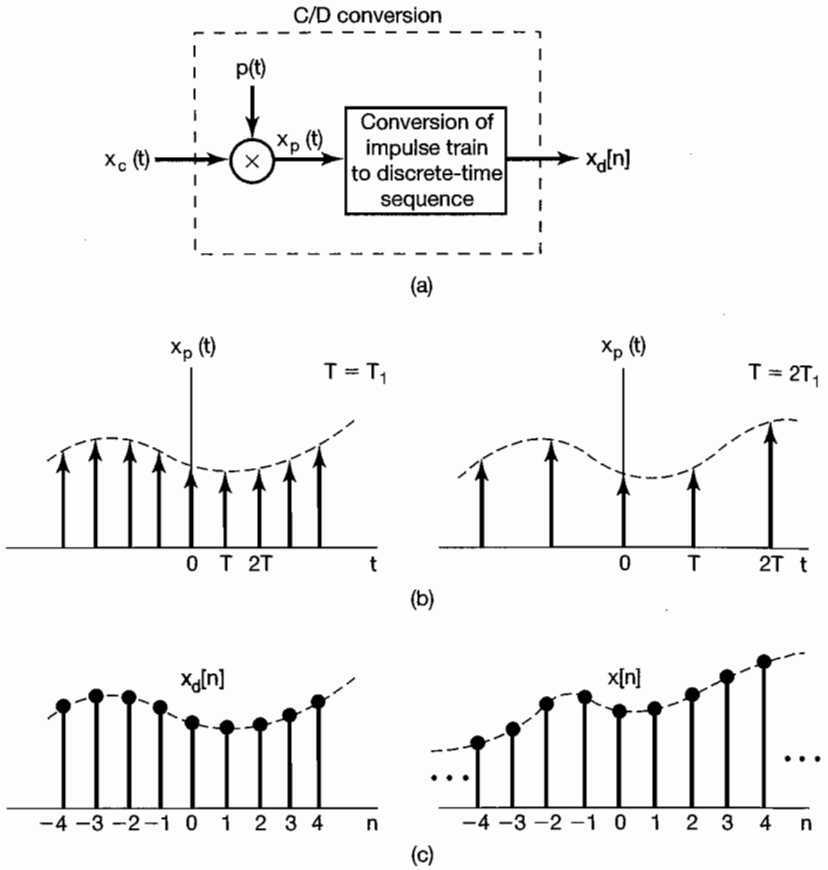


Figure 7.21 Sampling with a periodic impulse train followed by conversion to a discrete-time sequence: (a) overall system; (b) $x_p(t)$ for two sampling rates. The dashed envelope represents $x_c(t)$; (c) the output sequence for the two different sampling rates.

To begin let us express $X_p(j\omega)$, the continuous-time Fourier transform of $x_p(t)$, in terms of the sample values of $x_c(t)$ by applying the Fourier transform to eq. (7.3). Since

$$x_p(t) = \sum_{n=-\infty}^{+\infty} x_c(nT)\delta(t - nT), \tag{7.17}$$

and since the transform of $\delta(t - nT)$ is $e^{-j\omega nT}$, it follows that

$$X_p(j\omega) = \sum_{n=-\infty}^{+\infty} x_c(nT)e^{-j\omega nT} \tag{7.18}$$

Now consider the discrete-time Fourier transform of $x_d[n]$, that is,

$$X_d(e^{j\Omega}) = \sum_{n=-\infty}^{+\infty} x_d[n]e^{-j\Omega n}, \tag{7.19}$$

or, using eq. (7.16),

$$X_d(e^{j\Omega}) = \sum_{n=-\infty}^{+\infty} x_c(nT)e^{-j\Omega n}. \tag{7.20}$$

Comparing eqs. (7.18) and (7.20), we see that $X_d(e^{j\Omega})$ and $X_p(j\omega)$ are related through

$$X_d(e^{j\Omega}) = X_p(j\Omega/T). \tag{7.21}$$

Also, recall that, as developed in eq. (7.6) and illustrated in Figure 7.3,

$$X_p(j\omega) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X_c(j(\omega - k\omega_s)). \tag{7.22}$$

Consequently,

$$X_d(e^{j\Omega}) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} X_c(j(\Omega - 2\pi k)/T). \tag{7.23}$$

The relationship among $X_c(j\omega)$, $X_p(j\omega)$, and $X_d(e^{j\Omega})$ is illustrated in Figure 7.22 for two different sampling rates. Here, $X_d(e^{j\Omega})$ is a frequency-scaled version of $X_p(j\omega)$

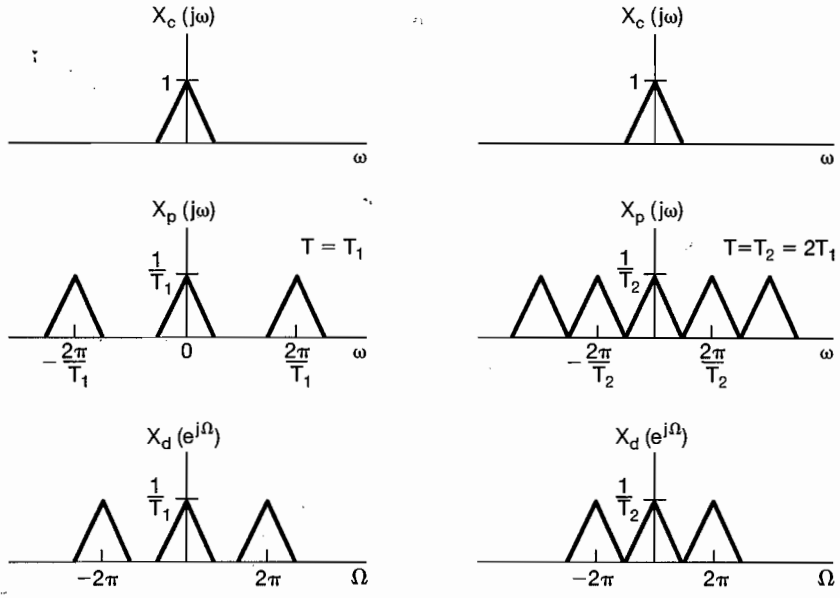


Figure 7.22 Relationship between $X_c(j\omega)$, $X_p(j\omega)$, and $X_d(e^{j\Omega})$ for two different sampling rates.

and, in particular, is periodic in Ω with period 2π. This periodicity is, of course, characteristic of any discrete-time Fourier transform. The spectrum of $x_d[n]$ is related to that of $x_c(t)$ through periodic replication, represented by eq. (7.22), followed by linear frequency scaling, represented by eq. (7.21). The periodic replication is a consequence of the first step in the conversion process in Figure 7.21, namely, the impulse-train sampling. The linear frequency scaling in eq. (7.21) can be thought of informally as a consequence of the normalization in time introduced by converting from the impulse train $x_p(t)$ to the discrete-time sequence $x_d[n]$. From the time-scaling property of the Fourier transform in Section 4.3.5, scaling of the time axis by $1/T$ will introduce a scaling of the frequency axis by T. Thus, the relationship $\Omega = \omega T$ is consistent with the notion that, in converting from $x_p(t)$ to $x_d[n]$, the time axis is scaled by $1/T$.

In the overall system of Figure 7.19, after processing with a discrete-time system, the resulting sequence is converted back to a continuous-time signal. This process is the reverse of the steps in Figure 7.21. Specifically, from the sequence $y_d[n]$, a continuous-time impulse train $y_p(t)$ can be generated. Recovery of the continuous-time signal $y_c(t)$ from this impulse train is then accomplished by means of lowpass filtering, as illustrated in Figure 7.23.

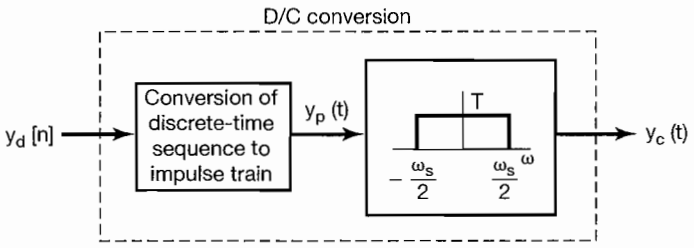


Figure 7.23 Conversion of a discrete-time sequence to a continuous-time signal.

Now let us consider the overall system of Figure 7.19, represented as shown in Figure 7.24. Clearly, if the discrete-time system is an identity system (i.e., $x_d[n] = y_d[n]$), then, assuming that the conditions of the sampling theorem are met, the overall system will be an identity system. The characteristics of the overall system with a more general frequency response $H_d(e^{j\Omega})$ are perhaps best understood by examining the representative example depicted in Figure 7.25. On the left-hand side of the figure are the representative

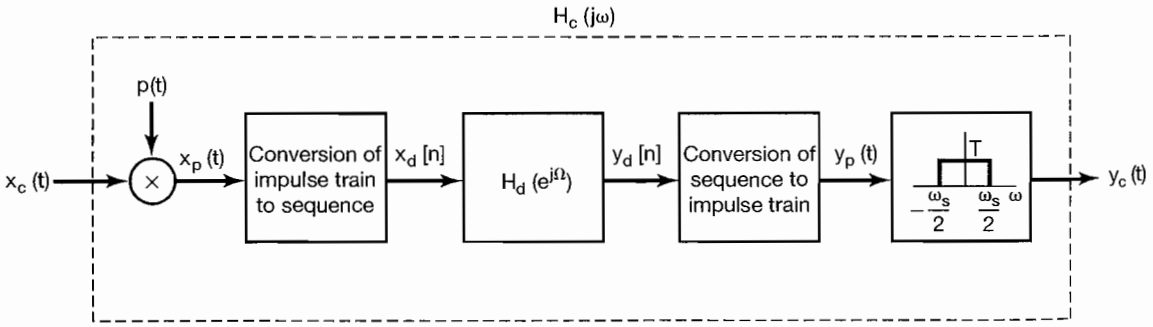


Figure 7.24 Overall system for filtering a continuous-time signal using a discrete-time filter.

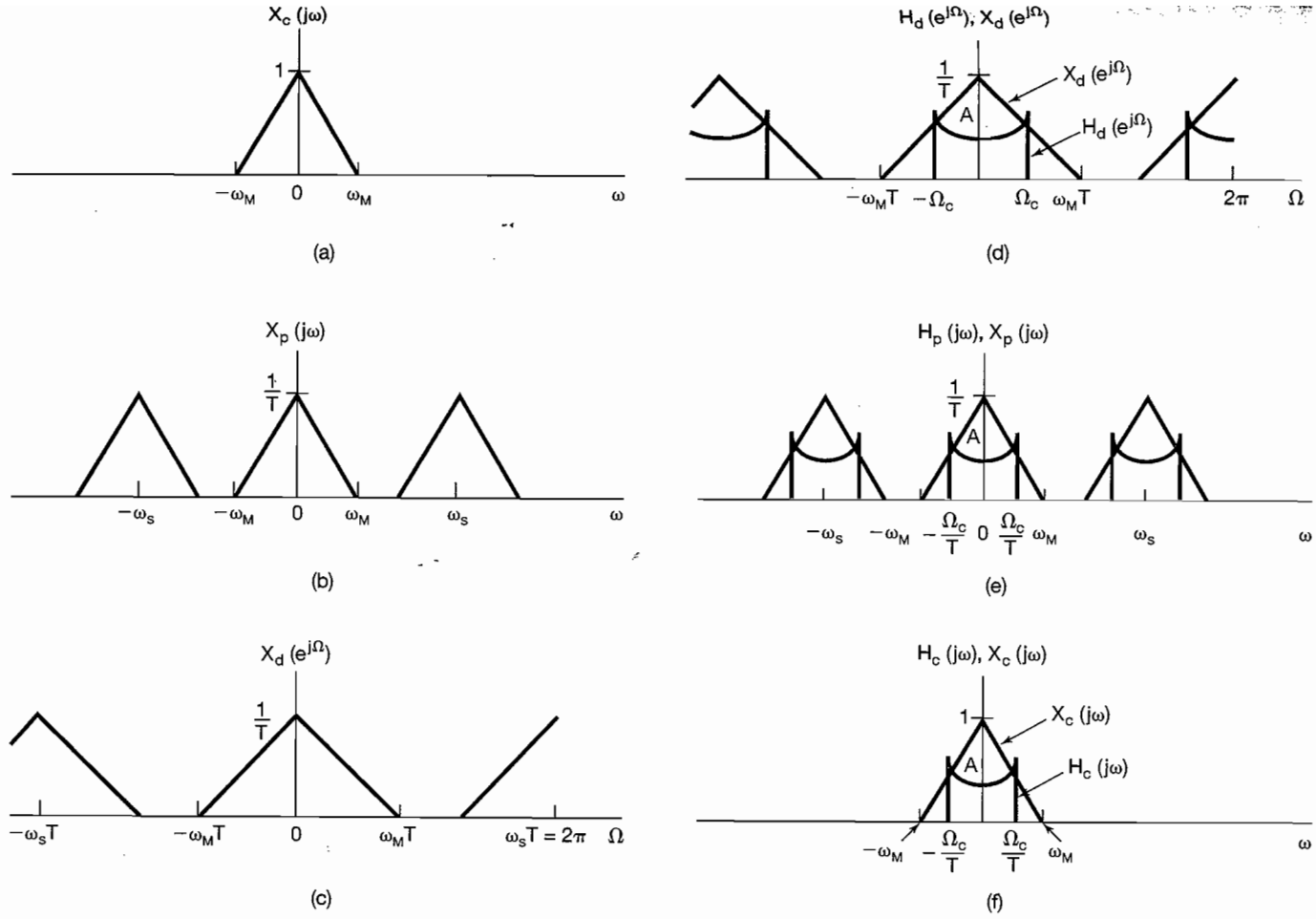


Figure 7.25 Frequency-domain illustration of the system of Figure 7.24: (a) continuous-time spectrum $X_c(j\omega)$; (b) spectrum after impulse-train sampling; (c) spectrum of discrete-time sequence $x_d[n]$; (d) $H_d(e^{j\Omega})$ and $X_d(e^{j\Omega})$ that are multiplied to form $Y_d(e^{j\Omega})$; (e) spectra that are multiplied to form $Y_p(j\omega)$; (f) spectra that are multiplied to form $Y_c(j\omega)$.

spectra $X_c(j\omega)$, $X_p(j\omega)$, and $X_d(e^{j\Omega})$, where we assume that $\omega_M < \omega_s/2$, so that there is no aliasing. The spectrum $Y_d(e^{j\Omega})$ corresponding to the output of the discrete-time filter is the product of $X_d(e^{j\Omega})$ and $H_d(e^{j\Omega})$, and this is depicted in Figure 7.25(d) by overlaying $H_d(e^{j\Omega})$ and $X_d(e^{j\Omega})$. The transformation to $Y_c(j\omega)$ then corresponds to applying a frequency scaling and lowpass filtering, resulting in the spectra indicated in Figure 7.25(e) and (f). Since $Y_d(e^{j\Omega})$ is the product of the two overlaid spectra in Figure 7.25(d), the frequency scaling and lowpass filtering are applied to both. In comparing Figures 7.25(a) and (f), we see that

$$Y_c(j\omega) = X_c(j\omega)H_d(e^{j\omega T}). \tag{7.24}$$

Consequently, for inputs that are sufficiently band limited, so that the sampling theorem is satisfied, the overall system of Figure 7.24 is, in fact, equivalent to a continuous-time LTI system with frequency response $H_c(j\omega)$ which is related to the discrete-time frequency response $H_d(e^{j\Omega})$ through

$$H_c(j\omega) = \begin{cases} H_d(e^{j\omega T}), & |\omega| < \omega_s/2 \\ 0, & |\omega| > \omega_s/2 \end{cases}. \tag{7.25}$$

The equivalent frequency response for this continuous-time filter is one period of the frequency response of the discrete-time filter with a linear scale change applied to the frequency axis. This relationship between the discrete-time frequency response and the equivalent continuous-time frequency response is illustrated in Figure 7.26.

The equivalence of the overall system of Figure 7.24 to an LTI system is somewhat surprising in view of the fact that multiplication by an impulse train is *not* a time-invariant operation. In fact, the overall system of Figure 7.24 is not time invariant for arbitrary inputs. For example, if $x_c(t)$ was a narrow rectangular pulse of duration less than T, then a time shift of $x_c(t)$ could generate a sequence $x[n]$ that either had all zero values or had one nonzero value, depending on the alignment of the rectangular pulse relative to the

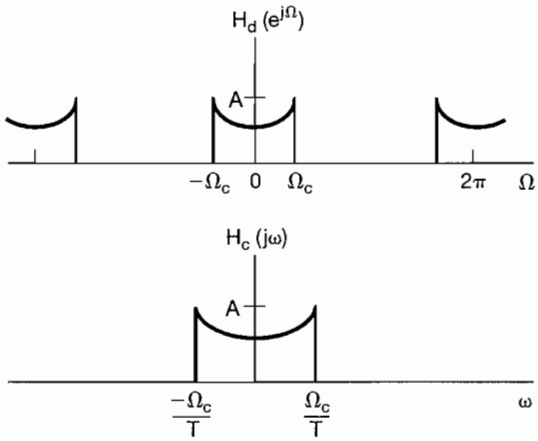


Figure 7.26 Discrete-time frequency response and the equivalent continuous-time frequency response for the system of Figure 7.24.

sampling impulse train. However, as suggested by the spectra of Figure 7.25, for *band-limited input signals* with a sampling rate sufficiently high so as to avoid aliasing, the system of Figure 7.24 *is* equivalent to a continuous-time LTI system. For such inputs, Figure 7.24 and eq. (7.25) provide the conceptual basis for continuous-time processing using discrete-time filters. This is now explored further in the context of some examples.

7.4.1 Digital Differentiator

Consider the discrete-time implementation of a continuous-time band-limited differentiating filter. As discussed in Section 3.9.1, the frequency response of a continuous-time differentiating filter is

$$H_c(j\omega) = j\omega, \tag{7.26}$$

and that of a band-limited differentiator with cutoff frequency ω_c is

$$H_c(j\omega) = \begin{cases} j\omega, & |\omega| < \omega_c \\ 0, & |\omega| > \omega_c \end{cases}, \tag{7.27}$$

as sketched in Figure 7.27. Using eq. (7.25) with a sampling frequency $\omega_s = 2\omega_c$, we see that the corresponding discrete-time transfer function is

$$H_d(e^{j\Omega}) = j\left(\frac{\Omega}{T}\right), \quad |\Omega| < \pi, \tag{7.28}$$

as sketched in Figure 7.28. With this discrete-time transfer function, $y_c(t)$ in Figure 7.24 will be the derivative of $x_c(t)$, assuming that there is no aliasing in sampling $x_c(t)$.

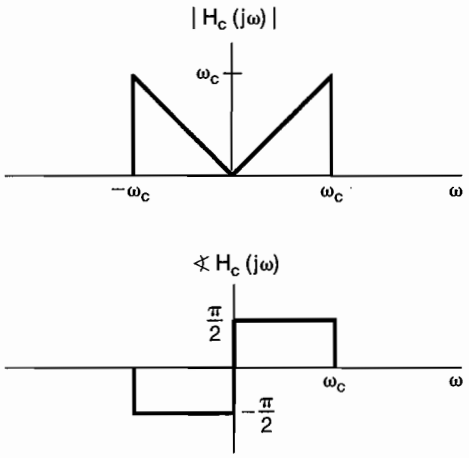


Figure 7.27 Frequency response of a continuous-time ideal band-limited differentiator $H_c(j\omega) = j\omega$, $|\omega| < \omega_c$.

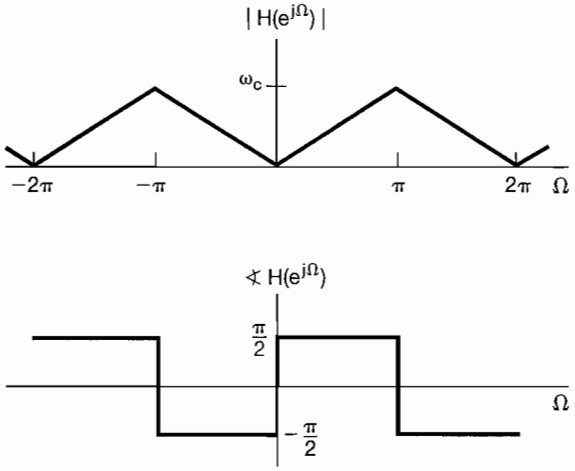


Figure 7.28 Frequency response of discrete-time filter used to implement a continuous-time band-limited differentiator.

Example 7.2

By considering the output of the digital differentiator for a continuous-time sinc input, we may conveniently determine the impulse response $h_d[n]$ of the discrete-time filter in the implementation of the digital differentiator. With reference to Figure 7.24, let

$$x_c(t) = \frac{\sin(\pi t/T)}{\pi t}, \tag{7.29}$$

where T is the sampling period. Then

$$X_c(j\omega) = \begin{cases} 1, & |\omega| < \pi/T \\ 0, & \text{otherwise} \end{cases},$$

which is sufficiently band limited to ensure that sampling $x_c(t)$ at frequency $\omega_s = 2\pi/T$ does not give rise to any aliasing. It follows that the output of the digital differentiator is

$$y_c(t) = \frac{d}{dt}x_c(t) = \frac{\cos(\pi t/T)}{Tt} - \frac{\sin(\pi t/T)}{\pi t^2}. \tag{7.30}$$

For $x_c(t)$ as given by eq. (7.29), the corresponding signal $x_d[n]$ in Figure 7.24 may be expressed as

$$x_d[n] = x_c(nT) = \frac{1}{T}\delta[n]. \tag{7.31}$$

That is, for $n \neq 0$, $x_c(nT) = 0$, while

$$x_d[0] = x_c(0) = \frac{1}{T}$$

which can be verified by l'Hôpital's rule. We can similarly evaluate $y_d[n]$ in Figure 7.24 corresponding to $y_c(t)$ in eq. (7.30). Specifically

$$y_d[n] = y_c(nT) = \begin{cases} \dfrac{(-1)^n}{nT^2}, & n \neq 0 \\ 0 & n = 0 \end{cases}. \tag{7.32}$$

which can be verified for $n \neq 0$ by direct substitution into eq. (7.30) and for $n = 0$ by application of l'Hôpital's rule.

Thus when the input to the discrete-time filter given by eq. (7.28) is the scaled unit impulse in eq. (7.31), the resulting output is given by eq. (7.32). We then conclude that the impulse response of this filter is given by

$$h_d[n] = \begin{cases} \dfrac{(-1)^n}{nT}, & n \neq 0 \\ 0, & n = 0 \end{cases}$$

7.4.2 Half-Sample Delay

In this section, we consider the implementation of a time shift (delay) of a continuous-time signal through the use of a system in the form of Figure 7.19. Thus, we require that the input and output of the overall system be related by

$$y_c(t) = x_c(t - \Delta) \tag{7.33}$$

when the input $x_c(t)$ is band limited and the sampling rate is high enough to avoid aliasing and where Δ represents the delay time. From the time-shifting property derived in Section 4.3.2

$$Y_c(j\omega) = e^{-j\omega\Delta} X_c(j\omega).$$

From eq. (7.25), the equivalent continuous-time system to be implemented must be band limited. Therefore, we take

$$H_c(j\omega) = \begin{cases} e^{-j\omega\Delta}, & |\omega| < \omega_c \\ 0, & \text{otherwise} \end{cases}, \tag{7.34}$$

where ω_c is the cutoff frequency of the continuous-time filter. That is, $H_c(j\omega)$ corresponds to a time shift as in eq. (7.33) for band-limited signals and rejects all frequencies greater than ω_c. The magnitude and phase of the frequency response are shown in Figure 7.29(a). With the sampling frequency ω_s taken as $\omega_s = 2\omega_c$, the corresponding discrete-time

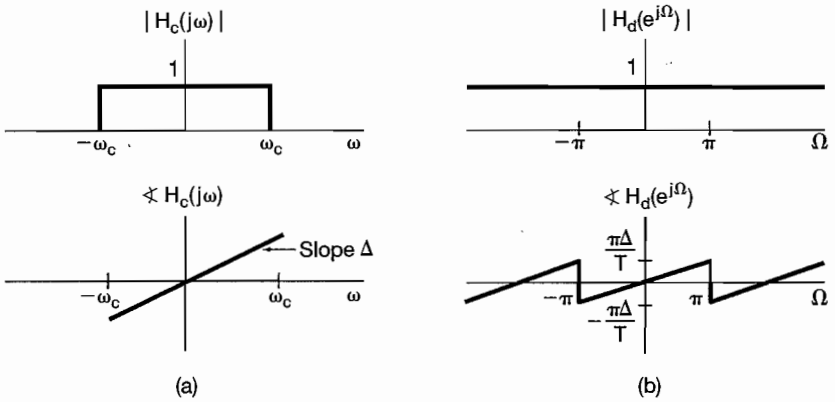


Figure 7.29 (a) Magnitude and phase of the frequency response for a continuous-time delay; (b) magnitude and phase of the frequency response for the corresponding discrete-time delay.

frequency response is

$$H_d(e^{j\Omega}) = e^{-j\Omega\Delta/T}, \quad |\Omega| < \pi, \tag{7.35}$$

and is shown in Figure 7.29(b).

For appropriately band-limited inputs, the output of the system of Figure 7.24 with $H_d(e^{j\Omega})$ as in eq. (7.35) is a delayed replica of the input. For Δ/T an integer, the sequence $y_d[n]$ is a delayed replica of $x_d[n]$; that is,

$$y_d[n] = x_d\left[n - \frac{\Delta}{T}\right]. \tag{7.36}$$

For Δ/T not an integer, eq. (7.36), as written, has no meaning, since sequences are defined only at integer values of the index. However, we can interpret the relationship between $x_d[n]$ and $y_d[n]$ in these cases in terms of band-limited interpolation. The signals $x_c(t)$ and $x_d[n]$ are related through sampling and band-limited interpolation, as are $y_c(t)$ and $y_d[n]$. With $H_d(e^{j\Omega})$ in eq. (7.35), $y_d[n]$ is equal to samples of a shifted version of the band-limited interpolation of the sequence $x_d[n]$. This is illustrated in Figure 7.30 with $\Delta/T = 1/2$, which is sometimes referred to as a half-sample delay.

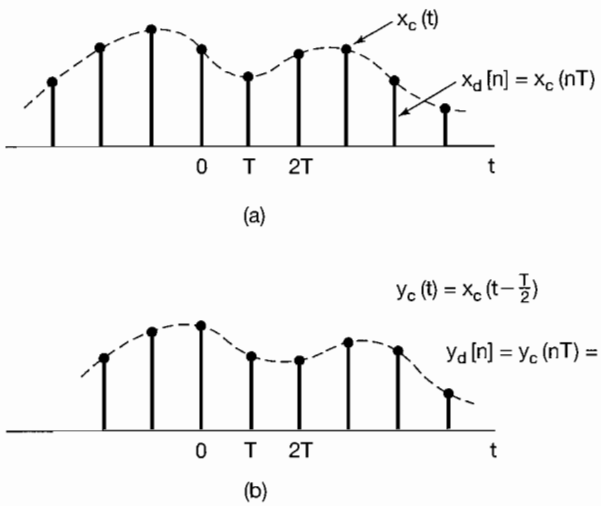


Figure 7.30 (a) Sequence of samples of a continuous-time signal $x_c(t)$; (b) sequence in (a) with a half-sample delay.

Example 7.3

The approach in Example 7.2 is also applicable to determining the impulse response $h_d[n]$ of the discrete-time filter in the half-sample delay system. With reference to Figure 7.24, let

$$x_c(t) = \frac{\sin(\pi t/T)}{\pi t}. \tag{7.37}$$

It follows from Example 7.2 that

$$x_d[n] = x_c(nT) = \frac{1}{T}\delta[n].$$

Also, since there is no aliasing for the band-limited input in eq. (7.37), the output of the half-sample delay system is

$$y_c(t) = x_c(t - T/2) = \frac{\sin(\pi(t - T/2)/T)}{\pi(t - T/2)},$$

and the sequence $y_d[n]$ in Figure 7.24 is

$$y_d[n] = y_c(nT) = \frac{\sin(\pi(n - \frac{1}{2}))}{T\pi(n - \frac{1}{2})}.$$

We conclude that

$$h[n] = \frac{\sin(\pi(n - \frac{1}{2}))}{\pi(n - \frac{1}{2})}.$$

7.5 SAMPLING OF DISCRETE-TIME SIGNALS

Thus far in this chapter, we have considered the sampling of continuous-time signals, and in addition to developing the analysis necessary to understand continuous-time sampling, we have introduced a number of its applications. As we will see in this section, a very similar set of properties and results with a number of important applications can be developed for sampling of discrete-time signals.

7.5.1 Impulse-Train Sampling

In analogy with continuous-time sampling as carried out using the system of Figure 7.2, sampling of a discrete-time signal can be represented as shown in Figure 7.31. Here, the new sequence $x_p[n]$ resulting from the sampling process is equal to the original sequence $x[n]$ at integer multiples of the sampling period N and is zero at the intermediate samples; that is,

$$x_p[n] = \begin{cases} x[n], & \text{if } n = \text{an integer multiple of } N \\ 0, & \text{otherwise} \end{cases}. \tag{7.38}$$

As with continuous-time sampling in Section 7.1, the effect in the frequency domain of discrete-time sampling is seen by using the multiplication property developed in Section 5.5. Thus, with

$$x_p[n] = x[n]p[n] = \sum_{k=-\infty}^{+\infty} x[kN]\delta[n - kN], \tag{7.39}$$

we have, in the frequency domain,

$$X_p(e^{j\omega}) = \frac{1}{2\pi}\int_{2\pi} P(e^{j\theta})X(e^{j(\omega-\theta)})d\theta. \tag{7.40}$$