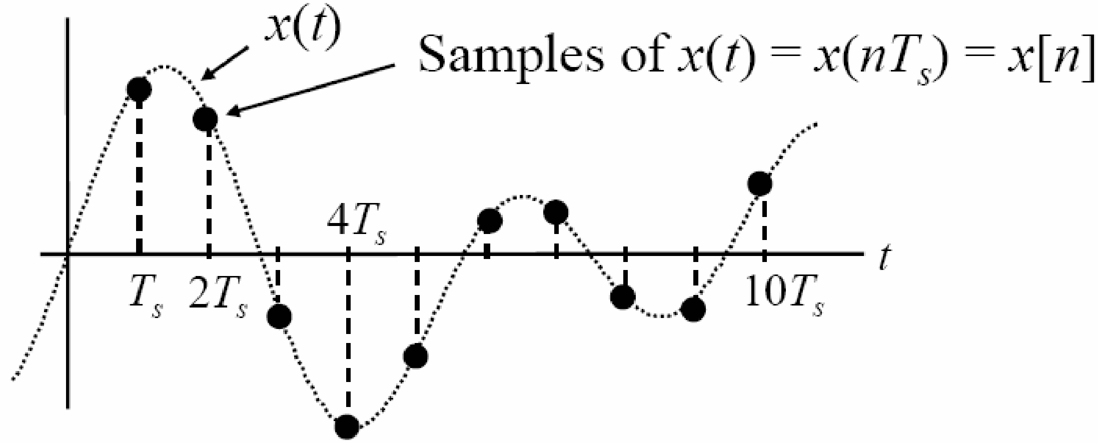
Section 2 - Digital Representation of Analog Signals

Module 1 - Sampling

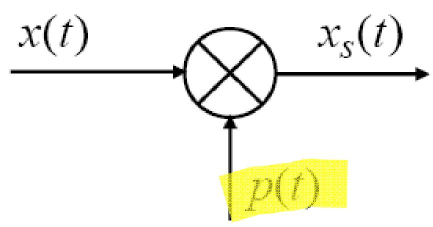
Set 2 - Lowpass Sampling Theorem



To sample a waveform is to represent that waveform at discrete points in time. These points are assumed to be periodic in time. The sampling period is T and the reciprocal is the sampling frequency.

$$\boxed{f_s = \frac{1}{T_s}}$$

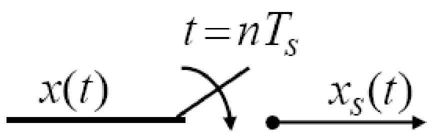
The sampling operation is often modeled as multiplying the signal to be sampled, $x(t)$, by a sampling waveform, $p(t)$, which is periodic at the sampling frequency.



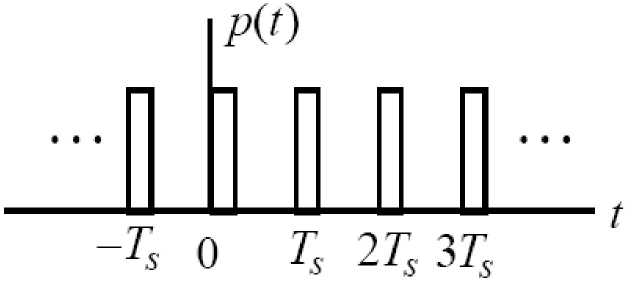
We will assume impulse-function sampling. For this case

$$\boxed{p(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s)}$$

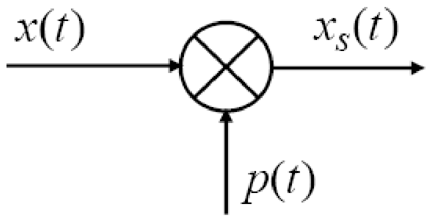
Two Choices for the Sampling Function



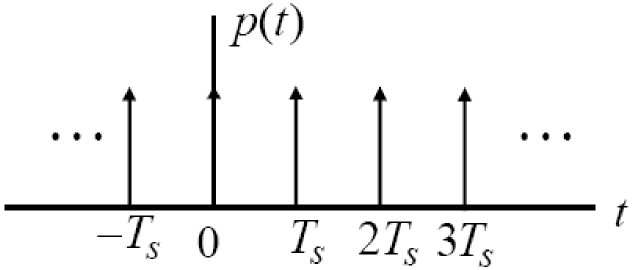
Pulse sampling:



The amplitude of the pulse carries the sample value. This suggests the closing of a switch for short instants (PAM).



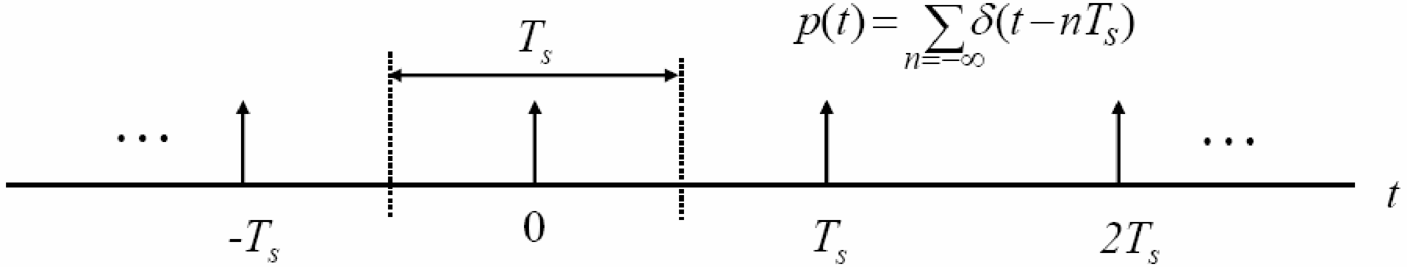
Impulse-function sampling:



The weight of the impulse carries the sample value. For this case, we use the notation $x_s(t) = x_\delta(t)$

Impulse Function Sampling

$$p(t)=\sum_{n=-\infty}^{\infty}\delta(t-nT_s)$$



Since the sampling waveform is periodic, it can be expressed in a Fourier series.

$$\boxed{p(t)=\sum_{n=-\infty}^{\infty}c_n\exp\left[j2\pi nf_s t\right]}$$

The Fourier coefficients are

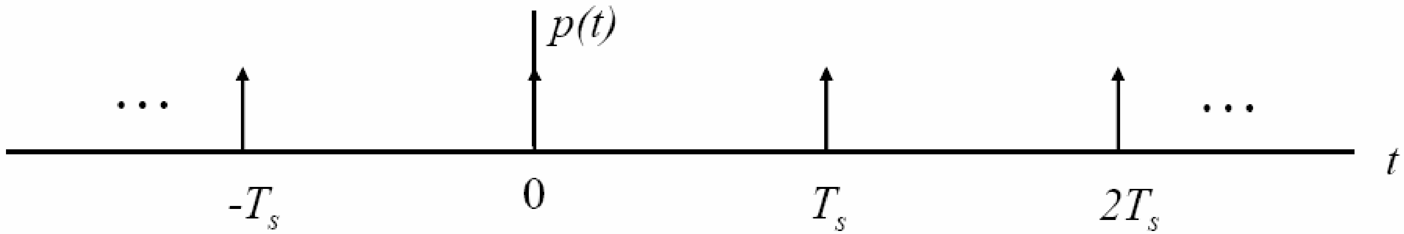
$$c_n=\frac{1}{T_S}\int_{T_S}\delta(t)\exp\left[-j2\pi n f_s t\right]dt=\frac{1}{T_S}=f_s$$

Thus

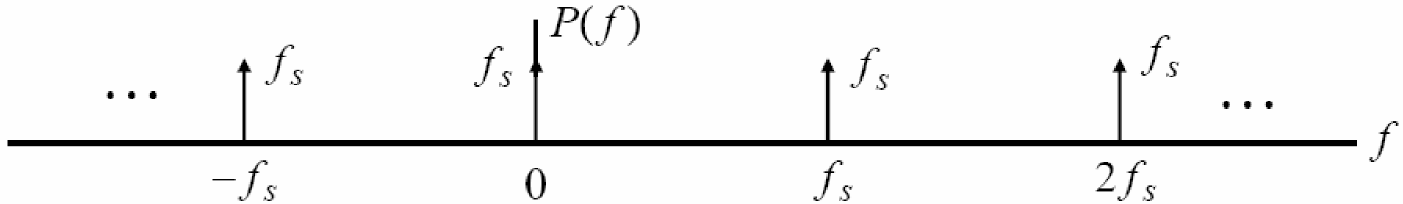
$$\boxed{p(t)=f_s\sum_{n=-\infty}^{\infty}\exp(j2\pi nf_s t)}$$

Impulse Function Sampling

Thus $\boxed{p(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s)}$



and $\boxed{P(f) = f_s \sum_{n=-\infty}^{\infty} \delta(f - nf_s)}$

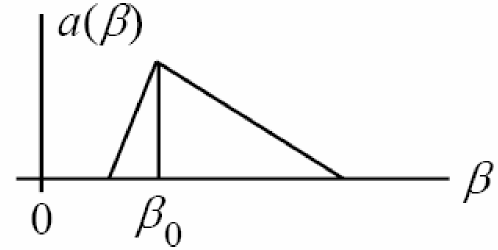


are a Fourier transform pair.

Convolution with an Impulse

$$c(\beta)=a(\beta)*b(\beta)$$

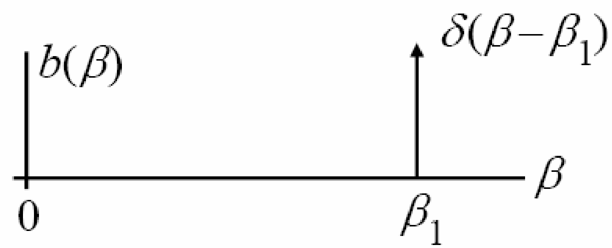
$$c(\beta)=\int_{-\infty}^{\infty} a(\lambda)b(\beta-\lambda)d\lambda$$



with

$$b(\beta)=\delta(\beta-\beta_1)$$

$$b(\beta-\lambda)=\delta(\beta-\lambda-\beta_1)$$

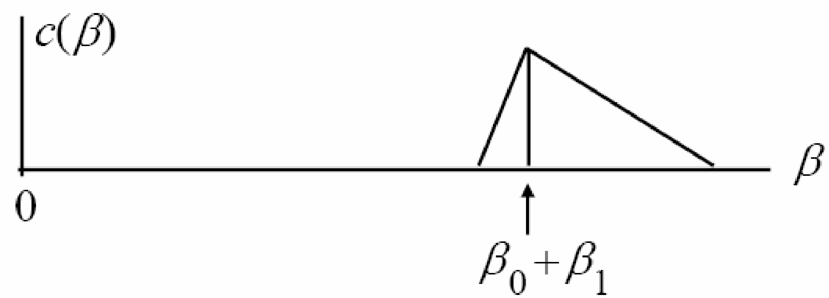


then

$$c(\beta)=\int_{-\infty}^{\infty} a(\lambda)\delta(\beta-\lambda-\beta_1)d\lambda$$

$$c(\beta)=\int_{-\infty}^{\infty} a(\lambda)\delta\left[(\beta-\beta_1)-\lambda\right]d\lambda$$

$$c(\beta)=a(\beta-\beta_1)$$



Time/Frequency Domain Sampling

Sampling in the time domain

$$x_\delta(t) = x(t)\sum_{n=-\infty}^{\infty}\delta(t-nT_s) = \sum_{n=-\infty}^{\infty}x(t)\delta(t-nT_s)$$

$$x_\delta(t) = \sum_{n=-\infty}^{\infty}x(nT_s)\delta(t-nT_s) = \sum_{n=-\infty}^{\infty}x[n]\delta(t-nT_s)$$

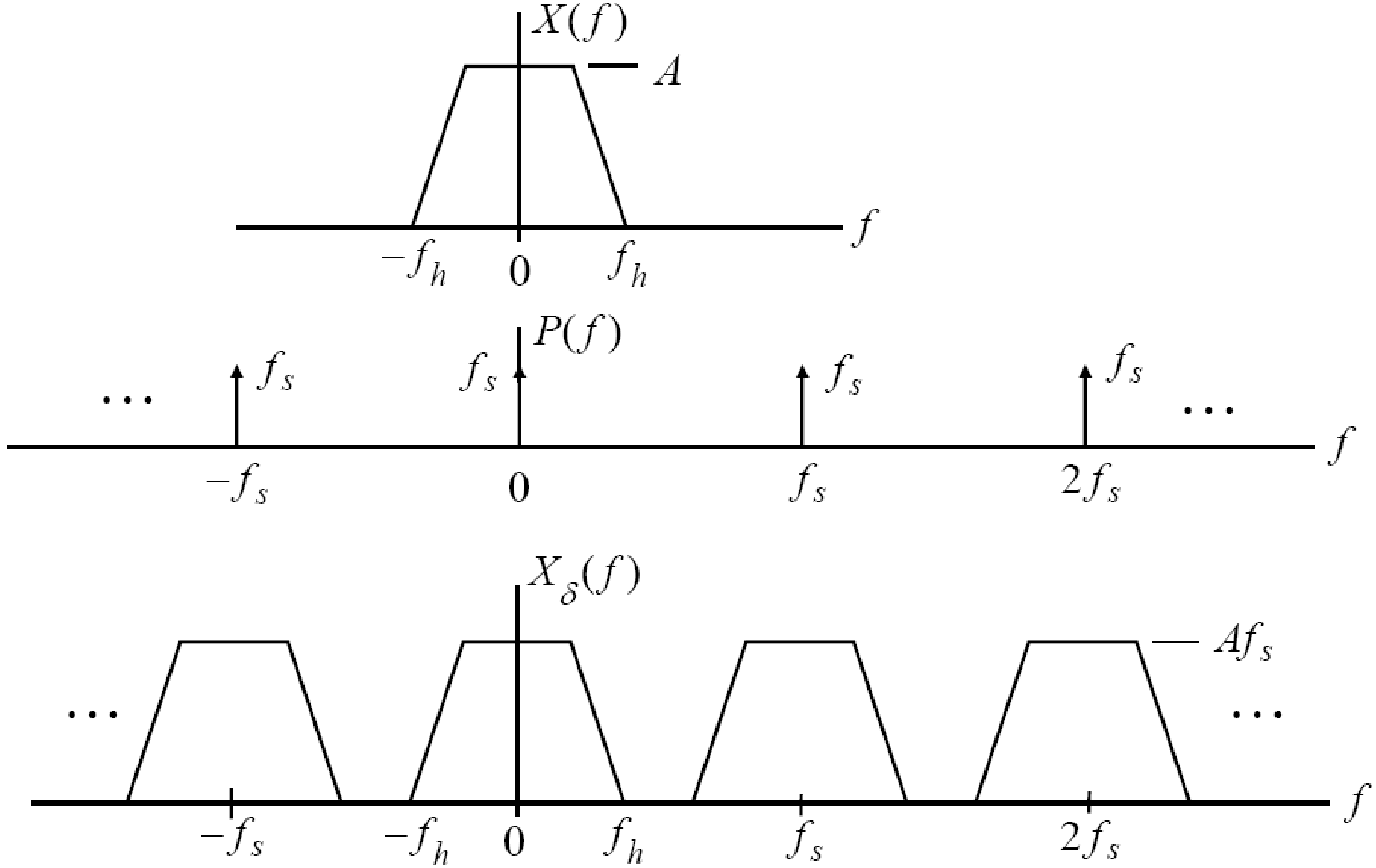
Note that

$$x[n] = x(nT_s)$$

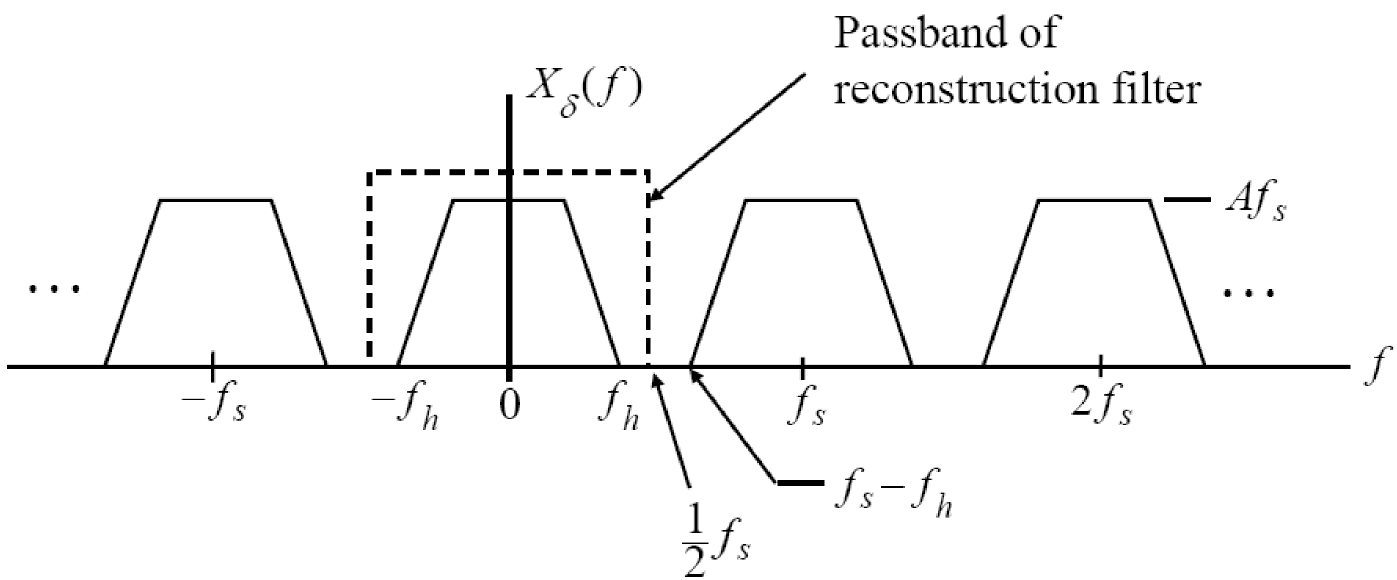
Sampling in the frequency domain

$$X_\delta(f) = X(f) * f_s\sum_{n=-\infty}^{\infty}\delta(f-nf_s) = f_s\sum_{n=-\infty}^{\infty}X(f-nf_s)$$

Sampling - Frequency Domain



Lowpass Sampling Theorem



For reconstruction without error:

$$f_s - f_h > f_h$$
$$f_s > 2f_h$$

Lowpass Sampling Theorem - 2

We have therefore illustrated the lowpass sampling theorem:

A bandlimited lowpass signal may be sampled and reconstructed without error from the samples if the sampling frequency exceeds $2f_h$ where f_h is the highest frequency in the signal being sampled.

Reconstruction and Interpolation

$$x_s(kT_s) \longrightarrow \boxed{\text{Reconstruction Filter}} \longrightarrow x_r(t)$$

The reconstructed signal is given by

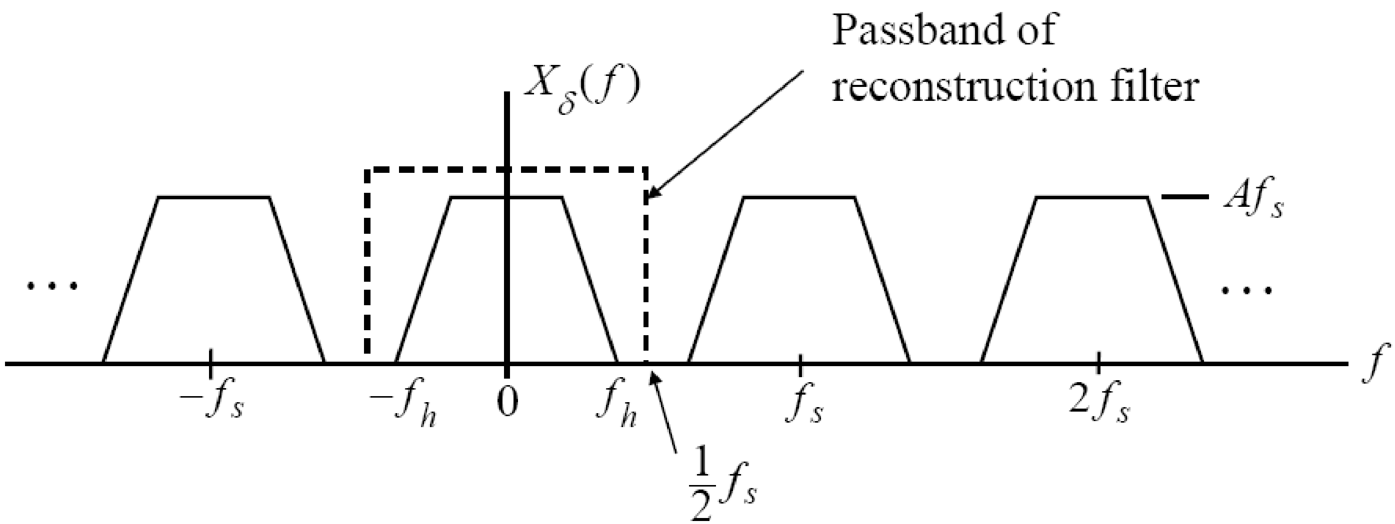
$$x_r(t) = \left\{ \sum_{k=-\infty}^{\infty} x(kT_s)\delta(t - kT_s) \right\} * h(t)$$

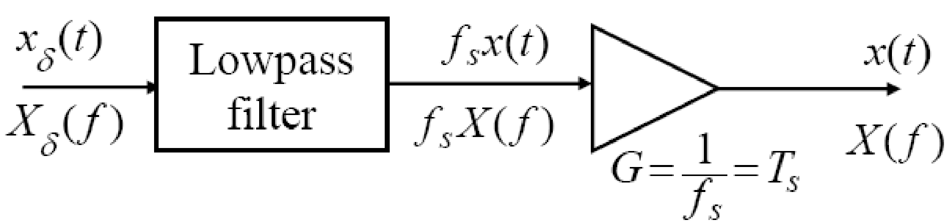
where $h(t)$ is the impulse response of the reconstruction filter.
This gives

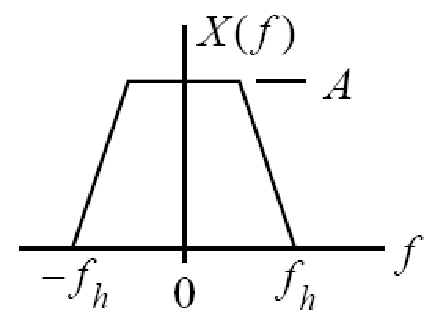
$$x_r(t) = \sum_{k=-\infty}^{\infty} x(kT_s)h(t - kT_s)$$

The problem is to find an appropriate $h(t)$.

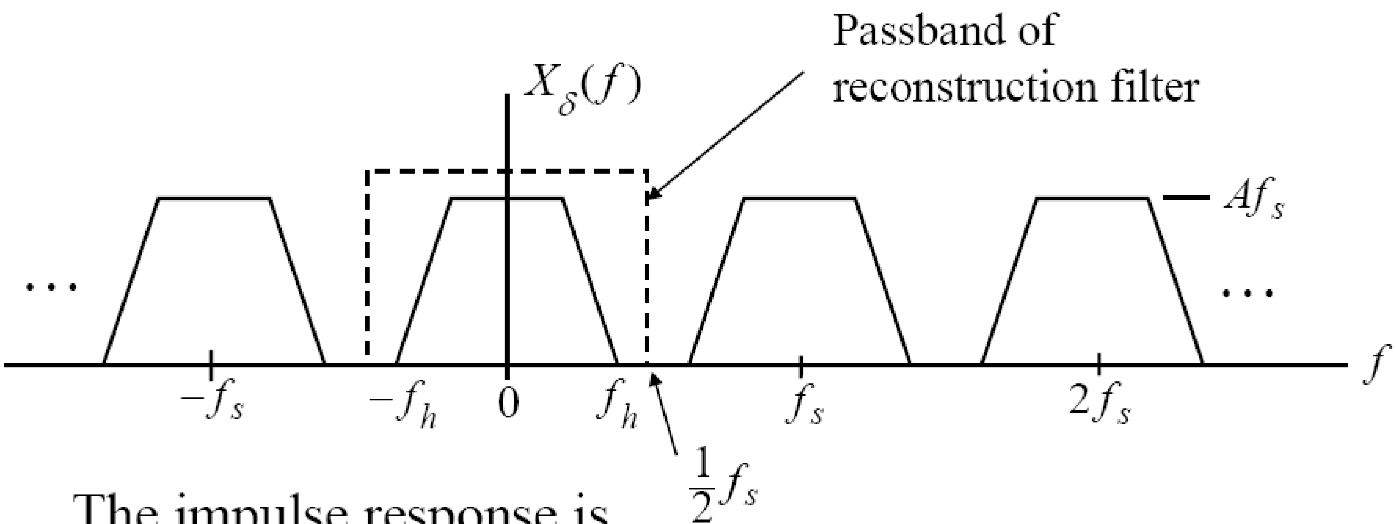
Reconstruction







Perfect Reconstruction



The impulse response is

$$h(t) = T_s \int_{-f_s/2}^{f_s/2} \exp(j2\pi f t)\,df = \frac{T_s}{j2\pi t}\left[\exp(j\pi f_s t) - \exp(-j2\pi f_s t)\right]$$

$$h(t) = \frac{T_s}{\pi t}\sin(\pi f_s t) = \mathrm{sinc}(f_s t)$$

Perfect Reconstruction

From

$$x_r(t) = \sum_{k=-\infty}^{\infty} x(kT_s) h(t - kT_s)$$

and

$$h(t) = \frac{T_s}{\pi t}\sin(\pi f_s t) = \text{sinc}(f_s t)$$

we have

$$\boxed{x_r(t) = \sum_{k=-\infty}^{\infty} x(kT_s)\,\text{sinc}[f_s(t - kT_s)] = \sum_{k=-\infty}^{\infty} x(kT_s)\,\text{sinc}\left[\frac{t}{T_s} - k\right]}$$

This is not practical since sinc(x) is infinite in extent and it therefore takes an infinite number of samples to interpolate a single point.

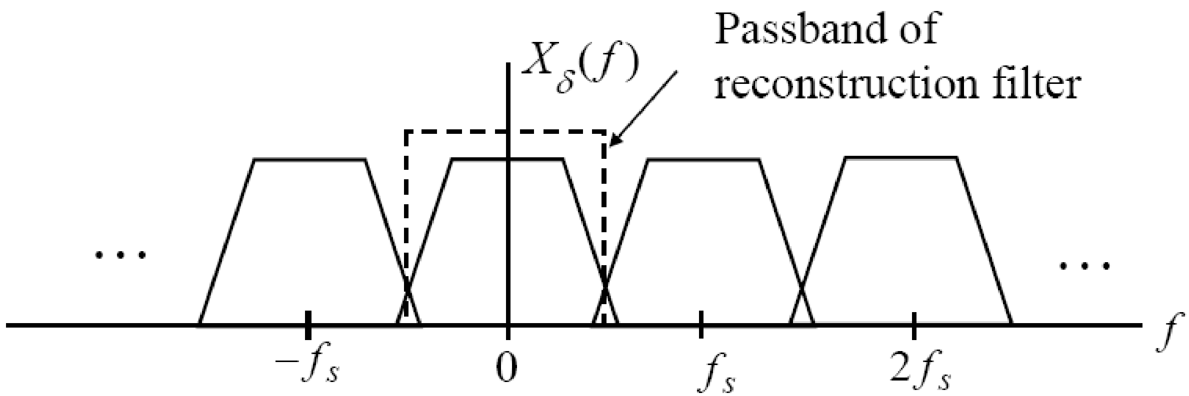
Perfect Reconstruction - 2

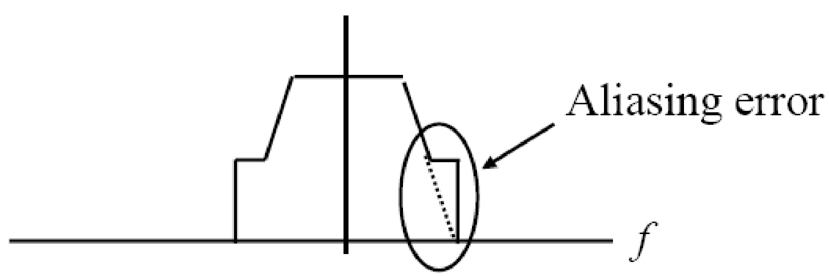
A practical solution is obtained by truncating the series to $2N+1$ terms

$$\boxed{x_r(t) \approx \sum_{k=-N}^{N} x(kT_s)\,\mathrm{sinc}[f_s(t-kT_s)] = \sum_{k=-N}^{N} x(kT_s)\,\mathrm{sinc}\left(\frac{t}{T_s}-k\right)}$$

The appropriate value for N is a tradeoff between accuracy and computational burden.

Aliasing Error





Section 2 - Digital Representation of Analog Signals

Module 1 - Sampling

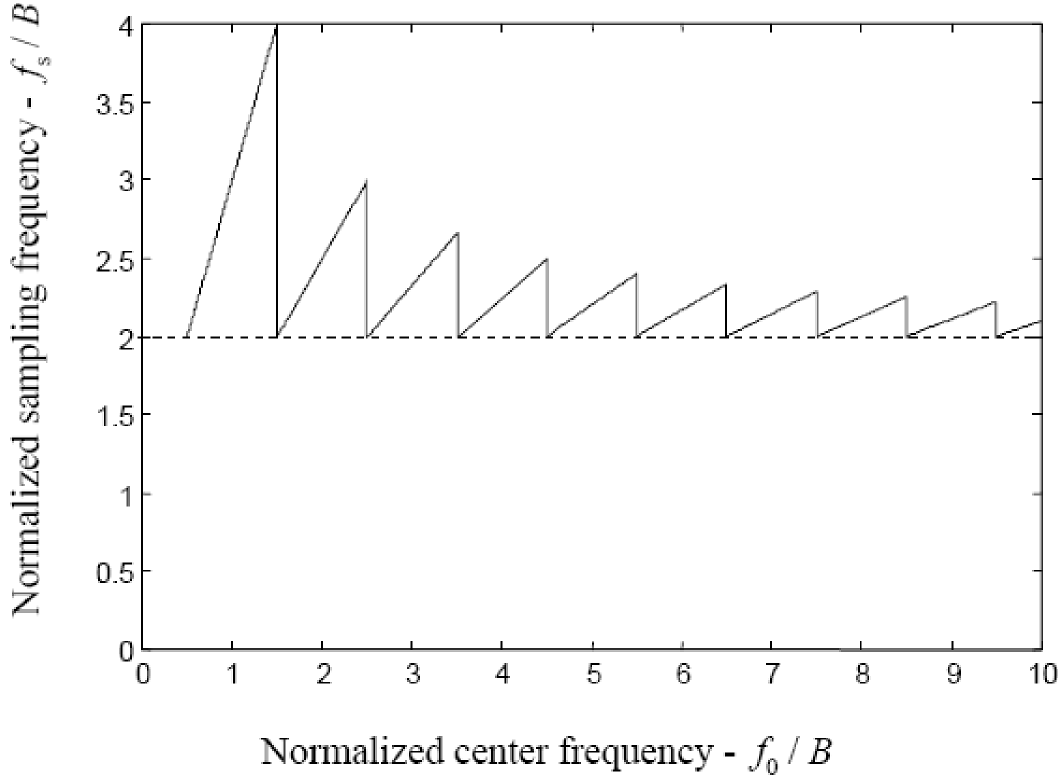
Set 3 - Bandpass Sampling Theorem

The Bandpass Sampling Theorem

We consider the bandpass sampling theorem for completeness. In the work to follow in this course we will for the most part be concerned with the lowpass sampling theorem.

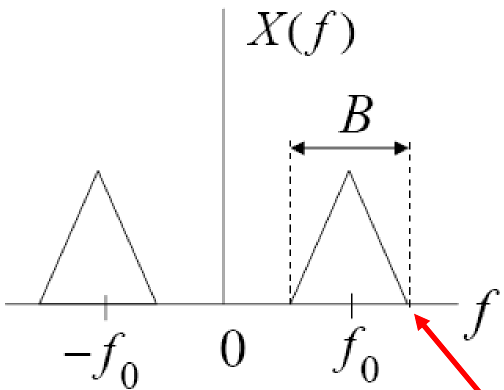
Thus, our treatment of the bandpass sampling theorem is very brief.

If a bandpass signal has bandwidth B and highest frequency f_h the signal can be sampled and reconstructed using a sampling frequency of $f_s = 2 f_h / m$ where m is the largest integer not exceeding f_h / B. All higher sampling frequencies are not necessarily usable unless they exceed $2 f_h$.

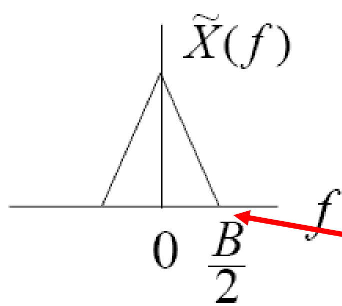


Note that $2B \leq f_s \leq 4B$ and that $f_s \rightarrow 2B$ as $f_0 / B \rightarrow \infty$

Complex Envelope



Bandpass signal



Complex envelope

$$x(t) = m(t)\cos\left(2\pi f_0 t\right)$$

$$x(t) = \mathrm{Re}\left\{\tilde{x}(t)\exp\left(j2\pi f_0 t\right)\right\}$$

$$\tilde{x}(t) = x_d(t) + jx_q(t) \quad \text{(complex lowpass)}$$

Minimum sampling frequency for bandpass signal $= 2\left(f_0 + \frac{B}{2}\right) = 2f_0 + B$

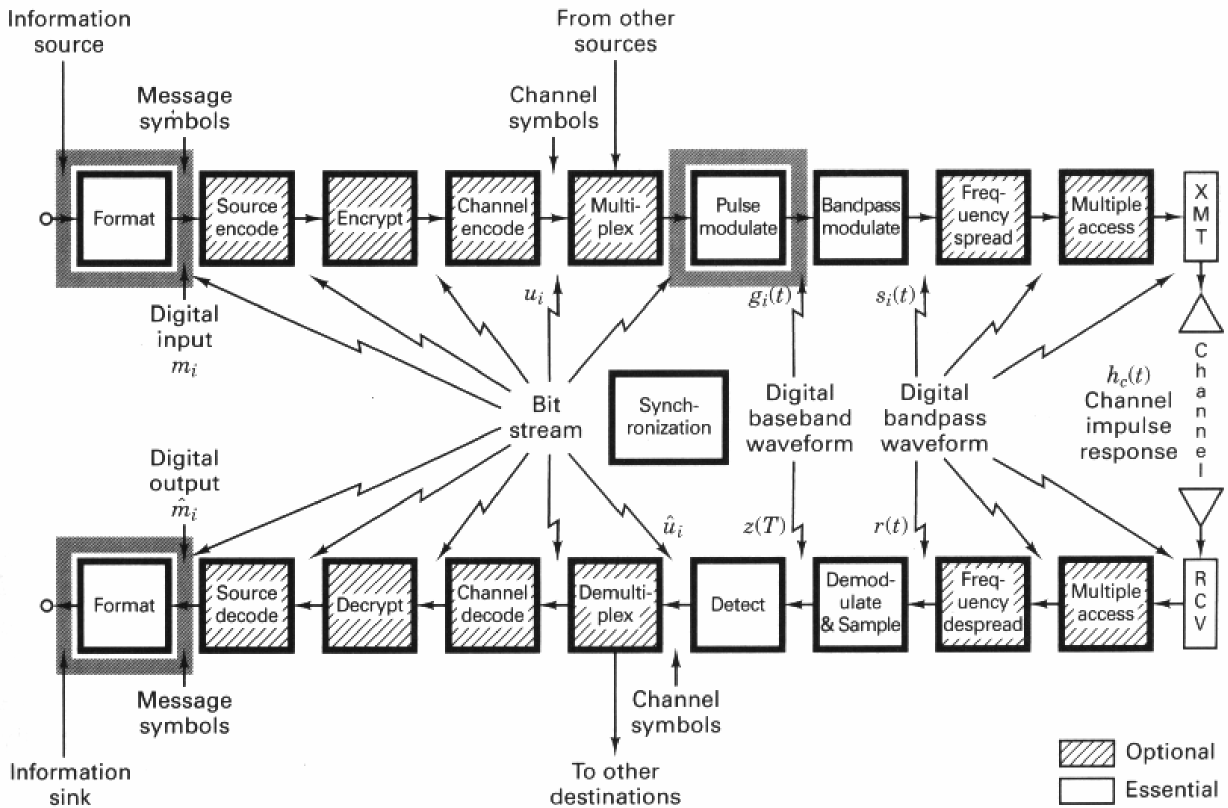
Minimum sampling frequency for complex envelope $= B + B = 2B$

We will have much more to say about the complex envelope representations for bandpass signals later. For now we simply observe that the minimum sampling frequency is essentially independent of the representation for the bandpass signal.

1. Bandpass signal: $f_s > 2B$
2. Complex envelope (bandpass) signal: $f_s > 2B$

Extra material

Formatting and Baseband Modulation



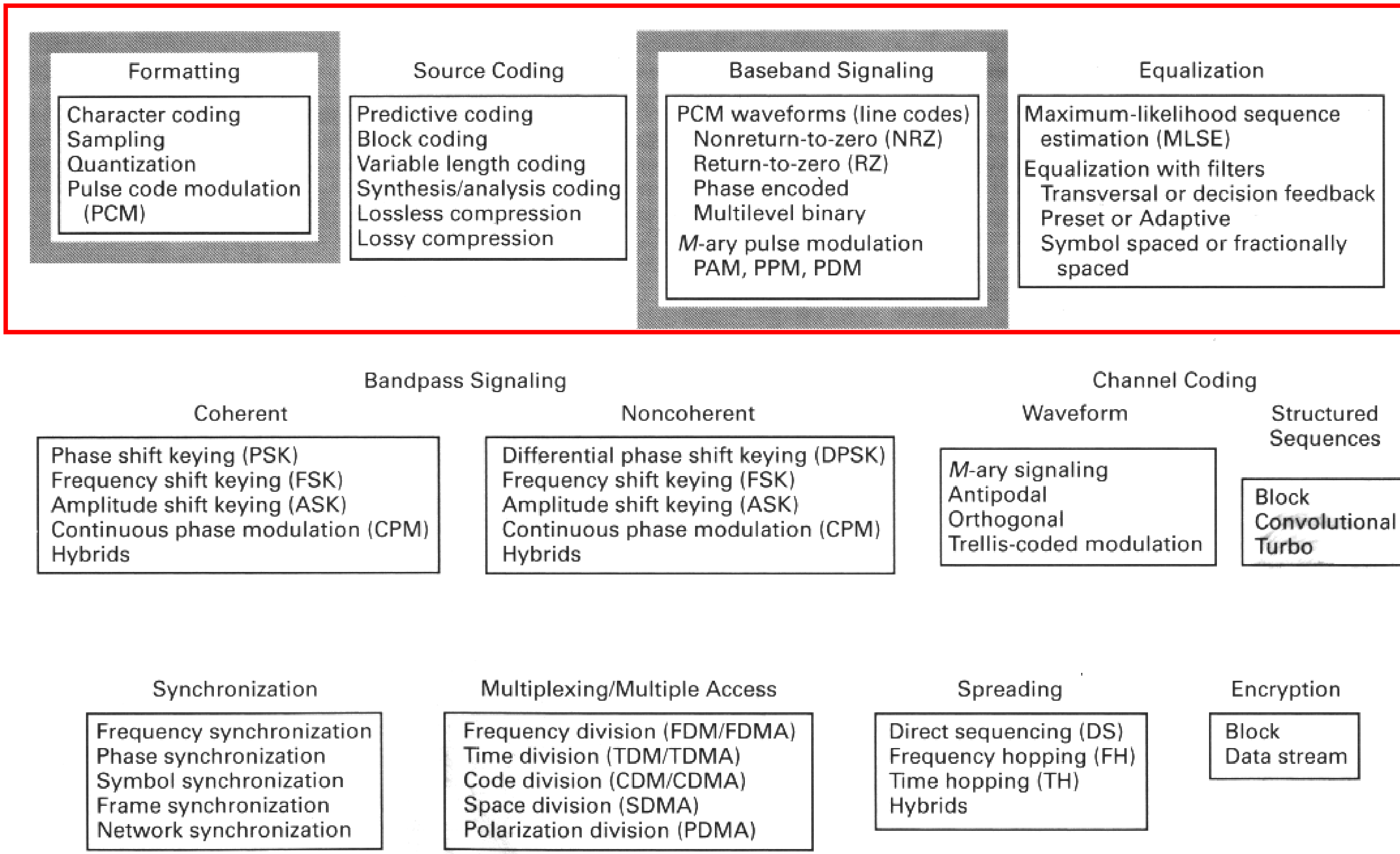
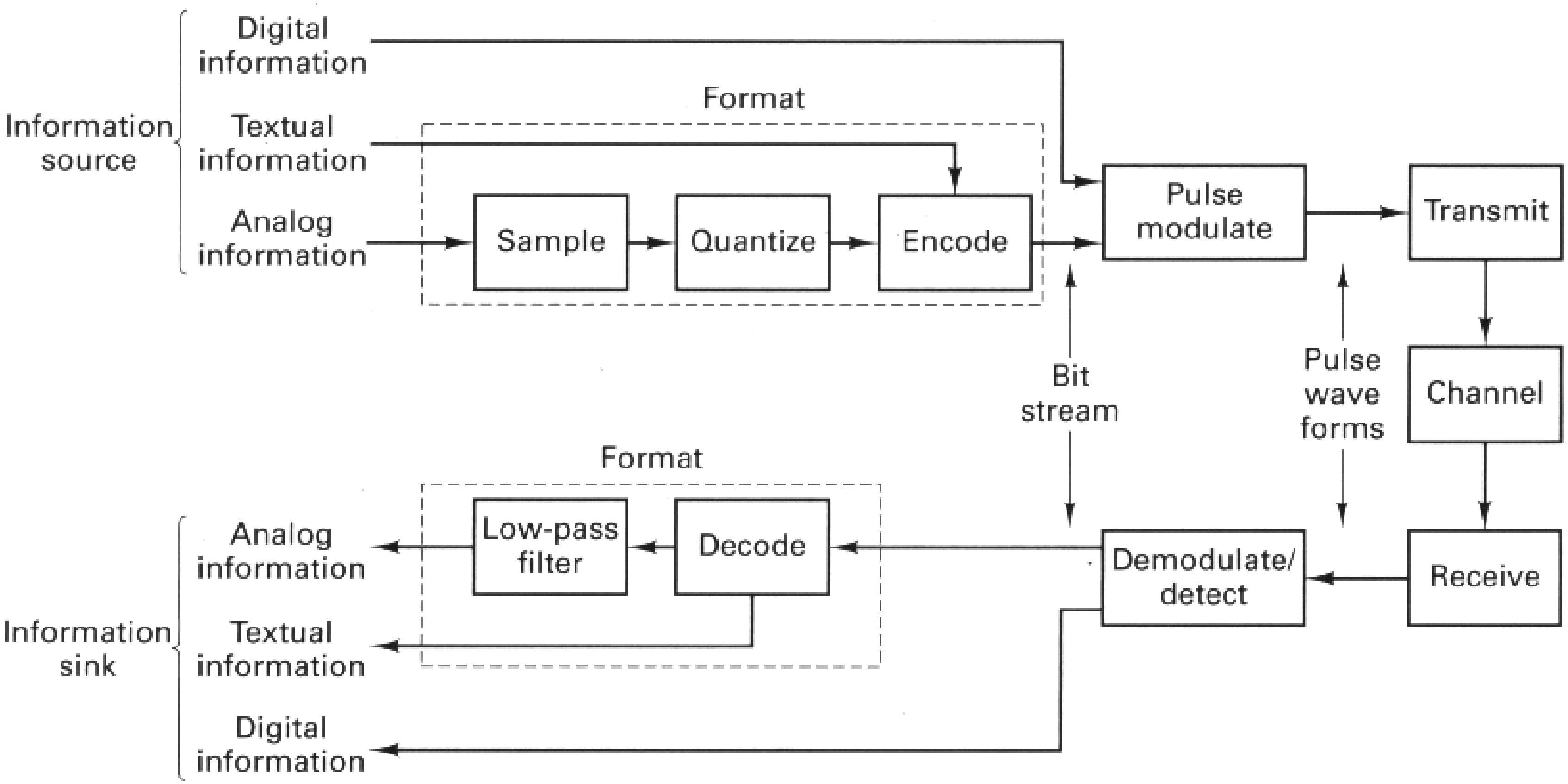


Figure 2.1 Basic digital communication transformations



Information source
Digital information
Textual information
Analog information
Format
Sample
Quantize
Encode
Pulse modulate
Transmit
Channel
Receive
Bit stream
Pulse wave forms
Demodulate/ detect
Format
Low-pass filter
Decode
Information sink
Analog information
Textual information
Digital information

2.4 FORMATTING ANALOG INFORMATION

If the information is analog, it cannot be character encoded as in the case of textual data; the information must first be transformed into a digital format. The process of transforming an analog waveform into a form that is compatible with a digital communication system starts with sampling the waveform to produce a discrete pulse-amplitude-modulated waveform, as described below.

2.4.1 The Sampling Theorem

The link between an analog waveform and its sampled version is provided by what is known as the *sampling process.* This process can be implemented in several ways, the most popular being the *sample-and-hold* operation. In this operation, a switch and storage mechanism (such as a transistor and a capacitor, or a shutter and a filmstrip) form a sequence of samples of the continuous input waveform. The output of the sampling process is called *pulse amplitude modulation* (PAM) because the successive output intervals can be described as a sequence of pulses with amplitudes derived from the input waveform samples. The analog waveform can be approximately retrieved from a PAM waveform by simple low-pass filtering. An important question: how closely can a filtered PAM waveform approximate the original input waveform? This question can be answered by reviewing the *sampling theorem,* which states the following [1]: A bandlimited signal having no spectral components above f_m hertz can be determined uniquely by values sampled at uniform intervals of

$$\boxed{T_s \leq \frac{1}{2f_m} \text{ sec}} \tag{2.1}$$

This particular statement is also known as the *uniform sampling theorem.* Stated another way, the upper limit on T_s can be expressed in terms of the sampling rate, denoted $f_s = 1/T_s$. The restriction, stated in terms of the sampling rate, is known as the *Nyquist criterion.* The statement is

$$\boxed{f_s \geq 2f_m} \tag{2.2}$$

The sampling rate $f_s = 2f_m$ is also called the *Nyquist rate.* The Nyquist criterion is a theoretically sufficient condition to allow an analog signal to be *reconstructed completely* from a set of uniformly spaced discrete-time samples. In the sections that follow, the validity of the sampling theorem is demonstrated using different sampling approaches.

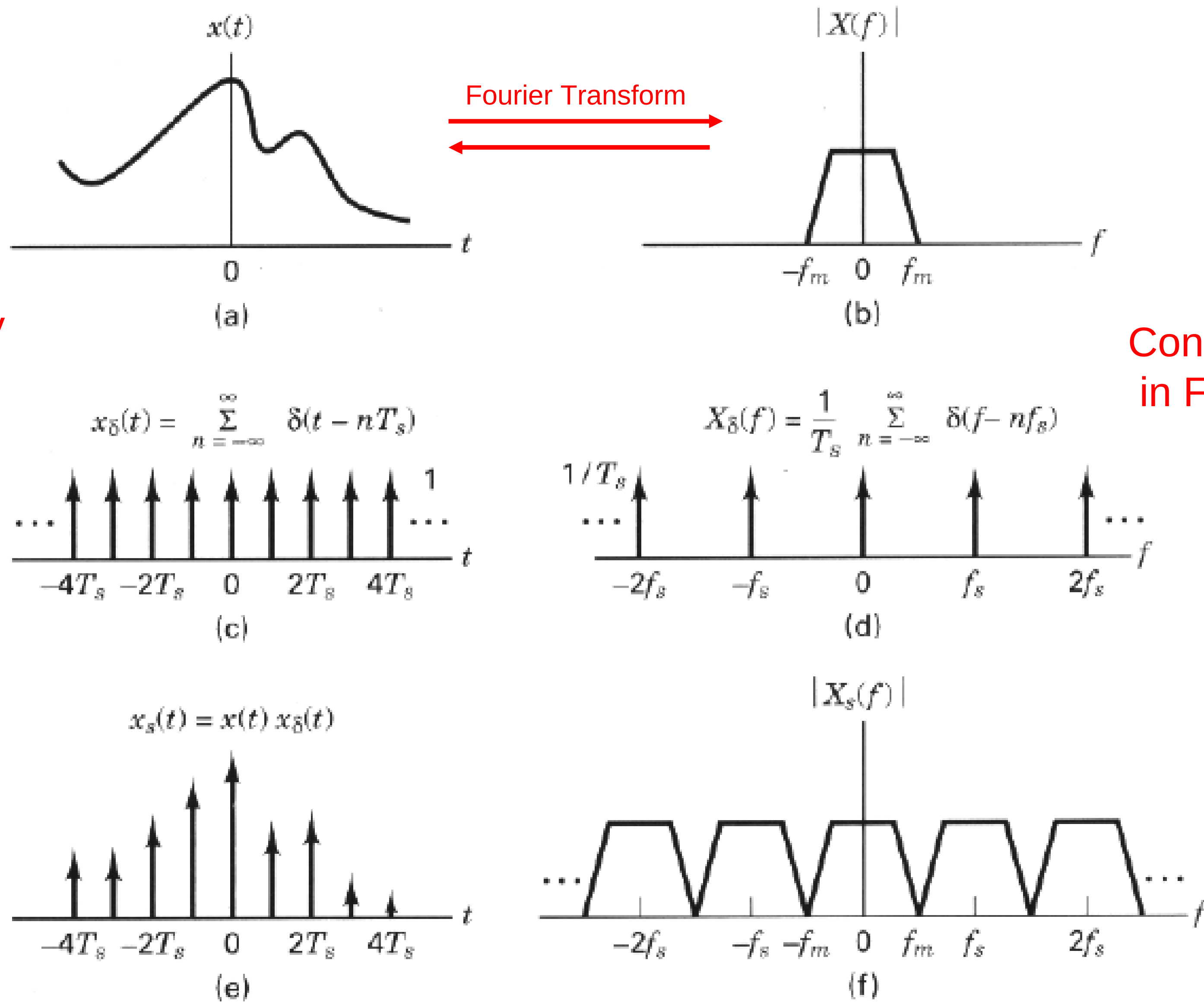
2.4.1.1 Impulse Sampling

Here we demonstrate the validity of the sampling theorem using the frequency convolution property of the Fourier transform. Let us first examine the case of *ideal sampling* with a sequence of unit impulse functions. Assume an analog waveform, $x(t)$, as shown in Figure 2.6a, with a Fourier transform, $X(f)$, which is zero outside the interval $(-f_m < f < f_m)$, as shown in Figure 2.6b. The sampling of $x(t)$ can be viewed as the product of $x(t)$ with a periodic train of unit impulse functions $x_\delta(t)$, shown in Figure 2.6c and defined as

$$\boxed{x_\delta(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s)} \qquad (2.3)$$

Multiply in time

Convolution in Frequency



where T_s is the sampling period and $\delta(t)$ is the unit impulse or Dirac delta function defined in Section 1.2.5. Let us choose $T_s = 1/2f_m$, so that the Nyquist criterion is just satisfied.

The *sifting property* of the impulse function (see Section A.4.1) states that

$$\boxed{x(t)\delta(t - t_0) = x(t_0)\delta(t - t_0)} \tag{2.4}$$

Using this property, we can see that $x_s(t)$, the sampled version of $x(t)$ shown in Figure 2.6e, is given by

$$\boxed{\begin{aligned} x_s(t) = x(t)x_\delta(t) &= \sum_{n=-\infty}^{\infty} x(t)\delta(t - nT_s) \\ &= \sum_{n=-\infty}^{\infty} x(nT_s)\delta(t - nT_s) \end{aligned}} \tag{2.5}$$

Using the *frequency convolution property* of the Fourier transform (see Section A.5.3), we can transform the time-domain product $x(t)x_\delta(t)$ of Equation (2.5) to the frequency-domain convolution $X(f) * X_\delta(f)$, where

$$X_\delta(f) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} \delta(f - nf_s) \tag{2.6}$$

is the Fourier transform of the impulse train $x_\delta(t)$ and where $f_s = 1/T_s$ is the sampling frequency. Notice that the Fourier transform of an impulse train is another impulse train; the values of the periods of the two trains are reciprocally related to one another. Figures 2.6c and d illustrate the impulse train $x_\delta(t)$ and its Fourier transform $X_\delta(f)$, respectively.

Convolution with an impulse function simply shifts the original function as follows:

$$X(f) * \delta(f - nf_s) = X(f - nf_s) \tag{2.7}$$

We can now solve for the transform $X_s(f)$ of the sampled waveform:

$$\begin{aligned} X_s(f) = X(f) * X_\delta(f) &= X(f) * \left[\frac{1}{T_s} \sum_{n=-\infty}^{\infty} \delta(f - nf_s) \right] \\ &= \frac{1}{T_s} \sum_{n=-\infty}^{\infty} X(f - nf_s) \end{aligned} \tag{2.8}$$

We therefore conclude that within the original bandwidth, the spectrum $X_s(f)$ of the sampled signal $x_s(t)$ is, to within a constant factor $(1/T_s)$, exactly the same as that of $x(t)$. In addition, the spectrum repeats itself periodically in frequency every f_s hertz. The sifting property of an impulse function makes the convolving of an impulse train with another function easy to visualize. The impulses act as sampling functions. Hence, convolution can be performed graphically by sweeping the impulse train $X_\delta(f)$ in Figure 2.6d past the transform $|X(f)|$ in Figure 2.6b. This sampling of $|X(f)|$ at each step in the sweep replicates $|X(f)|$ at each of the frequency positions of the impulse train, resulting in $|X_s(f)|$, shown in Figure 2.6f.

When the sampling rate is chosen, as it has been here, such that $\boxed{f_s = 2f_m}$, each spectral replicate is separated from each of its neighbors by a frequency band exactly equal to f_s hertz, and the analog waveform can theoretically be completely recovered from the samples, by the use of filtering. However, a filter with infinitely steep sides would be required. It should be clear that if $\boxed{f_s > 2f_m,}$ the replications will move farther apart in frequency, as shown in Figure 2.7a, making it easier to perform the filtering operation. A typical low-pass filter characteristic that might be used to separate the baseband spectrum from those at higher frequencies is shown in the figure. When the sampling rate is reduced, such that $\boxed{f_s < 2f_m,}$ the replications will overlap, as shown in Figure 2.7b, and some information will be lost. The phenomenon, the result of undersampling (sampling at too low a rate), is called *aliasing*. The Nyquist rate, $f_s = 2f_m$, is the sampling rate below which aliasing occurs; to avoid aliasing, the Nyquist criterion, $f_s \geq 2f_m$, must be satisfied.

As a matter of practical consideration, neither waveforms of engineering interest nor realizable bandlimiting filters are strictly bandlimited. Perfectly bandlimited signals do not occur in nature (see Section 1.7.2); thus, realizable signals, even though we may think of them as bandlimited, always contain some aliasing. These signals and filters can, however, be considered to be "essentially" bandlimited. By this we mean that a bandwidth can be determined beyond which the spectral components are attenuated to a level that is considered negligible.

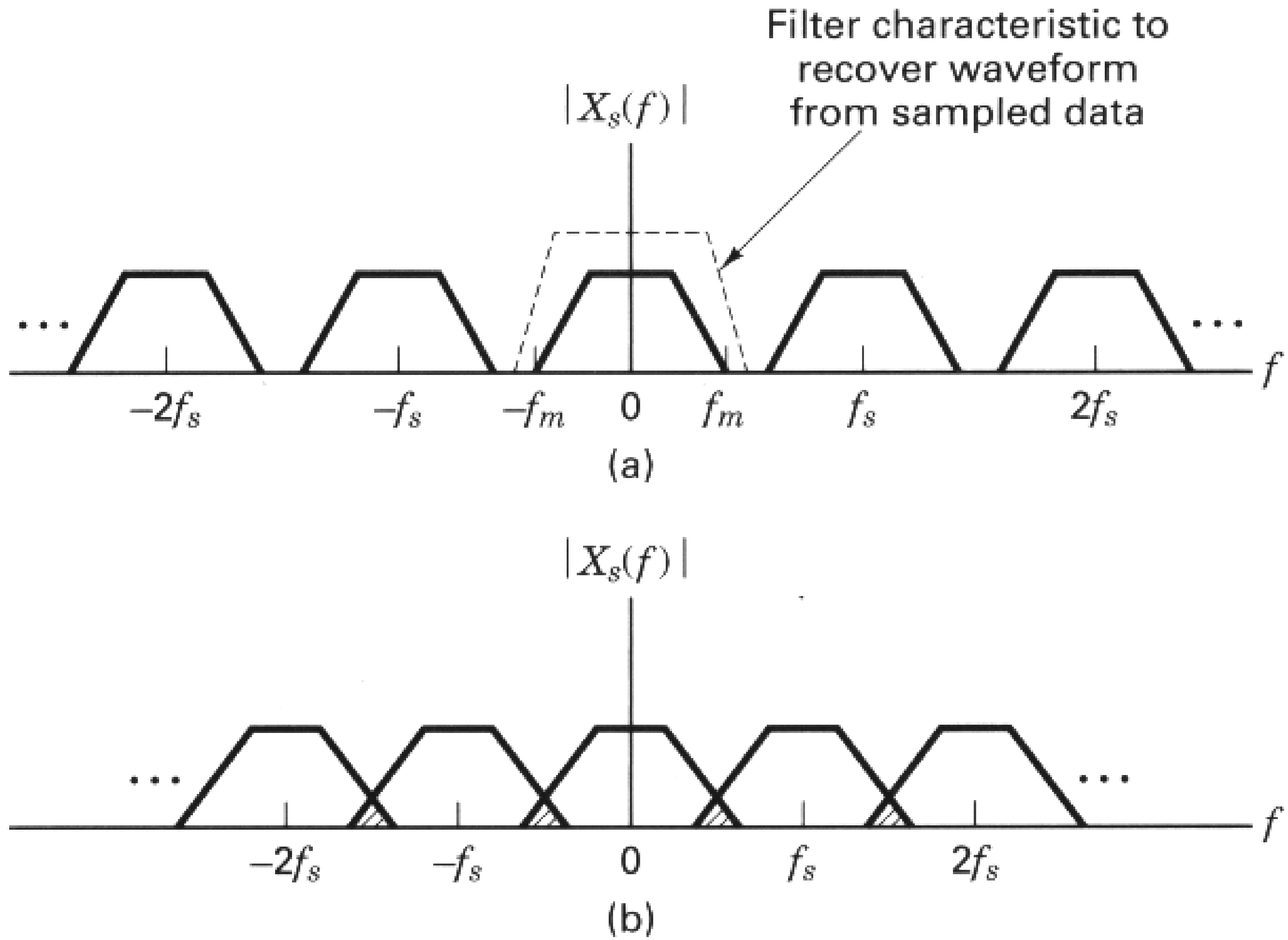


Figure 2.7 Spectra for various sampling rates. (a) Sampled spectrum ($f_s > 2f_m$). (b) Sampled spectrum ($f_s < 2f_m$).

this we mean that a bandwidth can be determined beyond which the spectral components are attenuated to a level that is considered negligible.

2.4.1.2 Natural Sampling

Here we demonstrate the validity of the sampling theorem using the frequency shifting property of the Fourier transform. Although instantaneous sampling is a convenient model, a more practical way of accomplishing the sampling of a bandlimited analog signal $x(t)$ is to multiply $x(t)$, shown in Figure 2.8a, by the pulse train or switching waveform $x_p(t)$, shown in Figure 2.8c. Each pulse in $x_p(t)$ has width T and amplitude $1/T$. Multiplication by $x_p(t)$ can be viewed as the opening and closing of a switch. As before, the sampling frequency is designated f_s, and its reciprocal, the time period between samples, is designated T_s. The resulting sampled-data sequence, $x_s(t)$, is illustrated in Figure 2.8e and is expressed as

$$x_s(t) = x(t)x_p(t) \tag{2.9}$$

The sampling here is termed *natural sampling,* since the top of each pulse in the $x_s(t)$ sequence retains the shape of its corresponding analog segment during the pulse interval. Using Equation (A.13), we can express the periodic pulse train as a Fourier series in the form

$$x_p(t) = \sum_{n=-\infty}^{\infty} c_n e^{j\,2\pi n f_s t} \tag{2.10}$$

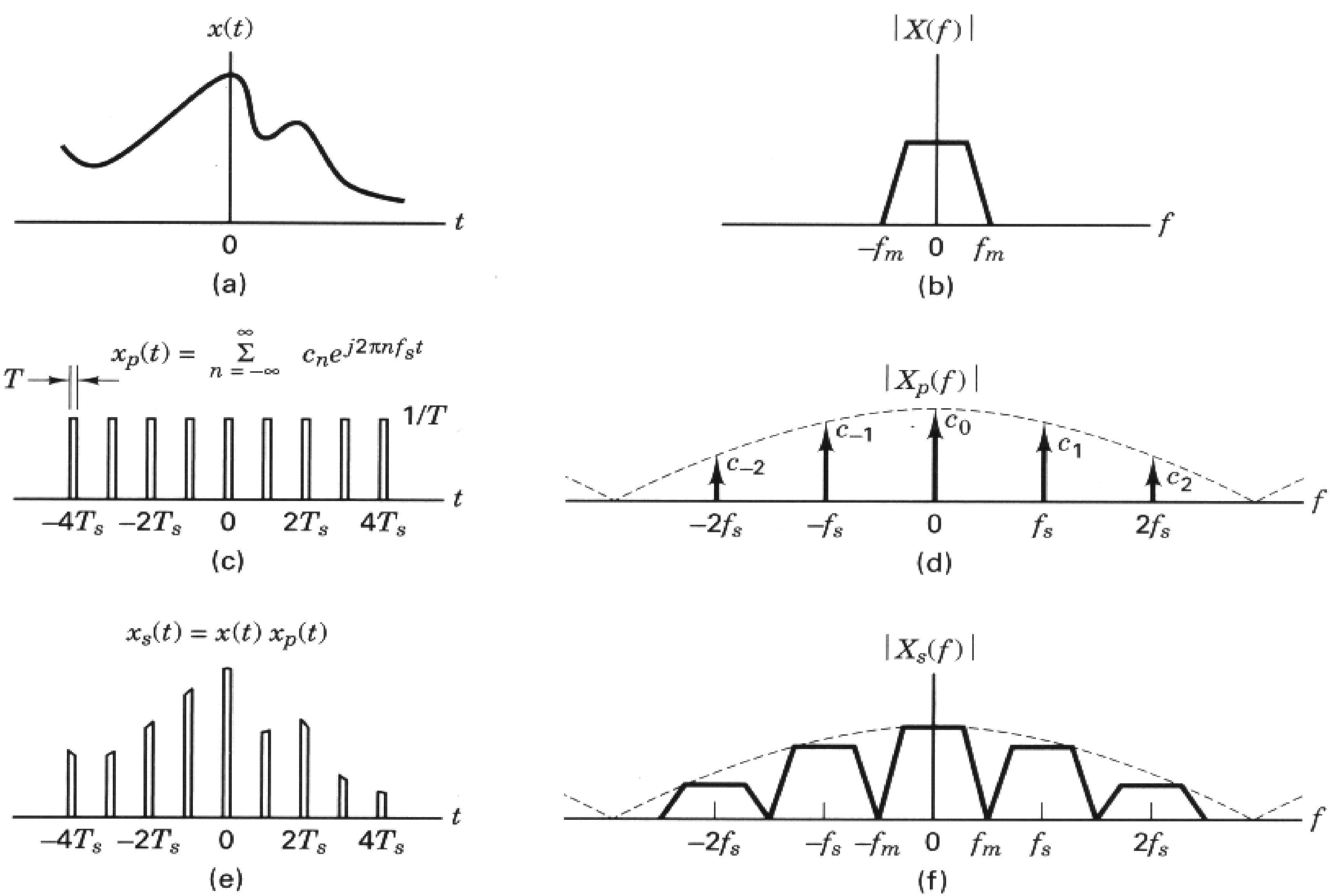


Figure 2.8 Sampling theorem using the frequency shifting property of the Fourier transform.

where the sampling rate, $f_s = 1/T_s$, is chosen equal to $2f_m$, so that the Nyquist criterion is just satisfied. From Equation (A.24), $\boxed{c_n = (1/T_s)\ \text{sinc}\ (nT/T_s)}$, where T is the pulse width, $1/T$ is the pulse amplitude, and

$$\text{sinc}\ y = \frac{\sin \pi y}{\pi y}$$

The envelope of the magnitude spectrum of the pulse train, seen as a dashed line in Figure 2.8d, has the characteristic sinc shape. Combining Equations (2.9) and (2.10) yields

$$x_s(t) = x(t) \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_s t} \tag{2.11}$$

The transform $X_s(f)$ of the sampled waveform is found as follows:

$$\boxed{X_s(f) = \mathcal{F}\left\{x(t) \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_s t}\right\}} \tag{2.12}$$

For linear systems, we can interchange the operations of summation and Fourier transformation. Therefore, we can write

$$X_s(f) = \sum_{n=-\infty}^{\infty} c_n \mathscr{F}\{x(t)e^{j2\pi n f_s t}\} \tag{2.13}$$

Using the *frequency translation* property of the Fourier transform (see Section A.3.2), we solve for $X_s(f)$ as follows:

$$\boxed{X_s(f) = \sum_{n=-\infty}^{\infty} c_n X(f - nf_s)} \tag{2.14}$$

Similar to the unit impulse sampling case, Equation (2.14) and Figure 2.8f illustrate that $X_s(f)$ is a replication of $X(f)$, periodically repeated in frequency every f_s hertz. In this natural-sampled case, however, we see that $X_s(f)$ is weighted by the Fourier series coefficients of the pulse train, compared with a constant value in the impulse-sampled case. It is satisfying to note that *in the limit,* as the pulse width, T, approaches zero, c_n approaches $1/T_s$ for all n (see the example that follows), and Equation (2.14) converges to Equation (2.8).

Example 2.1 Comparison of Impulse Sampling and Natural Sampling

Consider a given waveform $x(t)$ with Fourier transform $X(f)$. Let $X_{s1}(f)$ be the spectrum of $x_{s1}(t)$, which is the result of sampling $x(t)$ with a unit impulse train $x_\delta(t)$. Let $X_{s2}(f)$ be the spectrum of $x_{s2}(t)$, the result of sampling $x(t)$ with a pulse train $x_p(t)$ with pulse width T, amplitude $1/T$, and period T_s. Show that in the limit, as T approaches zero, $X_{s1}(f) = X_{s2}(f)$.

Solution

From Equation (2.8),

$$X_{s1}(f) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} X(f - nf_s)$$

and from Equation (2.14),

$$X_{s2}(f) = \sum_{n=-\infty}^{\infty} c_n X(f - nf_s)$$

As the pulse with $T \to 0$, and the pulse amplitude approaches infinity (the area of the pulse remains unity), $x_p(t) \to x_\delta(t)$. Using Equation (A.14), we can solve for c_n in the limit as follows:

$$\begin{aligned} c_n &= \lim_{T \to 0} \frac{1}{T_s} \int_{-T_s/2}^{T_s/2} x_p(t) e^{-j2\pi n f_s t}\, dt \\ &= \frac{1}{T_s} \int_{-T_s/2}^{T_s/2} x_\delta(t) e^{-j2\pi n f_s t}\, dt \end{aligned}$$

Since, within the range of integration, $-T_s/2$ to $T_s/2$, the only contribution of $x_\delta(t)$ is that due to the impulse at the origin, we can write

$$c_n = \frac{1}{T_s} \int_{-T_s/2}^{T_s/2} \delta(t) e^{-j2\pi n f_s t}\, dt = \frac{1}{T_s}$$

Therefore, in the limit, $X_{s1}(f) = X_{s2}(f)$ for all n.

2.4.1.3 Sample-and-Hold Operation

The simplest and thus most popular sampling method, *sample and hold,* can be described by the convolution of the sampled pulse train, $[x(t)x_\delta(t)]$, shown in Figure 2.6e, with a unity amplitude rectangular pulse $p(t)$ of pulse width T_s. This time, convolution results in the *flattop* sampled sequence

$$\boxed{\begin{aligned} x_s(t) &= p(t) * [x(t)x_\delta(t)] \\ &= p(t) * \left[x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \right] \end{aligned}} \tag{2.15}$$

The Fourier transform, $X_s(f)$, of the time convolution in Equation (2.15) is the frequency-domain product of the transform $P(f)$ of the rectangular pulse and the periodic spectrum, shown in Figure 2.6f, of the impulse-sampled data:

$$\boxed{\begin{aligned} X_s(f) &= P(f)\mathcal{F}\left\{ x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \right\} \\ &= P(f)\left\{ X(f) * \left[\frac{1}{T_s} \sum_{n=-\infty}^{\infty} \delta(f - nf_s) \right] \right\} \\ &= P(f) \frac{1}{T_s} \sum_{n=-\infty}^{\infty} X(f - nf_s) \end{aligned}} \tag{2.16}$$

Here, $P(f)$ is of the form $T_s \operatorname{sinc} fT_s$. The effect of this product operation results in a spectrum similar in appearance to the natural-sampled example presented in Figure 2.8f. The most obvious effect of the hold operation is the significant attenuation of the higher-frequency spectral replicates (compare Figure 2.8f to Figure 2.6f), which is a desired effect. Additional analog postfiltering is usually required to finish the filtering process by further attenuating the residual spectral components located at the multiples of the sample rate. A secondary effect of the hold operation is the nonuniform spectral gain $P(f)$ applied to the desired baseband spectrum shown in Equation (2.16). The postfiltering operation can compensate for this attenuation by incorporating the inverse of $P(f)$ over the signal passband.

2.4.2 Aliasing

Figure 2.9 is a detailed view of the positive half of the baseband spectrum and one of the replicates from Figure 2.7b. It illustrates aliasing in the frequency domain. The overlapped region, shown in Figure 2.9b, contains that part of the spectrum which is aliased due to *undersampling*. The aliased spectral components represent ambiguous data that appear in the frequency band between $(f_s - f_m)$ and f_m. Figure 2.10 illustrates that a higher sampling rate f'_s, can eliminate the aliasing by separating the spectral replicates; the resulting spectrum in Figure 2.10b corresponds to the case in Figure 2.7a. Figures 2.11 and 2.12 illustrate two ways of eliminating aliasing using *antialiasing filters*. In Figure 2.11 the analog signal is *prefiltered* so that the new maximum frequency, f'_m, is reduced to $f_s/2$ or less. Thus there are no aliased components seen in Figure 2.11b, since $f_s > 2f'_m$. Eliminating the aliasing terms prior to sampling is good engineering practice. When the signal structure is well known, the aliased terms can be eliminated after sampling, with a low-pass filter operating on the sampled data [2]. In Figure 2.12 the aliased components are removed by *postfiltering* after sampling; the filter cutoff frequency, f''_m, removes the aliased components; f''_m needs to be less than $(f_s - f_m)$. Notice that the filtering techniques for eliminating the aliased portion of the spectrum in Figures 2.11 and 2.12 *will result in a loss* of some of the signal information. For this reason, the sample rate, cutoff bandwidth, and filter type selected for a particular signal bandwidth are all interrelated.

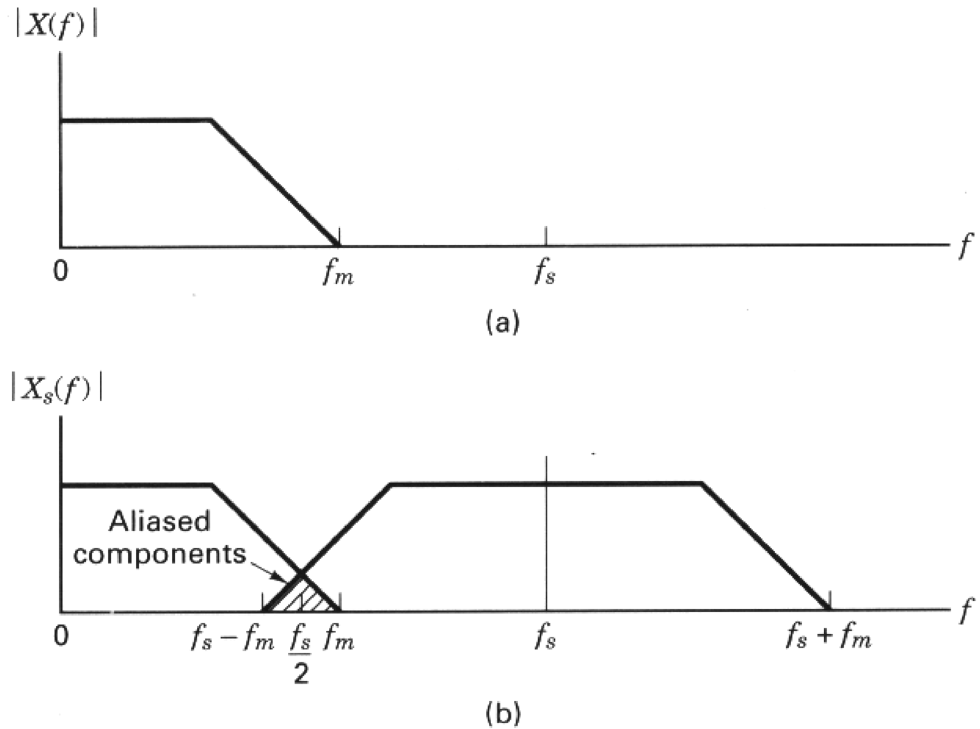


Figure 2.9 Aliasing in the frequency domain. (a) Continuous signal spectrum. (b) Sampled signal spectrum.

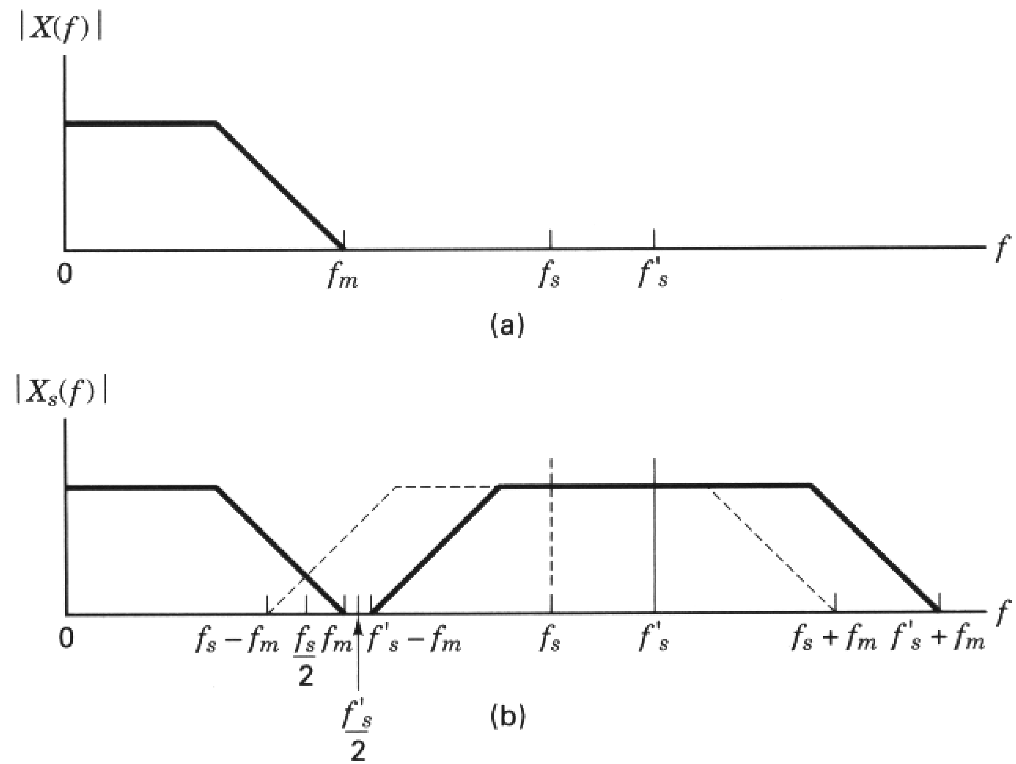


Figure 2.10 Higher sampling rate eliminates aliasing. (a) Continuous signal spectrum. (b) Sampled signal spectrum.

Realizable filters require a nonzero bandwidth for the transition between the passband and the required out-of-band attenuation. This is called the *transition bandwidth*. To minimize the system sample rate, we desire that the antialiasing filter have a small transition bandwidth. Filter complexity and cost rise sharply with narrower transition bandwidth, so a trade-off is required between the cost of a small transition bandwidth and the costs of the higher sampling rate, which are those of more storage and higher transmission rates. In many systems the answer has been to make the transition bandwidth between 10 and 20% of the signal band-

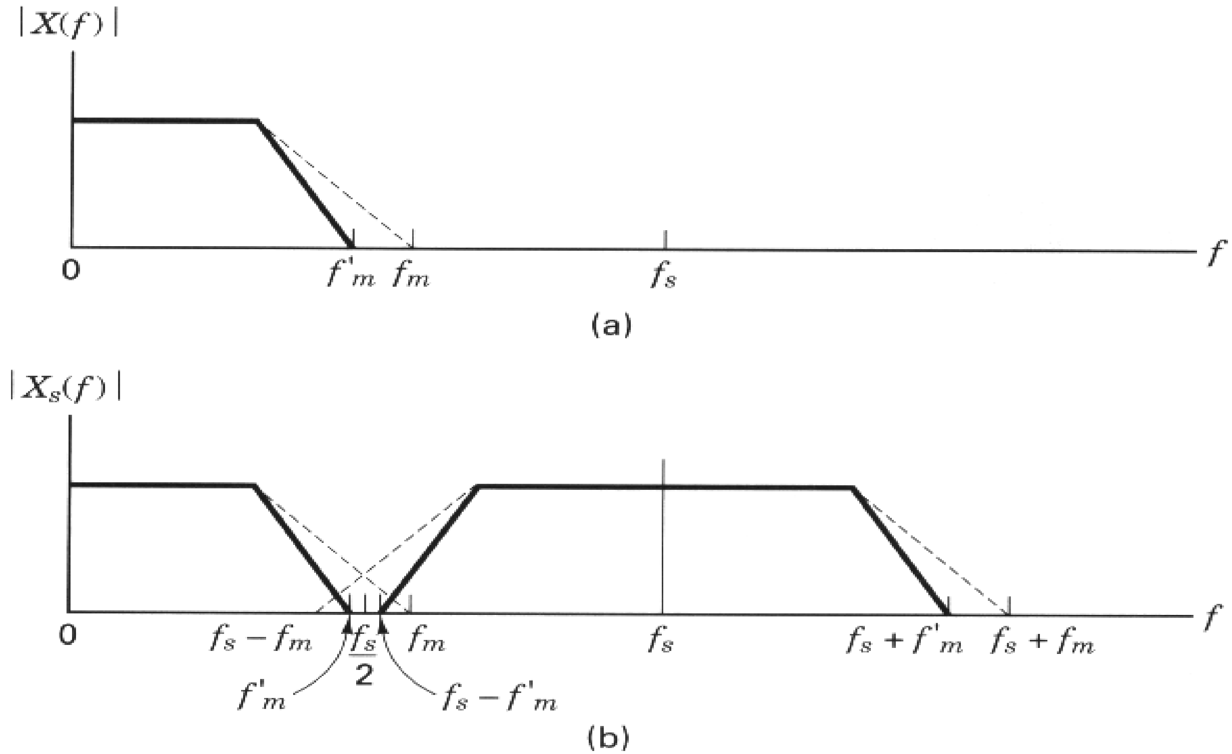


Figure 2.11 Sharper-cutoff filters eliminate aliasing. (a) Continuous signal spectrum. (b) Sampled signal spectrum.

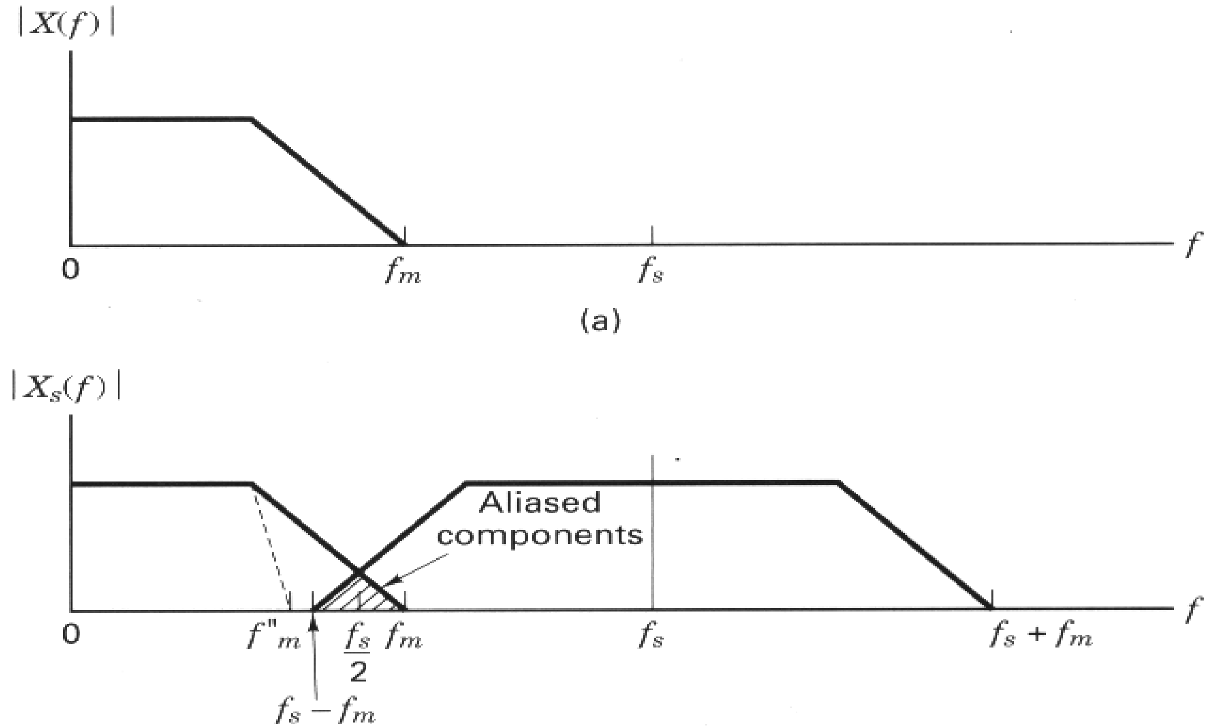


Figure 2.12 Postfilter eliminates aliased portion of spectrum. (a) Continuous signal spectrum. (b) Sampled signal spectrum.

width. If we account for the 20% transition bandwidth of the antialiasing filter, we have an *engineer's version* of the Nyquist sampling rate:

$$f_s \geq 2.2 f_m \tag{2.17}$$

Figure 2.13 provides some insight into aliasing as seen in the time domain. The sampling instants of the solid-line sinusoid have been chosen so that the sinusoidal signal is undersampled. Notice that the resulting ambiguity allows one to draw a totally different (dashed-line) sinusoid, following the undersampled points.

2.4.4 Signal Interface for a Digital System

Let us examine four ways in which analog source information can be described. Figure 2.14 illustrates the choices. Let us refer to the waveform in Figure 2.14a as the *original analog waveform.* Figure 2.14b represents a sampled version of the original waveform, typically referred to as *natural-sampled data* or PAM (pulse amplitude modulation). Do you suppose that the sampled data in Figure 2.14b are compatible with a digital system? No, they are not, because the amplitude of each natural sample still has an infinite number of possible values; a digital system deals with a finite number of values. Even if the sampling is flat-top sampling, the possible pulse values form an infinite set, since they reflect all the possible values of the continuous analog waveform. Figure 2.14c illustrates the original waveform represented by discrete pulses. Here the pulses have flat tops *and* the pulse amplitude values are limited to a finite set. Each pulse is expressed as a level from a finite number of predetermined levels; each such level can be represented by a symbol from a finite alphabet. The pulses in Figure 2.14c are referred to as *quantized samples;* such a format is the obvious choice for interfacing with a digital system. The format in Figure 2.14d may be construed as the output of a sample-and-hold circuit. When the sample values are quantized to a finite set, this format can also interface with a digital system. After quantization, the analog waveform can still be recovered, but not precisely; improved reconstruction fidelity of the analog waveform can be achieved by increasing the number of quantization levels (requiring increased system bandwidth). Signal distortion due to quantization is treated in the following sections (and later in

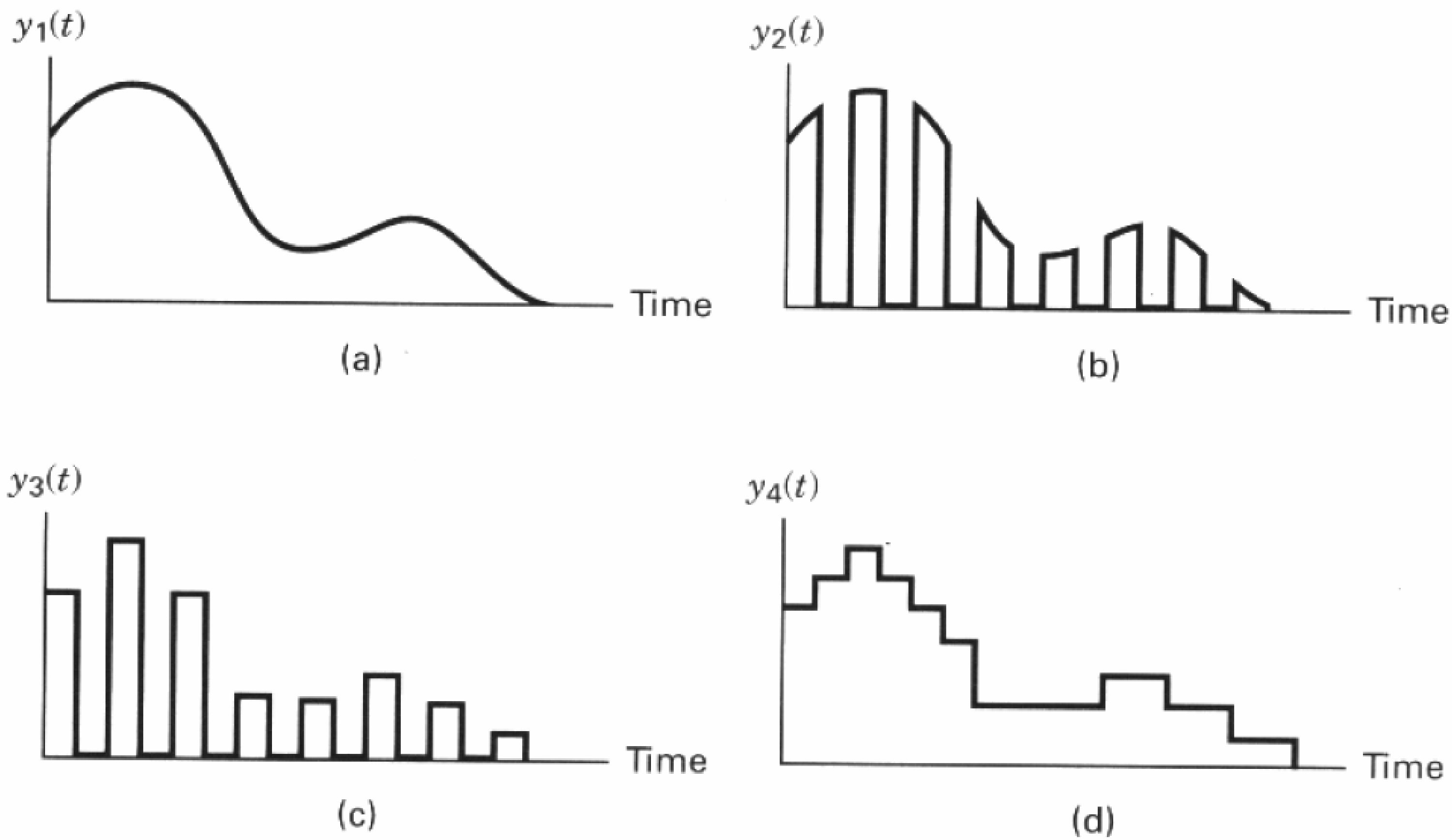
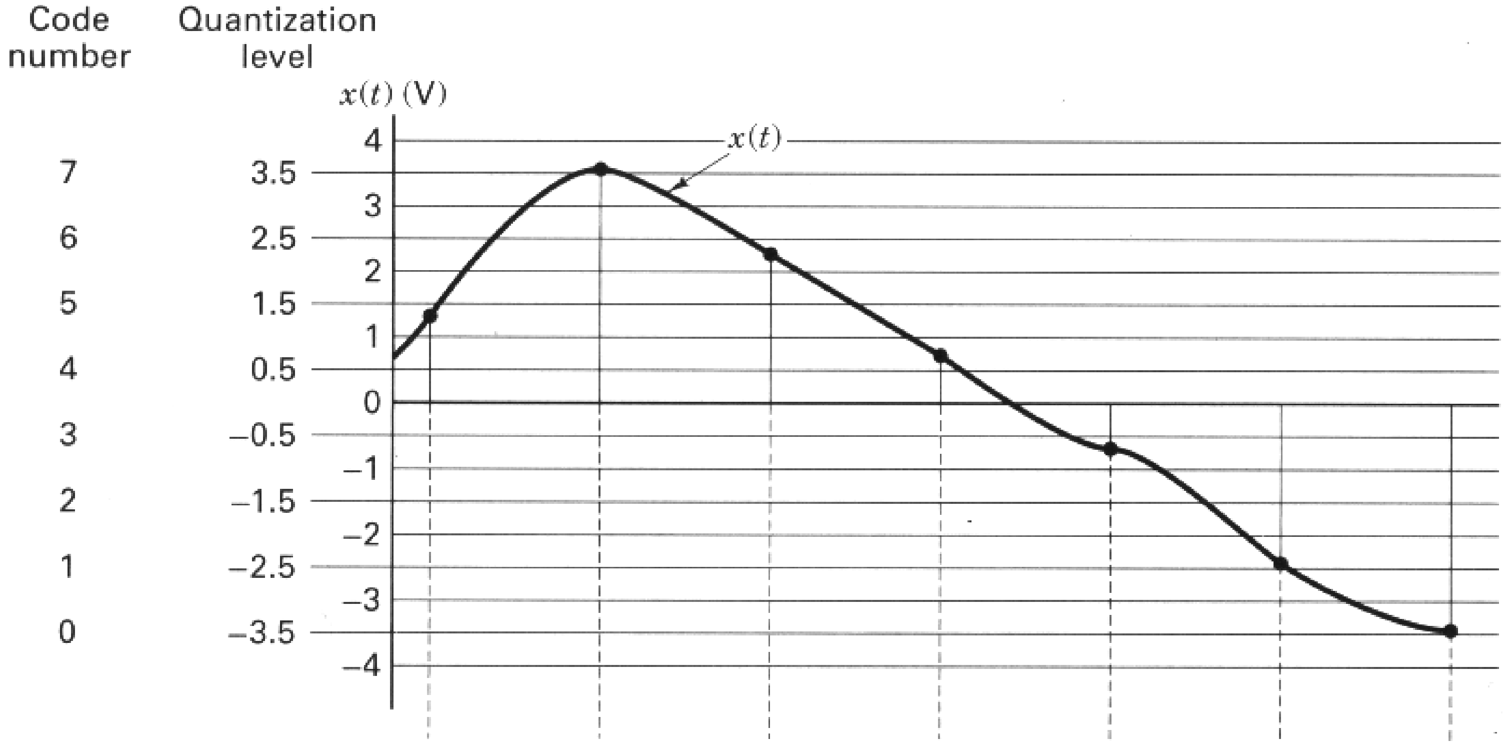


Figure 2.14 Amplitude and time coordinates of source data. (a) Original analog waveform. (b) Natural-sampled data. (c) Quantized samples. (d) Sample and hold.

2.6 PULSE CODE MODULATION

Pulse code modulation (PCM) is the name given to the class of baseband signals obtained from the quantized PAM signals by encoding each quantized sample into a *digital word* [3]. The source information is sampled and quantized to one of L levels; then each quantized sample is digitally encoded into an ℓ-bit ($\ell = \log_2 L$) codeword. For baseband transmission, the codeword bits will then be transformed to pulse waveforms. The essential features of binary PCM are shown in Figure 2.16. Assume that an analog signal $x(t)$ is limited in its excursions to the range –4 to +4 V. The step size between quantization levels has been set at 1 V. Thus, eight quantization levels are employed; these are located at –3.5, –2.5, . . . , +3.5 V. We assign the code number 0 to the level at –3.5 V, the code number 1 to the level at –2.5 V, and so on, until the level at 3.5 V, which is assigned the code number 7. Each code number has its representation in binary arithmetic, ranging from 000 for code number 0 to 111 for code number 7. Why have the voltage levels been chosen in this manner, compared with using a sequence of consecutive integers, 1, 2, 3, . . . ? The choice of voltage levels is guided by two constraints. First, the quantile intervals between the levels should be equal; and second, it is convenient for the levels to be symmetrical about zero.



Natural sample value	1.3	3.6	2.3	0.7	−0.7	−2.4	−3.4
Quantized sample value	1.5	3.5	2.5	0.5	−0.5	−2.5	−3.5
Code number	5	7	6	4	3	1	0
PCM sequence	101	111	110	100	011	001	000

Figure 2.16 Natural samples, quantized samples, and pulse code modulation. (Reprinted with permission from Taub and Schilling, *Principles of Communications Systems,* McGraw-Hill Book Company, New York, 1971, Fig. 6.5-1, p. 205.)

The ordinate in Figure 2.16 is labeled with quantization levels and their code numbers. Each sample of the analog signal is assigned to the quantization level closest to the value of the sample. Beneath the analog waveform $x(t)$ are seen four representations of $x(t)$, as follows: the natural sample values, the quantized sample values, the code numbers, and the PCM sequence.

Note, that in the example of Figure 2.16, each sample is assigned to one of eight levels or a three-bit PCM sequence. Suppose that the analog signal is a musical passage, which is sampled at the Nyquist rate. And, suppose that when we listen to the music in digital form, it sounds terrible. What could we do to improve the fidelity? Recall that the process of quantization replaces the true signal with an approximation (i.e., adds quantization noise). Thus, increasing the number of levels will reduce the quantization noise. If we double the number of levels to 16, what are the consequences? In that case, each analog sample will be represented as a four-bit PCM sequence. Will that cost anything? In a real-time communication system, the messages must not be delayed. Hence, the transmission time for each sample must be the same, regardless of how many bits represent the sample. Hence, when there are more bits per sample, the bits must move faster; in other words, they must be replaced by "skinnier" bits. The data rate is thus increased, and the cost is a greater transmission bandwidth. This explains how one can generally obtain better fidelity at the cost of more transmission bandwidth.

2.7 UNIFORM AND NONUNIFORM QUANTIZATION

2.7.1 Statistics of Speech Amplitudes

Speech communication is a very important and specialized area of digital communications. Human speech is characterized by unique statistical properties; one such property is illustrated in Figure 2.17. The abscissa represents speech signal magnitudes, normalized to the root-mean-square (rms) value of such magnitudes through a typical communication channel, and the ordinate is probability. For most voice

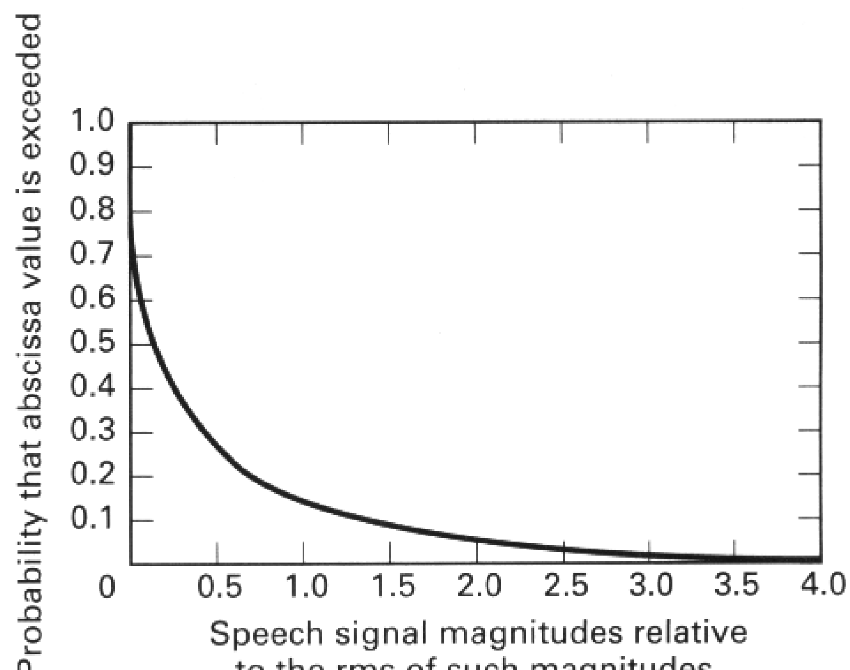


Figure 2.17 Statistical distribution of single-talker speech signal magnitudes.

communication channels, very low speech volumes predominate; 50% of the time, the voltage characterizing detected speech energy is less than one-fourth of the rms value. Large amplitude values are relatively rare; only 15% of the time does the voltage exceed the rms value. We see from Equation (2.18b) that the quantization noise depends on the step size (size of the quantile interval). When the steps are uniform in size the quantization is known as *uniform quantization.* Such a system would be wasteful for speech signals; many of the quantizing steps would rarely be used. In a system that uses equally spaced quantization levels, the quantization noise is the same for all signal magnitudes. Therefore, with uniform quantization, the signal-to-noise (SNR) is worse for low-level signals than for high-level signals. *Nonuniform quantization* can provide fine quantization of the weak signals and coarse quantization of the strong signals. Thus in the case of nonuniform quantization, quantization noise can be made proportional to signal size. The effect is to improve the overall SNR by reducing the noise for the predominant weak signals, at the expense of an increase in noise for the rarely occurring strong signals. Figure 2.18 compares the quantization of a strong versus a weak signal for uniform and nonuniform quantization. The staircase-like waveforms represent the approximations to the analog waveforms (after quantization distortion has been introduced). The SNR improvement that nonuniform quantization provides for the weak signal should be apparent. Nonuniform quantization can be used to make the SNR a constant for all signals within the input range. For voice signals, the typical input signal

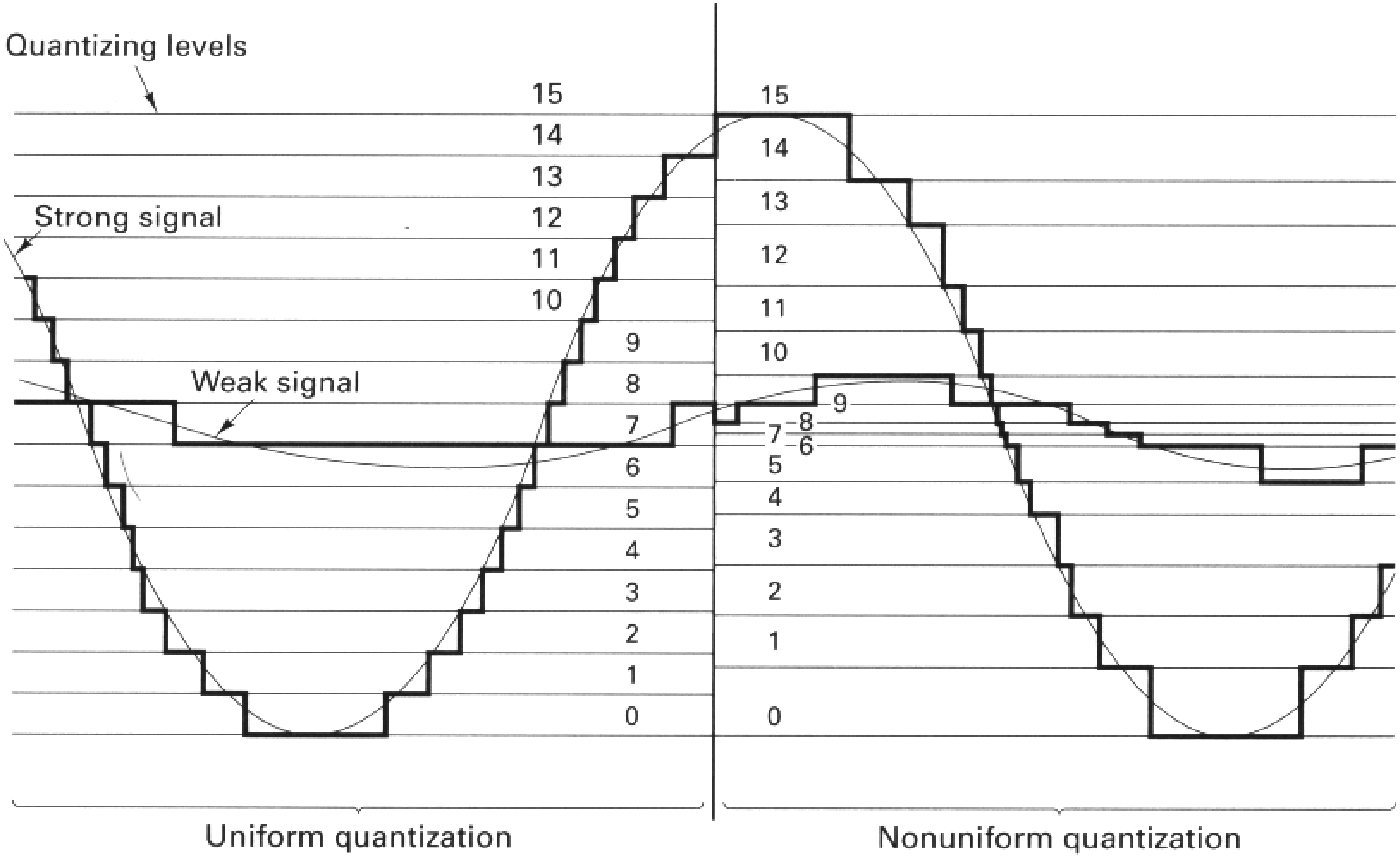


Figure 2.18 Uniform and nonuniform quantization of signals.

2.8.2 PCM Waveform Types

When pulse modulation is applied to a *binary* symbol, the resulting binary waveform is called a pulse-code modulation (PCM) waveform. There are several types of PCM waveforms that are described below and illustrated in Figure 2.22; in telephony applications, these waveforms are often called *line codes.* When pulse modulation is applied to a *nonbinary* symbol, the resulting waveform is called an M-ary

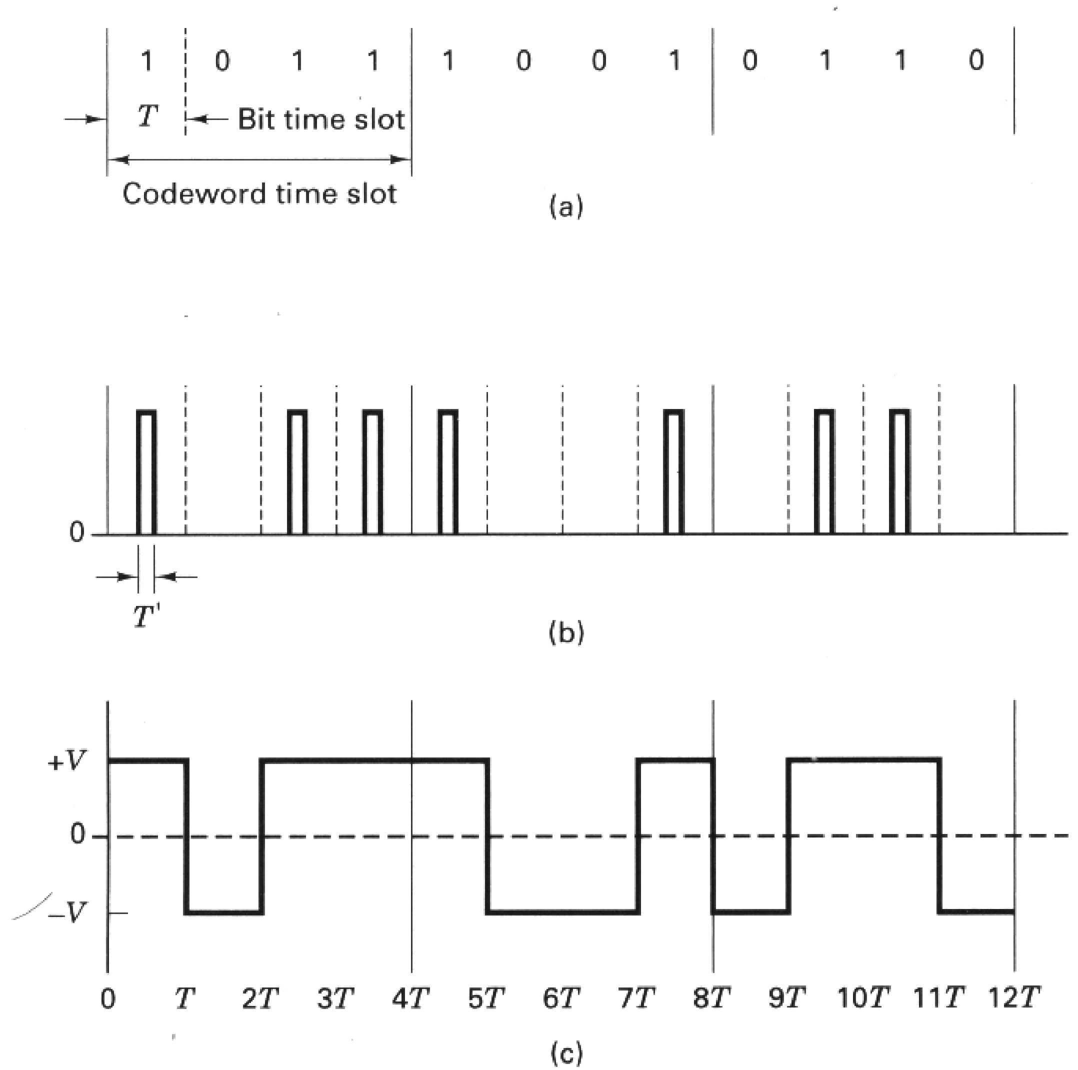


Figure 2.21 Example of waveform representation of binary digits. (a) PCM sequence. (b) Pulse representation of PCM. (c) Pulse waveform (transition between two levels).

pulse-modulation waveform, of which there are several types. They are described in Section 2.8.5, where one of them, called pulse-amplitude modulation (PAM), is emphasized. In Figure 2.1, the highlighted block, labeled *Baseband Signaling*, shows the basic classification of the PCM waveforms and the M-ary pulse waveforms. The PCM waveforms fall into the following four groups.

1. Nonreturn-to-zero (NRZ)
2. Return-to-zero (RZ)
3. Phase encoded
4. Multilevel binary

The NRZ group is probably the most commonly used PCM waveform. It can be partitioned into the following subgroups: NRZ-L (L for level), NRZ-M (M for mark), and NRZ-S (S for space). NRZ-L is used extensively in digital logic circuits.

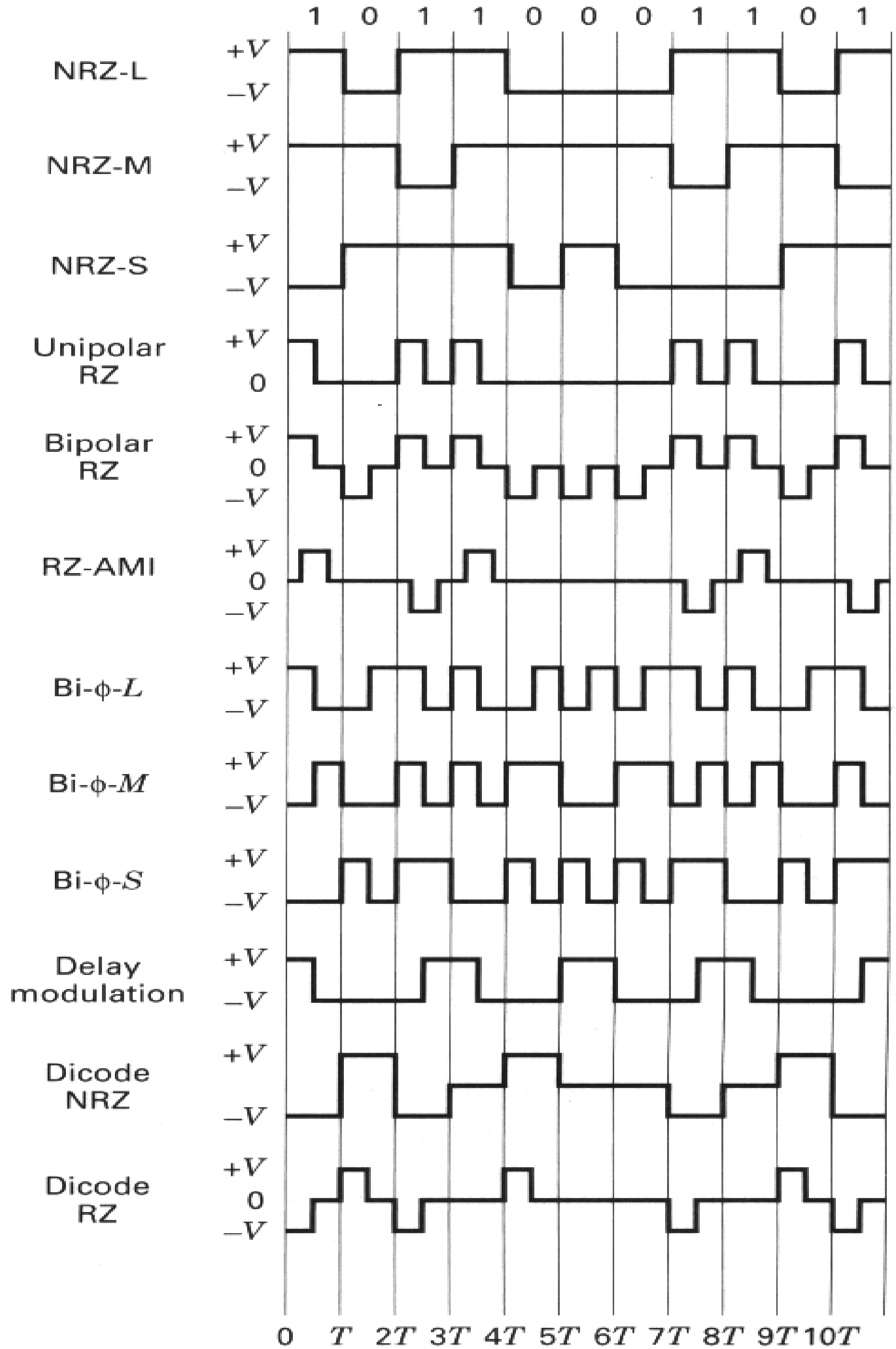


Figure 2.22 Various PCM waveforms.

A binary one is represented by one voltage level and a binary zero is represented by another voltage level. There is a change in level whenever the data change from a one to a zero or from a zero to a one. With NRZ-M, the one, or *mark,* is represented by a change in level, and the zero, or *space,* is represented by no change in level. This is often referred to as *differential encoding.* NRZ-M is used primarily in magnetic tape recording. NRZ-S is the complement of NRZ-M: A one is represented by no change in level, and a zero is represented by a change in level.

The RZ waveforms consist of unipolar-RZ, bipolar-RZ, and RZ-AMI. These codes find application in baseband data transmission and in magnetic recording. With unipolar-RZ, a one is represented by a half-bit-wide pulse, and a zero is represented by the absence of a pulse. With bipolar-RZ, the ones and zeros are represented by opposite-level pulses that are one-half bit wide. There is a pulse present in each bit interval. RZ-AMI (AMI for "alternate mark inversion") is a signaling scheme used in telephone systems. The ones are represented by equal-amplitude alternating pulses. The zeros are represented by the absence of pulses.

The phase-encoded group consists of bi-ϕ-L (bi-phase-level), better known as *Manchester coding;* bi-ϕ-M (bi-phase-mark); bi-ϕ-S (bi-phase-space); and *delay modulation* (DM), or *Miller coding.* The phase-encoding schemes are used in magnetic recording systems and optical communications and in some satellite telemetry links. With bi-ϕ-L, a one is represented by a half-bit-wide pulse positioned during the first half of the bit interval; a zero is represented by a half-bit-wide pulse positioned during the second half of the bit interval. With bi-ϕ-M, a transition occurs at the beginning of every bit interval. A one is represented by a second transition one-half bit interval later; a zero is represented by no second transition. With bi-ϕ-S, a transition also occurs at the beginning of every bit interval. A one is represented by no second transition; a zero is represented by a second transition one-half bit interval later. With delay modulation [4], a one is represented by a transition at the midpoint of the bit interval. A zero is represented by no transition, unless it is followed by another zero. In this case, a transition is placed at the end of the bit interval of the first zero. Reference to the illustration in Figure 2.22 should help to make these descriptions clear.

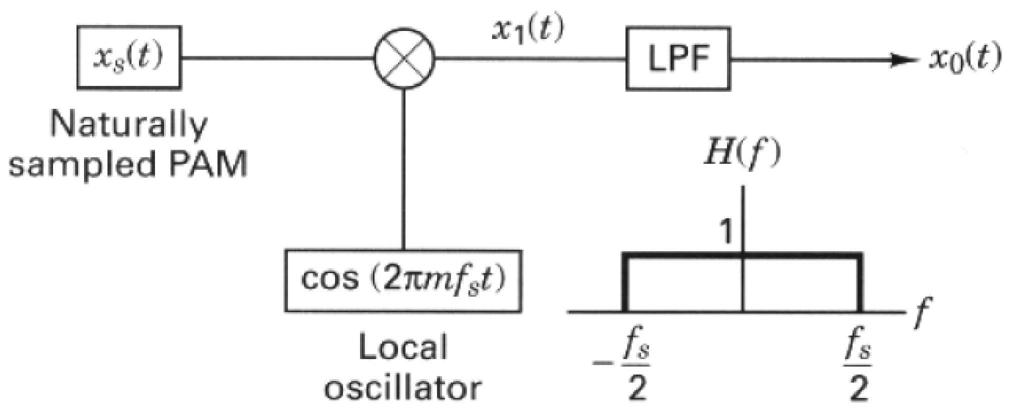
Many binary waveforms use three levels, instead of two, to encode the binary data. Bipolar RZ and RZ-AMI belong to this group. The group also contains formats called *dicode* and *duobinary*. With dicode-NRZ, the one-to-zero or zero-to-one data transition changes the pulse polarity; without a data transition, the zero level is sent. With dicode-RZ, the one-to-zero or zero-to-one transition produces a half-duration polarity change; otherwise, a zero level is sent. The three-level duobinary signaling scheme is treated in Section 2.9.

One might ask why there are so many PCM waveforms. Are there really so many unique applications necessitating such a variety of waveforms to represent digits? The reason for the large selection relates to the differences in performance that characterize each waveform [5]. In choosing a PCM waveform for a particular application, some of the parameters worth examining are the following:

1. *Dc component.* Eliminating the dc energy from the signal's power spectrum enables the system to be ac coupled. Magnetic recording systems, or systems using transformer coupling, have little sensitivity to very low frequency signal components. Thus low-frequency information could be lost.
2. *Self-Clocking.* Symbol or bit synchronization is required for any digital communication system. Some PCM coding schemes have inherent synchronizing

HW2
Due date March 17th

1. Given an analog waveform that has been sampled at its Nyquist rate, f_s, using natural sampling, prove that a waveform (proportional to the original waveform) can be recovered from the samples, using the recovery techniques shown in Figure P2.1. The parameter mf_s is the frequency of the local oscillator, where m is an integer.



2. Determine the number of quantization levels that are implied if the number of bits per sample in a given PCM code is (a) 5; (b) 8; (c) x.

3. Determine the minimum sampling rate necessary to sample and perfectly reconstruct the signal $x(t) = \sin(6280t)/(6280t)$.

4. **(a)** A waveform that is bandlimited to 50 kHz is sampled every 10 μs. Show graphically that these samples uniquely characterize the waveform. (Use a sinusoidal example for simplicity. Avoid sampling at points where the waveform equals zero.)

   (b) If samples are taken 30 μs apart instead of 10 μs, show graphically that waveforms other than the original can be characterized by the samples.