Chapter 8
PULSE MODULATION AND TIME-DIVISION MULTIPLEXING

8-1 Introduction and Objectives

All modulated waveforms considered thus far in the text have been continuous-time signals or, in more casual terminology, analog modulated signals. We will now begin the study of the classes of modulation that involve pulse and/or digital modulation techniques. Pulse and digital modulation signals are characterized by a series of samples of the signal rather than the complete signal. For pulse modulation techniques, the samples are encoded as the amplitude, width, or relative position of a sequence of pulse. For digital modulation techniques, the samples are encoded as digital numbers. The resulting signals are then defined only at discrete values of time, and portions of the signal are obviously "missed" by this process. However, it will be shown that if the signal is band-limited, and if a certain minimum number of samples are taken, the resulting signal can theoretically be completely reconstructed at the receiver.

The next several chapters will be devoted exclusively to those methods usually designated as "digital" communications, while this chapter will deal with those that are more commonly designated as "pulse" communications. However, both general categories have the same theoretical basis, and that background will be established in this chapter.

Objectives

After completing this chapter, the reader should be able to

1. Discuss the concept of a sampled-data signal and sketch its form.
2. Display the form of the spectrum of sampled signal.
3. State the sampling theorem and discuss its significance.
4. Define aliasing and discuss its effects.
5. Define the Nyquist rate and determine the folding frequency.
6. Discuss the form of an ideal impulse sampled signal and display the form of its spectrum.
7. Discuss the concept of a pulse-amplitude modulated (PAM) signal.
8. Compare the properties of a natural-sampled PAM signal and a flat-top PAM signal.
9. Discuss the process of the reconstruction of a PAM signal, including the use of a holding circuit.
10. Discuss the concept of time-division multiplexing (TDM) and show how it is achieved.
11. Discuss the forms of pulse-time modulation including pulse-width modulation (PWM) and pulse-position modulation (PPM).
8-2 Sampling Theorem

The starting point for all pulse and digital methods is the concept of a sampled-data signal. A sampled-data signal (also called a sampled signal) is one that consists of a regular sequence of encoded samples of a reference continuous-time or analog signal.

Consider the arbitrary analog signal \( v(t) \) shown in Figure 8-1(a). The form of a sampled data signal \( v_s(t) \) representing the analog signal is shown in Figure 8-1(c). The sampled signal is obtained by observing \( v(t) \) during short intervals of time of width \( \delta \) seconds. The sampling rate is designated as \( f_s \) and is given by

\[
 f_s = \frac{1}{T}
\]

where \( T \) is the sampling period and is the time between the beginning of one sample and the beginning of the next sample. During the interval between samples, \( v(t) \) is not observed at all. However, samples of other signals could be inserted in the "open space" as will be discussed later.

Note that since the sampling rate is a frequency, it is proper to express it in hertz. One hertz is interpreted as one sample per second in this context. While hertz will generally be employed for sampling rates, the term "samples per second" will be used when it adds clarity to a discussion.

The form of the signal generated by the process shown in Figure 8-1 can be described as a nonzero width, natural sampled signal. The "non-zero width" description represents the fact that the sampling pulses occupy some width, however small, and this is always the case with actual samples of analog signals. However, the concept of impulse sampling, in which the samples are assumed to have zero width, is very useful in certain analytical developments and will be considered in the next section. The "natural sampled" description refers to the fact that the top of each sample in Figure 8-1 follows the analog signal during the short interval \( \delta \). This form is the easiest to analyze mathematically and serves as a starting point for other forms.

It is convenient to express \( v_s(t) \) as the product of \( v(t) \) and the periodic pulse train \( p(t) \) shown in Figure 8-1(b). Since the pulses assume only the values 0 and 1, multiplication of \( v(t) \) by \( p(t) \) is equivalent to the sampling operation shown; i. e.

\[
 v_s(t) = v(t) p(t)
\]

Spectrum of Sampled Signal

It is very important to understand the form of the spectrum of the sampled-data signal. We will assume a baseband signal for the purpose of this development. An arbitrary baseband amplitude spectrum \( V(f) \) is shown in Figure 8-2(a). The spectrum of the periodic baseband pulse train \( p(t) \) is shown in Figure 8-2(b). This latter spectrum contains a dc component, a fundamental component at \( f_s = 1/T \), and an infinite number of harmonics at integer multiples of the sampling frequency. Based on the work of Chapter 2, the exponential form of the Fourier series of the pulse train can be expressed as

\[
p(t) = \sum_{n=-\infty}^{\infty} \tilde{P}_n e^{j2\pi nf}\]

The amplitude spectrum corresponding to \( \tilde{P}_n \) is given by
\[ \hat{P}_n = d \frac{\sin n\pi d}{n\pi d} \]  
(8-4)

where \( d = \tau / T \) is the duty cycle.

With the Fourier series form of (8-3) substituted in (8-2), the sampled signal can be expressed as

\[ v_s(t) = v(t) \sum_{-\infty}^{\infty} \hat{P}_n e^{jn\omega_0 t} = \sum_{-\infty}^{\infty} \hat{P}_n v(t) e^{jn\omega_0 t} \]  
(8-5)

Fourier transformation of equation (8-5) utilizing the modulation theorem yields

\[ \tilde{V}_s(f) = \sum_{-\infty}^{\infty} \hat{P}_n \tilde{V}(f - nf_s) \]  
(8-6)

The form and relative magnitude (but not necessarily actual magnitude levels) of \( \tilde{V}_s(f) \) are illustrated in Figure 8-2(c). (Only a limited portion of the negative frequency range is shown.) The spectrum of the sampled-data signal consists of the original baseband spectrum plus an infinite number of shifted or translated versions of the original spectrum. The translated components are shifted in frequency by increments equal to the sampling frequency and its harmonics. The magnitudes of the spectral components are multiplied by the \( P_n \) coefficients and they gradually diminish with increasing frequency. However, for a very short duty cycle (\( \tau \ll T \)), the relative magnitudes of the components diminish very slowly and the overall spectrum can be quite broad in form. (The actual magnitudes are small for a short duty cycle, but the relative magnitudes remain nearly the same for a wide frequency range.)

From the nature of the spectrum of the sampled signal, it is clear that the form of the original baseband spectrum is preserved in the component corresponding to \( n = 0 \) in (8-6). This component is multiplied by \( P_0 \), which may be small compared with unity if \( d \ll 1 \), so the level may be reduced. However, since the shape is preserved, the form of the spectrum is maintained in the frequency range from dc to \( W \) provided that there is no overlap in the spectrum from any of the other components.

**Minimum Sampling Rate**

An intuitive explanation of the sampling theorem will now be developed. From Figure 8-2(c), it is observed that the original spectrum extends from dc to \( W \) Hz, while the lowest portion of the first shifted component is at a frequency \( f_s - W \). To be able to recover the original signal from the sampled signal, it is necessary that no portion of the first translated component overlap the original spectrum, which means that

\[ f_s - W \geq W \]  
(8-7)

or

\[ f_s \geq 2W \]  
(8-8)

Equation (8-8) is a statement of the **sampling theorem**, which serves as the foundation of all sampled-data, pulse, and digital signal and modulation systems. In words, it states that a baseband signal must be uniformly sampled at a rate at least as high as twice the highest frequency in a spectrum in order to be recoverable by direct low-pass filtering. There are some special ways in which the strict requirement can be relaxed for band-pass signals, but for the more common baseband signal case, the requirement stated in (8-8) will be assumed.
Although the "greater than or equal" statement of (8-8) has theoretical significance, in actual applications of the theorem for baseband signals, a sampling rate somewhat greater than the theoretical minimum is employed. The reason for this can be readily deduced from Figure 8-2(c) by noting that if \( f_s = 2W \), there is no gap in the spectrum between the original spectrum and the first translated component, and a perfect block filter will be required to separate the components. Thus, some frequency interval, called a guard band, should be provided in order that the first translated component (as well as all higher order components) can be rejected by a realistic filter. A typical example of actual sampling rates is that of some commercial voice-grade telephone sampling systems. Based on an assumed value of \( W = 3.3 \text{ kHz} \), a sampling rate \( f_s = 8 \text{ kHz} \) is used, while the theoretical minimum would be \( f_s = 6.6 \text{ kHz} \).

**Alternate Form**

An alternate form of the sampling theorem can be stated in terms of the time interval between successive samples. Let \( T = 1/f_s \) represent the time interval between successive samples. The alternate form is obtained by taking the reciprocal of both sides of (8-8). Because this is an inequality, however, the sense of the inequality is reversed and we have

\[
T \leq \frac{1}{2W} \quad (8-9)
\]

Stated in words, this form indicates that the spacing between samples can be no greater than \( 1/(2W) = 0.5/W \). Both forms are useful.

Although in practice, it is necessary to sample at a rate somewhat greater than the theoretical minimum (or equivalently, have the time interval between samples to be less than the theoretical maximum), some of the theoretical developments that follow will employ the minimum theoretical sample rate for convenience. This theoretical minimum rate \( 2W \) is called the Nyquist rate.

**Aliasing**

If the actual sampling rate is less than the theoretical minimum, some of the spectral components will overlap as illustrated by Figure 8-3. In this case, components of the original spectrum appear at the same locations as other components and cannot be uniquely determined or separated. This process is called aliasing, and as the name implies, spectral components may appear to be different than what they really are.

A convenient definition in sampling theory is the folding frequency \( f_0 \). It is given by

\[
f_0 = \frac{f_s}{2} = \frac{1}{2T} \quad (8-10)
\]

The folding frequency is simply the highest theoretical frequency that can be processed by a sampled-data system with a sampling rate \( f_s \) without aliasing. Thus, the Nyquist rate is twice the folding frequency.

It is important in a sampled-data system that the minimum sampling rate be satisfied for all spectral components in a given signal, even though some may be out of the frequency range of interest. One might erroneously believe that as long as the minimum sampling rate is met for all components of interest, that the desired portion of the signal could be reconstructed. However, if there are unwanted components having a higher frequency than the folding frequency, aliasing will occur, and portions of the unwanted spectrum will fold into the desired spectral range. This concept will be illustrated in Example 8-4.
Since it is not always possible to predict precisely the upper frequency limit of a given complex signal, a common practice employed in sampled-data baseband systems is to pass the analog signal through a low-pass filter prior to sampling. Such a filter is called an \textit{anti-aliasing filter} and should have a cutoff frequency less than, or at least no greater than, the folding frequency.

**Total Number of Samples**

Let \( N \) represent the total number of samples required to reproduce a signal of duration \( t_p \). Since there are \( f_s \) samples per second, the total number required is

\[
N = f_s t_p = \frac{t_p}{T}
\]  

(8-11)

where \( t_p \) and \( T \) are expressed in the same units and \( t_p \) is expressed in seconds if \( f_s \) is expressed in hertz or samples per second.
Example 8-1

A baseband signal has frequency components from dc to 5 kHz. Determine the (a) theoretical minimum sampling rate and (b) maximum time interval between successive samples.

Solution

(a) The minimum sampling rate is

\[ f_s = 2W = 2 \times 5 \text{ kHz} = 10 \text{ kHz} \]  \hspace{1cm} \text{(8-12)}

Thus, it is necessary to obtain no less than 10,000 samples per second.

(b) The maximum interval between samples is

\[ T = \frac{1}{f_s} = \frac{1}{10,000} = 1 \times 10^{-4} \text{ s} = 100 \mu \text{s} \]  \hspace{1cm} \text{(8-13)}

Example 8-2

To provide some guard band, assume that the signal in Example 8-1 is sampled at a rate 25% above the theoretical minimum. Determine the (a) the sampling rate and (b) time interval between successive samples.

Solution

(a) The sampling rate will now be set at

\[ f_s = 1.25 \times 2W = 1.25 \times 10,000 = 12,500 \text{ Hz} = 12.5 \text{ kHz} \]  \hspace{1cm} \text{(8-14)}

(b) The time interval between samples is

\[ T = \frac{1}{f_s} = \frac{1}{12,500} = 80 \times 10^{-6} \text{ s} = 80 \mu \text{s} \]  \hspace{1cm} \text{(8-15)}

Example 8-3

For the signal of Examples 8-1 and 8-2, and with the more practical sampling rate of Example 8-2, assume that the signal has a duration of 30 minutes. Determine the total number of samples that must be taken.

Solution

Assuming the most basic units, the sampling rate is 12,500 samples per second. Therefore, the total duration must be expressed in seconds, and that value is 30 minutes x 60 seconds/minute = 1800 seconds. Hence,

\[ N = f_s t_p = 12500 \times 1800 = 22.5 \times 10^6 \text{ samples} \]  \hspace{1cm} \text{(8-16)}
Example 8-4

A signal \( v(t) \) consists of two components \( v_1(t) \) and \( v_2(t) \) with \( v = v_1 + v_2 \). The portion of interest is \( v_1 \), and its amplitude spectrum \( V_1(f) \) is band-limited from dc to 800 Hz as shown in Figure 8-4(a). The amplitude spectrum \( V_2(f) \) of the second component occupies a spectrum from 1200 to 1500 Hz as shown. It is necessary to sample the signal so that it can be processed by a pulse or digital modulation system. An inexperienced young engineer believes that since the component of interest extends only to 800 Hz, a sampling rate greater than 1600 samples per second should suffice. To provide some guard band, he or she selects a sampling rate of 2000 samples per second. (a) Show by constructing a spectral diagram for the sampled signal that there is a fallacy in the reasoning and that it will not be possible to reconstruct the signal. (b) What steps can be taken to rectify the problem?

Solution

(a) Let \( v_s(t) \) represent the sampled version of the analog signal \( v(t) \). The amplitude spectrum \( V_s(f) \) is obtained by sketching the form of the original spectrum plus successive shifts on the original spectrum as indicated by (8-6). In most cases, enough information about the form of the composite spectrum can be deduced from one or two shifted components.

The forms of the baseband and sampled spectra over a reasonable frequency range are shown in Figure 8-4(b). (The amplitude scales of (a) and (b) are different.) Observe that the lower portion of the first translated component, which corresponds to \( V_2(f) \), overlaps the upper 300-Hz range of \( V_1(f) \). There is no way that these parts of the total spectrum can be separated. Thus, aliasing has occurred.

(b) There are two approaches to solving the problem. The first is to increase the sampling rate to a value greater than twice the highest frequency of the composite spectrum (i.e., greater than \( 2 \times 1500 = 3000 \text{ Hz} \)). However, this solution would probably not be the best approach unless there is some reason to preserve \( v_2 \). A better solution probably would be to pass the composite analog signal through a presampling anti-aliasing filter with a fairly sharp cutoff just above 800 Hz and in which the attenuation is quite high in the range from 1200 to 1500 Hz. This filter would essentially eliminate \( v_2 \), and then 2000 samples per second should be adequate.
8-3 Ideal Impulse Sampling

The form of the sampled-data signal in the last section was derived on the assumption that each of the samples had a non-zero width \( \delta \). We will now consider the limiting case as \( \delta \) approaches zero. In this case, the samples can be conveniently represented as a sequence of impulse functions, for which the weight of each is the value of the signal at that point. This type of signal will be referred to as an impulse sampled signal.

While we understand that no real-life analog pulses could have zero width, the concept of the impulse sampled signal is very important in the study of digital signal processing. When an analog signal is converted into a digital signal and subsequently processed on a computer or microprocessor, it may be considered as a sequence of numbers, for which an assumed "width" is somewhat meaningless. A very convenient way of modeling such a signal is by means of an impulse sampled signal.

The form of the ideal impulse sampled signal is illustrated in Figure 8-5. An arbitrary analog signal shown in Figure 8-5(a) is sampled (or modulated) by the periodic impulse train shown in Figure 8-5(b). The resulting ideal representation of the impulse sampled signal is shown in Figure 8-5(c).

Spectrum of Impulse Train

As a necessary prelude to developing the form of the spectrum of the impulse sampled signal, the properties of the impulse train shown in Figure 8-5(b) will be investigated. This function can be expressed as

\[
p_\delta(t) = \sum_{-\infty}^{\infty} \delta(t-nT)
\]

where \( \delta(t-nT) \) is the mathematical symbol for an impulse occurring at \( t=nT \).

Since the impulse train is periodic in the time domain, it can be expanded in a complex exponential Fourier series of the form

\[
p_\delta(t) = \sum_{-\infty}^{\infty} P_{\delta n} e^{j\omega n t}
\]

where \( P_{\delta n} \) represents the coefficients of the Fourier series expansion. Recall from Chapter 2 that the exponential series coefficients can be expressed as

\[
P_{\delta n} = \frac{1}{T} \int_{-T/2}^{T/2} p_\delta(t) e^{-j\omega_n t} \, dt
\]

In the interval of one cycle centered at \( t=0 \), \( p_\delta(t) \) contains only the single impulse \( \delta(t) \). Thus, from the basic definition of the impulse function given in Chapter 3, \( P_{\delta n} \) is obtained as

\[
P_{\delta n} = \frac{1}{T} \int_{-T/2}^{T/2} \delta(t) e^{-j\omega_n t} \, dt = \frac{1}{T} f_s
\]

From this result, it is deduced that the Fourier coefficients of the impulse train all have equal weight and there is no convergence. This is compatible with the nonperiodic single impulse function considered in Chapter 3, in which the spectrum was determined to be a constant at all frequencies. For the periodic
impulse train, however, the spectrum is a constant, but exists only at integer multiples of the sampling frequency.

To summarize the Fourier series for the periodic impulse train can be expressed as

\[
p_0(t) = \sum_{n=-\infty}^{\infty} \frac{1}{T} e^{j2\pi n t/T}
\]  
(8-21)

Spectrum of Impulse Sampled Signal

The impulse sampled signal will be denoted as \( v_\delta(t) \) and it can be expressed as

\[
v_\delta(t) = v(t) p_0(t)
\]  
(8-22)

Substitution of (8-21) in (8-22) results in

\[
v_\delta(t) = v(t) \sum_{n=-\infty}^{\infty} \frac{1}{T} e^{j2\pi n t/T}
\]  
(8-23)

or

\[
v_\delta(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} v(t) e^{j2\pi n t/T}
\]  
(8-24)

Fourier transformation of both sides of (8-24) using the modulation theorem, we obtain

\[
\tilde{V}_\delta(f) = \frac{1}{T} \sum_{n=-\infty}^{\infty} \tilde{V}(f - nf_s)
\]  
(8-25)

The frequency domain equivalents of the preceding operations are illustrated in Figure 8-6. An assumed arbitrary baseband amplitude spectrum \( V(f) \) is shown in Figure 8-6(a), and the discrete spectrum of the periodic pulse train in shown in Figure 8-6(b). The form of the amplitude spectrum of the ideal impulse sampled signal is shown in Figure 8-6(c). The general form is similar to the form derived from the nonzero width natural sampling process shown in Figure 8-2, and the basic sampling requirements developed in the last section apply here. However, a comparison of the figures and equations (8-4), (8-6), and (8-25) reveals one major difference between the spectral forms. The spectral components derived with nonzero pulse widths gradually diminish with frequency, and the magnitudes follow a \( \sin x / x \) function envelope. However, components derived from impulse sampling are all of equal magnitude and do not diminish with frequency.

An important deduction from this discussion is that the spectrum of an impulse sampled signal is a periodic function of frequency. The "period" in the frequency domain is equal to the sampling frequency \( f_s \). Thus, ideal impulse sampling in the time domain leads to a periodic spectrum in the frequency domain. This development is as far as we will go with the concept of impulse sampling. However, the reader who pursues the field of digital signal processing will likely encounter the concept again. For the remainder of this chapter, all pulses considered will be assumed to have non-zero width.
Example 8-6

A pure sinusoid with a frequency of 1 kHz is sampled at intervals of 0.1 ms and converted into digital numbers to be processed on a computer. List all positive frequencies below 45 kHz in the spectrum.

Solution

The basic sampling rate is

\[
f_s = \frac{1}{T} = \frac{1}{0.1 \times 10^{-3}} = 10 \text{ kHz}
\]  

(8-26)

The process may be considered as impulse sampling. The spectrum consists of the original component at 1 kHz, and the sum and difference frequencies about the sampling frequency and all its harmonics. The frequencies below 45 kHz are

1 kHz,  
9 kHz, 11 kHz  
19 kHz, 21 kHz  
29 kHz, 31 kHz  
39 kHz, 41 kHz
8-4 Pulse-Amplitude Modulation

The emphasis thus far in the chapter has been on establishing the theoretical basis for the sampling process. The sampling theorem applies to all forms of pulse and digital modulation whenever an analog signal is represented as a series of samples.

The first and most basic method for consideration is that of pulse-amplitude modulation (PAM). A PAM signal consists of samples of the analog signal in which the amplitude of each pulse is proportional to the analog signal at that point, but in which the pulses have a fixed width. We immediately recognize that the form of the signal used in the basic development of the sampled-data concept in Section 8-2 fits the definition. Thus, the waveform shown back in Figure 8-1(c) is a form of a PAM signal. This form, in which the top of the pulse follows the analog signal during the pulse width is called a natural-sampled PAM signal.

In actual practice, PAM signals (along with converted digital values) usually are based on a flat-topped PAM signal, and this form is illustrated in Figure 8-7. The flat-top representation is a constant value which, for example, could represent the analog signal at one particular point in the sampling interval (usually the beginning).

It can be shown that the spectrum of the flat-top signal contains some distortion with respect to the ideal natural sampled signal. However, if the widths of the pulses are small compared to the sampling interval, the resulting distortion is usually negligible. If the samples are converted to digital words, there is no spectral distortion since, as we have seen, impulse sampling can be assumed.

A PAM signal may be used as a specific type of desired signal for pulse encoding and transmission. Also, other forms of pulse and/or digital modulation start with a PAM signal, and the signal is then converted to some other form.

In considering the properties of PAM signals, as well as other pulse-modulated methods to be considered, there is one important difference in the spectral form as compared with the analog methods considered earlier in the text. All analog methods considered involved translating the modulating spectrum to a higher frequency, usually in the RF range, and eliminating all low-frequency components. However, the signal spectral forms shown in previous figures in this chapter show that the final PAM spectrum starts in the baseband range and may extend all the way down to dc. Thus, PAM is not suitable in its basic form for RF transmission. In the various forms of pulse modulation, the primary objective of the modulation process is to replace the continuous signal by a sequence of samples, with the primary goal of sharing the transmission medium with other signals.

Some pulse-modulated waveforms are transmitted directly at baseband over wire links. However, when RF transmission is required, the PAM signal could be applied as a complex baseband modulating signal to a high-frequency transmitter, and the resulting spectrum would be shifted up to a higher frequency range. The signal would have undergone two levels of modulation, one of which converted the continuous signal to a sequence of discrete samples, and the other of which shifted the spectrum of sampled signal to the RF range to enhance transmission.

PAM Bandwidth

A fundamental question with PAM concerns the bandwidth required to transmit the encoded signal. Referring back to Figure 8-2, it is observed again that not only does the original spectrum appear from dc to W, but shifted versions of the spectrum appear about all frequencies that are integer multiples of the sampling rate. The peak values of these shifted components slowly reduce in magnitude as a function of frequency since these peaks are proportional to the \( \frac{\sin x}{x} \) function. We would expect intuitively that a reasonable number would have to be transmitted in order to maintain the integrity of the PAM signal.

A more convenient way of looking at the PAM signal is through the concept of pulse transmission, as discussed in Chapter 4. To preserve the quality of a PAM signal, the pulse amplitudes must be preserved,
and they must not be allowed to spread excessively. On the other hand, perfect reproduction of the beginning and ending of the pulses has been found to require more bandwidth than desired for most of the types of applications for which PAM is used. Specifically, a "coarse" type of reproduction criteria has been found to be sufficient, and the baseband bandwidth $B_T$ can be expressed as

$$B_T = \frac{K_1}{\tau} \quad (8-27)$$

where $\delta$ is the width of each pulse sample.

The constant $K_1$ depends on the spacing between adjacent pulses, the level of acceptable adjacent pulse crosstalk, the sharpness of the cutoff rate, and other factors. In certain theoretical developments, the value $K_1 = 0.5$ is used. In this case, the result of (8-27) is in perfect agreement with the "coarse" pulse transmission criterion of Chapter 4.

The generation of a PAM signal may be achieved by gating on the analog signal periodically for an interval $\delta$ and turning it off for the remainder of each sampling period. A balanced modulator utilizing a pulse train for the second input could be used or an analog switch could be used. PAM signals are formed as a portion of a time-division multiplexing system, and further discussion will be given in Section 8-5.

Signal Reconstruction

The reconstruction of a continuous signal from a PAM signal represents a important process. Refer back to the spectral diagrams of Figure 8-2. It is seen that the portion of the spectrum of the sampled signal from dc to $W$ has exactly the same form as the original signal. (It may have a different magnitude level but the equivalence of the two spectral shapes is the important property.) Therefore, if all the components above $W$ are eliminated, the spectrum of the original unsampled signal will remain, and the form of the original signal will remain. Thus, passing a PAM signal through a low-pass filter restores the original analog signal. The filter should have a flat passband from dc to $W$, and it should display a sharp cutoff between $W$ and $f_s - W$. Thus, some reasonable guard band should be provided.

Holding Circuits

The actual level of the recovered signal may be quite small compared to the original signal due to the loss of energy resulting from filtering. Frequently, holding circuits are used in conjunction with the filter as a means of maintaining a reasonable level of energy in the filtered signal, as well as for easing a portion of the filtering requirements.

While there are different orders for holding circuits, we will focus on the zero-order form, which is the simplest. The operation of a zero-order holding circuit is illustrated by the waveforms of Figure 8-8. The waveform of Figure 8-8(a) represents a single flat-top PAM signal for which restoration is desired. The output of a zero-order holding circuit is shown in Figure 8-8(b). Observe that a given pulse establishes in a very short time an output proportional to the pulse level. This level is retained at this value until the next pulse arrives.

The signal at the output of the zero-order holding circuit is a type of "staircase" approximation of the original analog signal. Therefore, it contains much less high-frequency content than the PAM signal, but it clearly displays some distortion due to the "steps". Additional low-pass filtering is required to eliminate the remaining high-frequency content, and this provides smoothing between the various levels.
Example 8-6

An analog signal $v(t)$ whose amplitude spectrum is shown in Figure 8-9(a) is sampled at a rate $f_s = 5$ kHz. The width of each sample is $\delta = 40$ ns. (a) Sketch the form of the sampled spectrum from dc to the first zero crossing of the pulse train spectrum. (b) Compute the approximate transmission baseband bandwidth for both $K_1 = 0.5$ and $K_1 = 1$ in equation (8-27).

Solution

(a) The basic concept to remember is that the spectrum of the sampled signal is obtained by reproducing the form of the spectrum of the unmodulated signal plus translated versions at integer multiples of the sampling frequency. In this example, the relative convergence of the spectral components will be investigated.

First, we observe the spectrum of the pulse train, which may be considered to be multiplied by the continuous signal to yield the sampled signal. Since the sampling rate is $f_s = 5$ kHz, the time interval between successive pulses is $T = 1/f_s = 1/(5 \times 10^3) = 200 \mu s$. The duty cycle is $d = \tau / T = 40 \mu s / 200 \mu s = 0.2$. Recalling the spectrum of a periodic pulse train from Chapter 3, the form for $d = 0.2$ is shown in Figure 8-9(b) up to the vicinity of the first zero crossing, which occurs at a frequency $f = 1/\tau = 1/40 \mu s = 25$ kHz. Observe that the magnitudes of the terms reduce in accordance with the $\sin x/x$ function as the frequency increases.

The spectrum of the sampled signal is obtained by centering the spectrum of the unsampled signal about each of the pulse train spectral frequencies and multiplying by the respective coefficient amplitudes. The amplitude scale of the pulse spectrum and the sampled signal spectrum are both amplified in order to enhance the illustration.

(b) For a choice $K_1 = 0.5$ in equation (8-27), the bandwidth required is

$$B_T = \frac{0.5}{\tau} = \frac{0.5}{40 \times 10^{-6}} = 12.5 \text{ kHz} \quad (8-28)$$

This is the minimum bandwidth used in system-level calculations and can be achieved only under idealized filter conditions. In many systems, a constant closer to $K_1 = 1$ is used. For this choice, the bandwidth is

$$B_T = \frac{1}{\tau} = \frac{1}{40 \times 10^{-6}} = 25 \text{ kHz} \quad (8-29)$$
8-5 **Time-Division Multiplexing**

The primary motivation for PAM (as well as other pulse-modulation forms) is a process called *time-division multiplexing* (TDM). Time-division multiplexing is a process in which a number of separate signals can be sent over the same transmission link by alternately sampling the signals in a successive pattern.

The process of TDM is illustrated in Figure 8-10. A *commutator* is required at the input to the transmission medium and a *decommutator* is required at the output. In most systems, these two processes are achieved electronically, but it is easier to illustrate with mechanical components. At the input, the commutator sequentially samples each data signal in order. The decommutator at the receiving end is assumed to be synchronized with the one at the transmitter and routes each signal to its proper destination. Thus, as far as each individual output is concerned, the signal appears as a sampled PAM signal, and it can be reconstructed by proper filtering.

Each signal must be sampled at the minimum Nyquist rate for its particular bandwidth. For the moment, we will assume that each analog signal being sampled has the same bandwidth $W$. This means that the commutators must rotate at a rate no less than $2W$ revolutions per second.

A typical layout of a TDM composite signal is shown in Figure 8-11. While a variety of examples could be shown, this particular example has seven separate data signals all sampled at the same rate. A space equal to a sample pulse duration is maintained between any two successive pulses, and flat-top sampling is employed. A *frame* is one complete structured time interval in which each of the data pulses, plus any additional system information, is provided. Along with samples of the seven data signals, one synchronizing pulse is transmitted in each frame for this system. While a number of possible methods for synchronization can be used, in this system a dc bias level is added to all data channels so that they are always positive. The sync pulse is then transmitted as a negative pulse so that the receiver can readily recognize its presence. Each time a sync pulse appears, the receiver decommutator is realigned with the commutator at the transmitter by pulse selective circuits.

For the illustration shown, there is a short "dead space" between successive samples. The unused space is often desirable in order to allow some spreading of the pulses resulting from the finite bandwidth limitations of any realistic transmission system. Without some space, there may be a noticeable amount of *crosstalk* between adjacent channels.

**PAM Minimum Bandwidth**

A theoretical lower bound for the bandwidth of a TDM system utilizing PAM will now be developed. While unrealistic in practice, the result provides some insight into time-bandwidth interchange and it will later be compared with some analog results.

Assume that $k$ baseband signals, each with bandwidth $W$, are to be sampled and multiplexed over some transmission system. The following idealized assumptions will be made: (1) sampling set at minimum Nyquist rate, (2) no spacing between adjacent pulses, (3) no sync pulses, and (4) lower bandwidth estimate of $0.5/\delta$. The form of the assumed signal is shown in Figure 8-12. The value $T_f$ represents the *frame period*.

Each signal must be sampled no less than $2W$ times per second. Since $T_f$ represents the time between samples, $T_f \leq 1/2W$. Based on the minimum Nyquist rate

$$T_f = \frac{1}{2W}$$

(8-30)

A given frame contains $k$ samples with each having a width $\delta$, which means that
The minimum transmission bandwidth is then

$$\tau = \frac{T_f}{k}$$

(8-31)

Thus, the more signals to be shared on a line, the greater the bandwidth required. This is a basic limitation imposed by nature and will be more meaningful after we look at some other systems.

**Multirate Sampling**

Thus far in the study of multiplexing, we have considered all of the signals to have the same bandwidth. However, suppose that the different analog signals have different bandwidths. One obvious solution is to sample all signals at or above the Nyquist rate of the signal having the highest bandwidth. This solution is somewhat inefficient if there are wide differences in the bandwidths.

A more attractive solution in some cases is to devise a frame sequence which samples the higher-frequency signals more often than the lower-frequency channels. While the timing and synchronizing of the signals may be more complex, the overall benefits may be worth the extra complexity. An example of such an arrangement will be given in the next chapter.

**PAM Generation**

Generation of a PAM multiplexed signal is most easily implemented with a combination of analog switches and multiplexing circuits. A circuit diagram of a system to combine four channels is shown in Figure 8-13. The switches are special $P$-Channel JFET's, which are available as individual units or as a combination of several on a single IC chip. In the latter form, an extra compensating JFET (shown at the top) is also included.

A given $P$-channel JFET acts as an open circuit when the gate voltage is more positive than a minimum voltage (related to the pinch-off voltage). Conversely, when the gate voltage is nearly zero, the JFET is turned on and acts like a small series resistance (typically less than 100 $\Omega$). By connecting the gate terminals (designated as A, B, C, and D) to suitable counter-type circuits, the gates can be turned on in sequence in accordance with a master clock reference. The gate voltage levels of many analog switch circuits such as this are compatible with basic digital logic levels such as TTL, which minimize the interfacing problems. For example, a TTL logic 1 might open the switch and a TTL logic 0 might close the switch.

The samples of the different signals are combined in the analog operational amplifier summing circuit. The compensating switch acts as a series resistance approximately equal to the series resistance of either of the series switches when they are on, and this compensates for the uncertainty in the gain level due to the switch resistance.
**Example 8-7**

Consider a PAM time-division system with seven signals. Each signal has a baseband bandwidth of 1 kHz. Based on the idealized criteria discussed in this section, determine the minimum bandwidth.

**Solution**

The idealized theoretical minimum bandwidth is the number of channels times the bandwidth per channel as given by (8-32). Hence,

\[ B_T = kW = 7 \times 1 \text{ kHz} = 7 \text{ kHz} \]  \hspace{1cm} (8-33)

**Example 8-8**

Assume that the system of Example 8-7 is implemented in the format of the system of Figure 8-11 with the sync pulse added. Assume that the sampling rate is set to be 25% greater than the theoretical minimum Nyquist rate. (a) Based on the 0.5/\( \delta \) rule, determine the approximate baseband bandwidth required for the composite baseband signal. (b) If the composite PAM signal is used to amplitude modulate a high-frequency RF carrier, determine the RF bandwidth required for the high-frequency signal.

**Solution**

(a) The sampling rate is to be set at 1.25 times the theoretical minimum so the sampling rate for each signal is

\[ f_s = 1.25 \times 2W = 1.25 \times 2 \times 1 \text{ kHz} = 2.5 \text{ kHz} \]  \hspace{1cm} (8-34)

The frame time \( T_f \) is

\[ T_f = \frac{1}{f_s} = \frac{1}{2.5 \times 10^3} = 0.4 \text{ ms} \]  \hspace{1cm} (8-35)

The bandwidth is determined from the minimum pulse width. From Figure 8-11, it is noted that there are 16 intervals of width \( \delta \) in a given frame (7 data pulses, 1 sync pulse, and 8 open spaces). The pulse width \( \delta \) is

\[ \tau = \frac{T_f}{16} = \frac{0.4 \text{ ms}}{16} = 25 \mu\text{s} \]  \hspace{1cm} (8-36)

The composite baseband bandwidth is

\[ B_T = \frac{0.5}{\tau} = \frac{0.5}{25 \times 10^{-6}} = 20 \text{ kHz} \]  \hspace{1cm} (8-37)

This is considerably greater than the idealistic minimum bandwidth.

(b) If the composite baseband signal amplitude modulates a high-frequency carrier the RF bandwidth is

\[ B_T = 2 \times 20 \text{ kHz} = 40 \text{ kHz} \]  \hspace{1cm} (8-38)
8-6 Pulse-Time Modulation

PAM was studied because it establishes the basic foundation for the sampling process, a requirement for all pulse and digital modulation systems. It is obvious in today's world, and in the world of the future, digital technology will dominate the communications field. However, there are a few special applications where other forms of pulse modulation are used, so a brief introduction to some of these concepts is worthwhile. In the pulse-time methods to be surveyed in this section, the amplitude of the pulse is maintained at a constant level, but either the width or the position of the pulse is dependent on the modulating signal level.

Pulse-Width Modulation

*Pulse-Width Modulation* (PWM) is a sampled-data process in which the width of each pulse is varied directly in accordance with the amplitude of the modulating signal at the particular sample point. Pulse-width modulation is also denoted by the alternate title of *Pulse-Duration Modulation* (PDM). Both terms appear in the literature.

The PWM process is illustrated in Figure 8-14. The assumed analog signal shown in Figure 8-14(a) is converted by appropriate electronic circuits to the pulse train shown in Figure 8-14(b). Observe that the most positive peak corresponds to the widest pulse, and the most negative peak corresponds to the narrowest pulse.

Pulse Position Modulation

*Pulse-position modulation* (PPM) is a sampled-data process in which the position of each pulse is varied directly in accordance with the amplitude of the modulating signal at the particular sample point. This process is illustrated in Figure 8-14(c) for the analog signal shown in Figure 8-14(a). In this case, the most positive peak corresponds to the maximum shift from a reference point (the beginning in this case) in each pulse interval, while the most negative peak corresponds to no shift.

PWM and PPM Bandwidth

An exact spectral analysis of either a PWM or a PPM signal is rather complex, but for the purpose of this brief introduction, a reasonable estimation process will suffice. Since the exact width or position of the pulse is critical to the modulation process, the "coarse" bandwidth approximation utilized for PAM is inadequate. Instead a "fine" approximation for the bandwidth should be used. Said differently, the allowable rise time of the pulse, which serves to locate the exact width or position of the pulse, is the criterion of reproduction. This means that for a given analog bandwidth, the required transmission bandwidth for either PWM or PPM is much greater than for PAM.

Analog and Pulse Modulation Comparisons

Although the analogies are not perfect, there is some similarity between analog and pulse modulation processes in the following sense:

<table>
<thead>
<tr>
<th>Analog Modulation</th>
<th>Pulse Modulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM</td>
<td>PAM</td>
</tr>
<tr>
<td>FM</td>
<td>PWM</td>
</tr>
<tr>
<td>PM</td>
<td>PPM</td>
</tr>
</tbody>
</table>

The reader is advised not to take these analogies too literally because they are all totally different processes. However, this categorization is helpful in remembering the different methods.
Example 8-9

An analog signal having a baseband bandwidth of 10 kHz is to be sampled and converted to a PWM format. Specifications dictate that each pulse must be reproduced to a sufficient accuracy that the rise time cannot exceed 1% of the sample time interval. Determine the approximate transmission bandwidth. Assume the minimum Nyquist sampling rate.

Solution

The minimum sampling rate required is

\[ f_s = 2W = 2 \times 10 \text{ kHz} = 20 \text{ kHz} \]  
\[ (8-39) \]

The frame time is

\[ T_f = \frac{1}{f_s} = \frac{1}{20 \text{ kHz}} = 50 \mu \text{s} \]  
\[ (8-40) \]

The maximum rise time permitted is

\[ t_r = 0.01 \times 50 \mu \text{s} = 0.5 \mu \text{s} \]  
\[ (8-41) \]

The approximate transmission bandwidth is then

\[ B_t = \frac{0.5}{t_r} = \frac{0.5}{0.5 \mu \text{s}} = 1 \text{ MHz} \]  
\[ (8-42) \]
8-7 Brief Introduction to Pulse-Code Modulation

Much of the next several chapters will be devoted to a detailed analysis of various forms of digital communication systems. However, within the context of sampling, it is appropriate to introduce here the idea of representing samples of an analog signal in terms of digital numbers. The basic process is called pulse-code modulation (PCM).

The most basic form of PCM utilizes the binary number system, which assumes only two states. While the actual values depend on the types of circuits involved, the levels are usually referred to as 0 and 1. A given binary digit is called a bit.

With binary PCM, each analog sample of the signal is represented by an $N$-bit binary word. The number of levels $M$ associated with $N$ bits is

$$M = 2^N$$

(8-43)

With analog signals or with pulse modulation forms, it is theoretically possible to represent an infinite number of levels. However, with PCM, the number of levels is always finite as indicated by (8-43). However, this does not pose any realistic performance limitation since increasing the number of bits in each sample may increase the number of levels to any arbitrary number.

A common example of PCM is that of commercial music Compact Discs (CDs). The samples of the signals recorded on CDs utilize 16 bits, which permits the possibility of $2^{16} = 65,536$ levels! Anyone owning a CD player is well aware of the high quality of the sound.

One of the great advantages of binary PCM is the fact that the receiver need only distinguish between two possible levels. Thus, as long as the noise is within reasonable limits, it is possible to determine whether a given pulse is a 1 or a 0, and the noisy and distorted pulses can be "cleaned up" and restored by creating replicas of the original pulses.

Since each sample of the signal requires $N$ bits, there must be $N$ pulses produced within a given sample length. This increases the bandwidth of binary PCM by a factor of $N$ over that required for PAM.

This brief section has been placed in this chapter simply to alert the reader to the fact that digital communications is part of the overall picture relating to pulse and sampling theory. Because of its importance in modern communications systems, the next several chapters will devoted to a detailed development of various forms of digital communication techniques.
**PROBLEMS**

8-1 A baseband signal has frequency components from dc to 8 kHz. Determine the (a) theoretical minimum sampling rate and (b) maximum time interval between successive samples.

8-2 A baseband signal has frequency components from dc to 20 kHz. Determine the (a) theoretical minimum sampling rate and (b) maximum time interval between successive samples.

8-3 Assume that the signal of Example 8-1 is sampled at a rate 25% above the theoretical minimum. Determine the (a) sampling rate and (b) maximum time interval between successive samples.

8-4 Assume that the signal of Example 8-2 is sampled at a rate 40% above the theoretical minimum. Determine the (a) sampling rate and (b) maximum time interval between successive samples.

8-5 For the signal of Problem 8-1, determine the minimum sampling rate such that the guard band is 6 kHz. (The guard band is the frequency difference between the highest baseband frequency and the lowest frequency of the first shifted component.)

8-6 For the signal of Problem 8-2, determine the minimum sampling rate such that the guard band is 10 kHz. (See the definition of guard band in Problem 8-5.)

8-7 The sampling rate for guard band is 44.1 kHz. Determine the highest theoretical baseband frequency that could be reproduced.

8-8 A system has a sampling rate of 14 kHz. Determine the highest theoretical baseband frequency that could be reproduced.

8-9 An analog signal has a duration of 1 minute. The spectral content ranges from near dc to 500 Hz. The signal is to be sampled, converted to digital format, and stored in memory for subsequent processing. To assist in recovery, the sampling rate is chosen to be 25% above the theoretical minimum. (a) Determine the minimum number of samples that must be taken if reconstruction is desired. (b) Determine the time interval between successive samples.

8-10 An analog signal has a duration of 1 hour. The spectral content ranges from near dc to 4 kHz. The signal is to be sampled, converted to digital format, and stored in memory for subsequent processing. To assist in recovery, the sampling rate is chosen to be 50% above the theoretical minimum. (a) Determine the minimum number of samples that must be taken if reconstruction is desired. (b) Determine the time interval between successive samples.

8-11 A digital signal processing system operates at a sampling rate of 10,000 samples/s. Determine the theoretical highest frequency permitted in a baseband signal if a guard band of 4 kHz is to be established between the upper range of the baseband signal and the lower range of the first translated component.

8-12 A digital signal processing system operates at a sampling rate of 9000 samples/s. Determine the theoretical highest frequency permitted in a baseband signal if the lowest frequency of the first translated component is to be 25% higher than the highest frequency of the baseband component.

8-13 A sinusoid with a frequency of 2 kHz is sampled at a rate of 16 kHz and converted into digital numbers to be processed on a computer. List all positive frequencies in the spectrum below 50 kHz.

8-14 A sinusoid with a frequency of 500 Hz is sampled at intervals of 0.25 ms and converted into digital numbers to be processed on a computer. List all positive frequencies in the spectrum below 18 kHz.
9-15 A signal consists of two sinusoidal components at frequencies of 1 kHz and 2 kHz. It is sampled at intervals of 0.1 ms and converted into digital numbers to be processed on a computer. List all positive frequencies below 35 kHz.

8-16 A signal consists of two sinusoidal components at frequencies of 500 Hz and 1.5 kHz. It is sampled at a rate of 5 kHz and converted into digital numbers to be processed on a computer. List all positive frequencies below 18 kHz.

8-17 A PAM TDM system has six signals plus synchronization with a frame format defined as follows: Sync pulse occupies an interval $2\delta$, while each signal sample has a width $\delta$ and a space of width $\delta$ is placed between successive pulses. The receiver detects the wider sync pulse once per frame and realigns the timing sequence accordingly. (a) Determine the approximate baseband bandwidth required for the composite PAM signal if each channel has a 1-kHz bandwidth. (b) If the composite PAM signal is used to amplitude modulate a high-frequency carrier, determine the RF bandwidth required for the resulting high-frequency signal.

8-18 For the PAM TDM system of Example 8-8, repeat all computations if the sampling rate is increased to twice the theoretical minimum rate.

8-19 For the PWM TDM system of Example 8-9, determine the approximate transmission bandwidth if the sampling rate is set at 25% above the theoretical minimum.

8-20 For the PWM TDM system of Example 8-9, determine the approximate transmission bandwidth if the rise time specification is changed so that it cannot exceed 0.2% of the time allocated to the given channel. All other specifications are the same as in Example 8-9.