CHAPTER

TEN

SAMPLING AND PULSE MODULATION

Experimental data and mathematical functions are frequently displayed as *continuous* curves, even though a finite number of *discrete points* was used to construct the graphs. If these points or *samples* have sufficiently close spacing, a smooth curve drawn through them allows you to interpolate intermediate values to any reasonable degree of accuracy. It can therefore be said that the continuous curve is adequately described by the sample points alone.

In similar fashion, an electric *signal* satisfying certain requirements can be reproduced from an appropriate set of *instantaneous samples*. Sampling therefore makes it possible to transmit a message in the form of *pulse modulation*, rather than as a continuous signal. Usually the pulses are quite short compared to the time between them, so a pulse-modulated wave has the property of being "off" most of the time.

This property of pulse modulation offers two potential advantages over CW modulation. First, the transmitted power can be concentrated into short bursts instead of being generated continuously. The system designer then has greater latitude for equipment selection, and may choose devices such as lasers and high-power microwave tubes that operate only on a pulsed basis. Second, the time interval between pulses can be filled with sample values from other signals, a process called *time-division multiplexing* (TDM).

But pulse modulation has the disadvantage of requiring very large transmission bandwidth compared to the message bandwidth. Consequently, the methods of *analog* pulse modulation discussed in this chapter are used primarily as *message processing* for TDM and/or prior to CW modulation. *Digital* or *coded* pulse modulation has additional advantages that compensate for the increased bandwidth, as we'll see in Chap. 12.

10.1 SAMPLING THEORY AND PRACTICE

The theory of sampling presented here sets forth the conditions for signal sampling and reconstruction from sample values. We'll also examine practical implementation of the theory and some related applications.

Chopper Sampling

A simple but highly informative approach to sampling theory comes from the switching operation of Fig. 10.1-1*a*. The switch periodically shifts between two contacts at a rate of $f_s = 1/T_s$ Hz, dwelling on the input-signal contact for τ seconds and on the grounded contact for the remainder of each period. The output $x_s(t)$ then consists of short segments of the input $x(t)$, as shown in Fig. 10.1-1*b*. Figure 10.1-1*c* is an electronic version of Fig. 10.1-1*a*; the output voltage equals the input voltage except when the clock signal forward-biases the diodes and thereby clamps the output to zero. This operation variously called *single-ended* or *unipolar chopping*, is not instantaneous sampling in the strict sense. Nonetheless, $x_s(t)$ will be designated the sampled wave and f_s the sampling frequency.

We now ask: Are the sampled segments sufficient to describe the original input signal and, if so, how can $x(t)$ be retrieved from $x_s(t)$? The answer to this question lies in the frequency domain, in the spectrum of the sampled wave.

As a first step toward finding the spectrum, we introduce a *switching function* $s(t)$ such that

$$x_s(t) = x(t)s(t) \tag{1}$$

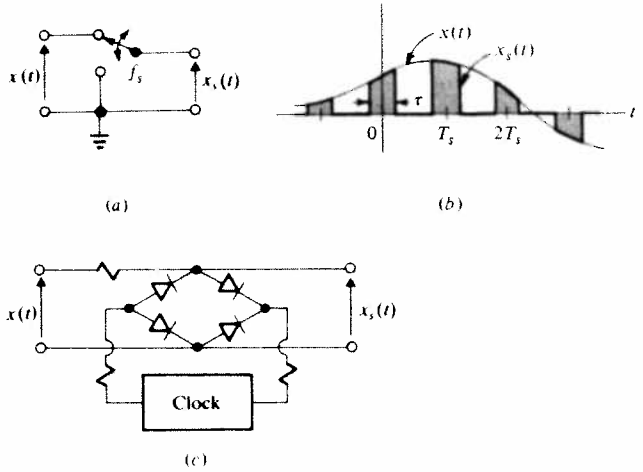


Figure 10.1-1 Switching sampler. (*a*) Functional diagram; (*b*) waveforms; (*c*) circuit realization with diode bridge.

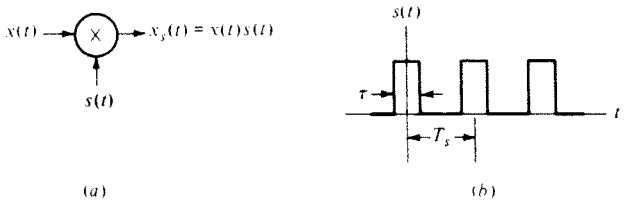


Figure 10.1-2 Sampling as multiplication. (*a*) Functional diagram; (*b*) switching function.

Thus the sampling operation becomes multiplication by $s(t)$, as indicated schematically in Fig. 10.1-2*a*, where $s(t)$ is nothing more than the periodic pulse train of Fig. 10.1-2*b*. Since $s(t)$ is periodic, it can be written as a Fourier series. Using the results of Example 2.1-1 we have

$$s(t) = \sum_{n=-\infty}^{\infty} f_s \tau \text{ sinc } nf_s\tau \, e^{j2\pi n f_s t} = c_0 + \sum_{n=1}^{\infty} 2c_n \cos n\omega_s t \tag{2}$$

where

$$c_n = f_s \tau \text{ sinc } nf_s\tau \qquad \omega_s = 2\pi f_s$$

Combining Eq. (2) with Eq. (1) yields the term-by-term expansion

$$x_s(t) = c_0\, x(t) + 2c_1 x(t) \cos \omega_s t + 2c_2\, x(t) \cos 2\omega_c t + \cdots \tag{3}$$

Thus, if the input message spectrum is $X(f) = \mathcal{F}[x(t)]$, the output spectrum is

$$\begin{aligned} X_s(f) = {} & c_0 X(f) + c_1[X(f - f_s) + X(f + f_s)] \\ & + c_2[X(f - 2f_s) + X(f + 2f_s)] \\ & + \cdots \end{aligned} \tag{4}$$

which follows directly from the modulation theorem.

While Eq. (4) appears rather messy, the spectrum of the sampled wave is readily sketched if the input signal is assumed to be *bandlimited*. Figure 10.1-3 shows a convenient $X(f)$ and the corresponding $X_s(f)$ for two cases, $f_s > 2W$ and $f_s < 2W$. This figure reveals something quite surprising: the sampling operation has left the message spectrum *intact*, merely repeating it periodically in the frequency domain with a spacing of f_s. We also note that the first term of Eq. (4) is precisely the message spectrum, attenuated by the *duty cycle* $c_0 = f_s\tau = \tau/T_s$.

If sampling preserves the message spectrum, it should be possible to recover or *reconstruct* $x(t)$ from the sampled wave $x_s(t)$. The reconstruction technique is not at all obvious from the time-domain relations in Eqs. (1) and (3). But referring again to Fig. 10.1-3, we see that $X(f)$ can be separated from $X_s(f)$ by *lowpass filtering*, provided that the spectral sidebands don't overlap. And if $X(f)$ alone is filtered from $X_s(f)$, we have recovered $x(t)$. Two conditions obviously are necessary to prevent overlapping spectral bands: the message must be bandlimited, and the

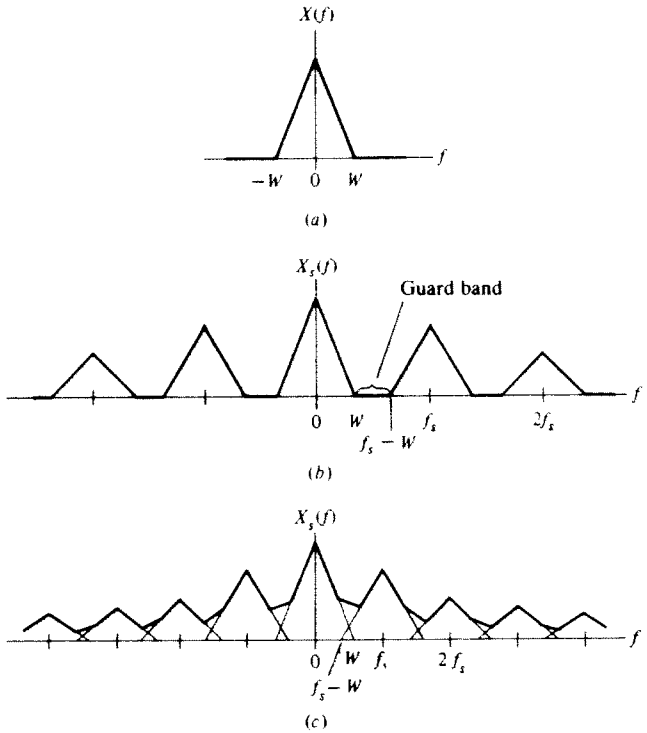


Figure 10.1-3 Spectra for switching sampling. (*a*) Message; (*b*) sampled message, $f_s > 2W$; (*c*) sampled message, $f_s < 2W$.

sampling frequency must be sufficiently great that $f_s - W \geq W$. Thus, we require

$$X(f) = 0 \qquad |f| > W$$

and

$$f_s \geq 2W \qquad \text{or} \qquad T_s \leq \frac{1}{2W} \tag{5}$$

The minimum sampling frequency $f_{s_{\min}} = 2W$ is called the *Nyquist rate*. When Eq. (5) is satisfied and $x_s(t)$ is filtered by an ideal LPF, the output signal will be proportional to $x(t)$; message reconstruction from the sampled signal therefore has been achieved. The exact value of the filter bandwidth B is unimportant as long as

$$W \leq B \leq f_s - W \tag{6}$$

so the filter passes $X(f)$ and rejects all higher components in Fig. 10.1-3*b*. Sampling at $f_s > 2W$ creates a *guard band* into which the transition region of a practical LPF can be fitted.

This analysis has shown that if a bandlimited signal is sampled at a frequency greater than the Nyquist rate, it can be *completely* reconstructed from the sampled wave. Reconstruction is accomplished by lowpass filtering. These conclusions may be difficult to believe at first exposure; they certainly test our faith in spectral analysis. Nonetheless, they are quite correct.

Finally, it should be pointed out that our results are independent of the sample-pulse duration, save as it appears in the duty cycle. If τ is made very small, $x_s(t)$ approaches a string of *instantaneous sample points*, which corresponds to *ideal sampling*. We'll pursue ideal sampling theory after a brief look at the *bipolar* chopper, which has $\tau = T_s/2$.

Example 10.1-1 Bipolar choppers Figure 10.1-4*a* depicts the circuit and waveforms for a bipolar chopper. The equivalent switching function is a *square wave* alternating between $s(t) = +1$ and -1. From the series expansion of $s(t)$ we get

$$x_s(t) = \frac{4}{\pi} x(t) \cos \omega_s t - \frac{4}{3\pi} x(t) \cos 3\omega_s t + \frac{4}{5\pi} x(t) \cos 5\omega_s t - \cdots \qquad (7)$$

whose spectrum is sketched in Fig. 10.1-4*b* for $f \geq 0$. Note that $X_s(f)$ contains no DC component and only the odd harmonics of f_s. Clearly, we can't recover $x(t)$ by lowpass filtering. Instead, the practical applications of bipolar choppers involve *bandpass* filtering.

If we apply $x_s(t)$ to a BPF centered at some odd harmonic nf_s, the output will be proportional to $x(t) \cos n\omega_s t$—a double-sideband suppressed-carrier waveform. Thus, a bipolar chopper serves as a *balanced modulator*. It also serves as a *synchronous detector* when the input is a DSB or SSB signal and

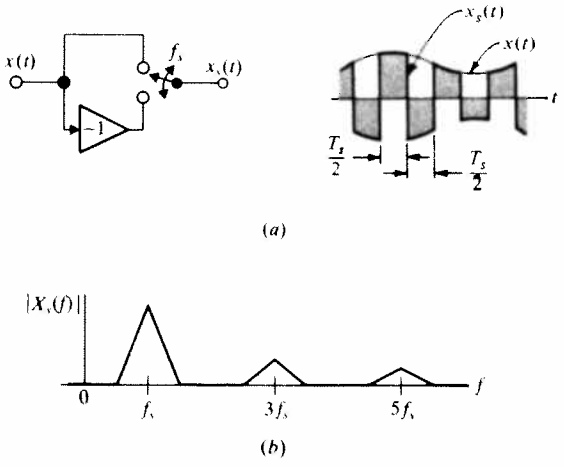


Figure 10.1-4 Bipolar chopper. (*a*) Circuit and waveforms; (*b*) spectrum.

the output is lowpass filtered. These properties are combined in the *chopper-stabilized amplifier*, which makes possible DC and low-frequency amplification using a high-gain AC amplifier. Additionally, Prob. 10.1-4 indicates how a bipolar chopper can be modified to produce the baseband mutliplexed signal for FM stereo.

Ideal Sampling and Reconstruction

By definition, ideal sampling is *instantaneous* sampling. The switching device of Fig. 10.1-1*a* yields instantaneous values only if $\tau \to 0$; but then $f_s \tau \to 0$, and so does $x_s(t)$. Conceptually, we overcome this difficulty by multiplying $x_s(t)$ by $1/\tau$ so that, as $\tau \to 0$ and $1/\tau \to \infty$, the sampled wave becomes a train of *impulses* whose *areas* equal the instantaneous sample values of the input signal. Formally, we write the rectangular pulse train as

$$s(t) = \sum_{k=-\infty}^{\infty} \Pi\left(\frac{t - kT_s}{\tau}\right)$$

from which we define the *ideal sampling function*

$$s_\delta(t) \triangleq \lim_{\tau \to 0} \frac{1}{\tau} s(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT_s) \qquad (8)$$

The *ideal sampled wave* is then

$$x_\delta(t) \triangleq x(t) s_\delta(t) \qquad (9a)$$

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

$$= \sum_{k=-\infty}^{\infty} x(kT_s)\, \delta(t - kT_s) \qquad (9b)$$

since $x(t)\, \delta(t - kT_s) = x(kT_s)\, \delta(t - kT_s)$.

To obtain the corresponding spectrum $X_\delta(f) = \mathcal{F}[x_\delta(t)]$ we note that $(1/\tau)x_s(t) \to x_\delta(t)$ as $\tau \to 0$ and, likewise, $(1/\tau)X_s(f) \to X_\delta(f)$. But each coefficient in Eq. (4) has the property $c_n/\tau = f_s \operatorname{sinc} nf_s\tau = f_s$ when $\tau = 0$. Therefore,

$$X_\delta(f) = f_s X(f) + f_s[X(f - f_s) + X(f + f_s)] + \cdots$$

$$= f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \qquad (10)$$

which is illustrated in Fig. 10.1-5 for the message spectrum of Fig. 10.1-3*a* taking $f_s > 2W$. We see that $X_\delta(f)$ is periodic in frequency with period f_s, a crucial observation in the study of sampled-data systems.

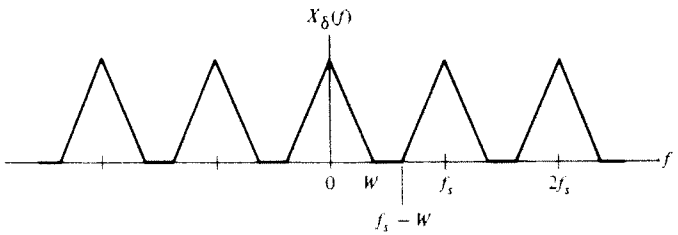


Figure 10.1-5 Spectrum of ideally sampled message.

Somewhat parenthetically, we can also develop an expression for $S_\delta(f) = \mathcal{F}[s_\delta(t)]$ as follows. From Eq. (9*a*) and the convolution theorem, $X_\delta(f) = X(f) * S_\delta(f)$ whereas Eq. (10) is equivalent to

$$X_\delta(f) = X(f) * \left[\sum_{n=-\infty}^{\infty} f_s \, \delta(f - nf_s) \right]$$

Therefore, we conclude that

$$S_\delta(f) = f_s \sum_{n=-\infty}^{\infty} \delta(f - nf_s) \tag{11}$$

so the spectrum of a periodic string of unit-weight impulses in the time domain is a periodic string of impulses in the frequency domain with spacing $f_s = 1/T_s$; in both domains we have a function that looks like a picket fence.

Returning to the main subject and Fig. 10.1-5, it's immediately apparent that if we invoke the same conditions as before — $x(t)$ bandlimited in W and $f_s \geq 2W$ — then a filter of suitable bandwidth will reconstruct $x(t)$ from the ideal sampled wave. Specifically, for an ideal LPF of gain K, time delay t_d, and bandwidth B, the transfer function is

$$H(f) = K\Pi\left(\frac{f}{2B}\right) e^{-j\omega t_d}$$

so filtering $x_\delta(t)$ produces the output spectrum

$$Y(f) = H(f)X_\delta(f) = Kf_s X(f) e^{-j\omega t_d}$$

assuming B satisfies Eq. (6). The output time function is then

$$y(t) = \mathcal{F}^{-1}[Y(f)] = Kf_s x(t - t_d) \tag{12}$$

which is the original signal amplified by Kf_s and delayed by t_d.

Further confidence in the sampling process can be gained by examining reconstruction in the time domain. The impulse response of the LPF is

$$h(t) = 2BK \text{ sinc } 2B(t - t_d)$$

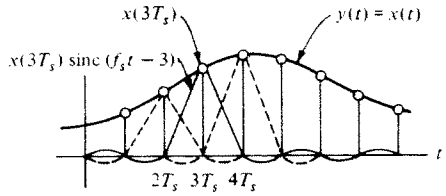


Figure 10.1-6 Ideal reconstruction.

And since the input $x_\delta(t)$ is a train of weighted impulses, the output is a train of weighted *impulse responses*, namely,

$$y(t) = h(t) * x_\delta(t) = \sum_k x(kT_s)h(t - kT_s)$$

$$= 2BK \sum_{k=-\infty}^{\infty} x(kT_s) \text{ sinc } 2B(t - t_d - kT_s) \tag{13}$$

Now suppose for simplicity that $B = f_s/2$, $K = 1/f_s$, and $t_d = 0$, so

$$y(t) = \sum_k x(kT_s) \text{ sinc } (f_s t - k)$$

We can then carry out the reconstruction process graphically, as shown in Fig. 10.1-6. Clearly the correct values are reconstructed at the sampling instants $t = kT_s$, for all sinc functions are zero at these times save one, and that one yields $x(kT_s)$. Between sampling instants $x(t)$ is *interpolated* by summing the precursors and postcursors from *all* the sinc functions. For this reason the LPF is often called an *interpolation filter*, and its impulse is called the *interpolation function*.

The above results are well summarized by stating the important theorem of *uniform (periodic) sampling*. While there are many variations of this theorem, the following form is best suited to our purposes.

> If a signal contains no frequency components for $|f| \geq W$, it is completely described by instantaneous sample values uniformly spaced in time with period $T_s \leq 1/2W$. If a signal has been sampled at the Nyquist rate or greater ($f_s \geq 2W$) and the sample values are represented as weighted impulses, the signal can be exactly reconstructed from its samples by an ideal LPF of bandwidth B, where $W \leq B \leq f_s - W$.

Another way to express the theorem comes from Eqs. (12) and (13) with $K = T_s$ and $t_d = 0$. Then $y(t) = x(t)$ and

$$x(t) = 2BT_s \sum_{k=-\infty}^{\infty} x(kT_s) \text{ sinc } 2B(t - kT_s) \tag{14}$$

provided $T_s \leq 1/2W$ and B satisfies Eq. (6). Therefore, just as a periodic signal is completely described by its Fourier series coefficients, a bandlimited signal is completely described by its instantaneous sample values *whether or not the signal actually is sampled.*

Several additional theorems pertaining to signal sampling have been developed. They are discussed in Black (1953, chap. 4) and Peebles (1976, chap. 7).

Exercise 10.1-1 Consider a sampling pulse train of the general form

$$s_p(t) = \sum_{k=-\infty}^{\infty} p(t - kT_s) \tag{15a}$$

whose pulse shape $p(t)$ equals zero for $|t| > T_s/2$ but is otherwise arbitrary. Use an exponential Fourier series and Eq. (21), Sect. 2.2, to show that

$$S_p(f) = f_s \sum_{n=-\infty}^{\infty} P(nf_s)\,\delta(f - nf_s) \tag{15b}$$

where $P(f) = \mathcal{F}[p(t)]$. Then let $p(t) = \delta(t)$ to obtain Eq. (11).

Practical Sampling and Aliasing

Practical sampling systems differ from ideal sampling in three obvious aspects:

1. The sampled wave consists of pulses having finite amplitude and duration, rather than impulses.
2. Practical reconstruction filters are not ideal filters.
3. The messages to be sampled are *timelimited* signals whose spectra are not and cannot be strictly bandlimited.

The first two differences may present minor problems, while the third leads to the more troublesome effect known as aliasing.

Regarding pulse-shape effects, our investigation of the unipolar chopper and the results of Exercise 10.1-1 correctly imply that almost any pulse shape $p(t)$ will do when sampling takes the form of a *multiplication* operation $x(t)s_p(t)$. Another operation produces *flat-top sampling*, described in the next section. This type of sampling may require equalization, but it does not alter our conclusion that pulse shapes are relatively inconsequential.

Regarding practical reconstruction filters, we consider the typical filter response superimposed on a sampled-wave spectrum in Fig. 10.1-7. If the filter is reasonably flat over the message band, its output will consist of $x(t)$ plus spurious frequency components at $|f| > f_s - W$, outside the message band. In audio systems these components would sound like high-frequency hissing or "noise."

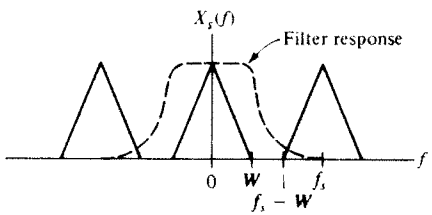


Figure 10.1-7 Practical reconstruction filter.

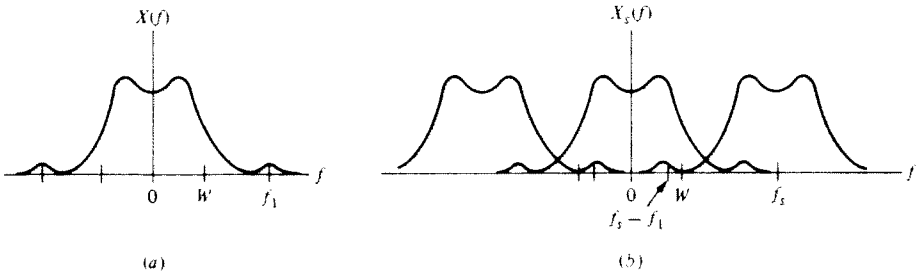


Figure 10.1-8 Aliasing effect. (*a*) Spectrum of timelimited message; (*b*) sampled message spectrum with overlaps.

However, they are considerably attenuated and their strength is proportional to $x(t)$, so they disappear when $x(t) = 0$. When $x(t) \neq 0$, the message tends to mask their presence and render them more tolerable. The combination of careful filter design and an adequate guard band created by taking $f_s > 2W$ makes practical reconstruction filtering nearly equivalent to ideal reconstruction.

Regarding the timelimited nature of real signals, a message spectrum like Fig. 10.1-8*a* may be viewed as a bandlimited spectrum if the frequency content above W is small and presumably unimportant for conveying the information. When such a message is sampled, there will be unavoidable overlapping of spectral components shown in Fig. 10.1-8*b*. In reconstruction, frequencies originally outside the nominal message band will appear at the filter output in the form of much *lower* frequencies. Thus, for example, $f_1 > W$ becomes $f_s - f_1 < W$, as indicated in the figure.

This phenomenon of downward frequency translation occurs whenever a frequency component is *undersampled*, so that $f_s < 2f_1$, and is given the descriptive name of *aliasing*. The aliasing effect is far more serious than spurious frequencies passed by nonideal reconstruction filters, for the latter fall *outside* the message band, whereas aliased components fall *within* the message band. Aliasing is combated by filtering the message as much as possible *before* sampling and, if necessary, sampling at much greater than the nominal Nyquist rate.

As illustration, the average voice spectrum (Fig. 3.2-1) extends well beyond 10 kHz, though most of the energy is concentrated in the range 100 to 600 Hz and a bandwidth of 3 kHz is sufficient for intelligibility. If a voice wave is prefiltered by a 3.3-kHz LPF and then sampled at $f_s = 8$ kHz, the aliased components are typically 30 dB below the desired signal and go unnoticed by the listener. Incidentally, these are the standard values for voice-telephone sampling.

Example 10.1-2 Sampling Oscilloscopes A practical application of aliasing occurs in the sampling oscilloscope, which exploits undersampling to display high-speed *periodic* waveforms that would otherwise be beyond the capability of the electronics. To illustrate the principle, consider the periodic waveform $x(t)$ with period $T_x + 1/f_x$ in Fig. 10.1-9*a*. If we use a sampling interval T_s

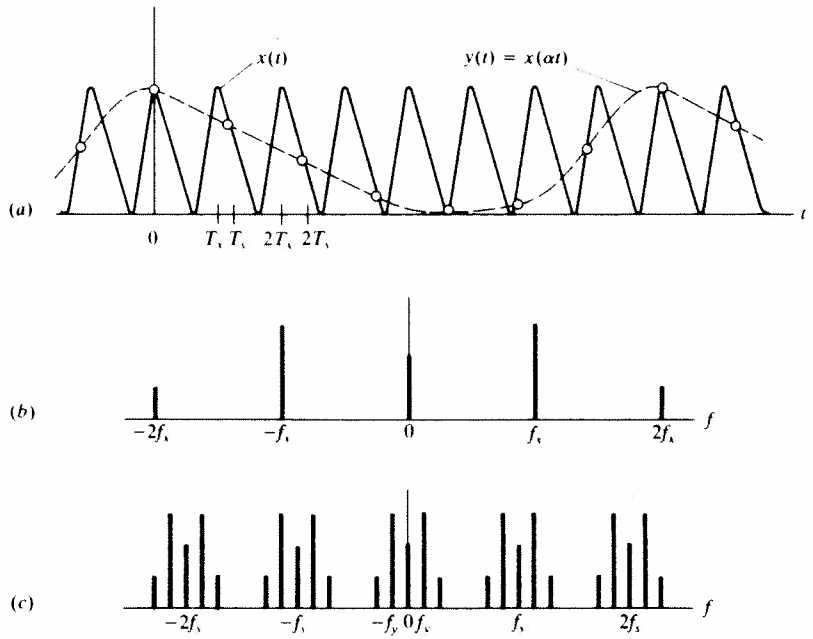


Figure 10.1-9 (*a*) Periodic waveform with undersampling; (*b*) spectrum of $x(t)$; (*c*) spectrum of $y(t) = x(\alpha t)$, $\alpha < 1$.

slightly greater than T_x and interpolate the sample points, we get the *expanded* waveform $y(t) = x(\alpha t)$ shown as a dashed curve. The corresponding sampling frequency is

$$f_s = (1 - \alpha)f_x \qquad 0 < \alpha < 1$$

so $f_s < f_x$ and even the fundamental frequency of $x(t)$ will be undersampled. Now let's find out if and how this system actually works by going to the frequency domain.

We assume that $x(t)$ has been prefiltered to remove any frequency components higher than the mth harmonic. Figure 10.1-9*b* shows a typical two-sided line spectrum of $x(t)$, taking $m = 2$ for simplicity. Since sampling translates all frequency components up and down by nf_s, the fundamental will appear in the spectrum of the sampled signal at

$$\pm f_y = \pm |f_x - f_s| = \pm \alpha f_x$$

as well as at $\pm f_x$ and at $f_x \pm nf_s = (1 + n)f_x \pm nf_y$. Similar translations applied to the DC component and second harmonic yield the spectrum in Fig. 10.1-9*c*, which contains a *compressed* image of the original spectrum centered at each multiple of f_s. Therefore, a lowpass filter with $B = f_s/2$ will con-

struct $y(t) = x(\alpha t)$ from $x_s(t)$ provided that

$$\alpha < \frac{1}{2m + 1}$$

which prevents spectral overlap.

Exercise 10.1-2 Demonstrate the aliasing effect for yourself by making a careful sketch of $\cos 2\pi 10t$ and $\cos 2\pi 70t$ for $0 \leq t \leq {}^1/_{10}$. Put both sketches on the same set of axes and find the sample values at $t = 0, {}^1/_{80}, {}^2/_{80}, \ldots, {}^8/_{80}$, which corresponds to $f_s = 80$. Also, convince yourself that no other waveform bandlimited in $10 < W < 40$ can be interpolated from the sample values of $\cos 2\pi 10t$.

10.2 ANALOG PULSE MODULATION

If a message waveform is adequately described by periodic sample values, it can be transmitted using analog pulse modulation wherein the sample values modulate some parameter of a pulse train. Pulse parameters suitable for modulation include amplitude, duration, and position. The corresponding processes are designated *pulse-amplitude modulation* (PAM), *pulse-duration modulation*† (PDM), and *pulse-position modulation* (PPM).

Figure 10.2-1 illustrates these types of pulse modulation with a representative message waveform. The modulated pulse parameter — amplitude, duration, or

† Also called pulse-*width* modulation (PWM).

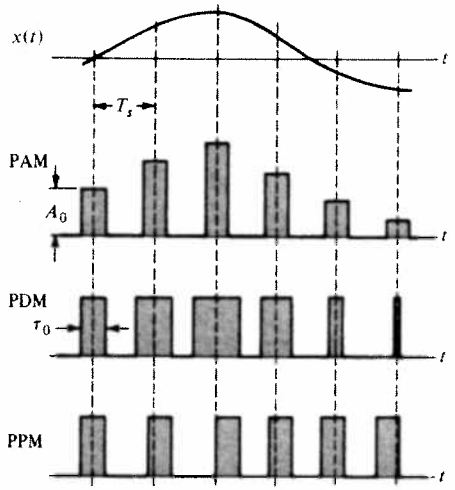


Figure 10.2-1 Types of analog pulse modulation.

relative position—varies in direct proportion to the sample values of $x(t)$. For clarity, the pulses are shown as rectangular and the pulse durations have been grossly exaggerated. Actual modulated waves would also be delayed slightly compared to the message, because the pulses can't be generated before the sampling instants.

It should be evident from these waveforms that a modulated pulse train has significant DC content and that the bandwidth required to preserve the pulse shape far exceeds the message bandwidth. Consequently, you seldom encounter a single-channel communication system with PAM, PDM, or PPM. But analog pulse modulation deserves attention for its major roles in time-division multiplexing, data telemetry, and instrumentation systems.

Flat-Top Sampling and PAM

Although a PAM wave could be obtained from a chopper circuit, a more popular method employs the *sample-and-hold* (S/H) technique. This operation produces flat-top pulses, as in Fig. 10.1-2, rather than curved-top chopper pulses. We therefore begin here with the properties of *flat-top sampling.*

A rudimentary S/H circuit consists of two FET switches and a capacitor, connected as shown in Fig. 10.2-2*a*. A gate pulse at G1 briefly closes the sampling switch and the capacitor holds the sampled voltage until discharged by a pulse applied to G2. (Commercial integrated-circuit S/H units have further refinements, including isolating op-amps at input and output.) Periodic gating of the sample-and-hold circuit generates the sampled wave

$$x_p(t) = \sum_k x(kT_s)p(t - kT_s) \tag{1}$$

illustrated by Fig. 10.2-2*b*. Note that each output pulse of duration τ represents a single instantaneous sample value.

To analyze flat-top sampling, we'll draw upon the relation $p(t - kT_s) = p(t) * \delta(t - kT_s)$ and write

$$x_p(t) = p(t) * \left[\sum_k x(kT_s)\,\delta(t - kT_s)\right] = p(t) * x_\delta(t)$$

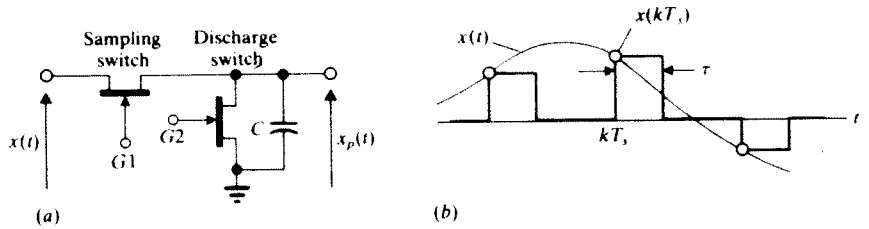


Figure 10.2-2 Flat-top sampling. (*a*) Sample-and-hold circuit; (*b*) waveforms.

Fourier transformation of this *convolution* operation yields

$$X_p(f) = P(f)\left[f_s \sum_n X(f - nf_s)\right] = P(f)X_\delta(f) \tag{2}$$

Figure 10.2-3 provides a graphical interpretation of Eq. (2), taking $X(f) = \Pi(f/2W)$. We see that flat-top sampling is equivalent to passing an ideal sampled wave through a network having the transfer function $P(f) = \mathcal{F}[p(t)]$.

The high-frequency rolloff characteristic of a typical $P(f)$ acts like a *lowpass* filter and attenuates the upper portion of the message spectrum. This loss of high-frequency content is called *aperture effect.* The larger the pulse duration or aperture τ, the larger the effect. Aperture effect can be corrected in reconstruction by including an *equalizer* with

$$H_{eq}(f) = Ke^{-j\omega t_d}/P(f) \tag{3}$$

However, little if any equalization is needed when $\tau/T_s \ll 1$.

Now consider a unipolar flat-top PAM signal defined by

$$x_p(t) = \sum_k A_0[1 + \mu x(kT_s)]p(t - kT_s) \tag{4}$$

The constant A_0 equals the unmodulated pulse amplitude, and the *modulation index* μ controls the amount of amplitude variation. The condition

$$1 + \mu x(t) > 0 \tag{5}$$

ensures a unipolar (single-polarity) waveform with no missing pulses. The resulting constant pulse rate f_s is particularly important for synchronization in time-division multiplexing.

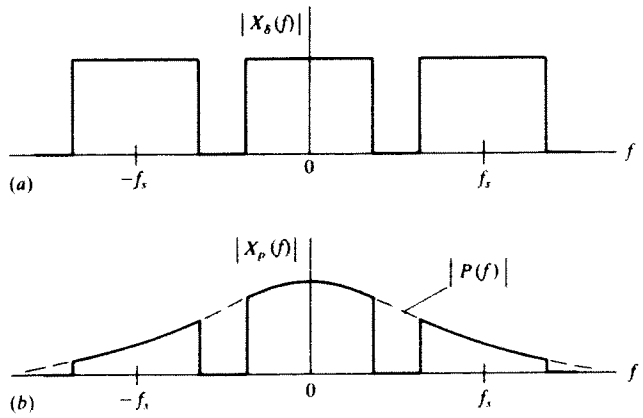


Figure 10.2-3 (*a*) Spectrum for ideal sampling when $X(f) = \Pi(f/2W)$; (*b*) aperture effect in flat-top sampling.

Comparison of Eqs. (1) and (4) shows that a PAM signal can be obtained from a sample-and-hold circuit with input $A_0[1 + \mu x(t)]$. Correspondingly, the PAM spectrum will look like Fig. 10.2-3*b* with $X(f)$ replaced by

$$\mathcal{F}\{A_0[1 + \mu x(t)]\} = A_0[\delta(f) + \mu X(f)],$$

which results in spectral impulses at all harmonics of f_s and at $f = 0$. Reconstruction of $x(t)$ from $x_p(t)$ therefore requires a *DC block* as well as lowpass filtering and equalization.

Clearly, PAM has many similarities to AM CW modulation—modulation index, spectral impulses, and DC blocks. (In fact, an AM wave could be derived from PAM by bandpass filtering.) But the PAM spectrum extends from DC up through several harmonics of f_s, and the estimate of required transmission bandwidth B_T must be based on time-domain considerations. For this purpose, we assume a small pulse duration τ compared to the time between pulses, so

$$\tau \ll T_s \leq \frac{1}{2W}$$

Adequate pulse resolution then requires

$$B_T \geq \frac{1}{2\tau} \gg W \tag{6}$$

Hence, practical applications of PAM are limited to those situations in which the advantages of a pulsed waveform outweigh the disadvantages of large bandwidth.

Exercise 10.2-1 Consider PAM transmission of a voice signal with $W \approx 3$ kHz. Calculate B_T if $f_s = 8$ kHz and $\tau = 0.1\ T_s$.

Pulse-Duration and Pulse-Position Modulation

We lump PDM and PPM together under one heading for two reasons. First, in both cases a *time* parameter of the pulse is being modulated, and the pulses have *constant amplitude*. Second, a close relationship exists between the modulation methods for PDM and PPM.

To demonstrate these points, Fig. 10.2-4 shows the block diagram and waveforms of a system that combines the sampling and modulation operations for either PDM or PPM. The system employs a comparator and a sawtooth-wave generator with period T_s. The output of the comparator is zero except when the message waveform $x(t)$ exceeds the sawtooth wave, in which case the output is a positive constant A. Hence, as seen in the figure, the comparator produces a PDM signal with *trailing-edge modulation* of the pulse duration. (Reversing the sawtooth results in leading-edge modulation, as in Fig. 7.3-3, while replacing the sawtooth with a triangular wave results in modulation on both edges.) Position modulation is obtained by applying the PDM signal to a monostable pulse generator that triggers on trailing edges at its input and produces short output pulses of fixed duration.

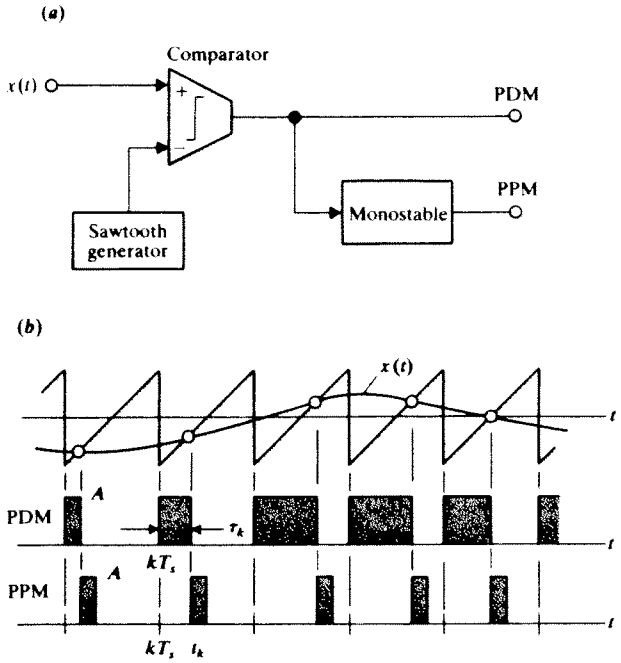


Figure 10.2-4 Generation of PDM or PPM. (*a*) Block diagram; (*b*) waveforms.

Careful examination of Fig. 10.2-4*b* reveals that the modulated duration or position depends on the message value at the time location t_k of the pulse edge, rather than the apparent sample time kT_s. Thus, the sample values are *nonuniformly* spaced. Inserting a sample-and-hold circuit at the input of the system gives uniform sampling if desired, but there's little difference between uniform and nonuniform sampling in the practical case of small amounts of time modulation such that $t_k - kT_s \ll T_s$.

If we assume nearly uniform sampling, the duration of the kth pulse in the PDM signal is

$$\tau_k = \tau_0[1 + \mu x(kT_s)] \tag{7}$$

in which the unmodulated duration τ_0 represents $x(kT_s) = 0$ and the modulation index μ controls the amount of duration modulation. Our prior condition on μ in Eq. (5) applies here to prevent missing pulses and "negative" durations when $x(kT_s) \leq 0$. The PPM pulses have fixed duration and amplitude so, unlike PAM and PDM, there's no potential problem of missing pulses. The kth pulse in a PPM signal begins at time

$$t_k = kT_s + t_d + t_0\, x(kT_s) \tag{8}$$

in which the unmodulated position $kT_s + t_d$ represents $x(kT_s) = 0$ and the constant t_0 controls the displacement of the modulated pulse.

The variable time parameters in Eqs. (7) and (8) make the expressions for $x_p(t)$ rather awkward. However, an informative approximation for the PDM waveform is derived by taking rectangular pulses with amplitude A centered at $t = kT_s$ and assuming that τ_k varies slowly from pulse to pulse. Series expansion then yields

$$x_p(t) \approx Af_s\tau_0[1 + \mu x(t)] + \sum_{n=1}^{\infty} \frac{2A}{\pi n} \sin n\phi(t) \cos n\omega_s t \tag{9}$$

where $\phi(t) = \pi f_s \tau_0[1 + \mu x(t)]$. Without attempting to sketch the corresponding spectrum, we see from Eq. (9) that the PDM signal contains the message $x(t)$ plus a DC component and *phase-modulated* waves at the harmonics of f_s. The phase modulation has negligible overlap in the message band when $\tau_0 \ll T_s$, so $x(t)$ can be recovered by lowpass filtering with a DC block.

Another message reconstruction technique converts pulse-*time* modulation into pulse-*amplitude* modulation, and works for PDM and PPM. To illustrate this technique the middle waveform in Fig. 10.2-5 is produced by a *ramp generator* that starts at time kT_s, stops at t_k, restarts at $(k + 1)T_s$, and so forth. Both the start and stop commands can be extracted from the edges of a PDM pulse, whereas PPM reconstruction must have an auxiliary *synchronization* signal for the start command.

Regardless of the particular details, demodulation of PDM or PPM requires received pulses with short *risetime* in order to preserve accurate message information. For a specified risetime $t_r \ll T_s$, the transmission bandwidth must satisfy

$$B_T \geq \frac{1}{2t_r} \tag{10}$$

which will be substantially greater than the PAM transmission bandwidth. In exchange for the extra bandwidth, we gain the benefit of constant-amplitude pulses that suffer no ill effects from *nonlinear distortion* in transmission since nonlinear distortion does not alter pulse duration or position.

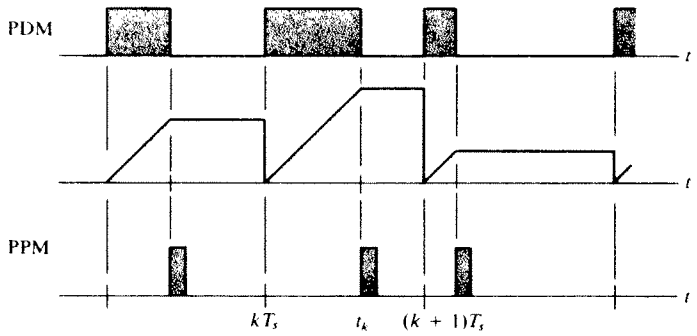


Figure 10.2-5 Conversion of PDM or PPM into PAM.

Additionally, like PM and FM CW modulation, PDM and PPM have the potential for *wideband noise reduction*—a potential more fully realized by PPM than by PDM. To appreciate why this is so, recall that the information resides in the time location of the pulse *edges*, not in the pulses themselves. Thus, somewhat like the carrier-frequency power of AM, the pulse power of pulse-time modulation is "wasted" power, and it would be more efficient to suppress the pulses and just transmit the edges! Of course we cannot transmit edges without transmitting pulses to define them. But we can send very short pulses indicating the position of the edges, a process equivalent to PPM. The reduced power required for PPM is a fundamental advantage over PDM, an advantage that becomes more apparent when we examine the signal-to-noise ratios.

Exercise 10.2-2 Derive Eq. (9) by the following procedure. First, assume constant pulse duration τ and write $x_p(t) = As(t)$ with $s(t)$ given by Eq. (2), Sect. 10.1. Then apply the *quasi-static approximation* $\tau \approx \tau_0[1 + \mu x(t)]$.

PPM Spectral Analysis★

Because PPM with nonuniform sampling is the most efficient type of analog pulse modulation for message transmission, we should take the time to analyze its spectrum. The analysis method itself is worthy of examination.

Let the kth pulse be centered at time t_k. If we ignore the constant time delay t_d in Eq. (8), nonuniform sampling extracts the sample value at t_k, rather than kT_s, so

$$t_k = kT_s + t_0 x(t_k) \tag{11}$$

By definition, the PPM wave is a summation of constant-amplitude position-modulated pulses, and can be written as

$$x_p(t) = \sum_k Ap(t - t_k) = Ap(t) * \left[\sum_k \delta(t - t_k)\right]$$

where A is the pulse amplitude and $p(t)$ the pulse shape. A simplification at this point is made possible by noting that $p(t)$ will (or should) have a very small duration compared to T_s. Hence, for our purposes, the pulse shape can be taken as impulsive, and

$$x_p(t) \approx A \sum_k \delta(t - t_k) \tag{12}$$

If desired, Eq. (12) can later be convolved with $p(t)$ to account for the non-impulsive shape.

In their present form, Eqs. (11) and (12) are unsuited to further manipulation the trouble is the position term t_k, which cannot be solved for explicitly. Fortunately, Rowe (1965, chap. 5) has devised a technique whereby t_k can be eliminated

entirely. Consider any function $g(t)$ having a single first-order zero at $t = \lambda$, such that $g(\lambda) = 0$, $g(t) \neq 0$ for $t \neq \lambda$, and $\dot{g}(t) \neq 0$ at $t = \lambda$. The distribution theory of impulses then shows that

$$\delta(t - \lambda) = |\dot{g}(t)|\, \delta[g(t)] \tag{13}$$

whose right-hand side is independent of λ. Equation (13) can therefore be used to remove t_k from $\delta(t - t_k)$ if we can find a function $g(t)$ that satisfies $g(t_k) = 0$ and the other conditions but does not contain t_k.

Suppose we take $g(t) = t - kT_s - t_0 x(t)$, which is zero at $t = kT_s + t_0 x(t)$. Now, for a given value of k, there is only one PPM pulse, and it occurs at $t_k = kT_s + t_0 x(t_k)$. Thus $g(t_k) = t_k - kT_s - t_0 x(t_k) = 0$, as desired. Inserting $\lambda = t_k$, $\dot{g}(t) = 1 - t_0 \dot{x}(t)$, etc., into Eq. (13) gives

$$\delta(t - t_k) = |1 - t_0 \dot{x}(t)|\, \delta[t - kT_s - t_0 x(t)]$$

and the PPM wave of Eq. (12) becomes

$$x_p(t) = A[1 - t_0 \dot{x}(t)] \sum_k \delta[t - t_0 x(t) - kT_s]$$

The absolute value is dropped since $|t_0 \dot{x}(t)| < 1$ for most signals of interest if $t_0 \ll T_s$. We then convert the sum of impulses to a sum of exponentials via

$$\sum_{k=-\infty}^{\infty} \delta(t - kT_s) = f_s \sum_{n=-\infty}^{\infty} e^{jn\omega_s t} \tag{14}$$

which is *Poisson's sum formula*. Thus, we finally obtain

$$x_p(t) = Af_s[1 - t_0 \dot{x}(t)] \sum_{n=-\infty}^{\infty} e^{jn\omega_s[t - t_0 x(t)]}$$

$$= Af_s[1 - t_0 \dot{x}(t)] \left\{1 + \sum_{n=1}^{\infty} 2 \cos [n\omega_s t - n\omega_s t_0 x(t)]\right\} \tag{15}$$

The derivation of Eq. (14) is considered in Prob. 10.2-11.

Interpreting Eq. (15), we see that PPM with nonuniform sampling is a combination of linear and exponential carrier modulation, for each harmonic of f_s is phase-modulated by the message $x(t)$ and amplitude-modulated by the derivative $\dot{x}(t)$. The spectrum therefore consists of AM and PM sidebands centered at all multiples of f_s, plus a DC impulse and the spectrum of $\dot{x}(t)$. Needless to say, sketching such a spectrum is a tedious exercise even for tone modulation. The leading term of Eq. (15) suggests that the message can be retrieved by lowpass filtering and *integrating*. However, the integration method does not take full advantage of the noise-reduction properties of PPM, so the usual procedure is conversion to PAM or PDM followed by lowpass filtering.

10.3 TIME-DIVISION MULTIPLEXING

A sampled waveform is "off" most of the time, leaving the time between samples available for other purposes. In particular, sample values from several different signals can be interlaced into a single waveform. This is the principle of *time-division multiplexing* (TDM) discussed here.

TDM Systems

The simplified system in Fig. 10.3-1 demonstrates the essential features of time-division multiplexing. Several input signals are prefiltered by the bank of input LPFs and sampled sequentially. The rotating sampling switch or *commutator* at the transmitter extracts one sample from each input per revolution. Hence, its output is a PAM waveform that contains the individual samples periodically interlaced in time. A similar rotary switch at the receiver, called a *decommutator* or *distributor*, separates the samples and distributes them to another bank of LPFs for reconstruction of the individual messages.

If all inputs have the same message bandwidth W, the commutator should rotate at the rate $f_s \geq 2W$ so that successive samples from any one input are spaced by $T_s = 1/f_s \leq 1/2W$. The time interval T_s containing one sample from each input is called a *frame*. If there are M input channels, the pulse-to-pulse spacing within a frame is $T_s/M = 1/Mf_s$. Thus, the total number of pulses per second will be

$$r = Mf_s \geq 2MW \tag{1}$$

which represents the pulse rate or *signaling rate* of the TDM signal.

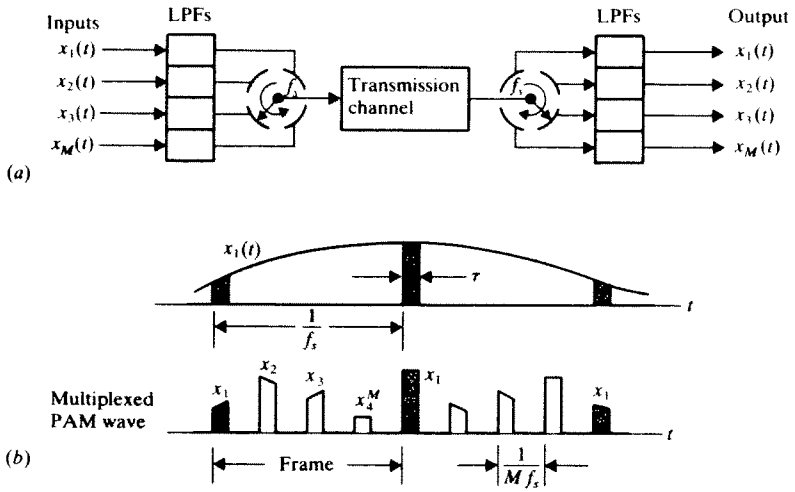
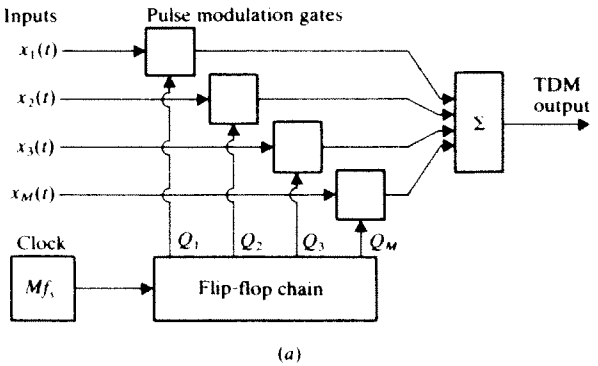


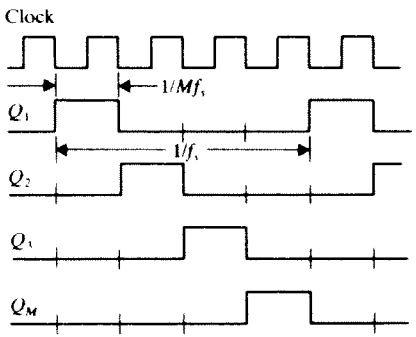
Figure 10.3-1 TDM system. (*a*) Block diagram; (*b*) waveforms.

Our primitive system shows mechanical switching to generate multiplexed PAM. But almost all practical TDM systems employ electronic switching. Furthermore, other types of pulse modulation can be used instead of PAM. Therefore, a more generalized commutator might have the structure diagrammed in Fig. 10.3-2, where pulse-modulation *gates* process the individual inputs to form the TDM output. The gate control signals come from a flip-flop chain (a broken-ring counter) driven by a digital clock at frequency Mf_s. The decommutator would have a similar structure.

Regardless of the type of pulse modulation, TDM systems require careful *synchronization* between commutator and decommutator. Synchronization is a critical consideration in TDM, because each pulse must be distributed to the correct output line at the appropriate time. A popular brute-force synchronization technique devotes one time slot per frame to a distinctive *marker* pulse or nonpulse, as illustrated in Fig. 10.3-3. These markers establish the frame frequency f_s at the receiver, but the number of signal channels is reduced to $M - 1$. Other synchronization methods involve auxiliary pilot tones or the statistical properties of the TDM signal itself.



(*a*)



(*b*)

Figure 10.3-2 (*a*) Electronic commutator for TDM; (*b*) timing diagram.

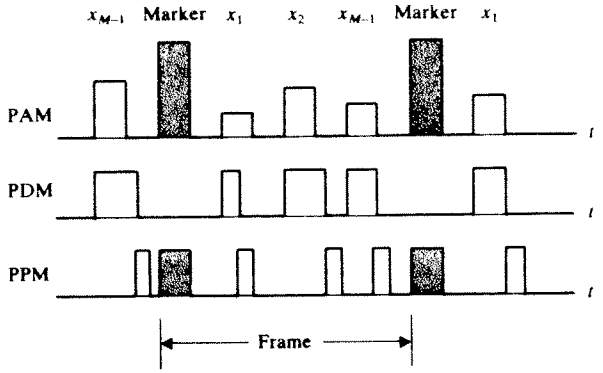


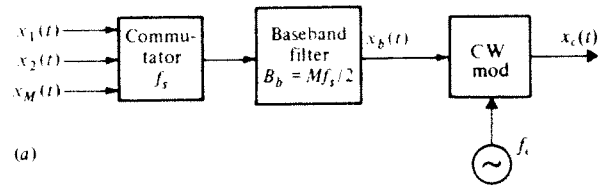
Figure 10.3-3 TDM synchronization markers.

Radio-frequency transmission of TDM necessitates the additional step of CW modulation to obtain a *bandpass* waveform. For instance, a TDM signal composed of duration or position modulated pulses could be applied to an AM transmitter with 100% modulation, thereby producing a train of constant-amplitude RF pulses. The compound process would be designated PDM/AM or PPM/AM, and the required transmission bandwidth would be twice that of the baseband TDM signal. The relative simplicity of this technique suits low-speed multichannel applications such as radio control for model airplanes.

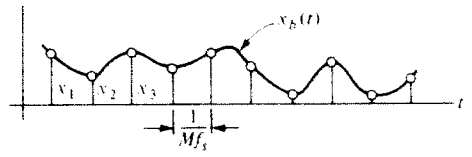
More sophisticated TDM systems may use PAM/SSB for bandwidth conservation or PAM/FM for wideband noise reduction. The complete transmitter diagram in Fig. 10.3-4*a* now includes a lowpass *baseband filter* with bandwidth

$$B_b = \tfrac{1}{2}r = \tfrac{1}{2}Mf_s \tag{2}$$

Baseband filtering prior to CW modulation produces a smooth modulating waveform $x_b(t)$ having the property that it passes through the individual sample values at the corresponding sample times, as portrayed in Fig. 10.3-4*b*. Since the



(*a*)



(*b*)

Figure 10.3-4 (*a*) TDM transmitter with baseband filtering; (*b*) baseband waveform.

interlaced sample spacing equals $1/Mf_s$, the baseband filter constructs $x_b(t)$ in the same way that an LPF with $B = f_s/2$ would reconstruct a waveform $x(t)$ from its periodic samples $x(kT_s)$ with $T_s = 1/f_s$.

If baseband filtering is employed, and if the sampling frequency is close to the Nyquist rate $f_{s_{min}} = 2W$ for the individual inputs, then the transmission bandwidth for PAM/SSB becomes

$$B_T = \tfrac{1}{2}M \times 2W = MW$$

Under these conditions, TDM approaches the theoretical *minimum bandwidth* of frequency-division multiplexing with SSB subcarrier modulation.

Although we've assumed so far that all input signals have the same bandwidth, this restriction is not essential and, moreover, would be unrealistic for the important case of analog data telemetry. The purpose of a telemetry system is to combine and transmit physical measurement data from different sources at some remote location. The sampling frequency required for a particular measurement depends on the physical process involved and can range from a fraction of one hertz up to several kilohertz. A typical telemetry system has a *main multiplexer* plus *submultiplexers* arranged to handle 100 or more data channels with various sampling rates.

Example 10.3-1 TDM Telemetry For the sake of illustration, suppose we need 5 data channels with minimum sampling rates of 3000, 700, 500, 300, and 200 Hz. If we used a 5-channel multiplexer with $f_s = 3000$ Hz for all channels, the TDM signaling rate would be $r = 5 \times 3000 = 15$ kHz—not including synchronization markers. A more efficient scheme involves an 8-channel main multiplexer with $f_s = 750$ Hz and a 2-channel submultiplexer with $f_s = 375$ Hz connected as shown in Fig. 10.3-5.

The two lowest-rate signals $x_4(t)$ and $x_5(t)$ are combined by the submultiplexer to create a pulse rate of $2 \times 375 = 750$ Hz for insertion into one channel of the main multiplexer. Hence, the samples of $x_4(t)$ and $x_5(t)$ will

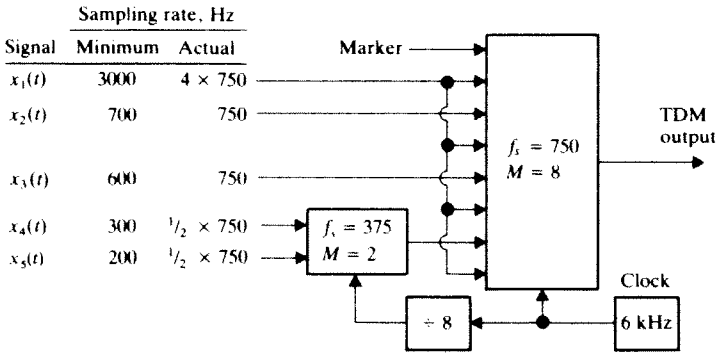


Figure 10.3-5 TDM telemetry system with main multiplexer and submultiplexer.

appear in *alternate* frames. On the other hand, the highest-rate signal $x_1(t)$ is applied to 4 inputs on the main multiplexer. Consequently, its samples appear in 4 equispaced slots within each frame, for an equivalent sampling rate of $4 \times 750 = 3000$ Hz. The total output signaling rate, including a marker, is $r = 8 \times 750$ Hz $= 6$ kHz. Baseband filtering would yield a smoothed signal whose bandwidth $B_b = 3$ kHz fits nicely into a voice telephone channel!

Exercise 10.3-1 Suppose the output in Fig. 10.3-5 is an unfiltered PAM signal with 50% duty cycle. Sketch the waveform for two successive frames, labeling each pulse with its source signal. Then calculate the required transmission bandwidth B_T from Eq. (6), Sect. 10.2.

Cross Talk and Guard Times

When a TDM system includes baseband filtering, the filter design must be done with extreme care to avoid interchannel *cross talk* from one sample value to the next in the frame. Digital signals suffer a similar problem called *intersymbol interference*, and we defer the treatment of baseband waveform shaping to Sect. 11.3.

A TDM signal without baseband filtering also has cross talk if the transmission channel results in pulses whose tails or postcursors overlap into the next time slot of the frame. Pulse overlap is controlled by establishing *guard times* between pulses, analogous to the guard bands between channels in an FDM system. Practical TDM systems have both guard times and guard bands, the former to suppress cross talk, the latter to facilitate message reconstruction with nonideal filters.

For a quantitative estimate of cross talk, let's assume that the transmission channel acts like a first-order lowpass filter with 3-dB bandwidth B. The response to a rectangular pulse then decays exponentially, as sketched in Fig. 10.3-6. The guard time T_g represents the *minimum* pulse spacing, so the pulse tail decays to a value no larger than $A_{ct} = Ae^{-2\pi BT_g}$ by the time the next pulse arrives. Accordingly, we define the *cross-talk reduction factor*

$$k_{ct} \triangleq 10 \log (A_{ct}/A)^2 \approx -54.5\, BT_g \quad \text{dB} \tag{3}$$

Keeping the cross talk below -30 dB calls for $T_g > 1/2B$.

Guard times are especially important in TDM with pulse-duration or pulse-position modulation because the pulse edges move around within their frame

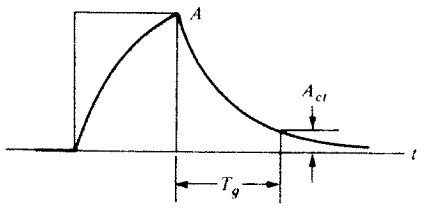


Figure 10.3-6 Cross talk in TDM.

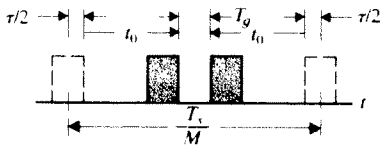


Figure 10.3-7 TDM/PPM with guard time.

slots. Consider the PPM case in Fig. 10.3-7: here, one pulse has been position-modulated forward by an amount t_0 and the next pulse backward by the same amount. The allowance for guard time T_g requires that $T_g + 2t_0 + 2(\tau/2) \leq T_s/M$ or

$$t_0 \leq \frac{1}{2}\left(\frac{T_s}{M} - \tau - T_g\right) \tag{4}$$

A similar modulation limit applies in the case of PDM.

Exercise 10.3-2 Nine voice signals plus a marker are to be transmitted via PPM on a channel having $B = 400$ kHz. Calculate T_g such that $k_{ct} \approx -60$ dB. Then find the maximum permitted value of t_0 if $f_s = 8$ kHz and $\tau = {}^1/_5(T_s/M)$.

Comparison of TDM and FDM

Time-division and frequency-division multiplexing accomplish the same end by different means. Indeed, they may be classified as *dual* techniques. The individual TDM channels are assigned to distinct *time slots* but jumbled together in the frequency domain; conversely, the individual FDM channels are assigned to distinct *frequency slots* but jumbled together in the time domain. What advantages, then, does TDM offer to the communications engineer?

First and foremost, TDM involves simpler instrumentation. Recall that FDM requires an analog subcarrier modulator, bandpass filter, and demodulator for every message channel, all of which are replaced by the TDM commutator and decommutator switching circuits. And TDM synchronization is but slightly more demanding than that of suppressed-carrier FDM.

Second, TDM is invulnerable to the usual causes of cross talk in FDM, namely, imperfect bandpass filtering and nonlinear cross modulation. However, TDM cross-talk immunity does depend on the transmission bandwidth and the absence of delay distortion.

Third, the use of submultiplexers allows a TDM system to accommodate different signals whose bandwidths or pulse rates may differ by more than an order of magnitude. This flexibility has particular value for multiplexing *digital* signals, as we'll see in Sect. 12.4.

Finally, TDM may or may not be advantageous when the transmission medium is subject to *fading*. Rapid wideband fading might strike only occasional pulses in a given TDM channel, whereas all FDM channels would be affected. But slow narrowband fading wipes out all the TDM channels, whereas it might hurt only one FDM channel.

10.4 NOISE IN PULSE MODULATION

Here we complete our study of analog pulse modulation by examining system performance in the presence of noise. We'll assume throughout the case of lowpass pulse transmission with independent additive gaussian white noise at the receiver. The specific topics at hand are: pulse measurements in noise, reconstruction from noisy samples, signal-to-noise ratios, and false-pulse threshold effect.

Pulse Measurements in Noise

Consider initially the problem of measuring some parameter of a *single* pulse $p(t)$ contaminated by noise, as represented in Fig. 10.4-1*a*. Let the pulse be more-or-less rectangular, with amplitude A, duration τ, and energy $E_p = A^2\tau$. Let the noise be white with power spectral density $G(f) = \eta/2$ and zero mean value. A reasonably selective lowpass filter having unit gain and noise bandwidth $B_N \geq 1/2\tau$ will remove excess noise while passing $p(t)$. The output $v(t) = p(t) + n(t)$ sketched in Fig. 10.4-1*b* consists of noise variations superimposed on a trapezoidal pulse shape with risetime $t_r \approx 1/2B_N$.

If you want to measure the pulse amplitude, you should do so at some instant t_a near the center of the pulse. A single measurement yields the random quantity

$$v(t_a) = A + n(t_a) = A + \epsilon_A$$

where $\epsilon_A = n(t_a)$ represents the *amplitude error*. The error variance is then

$$\sigma_A^2 = \overline{n^2} = \eta B_N \tag{1}$$

which should be small compared to A^2 for an accurate measurement. Since $A^2 = E_p/\tau$ and $B_N \geq 1/2\tau$, we can write the lower bound

$$\sigma_A^2 \geq \frac{\eta}{2\tau} = \frac{\eta}{2E_p} A^2 \tag{2}$$

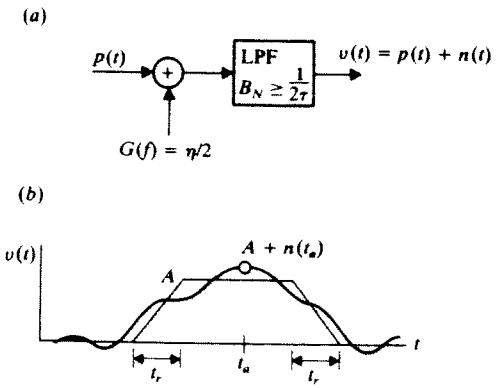


Figure 10.4-1 Pulse measurement in noise. (*a*) Model; (*b*) waveform.

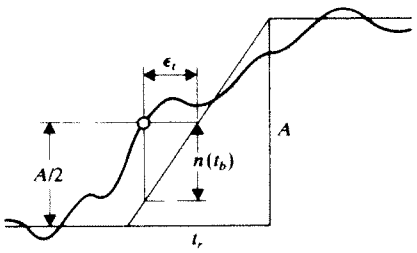


Figure 10.4-2 Time-position error caused by noise.

Any filter bandwidth less than about $1/2\tau$ would reduce the output pulse amplitude as well as the noise. Achieving the lower bound requires a *matched* filter; see Eq. (17), Sect. 5.4.

Measurements of pulse *position* or *duration* are usually carried out by detecting the instant t_b when $v(t)$ crosses some fixed level such as $A/2$. The noise perturbation $n(t_b)$ shown in the enlarged view of Fig. 10.4-2 causes a *time-position error* ϵ_t. From the similar triangles here, we see that $\epsilon_t/n(t_b) = t_r/A$ so $\epsilon_t = (t_r/A)n(t_b)$ and

$$\sigma_t^2 = (t_r/A)^2\overline{n^2} = (t_r/A)^2\eta B_N$$

Substituting $t_r \approx 1/2B_N$ and $A^2 = E_p/\tau$ yields

$$\sigma_t^2 \approx \frac{\eta}{4B_N A^2} = \frac{\eta\tau}{4B_N E_p} \tag{3}$$

which implies that we can make σ_t arbitrarily small by letting $B_N \to \infty$ so that $t_r \to 0$. But the received pulse actually has a nonzero risetime determined by the *transmission bandwidth* B_T. Hence

$$\sigma_t^2 \geq \frac{\eta}{4B_T A^2} = \frac{\eta\tau}{4B_T E_p} \tag{4}$$

and the lower bound is obtained with filter bandwidth $B_N \approx B_T$—in contrast to the lower bound on σ_A obtained with $B_N \approx 1/2\tau$.

Exercise 10.4-1 Suppose a 10-μs pulse is transmitted on a channel having $B_T = 800$ kHz and $\eta = E_p/50$. Calculate σ_A/A and σ_t/τ when: (*a*) $B_N = 1/2\tau$; (*b*) $B_N = B_T$.

Pulse Modulation With Noise

Regardless of the particular circuitry employed, demodulating a pulse-modulated wave boils down to message reconstruction from sample values. A generalized demodulator would therefore consist of a pulse converter that transforms the pulse-modulated wave into a train of weighted impulses from which an ideal LPF reconstructs the message. Figure 10.4-3 diagrams this demodulation model including additive noise.

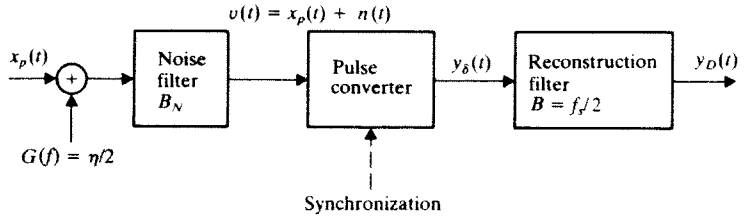


Figure 10.4-3 Model for demodulation of analog pulse modulation with noise.

A noise-limiting filter at the input passes the received signal plus noise

$$v(t) = x_p(t) + n(t)$$

The converter measures the amplitude, duration, or position of each pulse in $v(t)$ and generates

$$y_\delta(t) = \sum_k [\mu_p x(kT_s) + \epsilon_k]\,\delta(t - kT_s) \tag{5}$$

where μ_p is the modulation constant relating $x(kT_s)$ to $x_p(t)$ and ϵ_k is the measurement error induced by $n(t)$. We'll assume for convenience that the reconstruction filter has bandwidth $B = f_s/2$, gain $K = T_s = 1/2B$, and zero time delay. Hence, the impulse-train input $y_\delta(t)$ produces the final output

$$\begin{aligned} y_D(t) &= \sum_k [\mu_p x(kT_s) + \epsilon_k]\,\text{sinc}\,(f_s t - k) \\ &= \mu_p x(t) + n_D(t) \end{aligned} \tag{6a}$$

with

$$n_D(t) = \sum_k \epsilon_k\,\text{sinc}\,(f_s t - k) \tag{6b}$$

which represents the noise at the destination. The signal portion of Eq. (6*a*) follows from Eq. (14), Sect. 10.1.

Since the errors ϵ_k are proportional to sample values of $n(t)$ spaced by T_s, and since the noise-limiting filter has $B_N \geq 1/2\tau > 1/T_s$, the values of ϵ_k will be essentially uncorrelated and will have zero mean. We can therefore write the destination noise power as

$$N_D = \overline{n_D^2} = \overline{\epsilon_k^2} = \sigma^2$$

The signal power in terms of $S_x = \overline{x^2}$ is

$$S_D = \mu_p^2\,\overline{x^2} = \mu_p^2 S_x$$

where our normalization convention $|x(t)| \leq 1$ requires $S_x \leq 1$. Hence,

$$\left(\frac{S}{N}\right)_D = \frac{S_D}{N_D} = \frac{\mu_p^2}{\sigma^2} S_x \tag{7}$$

which expresses the destination signal-to-noise ratio in terms of the error variance σ^2 caused by reconstruction from noisy samples. Our next task is to determine μ_p^2 and σ^2 for specific types of pulse modulation.

A PAM signal contains the message samples in the modulated pulse amplitude $A_0[1 + \mu x(kT_s)]$, so the modulation constant is $\mu_p = A_0\mu \leq A_0$. The upper limit $\mu = 1$ comes from our signal normalization and Eq. (5), Sect. 10.2. The amplitude error variance is $\sigma^2 = \sigma_A^2 = \eta B_N$. Thus, under the best conditions of maximum modulation ($\mu = 1$) and minimum noise bandwidth ($B_N \approx 1/2\tau$), we have

$$\left(\frac{S}{N}\right)_D = \frac{2A_0^2\tau}{\eta} S_x$$

Notice that $A_0^2\tau$ equals the received signal energy in an unmodulated pulse.

The *average energy* per modulated pulse is

$$A_0^2\overline{[1 + x(kT_s)]^2}\tau = A_0^2(1 + S_x)\tau$$

when $\mu = 1$ and $x(t)$ has no DC component. Multiplying this average energy by the pulse rate f_s gives the *received signal power*

$$S_R = f_s A_0^2(1 + S_x)\tau$$

We now obtain our final result in the form

$$\left(\frac{S}{N}\right)_D = \frac{S_x}{1 + S_x}\frac{2S_R}{\eta f_s} = \frac{S_x}{1 + S_x}\left(\frac{2W}{f_s}\right)\gamma \qquad \text{PAM} \tag{8}$$

where we've introduced $\gamma = S_R/\eta W$. Equation (8) shows that $(S/N)_D \leq \gamma/2$, so PAM performance is at least 3 dB below unmodulated baseband transmission — just like AM CW modulation. The maximum value is seldom achieved in practice, nor is it sought after. The merit of PAM resides in its simplicity for multiplexing, not in its noise performance.

However, PPM and PDM do offer some improvement by virtue of wideband noise reduction. For if $B_N \approx B_T$, the time-position error variance is $\sigma^2 = \sigma_t^2 = \eta/(4B_T A^2)$. Since the pulse amplitude A is a constant, the received power can be written as $S_R = f_s A^2\tau_0$, where τ_0 denotes the *average* pulse duration in PDM or the fixed pulse duration in PPM. Equation (7) then becomes

$$\left(\frac{S}{N}\right)_D = \frac{4\mu_p^2 B_T A^2}{\eta} S_x = 4\mu_p^2 B_T \frac{S_R}{\eta f_s \tau_0} S_x$$

$$= 4\mu_p^2 B_T\left(\frac{W}{f_s\tau_0}\right)S_x\gamma \qquad \text{PDM or PPM} \tag{9}$$

This expression reveals that $(S/N)_D$ increases with increasing transmission bandwidth B_T. The underlying physical reason should be evident from Fig. 10.4-2 with $t_r \approx 1/2B_T$.

The PPM modulation constant is $\mu_p = t_0$, the maximum pulse displacement. The parameters t_0 and τ_0 are constrained by

$$t_0 \leq T_s/2 \qquad \tau_0 = \tau \geq 2t_r \approx 1/B_T$$

and $f_s = 1/T_s \geq 2W$. Taking all values to be optimum with respect to noise reduction, we obtain the upper bound

$$\left(\frac{S}{N}\right)_D \leq \frac{1}{8}\left(\frac{B_T}{W}\right)^2 S_x\gamma \qquad \text{PPM} \tag{10}$$

Hence, PPM performance improves as the *square* of the bandwidth ratio B_T/W. A similar optimization for PDM with $\mu_p = \mu\tau_0 \leq \tau_0$ yields the less-impressive result

$$\left(\frac{S}{N}\right)_D \leq \frac{1}{2}\frac{B_T}{W} S_x\gamma \qquad \text{PDM} \tag{11}$$

To approach the upper bound, a PDM wave must have a 50% duty cycle so that $\tau_0 \approx T_s/2$.

Practical PPM and PDM systems may fall short of the values predicted in Eqs. (10) and (11) by 10 dB or more. Consequently, the noise reduction does not measure up to that of wideband FM. But remember that the average power S_R comes from short-duration high-power pulses rather than being continuously delivered as in CW modulation. Power-supply considerations may therefore favor pulsed operation in some circumstances.

Exercise 10.4-2 Explain why a single-channel PDM system must have $\mu\tau_0 \leq 1/4W$. Then derive Eq. (11) from Eq. (9) with $\mu_p = \mu\tau_0$.

False-Pulse Threshold Effect

Suppose you try to increase the value of $(S/N)_D$ in a PDM or PPM system by making B_T very large. Since $\overline{n^2}$ increases with B_T, the noise variations in $v(t) = x_p(t) + n(t)$ will eventually dominate and be mistaken for signal pulses. If these *false pulses* occur often, the reconstructed waveform has no relationship to $x(t)$ and the message will have been completely lost. Hence, pulse-time modulation involves a false-pulse *threshold effect*, analogous to the threshold effect in wideband FM. This effect does not exist in PAM with synchronization because we always know when to measure the amplitude.

To determine the threshold level, we'll say that false pulses are sufficiently infrequent if $P(n \geq A) \leq 0.01$. For gaussian noise with $\sigma_n^2 = \overline{n^2} = \eta B_T$, the corresponding threshold condition is approximately

$$A \geq 2\sigma_n$$

so the pulse must be strong enough to "lift" the noise by at least twice its rms value. (This is the same condition as the *tangential sensitivity* in pulsed radar

systems.) Using the fact that $A^2 = S_R/\tau_0 f_s$, we have

$$\frac{S_R}{\tau_0 f_s} \geq 4\eta B_T$$

or

$$\gamma_{th} = \left(\frac{S_R}{\eta W}\right)_{min} = 4\tau_0 f_s \frac{B_T}{W} \geq 8 \qquad (12)$$

This threshold level is appreciably less than that of FM, so PPM could be advantageous for those situations where FM would be below its threshold point.

10.5 PROBLEMS

10.1-1 Consider the chopper-sampled waveform in Eq. (3) with $\tau = T_s/2$, $f_s = 100$ Hz, and $x(t) = 2 + 2\cos 2\pi 30t + \cos 2\pi 80t$. Draw and label the one-sided line spectrum of $x_s(t)$ for $0 \leq f \leq 300$ Hz. Then find the output waveform when $x_s(t)$ is applied to an ideal LPF with $B = 75$ Hz.

10.1-2 Do Prob. 10.1-1 with $x(t) = 2 + 2\cos 2\pi 30t + \cos 2\pi 140t$.

10.1-3 The usable frequency range of a certain amplifier is f_ℓ to $f_\ell + B$, with $B \gg f_\ell$. Devise a system that employs bipolar choppers and allows the amplifier to handle signals having significant DC content and bandwidth $W \ll B$.

10.1-4 The baseband signal for FM stereo (Fig. 8.2-5*b*) is

$$x_b(t) = [x_L(t) + x_R(t)] + [x_L(t) - x_R(t)]\cos \omega_s t + A \cos \omega_s t/2$$

with $f_s = 38$ kHz. The chopper system in Fig. P10.1-4 is intended to generate this signal. The LPF has gain K_1 for $|f| \leq 15$ kHz, gain K_2 for $23 \leq |f| \leq 53$ kHz, and rejects $|f| \geq 99$ kHz. Use a sketch to show that $x_s(t) = x_L(t)s(t) + x_R(t)[1 - s(t)]$, where $s(t)$ is a unipolar switching function with $\tau = T_s/2$. Then find the necessary values of K_1 and K_2.

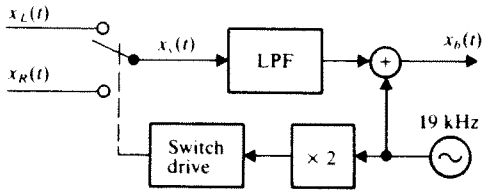


Figure P10.1-4

10.1-5 A popular *stereo decoder* circuit employs transistor switches to generate $v_L(t) = x_1(t) - Kx_2(t)$ and $v_R(t) = x_2(t) - Kx_1(t)$ where K is a constant, $x_1(t) = x_b(t)s(t)$, $x_2(t) = x_b(t)[1 - s(t)]$, $x_b(t)$ is the FM stereo baseband signal in Prob. 10.1-4, and $s(t)$ is a unipolar switching function with $\tau = T_s/2$. (*a*) Determine K such that lowpass filtering of $v_L(t)$ and $v_R(t)$ yields the desired left- and right-channel signals. (*b*) What's the disadvantage of a simpler switching circuit that has $K = 0$?

10.1-6 Derive Eq. (11) using Eq. (14), Sect. 2.4.

10.1-7 Suppose $x(t)$ has the spectrum in Fig. P10.1-7 with $f_u = 25$ kHz and $W = 10$ kHz. Sketch $x_\delta(f)$ for $f_s = 60$, 45, and 25 kHz. Comment in each case on the possible reconstruction of $x(t)$ from $x_\delta(t)$.

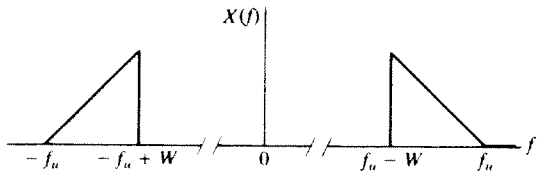


Figure P10.1-7

10.1-8 Consider the bandpass signal spectrum in Fig. P10.1-7 whose Nyquist rate is $f_s = 2f_u$. However, the *bandpass sampling theorem* states that $x(t)$ can be reconstructed from $x_\delta(t)$ by *bandpass* filtering if $f_s = 2f_u/m$ and the integer m satisfies $(f_u/W) - 1 < m \leq f_u/W$. (*a*) Find m and plot f_s/W versus f_u/W for $0 < f_u/W \leq 5$. (*b*) Check the theorem by plotting $X_\delta(f)$ when $f_u = 2.5W$ and $f_s = 2.5W$. Also show that the higher rate $f_s = 4W$ would not be acceptable.

10.1-9 The signal $x(t) = \text{sinc}^2 5t$ is ideally sampled at $t = 0, \pm 0.1, \pm 0.2, \ldots$, and reconstructed by an ideal LPF with $B = 5$, unit gain, and zero time delay. Carry out the reconstruction process graphically, as in Fig. 10.1-6, for $|t| \leq 0.2$.

10.1-10 A rectangular pulse with $\tau = 2$ is ideally sampled and reconstructed using an ideal LPF with $B = f_s/2$. Sketch the resulting output waveforms when $T_s = 0.8$ and 0.4, assuming one sample time is at the center of the pulse.

10.1-11 Suppose an ideally sampled wave is reconstructed using a *zero-order hold* (Example 3.1-3) with time delay $T = T_s$. (*a*) Find and sketch $y(t)$ to show that the reconstructed waveform is a *staircase* approximation of $x(t)$. (*b*) Sketch $|Y(f)|$ for $X(f) = \Pi(f/2W)$ with $W \ll f_s$. Comment on the significance of your result.

10.1-12‡ The reconstruction system in Fig. P10.1-12 is called a *first-order hold*. Each block labeled ZOH is a zero-order hold (Example 3.1-3) with time delay $T = T_s$. (*a*) Find $h(t)$ and sketch $y(t)$ to interpret the reconstruction operation. (*b*) Show that $H(f) = T_s(1 + j2\pi f T_s)(\text{sinc}^2 f T_s)\exp(-j2\pi f T_s)$. Then sketch $|Y(f)|$ for $X(f) = \Pi(f/2W)$ with $W < f_s/2$.

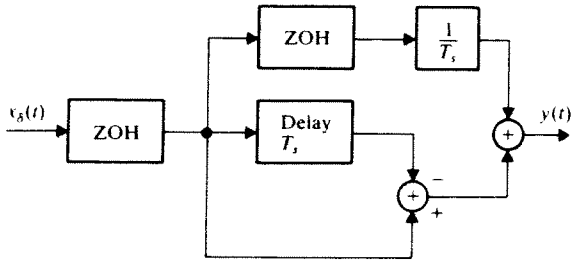


Figure P10.1-12

10.1-13‡ Use Parseval's theorem and Eq. (14) with $T_s = 1/2W$ and $B = W$ to show that the energy of a bandlimited signal is related to its sample values by

$$E = (1/2W) \sum_{k=-\infty}^{\infty} |x(k/2W)|^2$$

10.1-14 The *frequency-domain sampling theorem* says that if $x(t)$ is a *timelimited* signal, such that $x(t) = 0$ for $|t| \geq T$, then $X(f)$ is completely determined by its sample values $X(nf_0)$ with $f_0 \leq 1/2T$. Prove this theorem by writing the Fourier series for the periodic signal $v(t) = x(t) * [\Sigma_k \delta(t - kT_0)]$, where $T_0 \geq 2T$, and using the fact that $x(t) = v(t)\Pi(t/2T)$.

10.1-15 A signal with period $T_x = 0.08$ μs is to be displayed using a sampling oscilloscope whose internal high-frequency response cuts off at $B = 6$ MHz. Determine maximum values for the sampling frequency and the bandwidth of the pre-sampling LPF.