8 Sampling Process

We have already talked about the transmission of a discrete source over a discrete channel in the previous sections. However, if the message signal happens to be analog in nature, as in speech signal or video signal, then it has to be converted into digital form before it can be transmitted by digital means. The *sampling process* is the first step in *analog-to-digital conversion*. Two other processes, *quantizing* and *encoding*, are also involved in the conversion. These operations will be discussed in subsequent sections.

Sampling also provides the basis for the time division multiplexing of signals which is a method for simultaneous transmission of several signals through the same communication channel without mutual interference. This is explained in the following:

8.1 Time division multiplexing versus frequency division multiplexing

As we saw before, a modulated signal using a sinusoidal carrier has the following general form:

$$\phi(t) = a(t)\cos[\theta(t)] = a(t)\cos[\omega_c t + \gamma(t)] \tag{462}$$

where ω_c is called the carrier frequency.

In amplitude modulation, the phase $\gamma(t)$ in (462) is constant and the amplitude $a(t)$ is changed in proportion to the input signal. In angle modulation, the amplitude $a(t)$ in (462) is constant and the phase $\gamma(t)$ is changed in proportion to the input signal.

Modulation results in a translation of the frequency components of the input signal to higher frequencies around ω_c. The frequency translating property of modulation can be used to transmit a large number of signals at the same time without mutual interference. This is called the frequency division multiplexing and is based on using different carrier frequencies for different signals (refer to Fig. (56)). If the bandwidth of the signals is ω_m, then two subsequent modulating frequency should be at least $2\omega_m$ apart. In the receiver side, depending of the application, one can demodulate all the signals simultaneously or use a tunable bandpass filter to separate one of the signals.

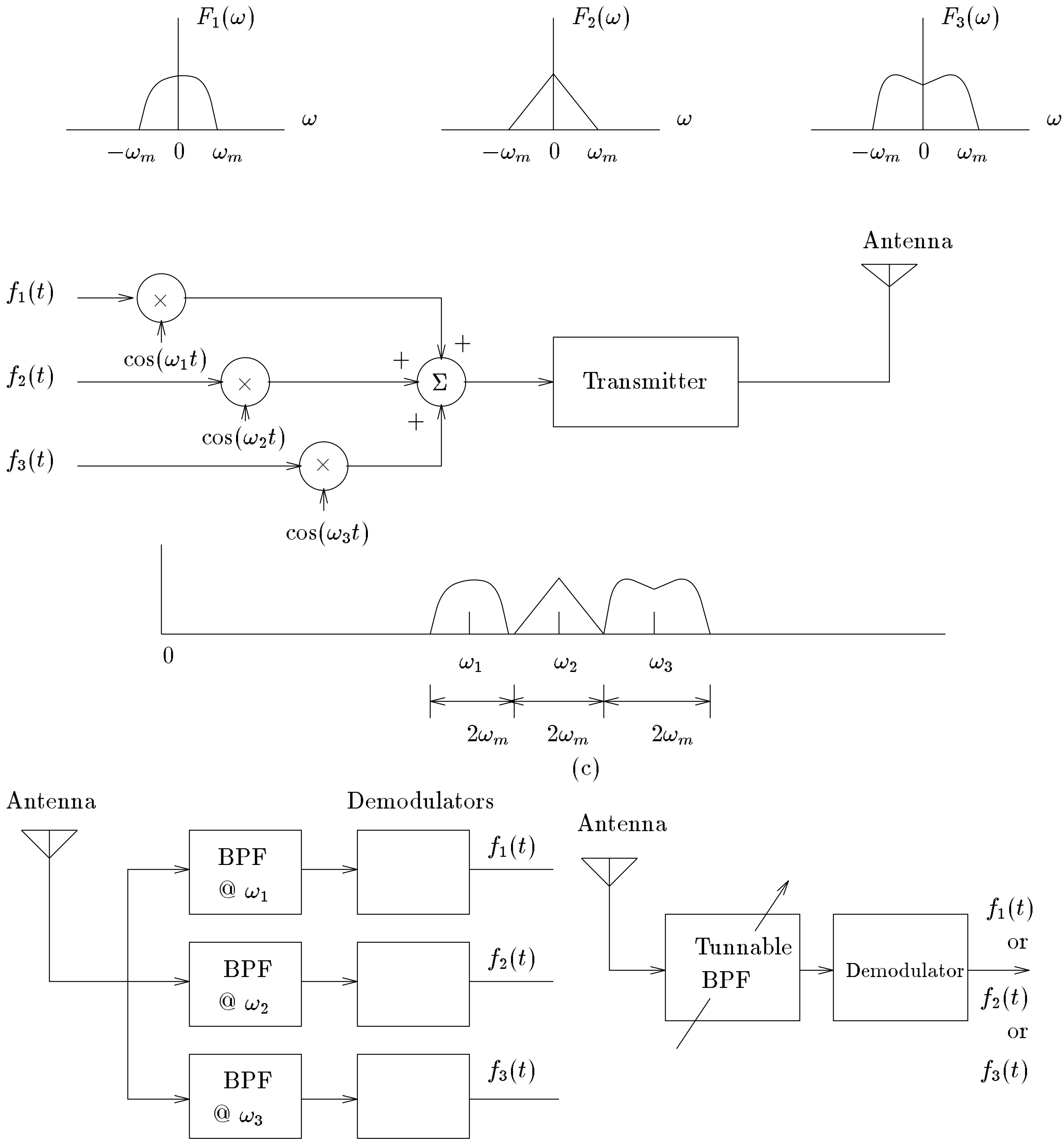


Figure 56: Frequency division multiplexing (FDM).

Time division multiplexing is an alternative method for the simultaneous transmission of different signals. It is based on dividing the time axis into nonoverlapping segments and assigning each segment to a different input signal. This is explained in Fig. (57). The first step in TDM process is sampling.

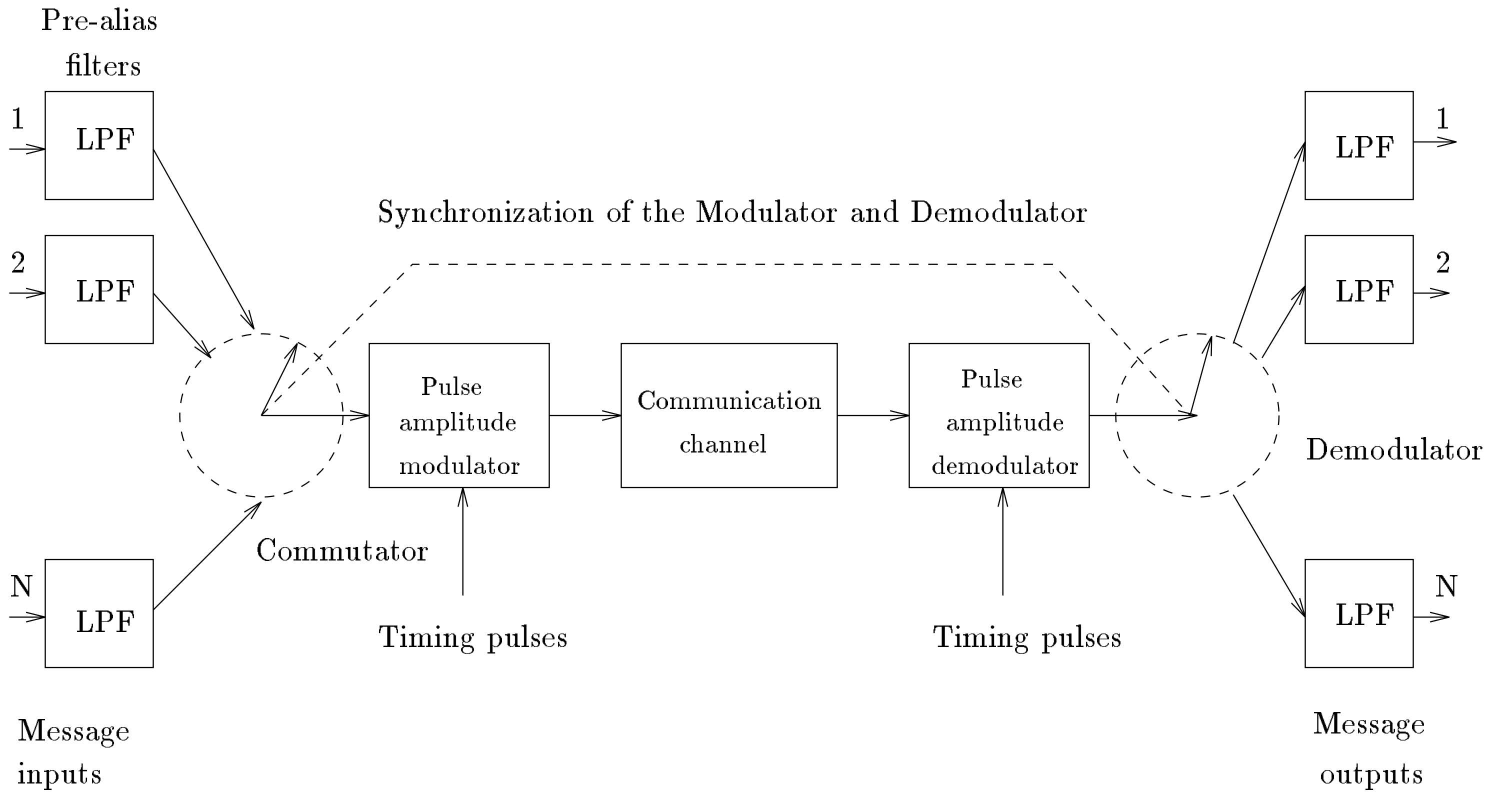
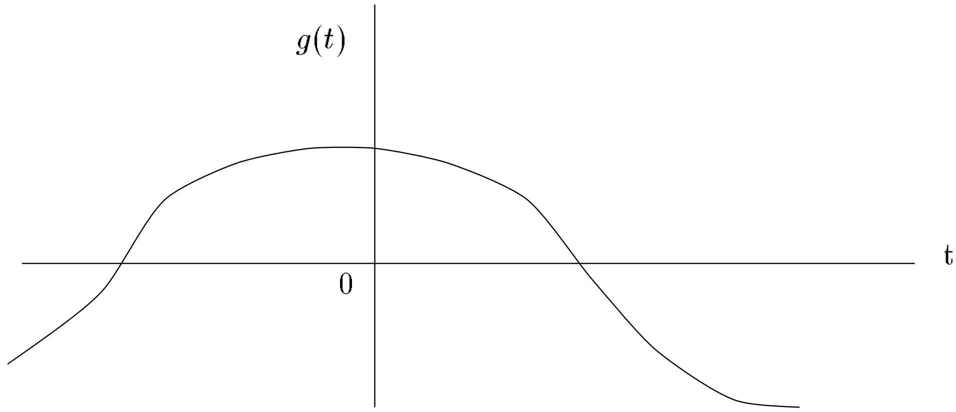


Figure 57: Time division multiplexing (TDM)

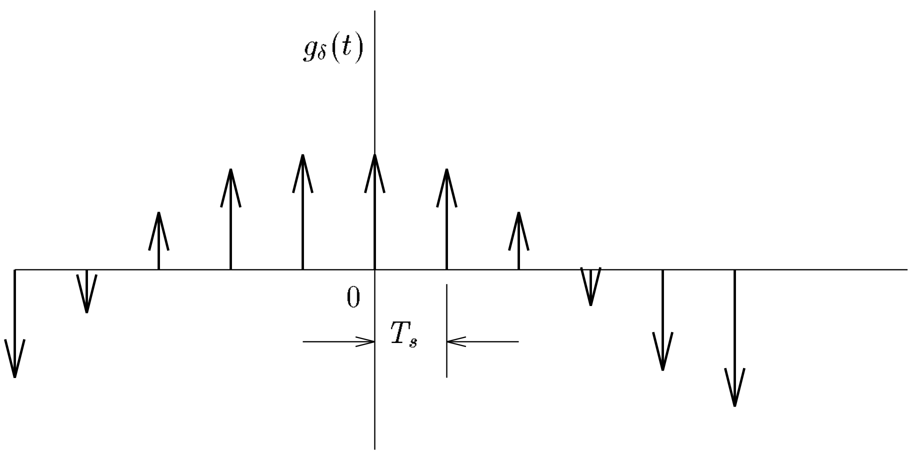
8.2 Sampling theory

In the sampling process, a continuous-time signal is converted into a discrete-time signal by measuring the signal at periodic instants of time. For the sampling process to be of practical utility, it is necessary that we choose the sampling rate properly, so that the discrete time signal resulting from the process uniquely defines the original continuous-time signal in an efficient way (using a small number of samples).

Consider an analog signal $g(t)$ that is continuous in both time and amplitude. We assume that $g(t)$ has infinite duration but finite energy. A segment of the signal $g(t)$ is depicted in Fig. (58). Let the sample values of the signal $g(t)$ at times $t = 0, \pm T_s, \pm 2T_s, \ldots$ be denoted by the series $\{g(nT_s), n = 0, \pm 1, \pm 2, \ldots\}$. We refer to T_s as the *sampling period* and to $f_s = 1/T_s$ as the *sampling rate*.



(a)



(b)

Figure 58: Illustration of the ideal sampling process. (a)Analog signal. (b) Discrete-time signal.

We define the discrete-time signal, $g_\delta(t)$, that results from the sampling process as

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} g(nT_s)\delta(t - nT_s) \tag{463}$$

where $\delta(t - nT_s)$ is a *Dirac delta function* located at time $t = nT_s$.

We have

$$g(t)\delta(t - nT_s) = g(nT_s)\delta(t - nT_s) \tag{464}$$

$$\begin{aligned} g_\delta(t) &= g(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \\ &= g(t)\delta_{T_s}(t) \end{aligned} \tag{465}$$

where $\delta_{T_s}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s)$ is the *Dirac Comb* or *ideal sampling function.*

From the properties of the Fourier transform, we know that the multiplication of the two time functions is equivalent to the convolution of their respective Fourier transforms. let $G(f)$ and $G_\delta(f)$ denote the Fourier transforms of $g(t)$ and $g_\delta(t)$, respectively.

It is known that for the Fourier transform of $\delta_{T_s}(t)$, we have

$$F[\delta_{T_s}(t)] = f_s \sum_{m=-\infty}^{\infty} \delta(f - mf_s) \tag{466}$$

where $f_s = 1/T_s$.

Transforming equation (465) into the frequency domain, we obtain

$$G_\delta(f) = G(f) * \left[f_s \sum_{m=-\infty}^{\infty} \delta(f - mf_s) \right] \tag{467}$$

where $*$ denotes convolution. Interchanging the order of summation and convolution yields

$$G_\delta(f) = f_s \sum_{m=-\infty}^{\infty} G(f) * \delta(f - mf_s) \tag{468}$$

$$G_\delta(f) = f_s \sum_{m=-\infty}^{\infty} G(f - mf_s) \tag{469}$$

Note that $G_\delta(f)$ represents a periodic extension of the original spectrum $G(f)$. This means that the process of uniformly sampling a signal in the time domain results in a periodic spectrum in the frequency domain with a period equal to the sampling rate.

Taking the Fourier transform of both sides of equation (463) and noting that the Fourier transform of the delta function $\delta(t - nT_s)$ is equal to $\exp(-j2\pi nfT_s)$ results in

$$G_\delta(f) = \sum_{n=-\infty}^{\infty} g(nT_s)\exp(-j2\pi nfT_s) \tag{470}$$

This relation may be viewed as a *complex Fourier series representation* of the periodic frequency function $G_\delta(f)$, with the sequence of samples $\{g(nT_s)\}$, defining the coefficients of the expansion.

Suppose that the signal is strictly band-limited, with no frequency components higher than W hertz, as illustrated in Fig. 59.

Suppose also that we choose the sampling period $T_s = 1/2W$. Then the corresponding spectrum $G_\delta(f)$ of the sampled signal $g_\delta(t)$ is as shown in Fig. 59b. Putting $T_s = 1/2W$ in equation (470) yields

$$G_\delta(f) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right)\exp\left(-\frac{j\pi nf}{W}\right) \tag{471}$$

Putting $f_s = 2W$ in Eq. 469, we have

$$G_\delta(f) = 2WG(f), \quad -W \leq f \leq W \tag{472}$$

or,

$$G(f) = \frac{1}{2W}G_\delta(f), \quad -W \leq f \leq W \tag{473}$$

It follows from equation (471) that we may also write

$$G(f) = \frac{1}{2W}\sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right)\exp\left(-\frac{j\pi nf}{W}\right), \quad -W \leq f \leq W \tag{474}$$

Therefore if the sample values $g(n/2W)$ of the signal $g(t)$ are specified for all time, then the Fourier transform $G(f)$ of the signal is uniquely determined by using the Fourier series of equation (474). In other words, the sequence $\{g(n/2W)\}$ contains all the information of $g(t)$.

Consider next the problem of reconstructing the signal $g(t)$ from the sequence of sample values $\{g(n/2W)\}$. We get

$$g(t) = \int_{-\infty}^{\infty} G(f)\exp(j2\pi ft)df$$

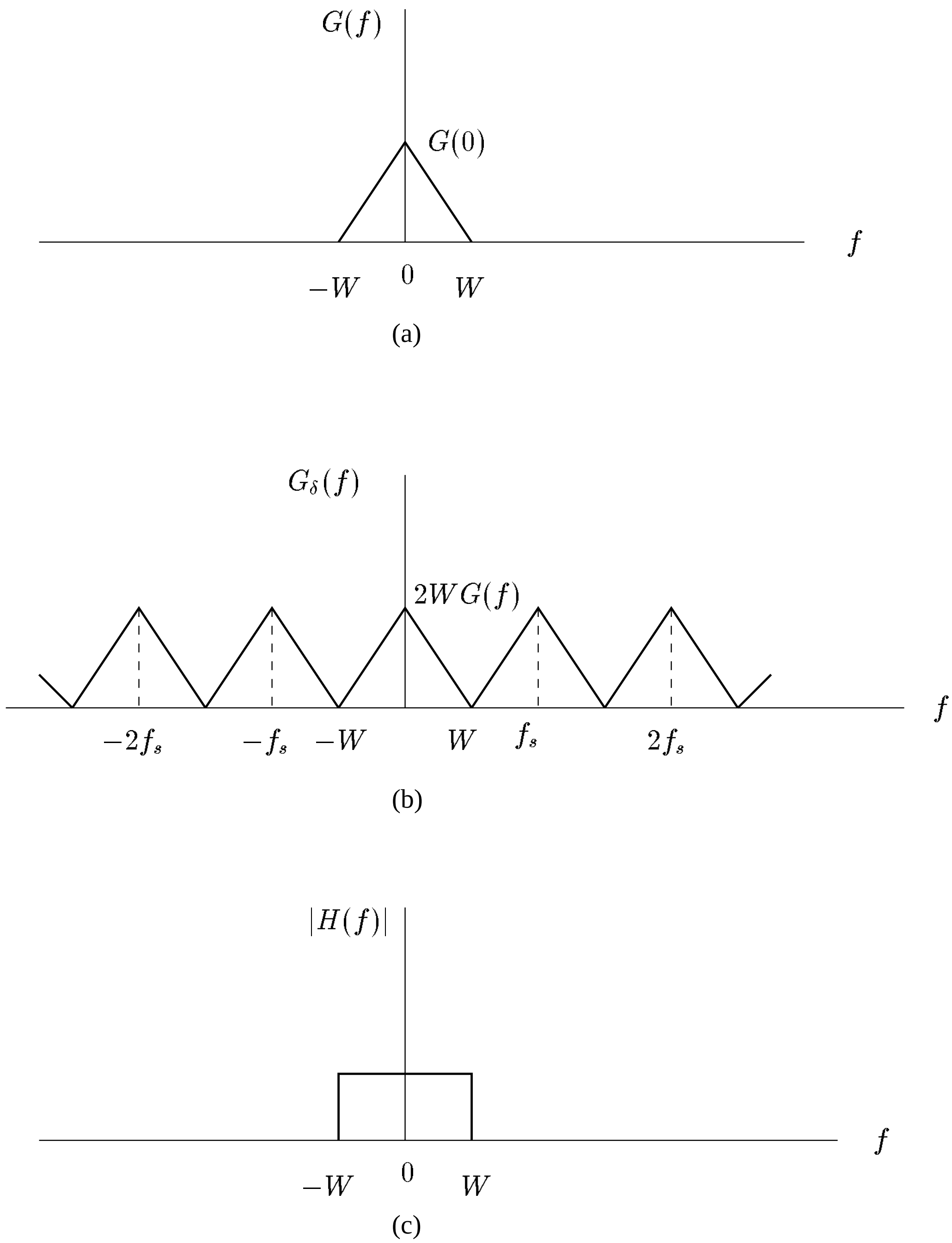


Figure 59: (a) Spectrum of signal $g(t)$. (b) Spectrum of sampled signal $g_\delta(t)$ for a sampling rate $f_s = 2W$. (c) Ideal amplitude response of reconstruction filter.

$$= \int_{-W}^{W} \frac{1}{2W} \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \exp\left(-\frac{j\pi n f}{W}\right) \exp(j2\pi f t) df$$

Interchanging the order of the summation and integration, we obtain,

$$g(t) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \frac{1}{2W} \int_{-W}^{W} \exp\left[j2\pi f\left(t - \frac{n}{2W}\right)\right] df \tag{475}$$

The integral term in Eq. (475) may be readily evaluated, yielding

$$g(t) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \frac{\sin(2\pi Wt - n\pi)}{(2\pi Wt - n\pi)} \tag{476}$$

Using the notation,

$$\text{sinc } x = \frac{\sin(\pi x)}{\pi x} \tag{477}$$

we obtain,

$$g(t) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \text{sinc}(2Wt - n) \tag{478}$$

The sinc function exhibits an important property known as the *interpolatory property*, which is described as follows:

$$\text{sinc } x = \begin{cases} 1 & x = 0 \\ 0 & x = \pm 1, \pm 2, \ldots \end{cases} \tag{479}$$

Considering this property, it is seen that equation (478) provides an interpolation formula for reconstructing the original signal $g(t)$ from the sequence of sample values $\{g(n/2W)\}$, with the sinc function $\text{sinc}(2Wt)$ playing the role of an interpolation function.

A practical method for the reconstruction of the time signal $g(t)$ from its samples is as follows: By inspection of the spectrum of Fig. (59b), we see that the original signal $g(t)$ may be recovered exactly from the sequence of samples $\{g(n/2W)\}$ by passing it through an ideal low-pass filter of bandwidth W. The ideal amplitude response of the *reconstruction filter* is shown in Fig. (59c).

8.3 Signal space interpolation

The function $\text{sinc}(2Wt - n)$, where n is an integer, is one of a family of shifted sinc functions that are mutually orthogonal. To prove this, we use the formula

$$\int_{-\infty}^{\infty} g_1(t)g_2^*(t)dt = \int_{-\infty}^{\infty} G_1(f)G_2^*(f)df \tag{480}$$

Put

$$g_1(t) = \text{sinc}(2Wt - n) = \text{sinc}\left[2W(t - \frac{n}{2W})\right] \tag{481}$$

and

$$g_2(t) = \text{sinc}(2Wt - m) = \text{sinc}\left[2W(t - \frac{m}{2W})\right] \tag{482}$$

we have the Fourier transform pair:

$$\text{sinc}(2Wt) \rightleftharpoons \frac{1}{2W}\text{rect}\left(\frac{f}{2W}\right) \tag{483}$$

$$\text{rect}(x) = \begin{cases} 1 & -\frac{1}{2} < x < \frac{1}{2} \\ 0 & |x| > \frac{1}{2} \end{cases} \tag{484}$$

Recall that if $x(t) \rightleftharpoons X(f)$ then, $x(t - t_0) \rightleftharpoons e^{-j2\pi f t_0}X(f)$. Using this fact we obtain,

$$G_1(f) = \frac{1}{2W}\text{rect}\left(\frac{f}{2W}\right)\exp\left(-\frac{j\pi n f}{W}\right) \tag{485}$$

and

$$G_2(f) = \frac{1}{2W}\text{rect}\left(\frac{f}{2W}\right)\exp\left(-\frac{j\pi m f}{W}\right) \tag{486}$$

Hence,

$$\begin{aligned} \int_{-\infty}^{\infty} \text{sinc}(2Wt - n)\text{sinc}(2Wt - m)dt &= \left(\frac{1}{2W}\right)^2 \int_{-W}^{W} \exp\left[-\frac{j\pi f}{W}(n - m)\right] df \\ &= \frac{\sin[\pi(n - m)]}{2W\pi(n - m)} \\ &= \frac{1}{2W}\text{sinc}(n - m) \end{aligned}$$

This result equals 1/2W when $n = m$, and zero when $n \neq m$. We therefore have

$$\int_{-\infty}^{\infty} \text{sinc}(2Wt - n)\text{sinc}(2Wt - m)dt = \begin{cases} \frac{1}{2W} & n = m \\ 0 & n \neq m \end{cases} \tag{487}$$

This proves the orthogonality of the sinc functions.

Equation (478) represents the expansion of the signal $g(t)$ as an infinite sum of orthogonal functions with the coefficients of the expansion, $g(n/2W)$, defined by

$$g\left(\frac{n}{2W}\right) = 2W \int_{-\infty}^{\infty} g(t)\text{sinc}(2Wt - n)dt \tag{488}$$

The minimum sampling rate of $2W$ samples per second, for a signal band-width of W hertz, is called the *Nyquist rate*. Correspondingly, the reciprocal $1/2W$ is called the *Nyquist interval*.

8.4 Quadrature sampling of band-pass signals

Consider a band-pass signal $g(t)$ (limited to the frequency band $[f_c - W, f_c + W]$ and its negative) whose spectrum is illustrated in Fig. (60a) .

Let $g_I(t)$ denote the in-phase component of the band-pass signal $g(t)$ and $g_Q(t)$ denote its quadrature component. We may then express $g(t)$ in terms of $g_I(t)$ and $g_Q(t)$ as follows:

$$g(t) = g_I(t)\cos(2\pi f_c t) - g_Q(t)sin(2\pi f_c t) \tag{489}$$

We know that the two signals $g_I(t)$, $g_Q(t)$ are low pass and limited to a frequency band of $[-W, W]$. This means that we can represent each of these two signals using $2W$ samples per second. This results in a total of $4W$ samples per second for the bandpass signal $g(t)$ which has a band width of $2W$. To reconstruct the original band-pass signal from its quadrature-sampled version, we first reconstruct $g_I(t)$, $g_Q(t)$ and then combine them using (489).

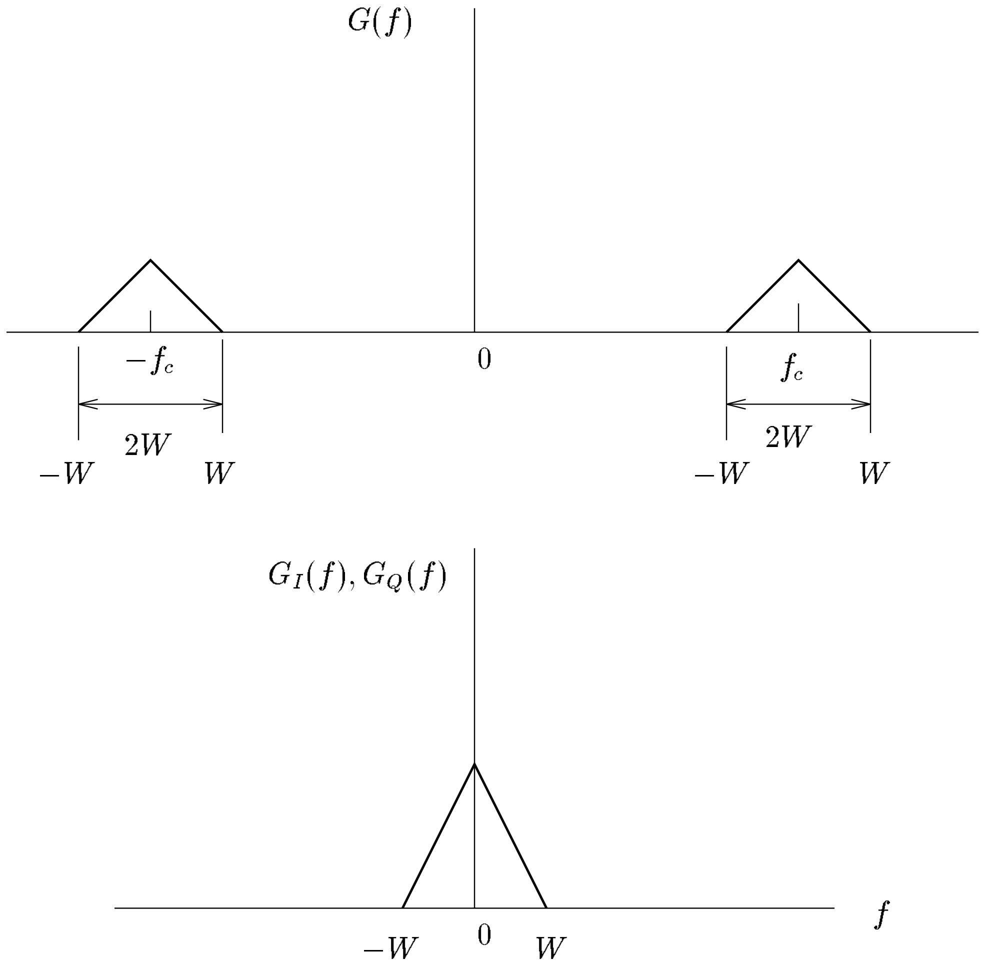


Figure 60: (a) Spectrum of band-pass signal $g(t)$. (b) Spectrum of low-pass in-phase component $g_I(t)$ and quadrature component $g_Q(t)$.

8.5 Sampling procedure

We know that a signal cannot be finite in both time and frequency. In practice we have to work with a finite segment of the signal, in which case the spectrum cannot be strictly band-limited. consequently when a signal of finite duration is sampled, an *error in the reconstruction* occurs as a result of the sampling process.

The spectrum $G_\delta(f)$ of the discrete-time signal $g_\delta(t)$, resulting from the use of the idealized sampling, is the sum of $G(f)$ and an infinite number of frequency-shifted replica of it. If $G(f)$ is not bandlimited, we find that points of the frequency-shifted replica are *folded over* inside the desired spectrum. This is called *aliasing* or *foldover*.

Prior to sampling, a low-pass pre-alias filter is used to attenuate those high frequency components of the signal. The filtered signal is sampled at a rate slightly higher than the Nyquist rate $2W$, where W is the cutoff frequency of the pre-alias filter.

The use of a sampling rate f_s higher than the Nyquist rate $2W$ has the desirable effect of making it somewhat easier to design the low-pass reconstruction filter so as to recover the original analog signal from its sampled version. With such a sampling rate, we find that there are gaps, each of width $f_s - 2W$ between the frequent shifted replica of $G(f)$. Accordingly, we may design the reconstruction filter with a higher degree of flexibility.

8.6 Practical aspects of sampling and signal recovery

8.6.1 Ordinary samples of finite duration

Consider the waveforms $g(t)$, $c(t)$, and $s(t)$ illustrated in parts (a), (b) and (c) of Fig. (61) respectively.

We have

$$s(t) = c(t)g(t) \tag{492}$$

However, $c(t)$ may be expressed in the form of a complex Fourier series as

$$c(t) = f_s T A \sum_{n=-\infty}^{\infty} \operatorname{sinc}(n f_s T) \exp(j2\pi n f_s t) \tag{493}$$

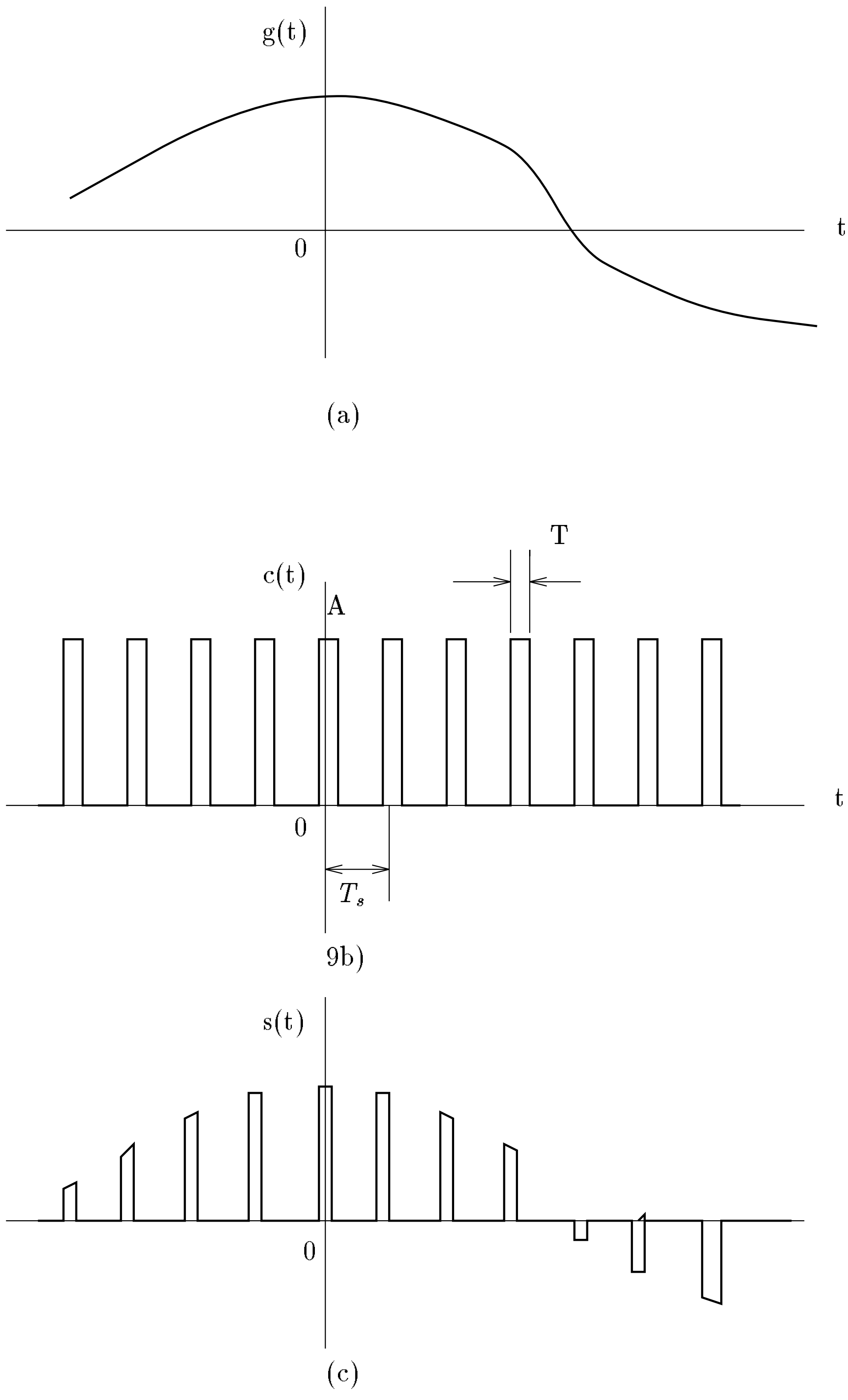


Figure 61: (a) Analog signal. (b) Sampling function. (c) Sampled signal.

where $T_s f_s = 1$ defines the sampling rate f_s.

$$s(t) = f_s T A \sum_{n=-\infty}^{\infty} \text{sinc}(n f_s T) \exp(j2\pi n f_s t) g(t) \tag{494}$$

Taking the Fourier transform, we get

$$S(f) = f_s T A \sum_{m=-\infty}^{\infty} \text{sinc}(m f_s T) G(f - m f_s) \tag{495}$$

Where $S(f) = F[s(t)]$ and $G(f) = F[g(t)]$.

The relation between the spectra $G(f)$ and $S(f)$ is illustrated in Fig. (62).

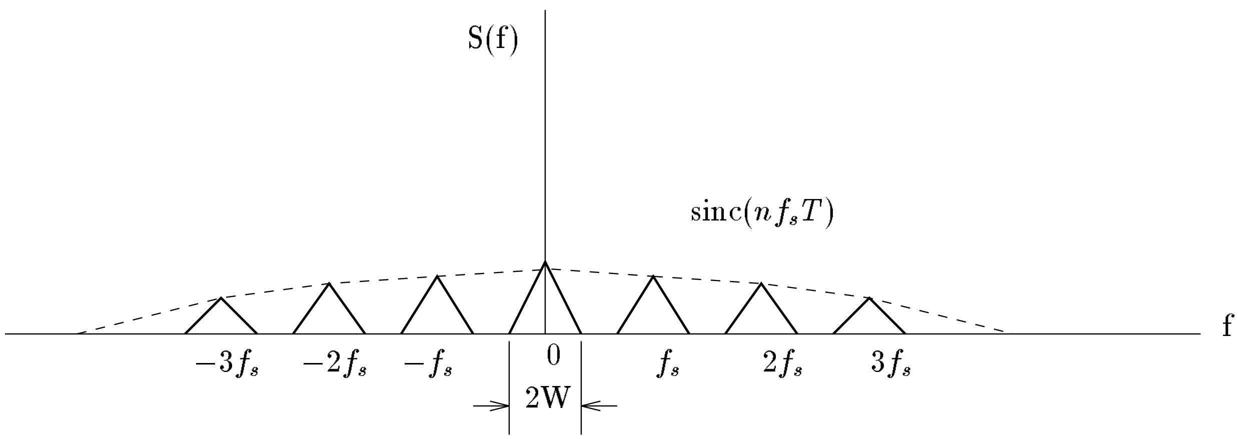


Figure 62: Illustrating the effect of using ordinary pulses of finite duration on the spectrum of a sampled signal

Signal $g(t)$ can be recovered from $s(t)$ with no distortion by passing $s(t)$ through an ideal low-pass filter.

8.6.2 Flat-top samples

Consider the situation illustrated in Fig. (63), we may write

$$s(t) = \sum_{n=-\infty}^{\infty} g(nT_s) h(t - nT_s) \tag{496}$$

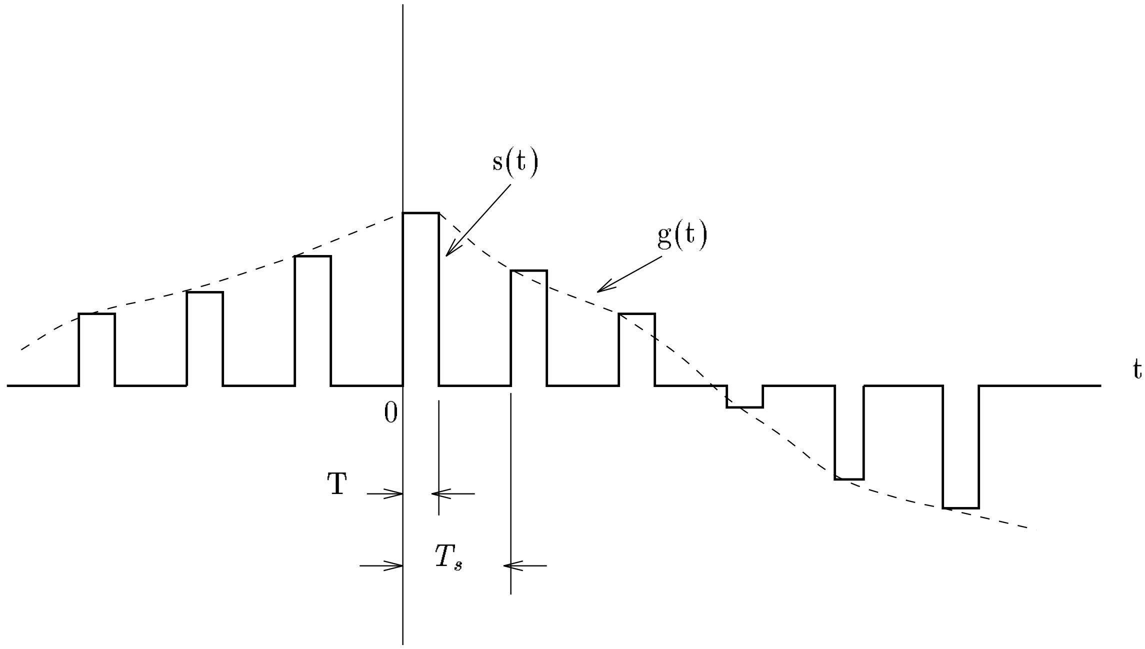


Figure 63: Flat-top Samples

$$h(t) = \begin{cases} 1 & 0 < t < T \\ 0 & t < 0,\ and\ t > T \end{cases}$$
$$= \text{rect}\left(\frac{t}{T} - \frac{1}{2}\right) \tag{497}$$

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} g(nT_s)\delta(t - nT_s) \tag{498}$$

$$\begin{aligned} g_\delta(t) * h(t) &= \int_{-\infty}^{\infty} g_\delta(\tau)h(t-\tau)d\tau \\ &= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g(nT_s)\delta(\tau - nT_s)h(t-\tau)d\tau \\ &= \sum_{n=-\infty}^{\infty} g(nT_s) \int_{-\infty}^{\infty} \delta(\tau - nT_s)h(t-\tau)d\tau \end{aligned} \tag{499}$$

$$g_\delta(t) * h(t) = \sum_{n=-\infty}^{\infty} g(nT_s)h(t - nT_s) \tag{500}$$

Therefore,

$$s(t) = g_\delta(t) * h(t) \tag{501}$$

Taking the Fourier transform, we get

$$S(f) = G_\delta(f)H(f) \tag{502}$$

Substituting of Eq. (469) into Eq. (502) yields

$$S(f) = f_s \sum_{m=-\infty}^{\infty} G(f - mf_s)H(f) \tag{503}$$

Suppose that $g(t)$ is strictly band-limited and that the sampling rate f_s is greater than the Nyquist rate. Then passing $s(t)$ through a low-pass reconstruction filter, we find that the spectrum of the resulting filter output is equal to $G(f)H(f)$.
From Eq. (497) we find that

$$H(f) = T\mathrm{sinc}(fT)\exp(-j\pi fT) \tag{504}$$

which is plotted in Fig (64b). Hence we see that by using the flat-top samples, we have introduced *amplitude distortion* as well as the *delay* of $T/2$.
The distortion caused by lengthening the samples is referred to as the *aperture effect*.
This distortion may be corrected by connecting an *equalizer* in cascade with the low-pass reconstruction filter. Ideally, the amplitude response of the equalizer is given by

$$\frac{1}{|H(f)|} = \frac{1}{T\mathrm{sinc}(fT)} = \frac{1}{T}\frac{\pi fT}{sin(\pi fT)} \tag{505}$$

8.7 Sample-and-hold Circuit for Signal Recovery

The output of the sample-and-hold circuit is given by

$$u(t) = \sum_{n=-\infty}^{\infty} g(nT_s)h(t - nT_s) \tag{506}$$

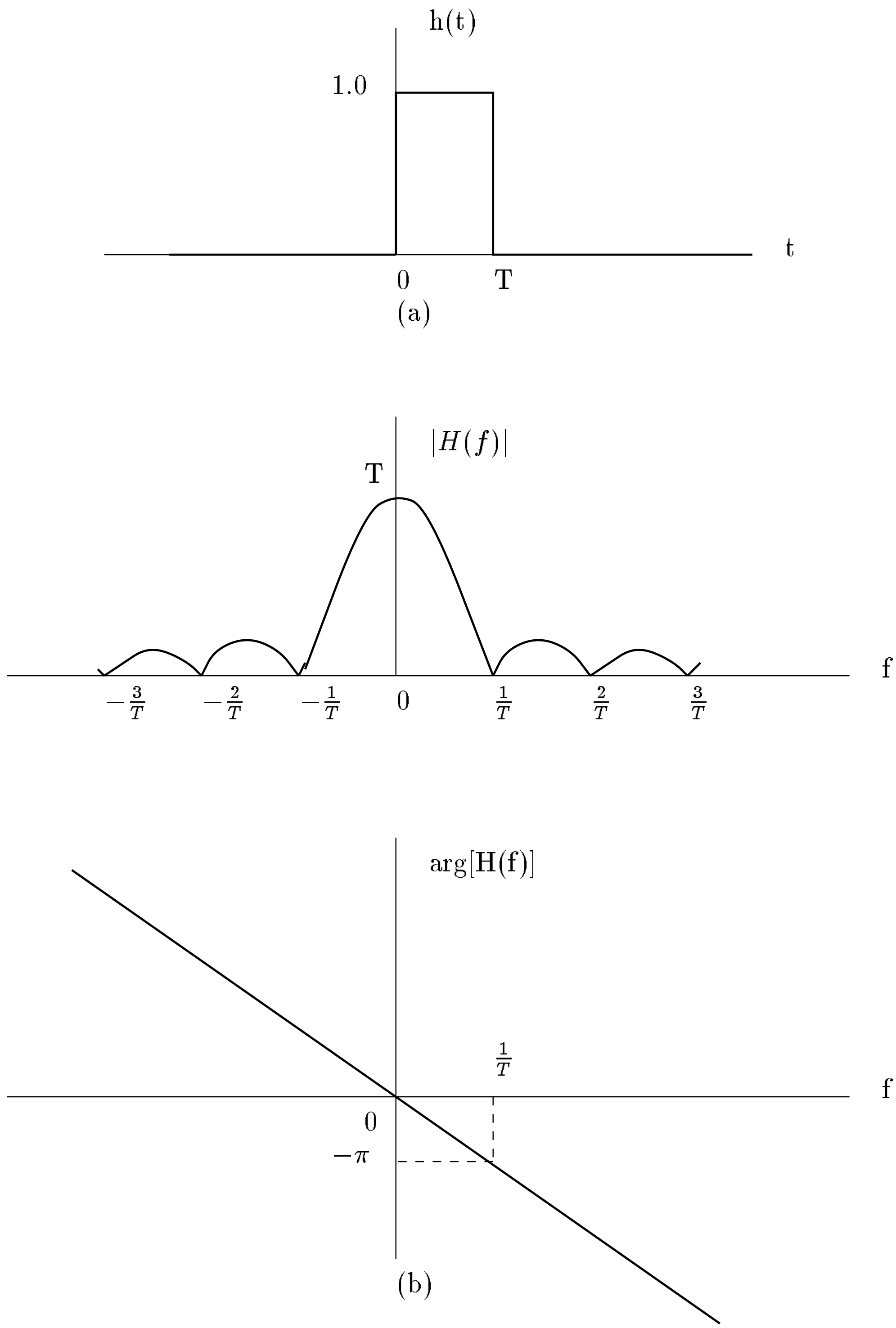


Figure 64: (a) Rectangular pulse h(t). (b) Spectrum H(f).

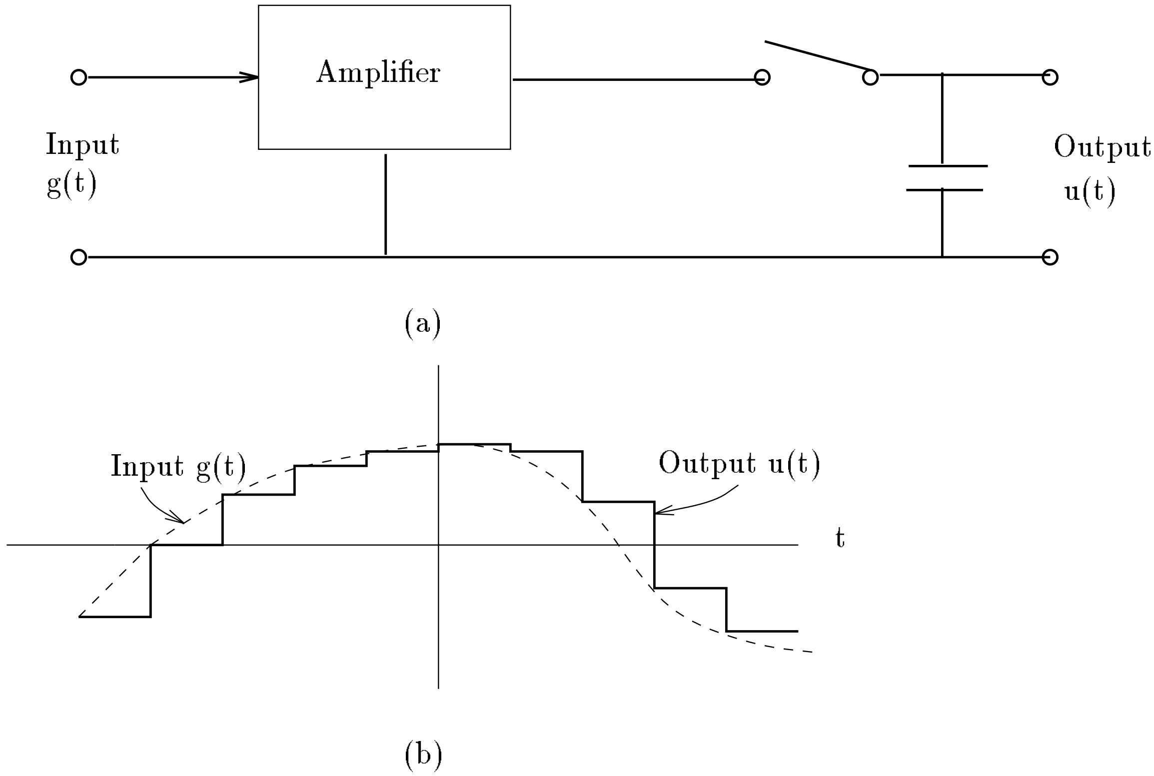


Figure 65: (a) Sample-and-hold circuit. (b) Idealized output waveform of the circuit.

where

$$h(t) = \begin{cases} 1 & 0 < t < T_s \\ 0 & t < 0, \ \ and \ t > T_s \end{cases} \tag{507}$$

$$U(f) = f_s \sum_{m=-\infty}^{\infty} H(f)G(f - mf_s) \tag{508}$$

where

$$H(f) = T_s \mathrm{sinc}(fT_s) \exp(-j\pi f T_s) \tag{509}$$

These operations are illustrated by the block diagram shown in Fig. (66).

8.8 Pulse-Amplitude Modulation

In pulse amplitude modulation (PAM), the amplitude of a carrier consisting of a periodic train of rectangular pulses is varied in proportion to sample values of a message signal.

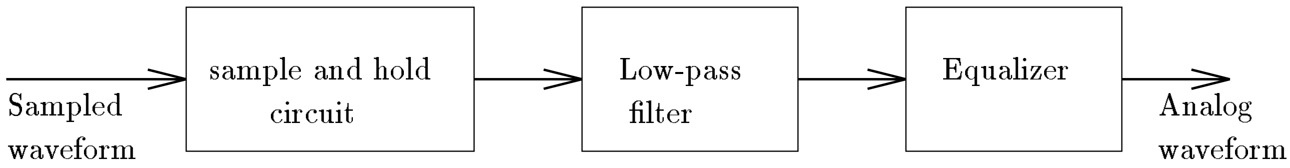


Figure 66: Components of a scheme for signal reconstruction

We find that the PAM so defined is exactly the same as *flat-top* sampling.

According to the definition of given before in terms of rectangular pulses, we would require a very wide band of frequencies to transmit *PAM*. However this need not be so if we were to formulate the definition of *PAM* in terms of a standard pulse, which the system is capable of transmitting. Let $\nu(t)$ denote such a pulse. We then define a *PAM* wave , $s(t)$ as follows

$$s(t) = \sum_{n=-\infty}^{\infty} g(nT_s)\nu(t - nT_s) \tag{510}$$

8.9 Time-Division Multiplexing (TDM)

The concept of **TDM** is illustrated by the block diagram shown in Fig. (57).

The function of the commutator is two-fold: (1) to take a narrow sample of each of the N input messages at a rate f_s (2) to sequentially interleave these N samples inside a sampling interval $T_s = 1/f_s$.

The use of time-division multiplexing introduces a *bandwidth expansion factor* N, because the scheme must squeeze N samples derived from N independent message signals into a time slot equal to one sampling interval.

9 Waveform Coding Techniques

Consider the problem of transmitting an analog source over a digital channel. There are three major operations involved in this transmission. These are the operations of (i) sampling (ii) quantization (analog-to-digital conversion), and (iii) encoding. Sampling changes the analog signal (which is a *continuous time*, *continuous amplitude* process) into a *discrete time*, *continuous amplitude* process. Then, quantization is achieved on this *discrete time*, *continuous amplitude* process to produce a *discrete time*, *discrete amplitude* process. We already talked about the sampling process. In this section, we address the problem of quantization. The encoding process will be discussed in a subsequent section.

9.1 Quantizing

In a linear system, the transfer characteristics between the input and the output is in the form of a straight line.

A quantizer is a nonlinear system based on a transfer characteristics looks like a staircase. An example is given in Fig. 67(a). In a quantizer, the range of the input sample values is divided into a finite set of *decision levels* or *decision thresholds* that are aligned with the risers of the staircase. The segment of the input axis located between two consequetive *decision levels* is denoted as a *quantization interval*. If the input is located in a given *quantization interval*, the output is assigned a discrete value which is aligned with the tread of the staircase. This discrete value is called the *representation level* or *reconstruction value* corresponding to the given quantization interval. In transfer characteristic shown in Fig. 67(a), the decision thresholds are at points $\pm\Delta/2, \pm 3\Delta/2, \pm 5\Delta/2, \ldots$ and the representation levels are located at points $0, \pm\Delta, \pm 2\Delta, \ldots$. This is called a *midtread type* quantizer. In transfer characteristic shown in Fig. 68(a), the decision thresholds are at points $0, \pm\Delta, \pm 2\Delta, \ldots$ and the representation levels are located at points $\pm\Delta/2, \pm 3\Delta/2, \pm 5\Delta/2, \ldots$. This is called a *midriser type* quantizer.

For an analog input sample that lies anywhere inside an interval of either transfer characteristics, the quantizer produces a discrete output equal to the midvalue of the pair of decision thresholds. A *quantization error* is introduced, the value of which equals the difference between the output and input values of the quantizer. Figures 67(b), 68(b) show examples of such a quantization error.

Definition: A quantizer is called *symmetric* if its transfer characteristic is symmetrical for positive and negative input values. Quantizers shown in Figs. 67(a), 68(a) are both symmetrical.

Definition: A quantizer is called *uniform* if the separation between its representation levels are the same with a common value is called the *step size*. Quantizers shown in Figs. 67(a), 68(a) are both uniform.

Definition: A quantizer is called *memoryless* in that the quantizer output is determined only by the value of a corresponding input sample, independent of the earlier analog samples.

9.2 Idle Channel Noise

In a quantizer of the midrise type, as in Fig. 68(a), zero input amplitude is encoded into one of the two innermost representations levels $\pm\Delta/2$. This results in a quantization error even if the input to the quantizer is equal to zero. This error is denoted as the *idle channel noise*. Assuming that the two representation levels $\pm\Delta/2$ are equiprobable, the idle channel noise for midriser quantizer has zero mean and a variance of $\Delta^2/4$. In a quantizer of the midtread type, as in Fig. 67(a), the output is zero for zero input and the idle channel noise is correspondingly zero.

9.3 Quantization Noise and Signal-to-Noise Ratio (SNR)

Consider a *symmetric, uniform, memoryless quantizer* with a total of L representation levels. Let x denote the quantizer input, and y denote the quantizer output. The transfer characteristic of the quantizer is shown by,

$$y = Q(x) \tag{511}$$

which is a staircase function. Suppose that we use the notation $\mathcal{J}_k$, $k = 1, 2, \ldots, L$, to shown the *quatization intervals*. The decision thresholds corresponding to the kth

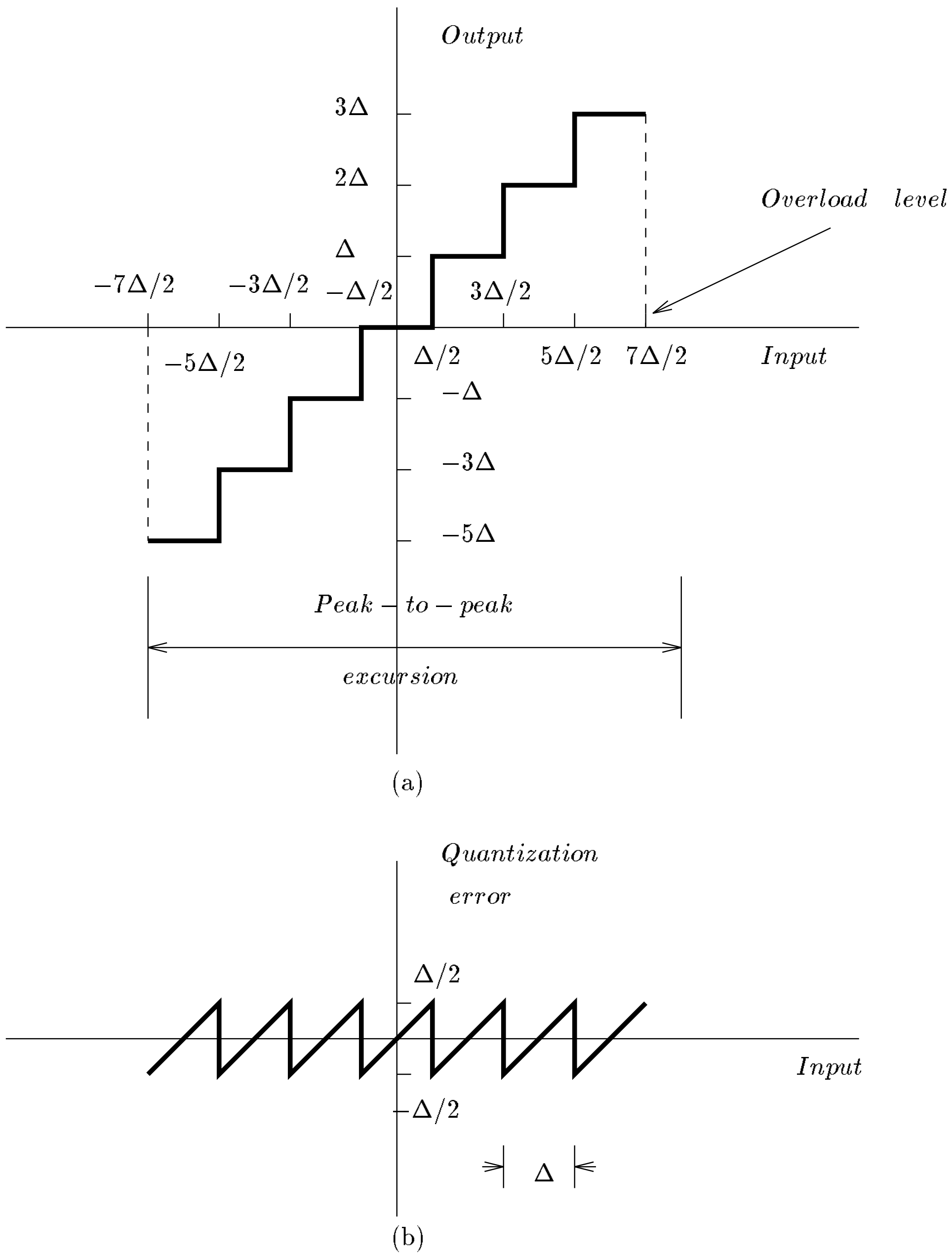


Figure 67: (a) Transfer characteristics of quantizer of midtread type. (b) Variation of the quantization error with input

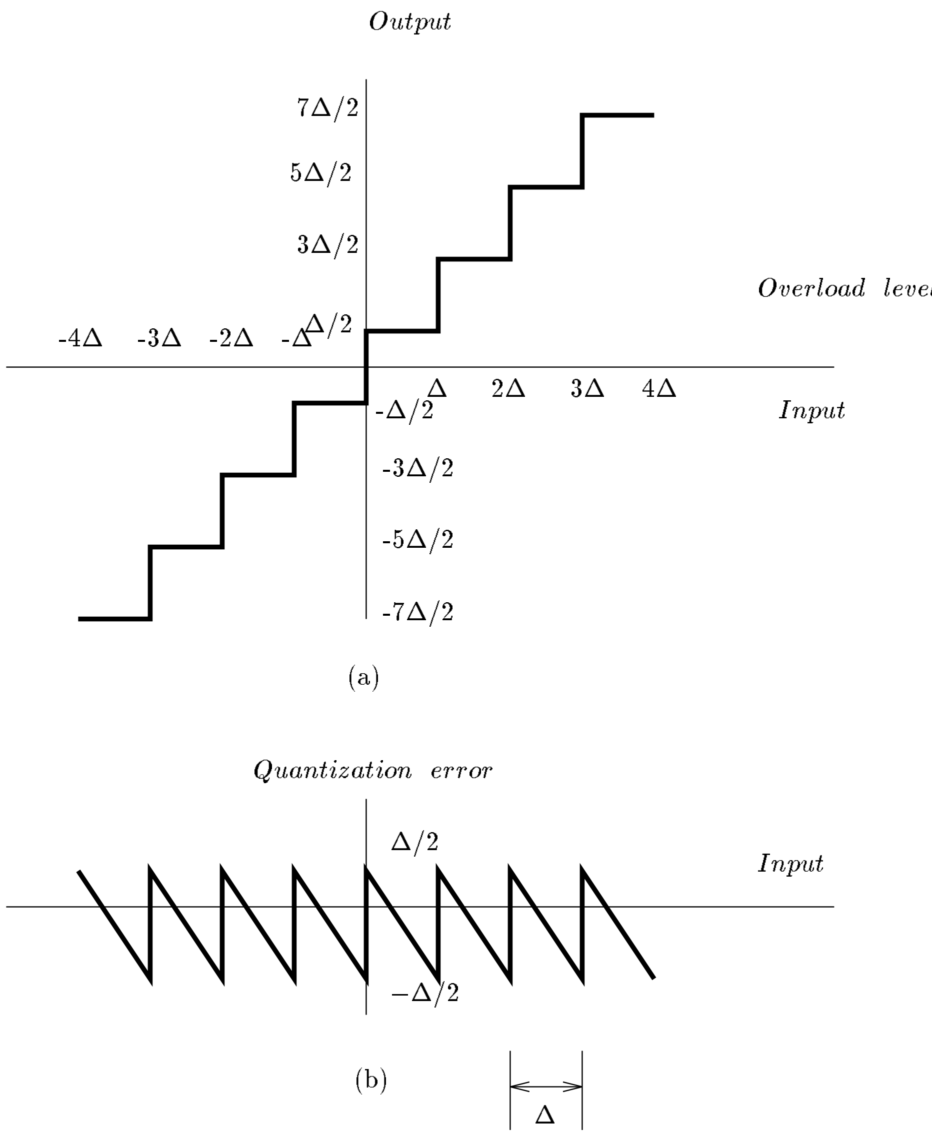


Figure 68: (a) transfer characteristics of quantizer of midrise type. (b) Variation of the quantization error with input

quatization interval are denoted as x_k and x_{k+1}. We have,

$$\mathcal{J}_k = \{x_k < x < x_{k+1}\} \qquad k = 1, 2, \ldots, L \tag{512}$$

If the representation level corresponding to the kth quatization interval is denoted as y_k, we have,

$$y = y_k, \qquad \text{if } x \text{ lies in the interval } \mathcal{J}_k \tag{513}$$

Let q denote the quantization error, $-\Delta/2 \leq q \leq \Delta/2$. We may write

$$y_k = x + q, \qquad \text{if } x \text{ lies in the interval } \mathcal{J}_k \tag{514}$$

If the quantizer is fine enough (small Δ), then the distortion produced by the quantization operation affects the performance of the transmission as if it were an *independent source of additive noise.* It is also found that the power spectral density of the quantization noise has a large bandwidth compared with the signal bandwidth. Thus, with the quantization noise uniformly distributed throughout the signal band, its interfering effect on a signal is similar to that of a *white noise.*

Assume that the quatization error (random variable q) is uniformly distributed over the possible range $-\Delta/2$ to $\Delta/2$. We have,

$$f_Q(q) = \begin{cases} \frac{1}{\Delta} & -\frac{\Delta}{2} \leq q \leq \frac{\Delta}{2} \\ 0 & \text{otherwise} \end{cases} \tag{515}$$

where $f_Q(q)$ is the probability density function of the quantization error. Then, the mean of the quantization error is zero, and its variance σ_Q^2 is equal to,

$$\begin{aligned} \sigma_Q^2 &= E[Q^2] \\ &= \int_{-\infty}^{\infty} q^2 f_Q(q) dq \\ &= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} q^2 dq \end{aligned} \tag{516}$$

$$= \frac{\Delta^2}{12} \tag{517}$$

Note that $\Delta^2/12$ can be viewed as the variance of quantization error conditioned on the interval $\mathcal{J}_k$.

Let the variance of the baseband signal $x(t)$ at the quantizer input be denoted by σ_X^2. We define an *output signal-to-noise quantization noise ratio* (SNR) as

$$(\text{SNR}) = \frac{\sigma_X^2}{\sigma_Q^2} = \frac{\sigma_X^2}{\Delta^2/12} \tag{518}$$

It is seen that the quantization SNR depends on the variance (power) of the input signal. It would be highly desirable from a practical viewpoint for the quantization SNR to remain constant for a wide range of input power levels. Such a quantizer is called a *robust quantizer*. This is discussed in the following.

9.4 Robust Quantization

The provision for a robust performance necessitates the use of a *nonuniform quantizer*. In this case, the range of the smaller input values are assigned more representation levels. This selection is justified by considering the following two facts (i) generally, smaller input values occur with higher probability and therefore should be represented with a higher precision, (ii) the characteristics of human hearing has a special characteristics that large signal amplitudes mask quantization noise to some extent.

Nonuniform quantization can be achieved by using a *compressor* followed by a *uniform quantizer*. By cascading this combination with an *expander* which acts as an inverse to the compressor, the original signal samples are restored to their correct values except for the effect of the quantization error. These operations are shown in Figures 70, 69. The combination of a *compressor* and an *expander* is called a *compander*.

As we are concerned with *symmetric, memoryless quantizers*, then the transfer characteristics of the compressor is represented by a memoryless nonlinearity $c(x)$, that has odd symmetry, i.e.,

$$c(-x) = -c(x) \tag{519}$$

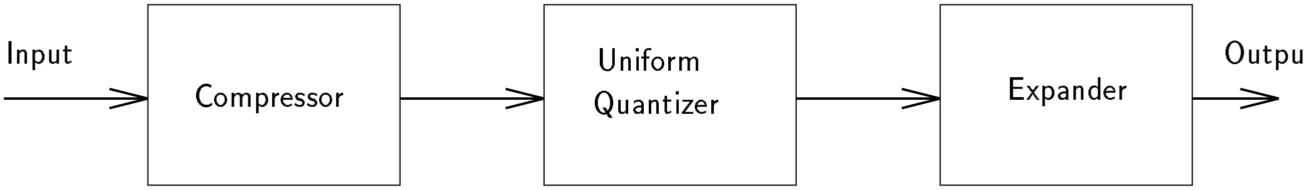


Figure 69: Model of nonuniform quantizer.

We also assume that,

$$c(x) = \begin{cases} x_{max} & x = x_{max} \\ 0 & x = 0 \\ -x_{max} & x = -x_{max} \end{cases} \tag{520}$$

The compression characteristic $c(x)$ relates nonuniform intervals at the compressor input to the uniform intervals at the compressor output.

The uniform intervals are of width $2x_{max}/L$ each, where L is the number of representation levels of the quantizer. It is assumed that L is a large number (fine quantizer). Under these conditions, the compression characteristic $c(x)$ in the kth interval, $\mathcal{J}_k$, in approximated by a straight-line segment with a slope equal to $2x_{max}/L\Delta_k$, where Δ_k is the width of the interval $\mathcal{J}_k$. This means that

$$\frac{dc(x)}{dx} \simeq \frac{2x_{max}}{L\Delta_k} \qquad k = 0, 1, \ldots, L-1 \tag{521}$$

Two other assumptions are also made:

1. The probability density function of input $f_X(x)$ is symmetric with respect to x.

2. In each interval $\mathcal{J}_k$, $k = 1, 2, \ldots, L$, the probability density function $f_X(x)$ is approximately constant. This assumption is expressed as,

$$f_X(x) \simeq f_X(y_k) \qquad x_k \leq X \leq x_{k+1} \tag{522}$$

We assume that the representation level y_k corresponding to the kth quantization interval $\mathcal{J}_k$ lies in the middle of $\mathcal{J}_k$, i.e.,

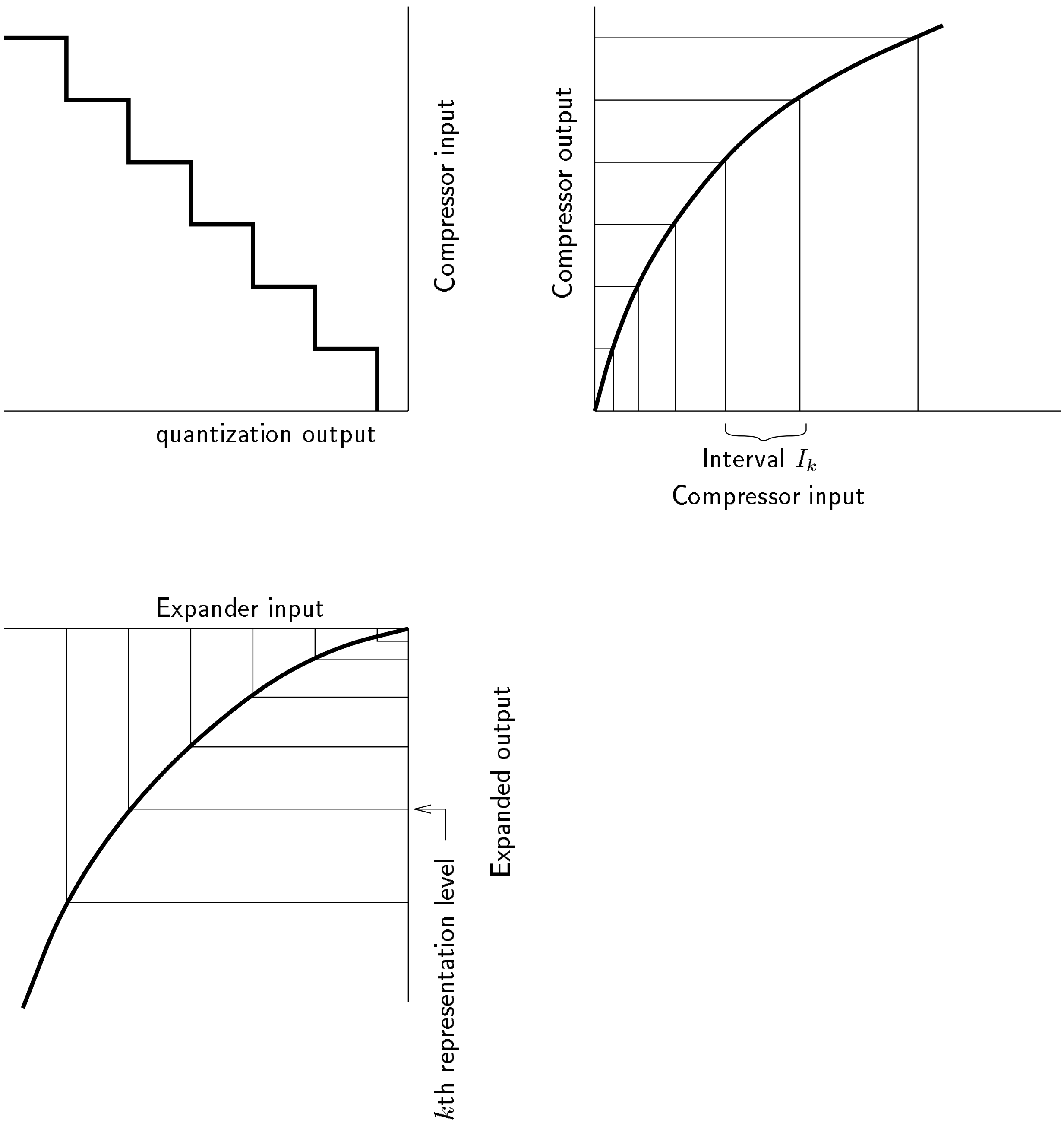


Figure 70: Transfer characteristics of compressor, uniform quantizer, and expander (shown for positive amplitudes only)

$$y_k = \frac{1}{2}(x_k + x_{k+1}) \qquad k = 0, 1, \ldots, L-1 \tag{523}$$

The width of the interval $\mathcal{J}_k$ equals,

$$\Delta_k = x_{k+1} - x_k \qquad x_k \leq X \leq x_{k+1} \tag{524}$$

Under these conditions, the probability that the input signal lies in the interval $\mathcal{J}_k$ is equal to,

$$\begin{aligned} p_k &= P(x_k \leq X \leq x_{k+1}) \\ &= f_X(y_k)\Delta_k \qquad k = 0, 1, \ldots, L-1 \end{aligned} \tag{525}$$

Let the random variable Q denote the quantization error

$$Q = y_k - X \qquad x_k \leq X \leq x_{k+1} \tag{526}$$

We have,

$$\begin{aligned} \sigma_Q^2 &= E[Q^2] \\ &= E[(X - y_k)^2] \\ &= \int_{-x_{max}}^{x_{max}} (y - x_k)^2 f_X(x) dx \end{aligned} \tag{527}$$

$$\sigma_Q^2 = \sum_{k=0}^{L-1} \frac{p_k}{\Delta_k} \int_{x_k}^{x_{k+1}} (x - y_k)^2 dx \tag{528}$$

Substituting Eq. 523, we get the result

$$\sigma_Q^2 = \frac{1}{12} \sum_{k=0}^{L-1} p_k \Delta_k^2 \tag{529}$$

As we saw in the previous section, the quantity $\Delta_k^2/12$ can be viewed as the variance of quantization error conditioned on the interval $\mathcal{J}_k$.

From (521), we obtain,

$$\Delta_k \simeq \frac{2x_{max}}{L}\left[\frac{dc(x)}{dx}\right]^{-1} \qquad k = 0, 1, \ldots, L-1 \tag{530}$$

Substituting Eq. 530 in Eq. 529, we get the result

$$\sigma_Q^2 \simeq \frac{x_{max}^2}{3L^2} \sum_{k=0}^{L-1} p_k \left[\frac{dc(x)}{dx}\right]^{-2} \tag{531}$$

Replacing for p_k from Eq. 525 and using integration instead of the summation over k, we obtain

$$\sigma_Q^2 \simeq \frac{x_{max}^2}{3L^2} \int_{-x_{max}}^{x_{max}} f_X(x) \left[\frac{dc(x)}{dx}\right]^{-2} dx \tag{532}$$

The output signal-to-quantization ratio is defined by

$$(\text{SNR}) = \frac{\sigma_X^2}{\sigma_Q^2} \tag{533}$$

where,

$$\sigma_X^2 = \int_{-x_{max}}^{x_{max}} x^2 f_X(x) dx$$

This results in,

$$(\text{SNR}) = \frac{3L^2}{x_{max}^2} \frac{\int_{-x_{max}}^{x_{max}} x^2 f_X(x) dx}{\int_{-x_{max}}^{x_{max}} f_X(x) \left[\frac{dc(x)}{dx}\right]^{-2} dx} \tag{534}$$

For a robust performance, the output signal-to-noise ratio should ideally be independent of the probability density function of the input random variable X. This requirement is met if

$$\frac{dc(x)}{dx} = \frac{K}{x} \qquad -x_{max} \leq x \leq x_{max} \tag{535}$$

resulting in,

$$c(x) = x_{max} + K \ln\left(\frac{x}{x_{max}}\right) \qquad x > 0 \tag{536}$$

Unfortunately, the characteristic function in (536) is unrealizable since $c(0)$ is not finite. Instead, we have to use an approximation to (536). Two widely used solutions to this problem are as follows:

μ-law companding

$$\frac{c(|x|)}{x_{xmax}} = \frac{\ln(1 + \mu|x|/x_{max})}{\ln(1+\mu)} \qquad 0 \leq \frac{|x|}{x_{max}} \leq 1 \tag{537}$$

A practical value for μ is 255. The $\mu-$law is used for **PCM** telephone systems in the United States, Canada, and Japan.

<u>A**-law companding**</u>

$$\frac{c(|x|)}{x_{xmax}} = \begin{cases} \frac{A|x|/x_{max}}{1+\ln A} & 0 \le \frac{|x|}{x_{max}} \le \frac{1}{A} \\ \\ \frac{1+\ln(A|x|/x_{max})}{1+\ln A} & \frac{1}{A} \le \frac{|x|}{x_{max}} \le 1 \end{cases} \tag{538}$$

A practical value for A is 87.56. The A-law companding is used for **PCM** telephone systems in Europe.

10 Differential Pulse-Code Modulation

In the use of **PCM** for the digitization of voice and video signal, the signal does not change rapidly from one sample to the next with the result that the difference between adjacent samples has a variance that is smaller than the variance of the signal itself.

In particular, if we know the past behavior of a signal up to a certain point in time, it is possible to make some inference about its future values. This provides motivation for the differential quantization scheme shown in Fig. 71(a).

The difference signal $e(nT_s)$ is called a *prediction error*, it is the amount by which the predictor fails to predict the input exactly.

$$\begin{aligned} v(nT_s) &= Q[e(nT_s)] \\ &= e(nT_s) + q(nT_s) \end{aligned} \tag{539}$$

where $q(nT_s)$ is the quantization error.

For the system in Fig. 71, we may write,

$$u(nT_s) = \hat{x}(nT_s) + v(nT_s) \tag{540}$$

$$u(nT_s) = \hat{x}(nT_s) + e(nT_s) + q(nT_s) \tag{541}$$

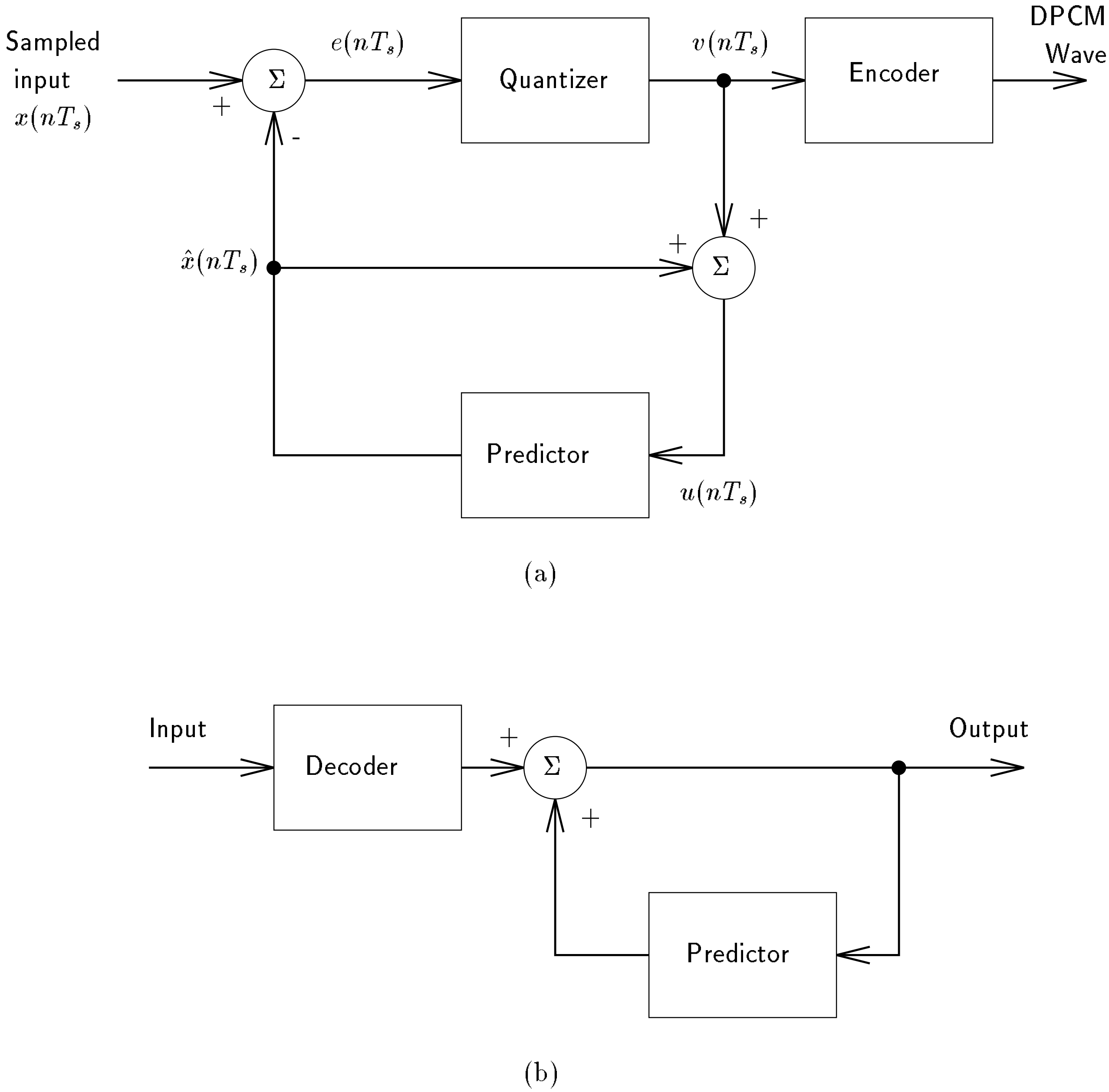


Figure 71: DPCM system. (a) Transmitter. (b) Receiver

Finally, replacing $x(nT_s) = \hat{x}(nT_s) + e(nT_s)$, we obtain,

$$u(nT_s) = x(nT_s) + q(nT_s) \tag{542}$$

This means that irrespective of the properties of the predictor, the quantized signal $u(nT_s)$ at the predictor input differs from the input signal $x(nT_s)$ by the quantization error.
In the absence of the channel noise the corresponding receiver output is equal to $u(nT_s)$. Obviously, the power of the process E is smaller than the power of the original signal X. The corresponding reduction in power is measures by the *prediction gain* which is defined as,

$$G_P = \frac{\sigma_X^2}{\sigma_E^2}$$

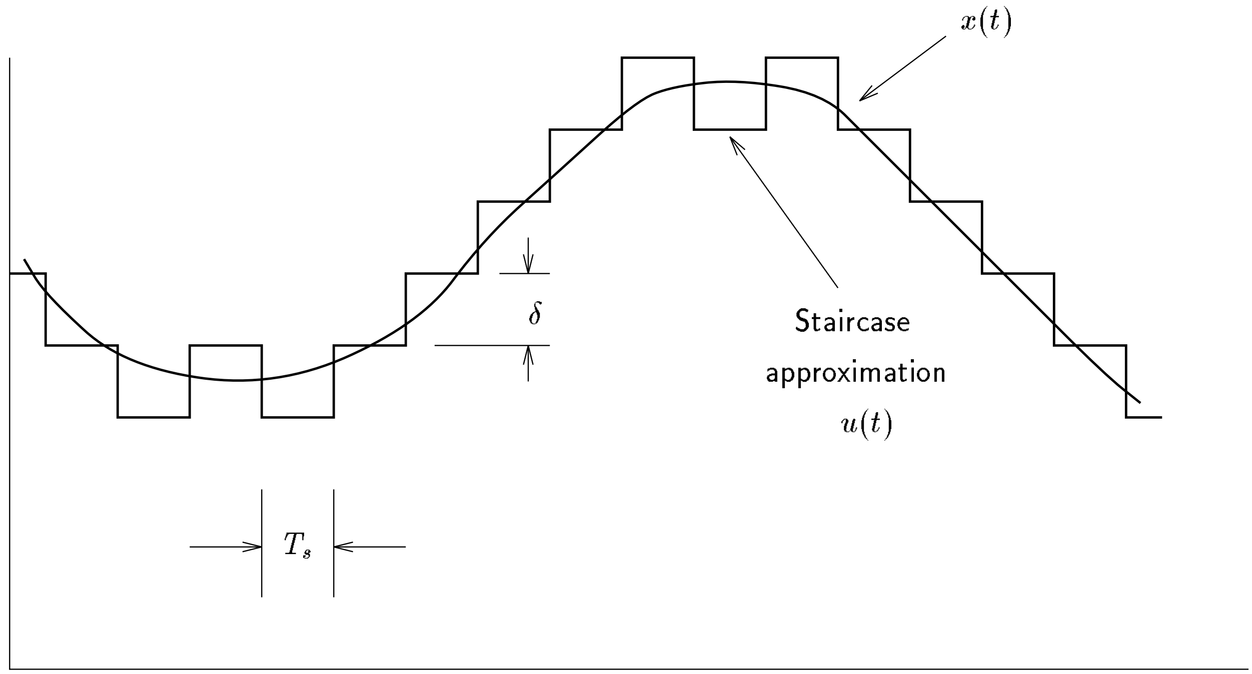
We have,

$$(\text{SNR})_0 \triangleq \left(\frac{\sigma_X^2}{\sigma_Q^2}\right) = \left(\frac{\sigma_X^2}{\sigma_E^2}\right)\left(\frac{\sigma_E^2}{\sigma_Q^2}\right) = G_P(\text{SNR})_P$$

where $(\text{SNR})_P$ is the SNR of the quantizer in Fig. 71. This means that we have an improvement in the overall SNR with a factor equal to G_P.

11 Delta Modulation

The use of DPCM suggest the following possibility: *Oversampling a signal (at a rate higher than the Nyquist rate) purposely to increase the correlation between adjacent samples of the signal, so as to permit the use of a simple quantizing strategy.*
This strategy results in a higher number of samples per second, however, exploiting the large dependency between successive samples, we are able to quantize the samples using a smaller number of quantization symbols. On the other hand, as we already discussed in the section on signal constellations, for a given channel bandwidth, having a larger number of symbols per second requires a higher average energy for transmission. The objective is to find an intermediate solution which provides a compromise between the sampling rate and the number of the quatization symbols resulting in a good overall performance. In the following section, we discuss an example of such a method acting in an extreme case of using a quantizer with only two levels.



(a)

Binary sequence at modulator output: 0 0 1 0 1 1 1 1 1 0 1 0 0 0 0 0

(b)

Figure 72: Illustration of delta modulation

Delta modulation (DM) is the *one-bit (or two level)* version of DPCM. DM provides a staircase approximation to the oversampled version of the input signal. The difference between the input and the approximation is quantized into only two levels, namely, $\pm\delta$. The step size Δ of the quantizer is

$$\Delta = 2\delta \tag{543}$$

Denote the input signal as $x(t)$ and the staircase approximation to it as $u(t)$. Then

$$\begin{aligned} e(nt_s) &= x(nT_s) - \hat{x}(nT_s) \\ &= x(nT_s) - u(nT_s - T_s) \end{aligned} \tag{544}$$

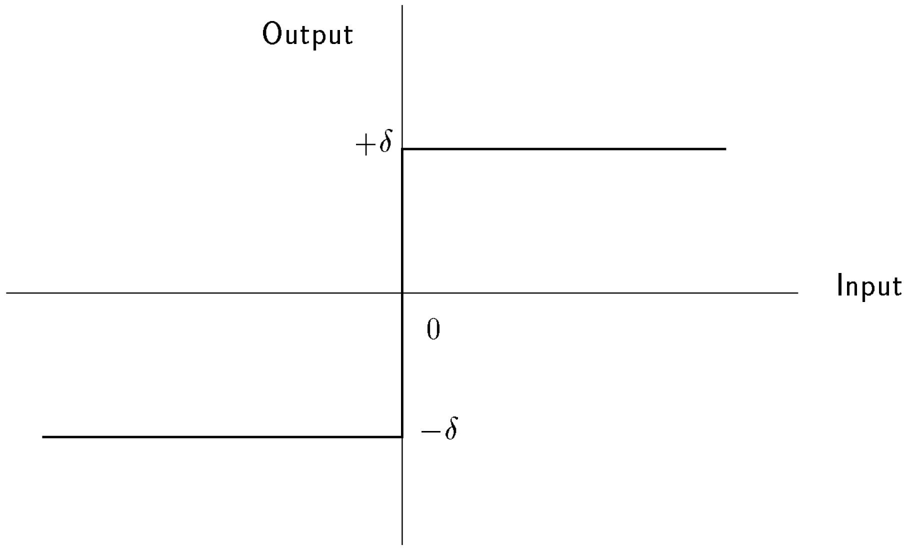


Figure 73: Input-output characteristic of two-level quantizer

$$b(nT_s) = \delta \ \ \mathrm{sgn}[e(nT_s)] \tag{545}$$

$$u(nT_s) = u(nT_s - T_s) + b(nT_s) \tag{546}$$

We assume that the accumulator is initially set to zero. Then

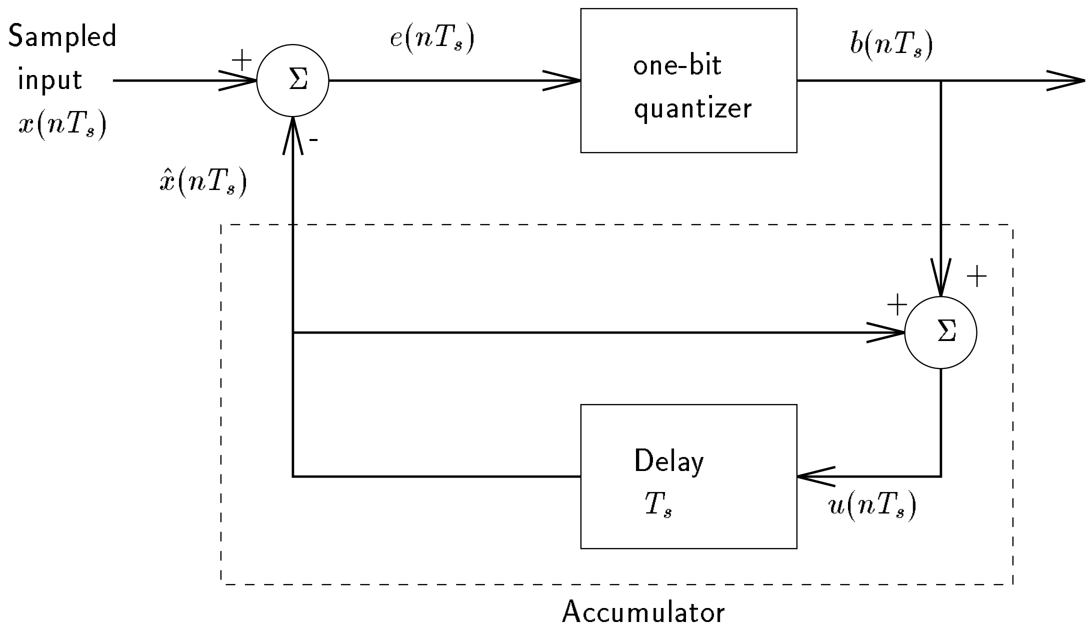
$$\begin{aligned} u(nT_s) &= \delta \sum_{i=1}^{n} \ \mathrm{sgn}[e(iT_s)] \\ &= \sum_{i=1}^{n} b(iT_s) \end{aligned} \tag{547}$$

11.1 Quantization Noise in Delta Modulation

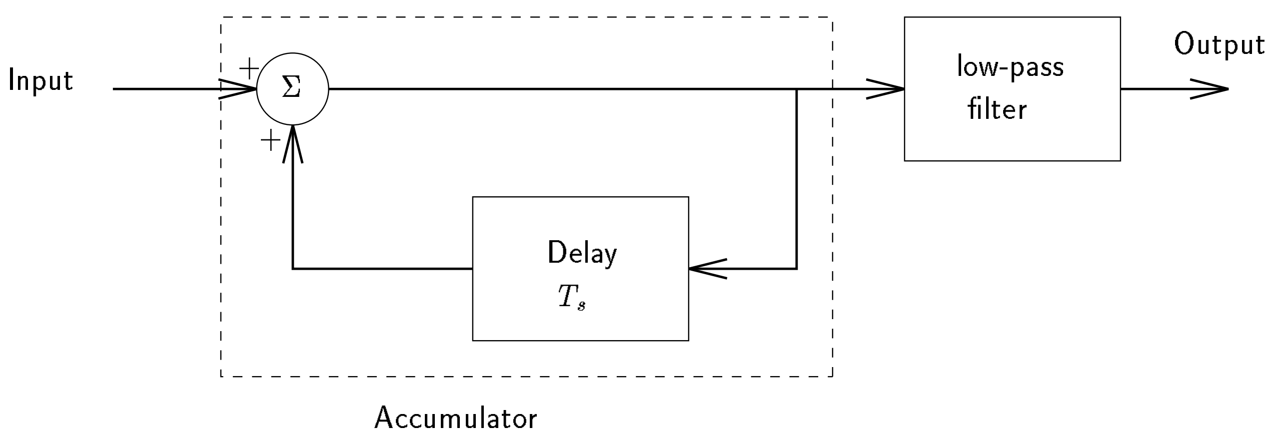
Delta modulation systems are subject to two types of quantization errors: *slope-overhead distortion*, and *granular noise.*

Let $q(nT_s)$ denote the *quantizing error*. We have,

$$u(nT_s) = x(nT_s) + q(nT_s) \tag{548}$$



(a)



(b)

Figure 74: DM system. (a) Transmitter. (b) Receiver

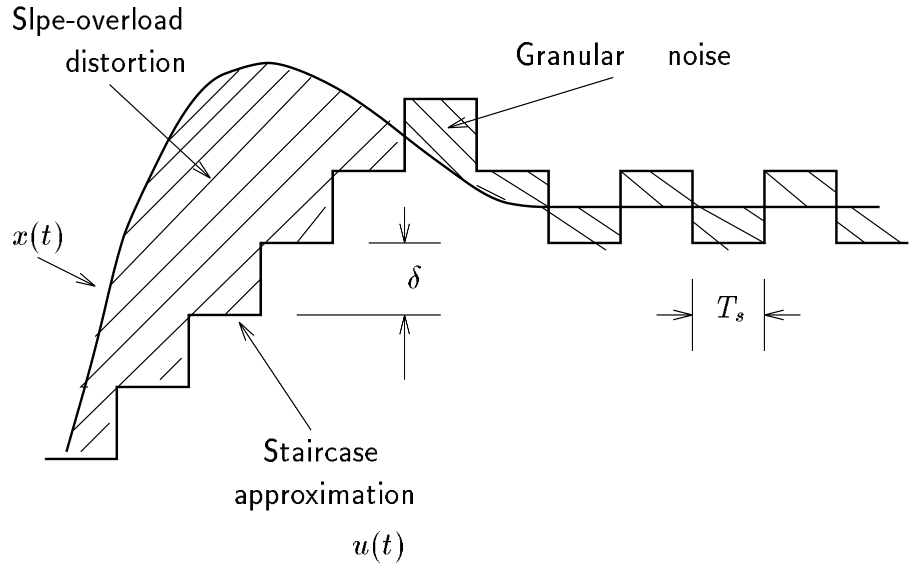


Figure 75: Illustration of quantization error in delta modulation

Accordingly, using Eq. 548 to eliminate $u(nT_s - T_s)$ from Eq. 544, we may express the prediction error $e(nT_s)$ as

$$e(nT_s) = x(nT_s) - x(nT_s - T_s) - q(nT_s - T_s) \tag{549}$$

Except for the quantization error $q(nT_s - T_s)$, the quantizer input may be viewed as a digital approximation to the derivative of the input signal. In order for the sequence of samples $\{u(nT_s)\}$ to increase as fast as the input sequence of samples $\{x(nT_s)\}$ in a region of maximum slope of $x(t)$, we require that the condition

$$\frac{\delta}{T_s} \geq \max \mid \frac{dx(t)}{dt} \mid \tag{550}$$

Otherwise the step size $\Delta = 2\delta$ is too small with the result that $u(t)$ falls behind $x(t)$, as illustrated in Fig. 75. This condition is called **slope-overhead**.

Granular noise occurs when the step size Δ is too large relative to the local slope characteristics of the input waveform $x(t)$.

The choice of the optimum step size that minimizes the mean-square value of the quantizing error in a linear delta modulator will be the result of a *compromise between slope overhead distortion and granular noise.*