Chapter 2

Pulse Modulation and Sampling Systems

Ken C. Pohlmann

Introduction

There are many ways in which information can be conveyed by a pulse modulation system. Pulse amplitude modulation, pulse position modulation, and pulse width modulation all offer various advantages. However, following its invention in 1937, pulse code modulation was regarded as an almost ideal modulation code for conversion, storage, and transmission of digital signals, with moderate bandwidth requirements and low coding errors. As with any pulse coding method, there are many benchmarks, such as efficiency and bit error probability, which permit evaluation of the format's suitability in an application. The nature of any sampling system is perhaps best described by its frequency domain properties. Fourier analysis clearly shows the way in which a rectangular time pulse modulates the amplitude of a carrier frequency. For example, it can be seen that transmission of narrow pulses necessitates use of a channel with a high bandwidth. Any sampling system is bound by the sampling theorem, which defines the relationship between message and sampling signal rates. In addition, the theorem dictates that the message be bandlimited. Quantization introduces an approximation error which can be predicted with great accuracy, and minimized with techniques such as dither.

Pulse Modulation

In many applications, including the transmission or recording of speech and music, analog information may be converted to a digital format using one of many types of pulse modulation. For example, a pulse's width or variable duration, or variable position in time may represent information about the signal; pulse width modulation (PWM) is an example of the former, and pulse position modulation (PPM) is an example of the latter. Pulse width modulation and pulse position modulation samples may be formed by comparing analog samples of the input signal with a sawtooth waveform. The time between the sampling signal and the intersection of the analog signal with the sawtooth determines the width of a pulse or its position. A signal may also be conveyed through an uncoded pulse modulation system; pulse amplitude modulation (PAM) is an example of this method. Samples are transmitted via pulses with amplitudes that represent the amplitude of the signal at sample time. Although PWM, PPM, and PAM may be used in the context of conversion, they are not directly suitable for transmission or recording because they are overly affected by noise in the transmission channel and in the storage media.

Pulse Code Modulation

The most commonly used modulation method is pulse code modulation (PCM). This method was devised in 1937 by Sir Alec Reeves while working as an engineer in the International Telephone and Telegraph Company (ITT) laboratories in France. In PCM, the input signal must undergo sampling, quantization, and coding. By coding the measured analog amplitude values of sampled information into a series of pulses, binary numbers may be used to represent the information. At the receiver the pulse code can be used to reconstruct an analog waveform. The binary words which represent the amplitude of the signal are directly coded into PCM waveforms as shown in Figure 1.

Given a converter with k-bit wordlength, 2^k unique code words are created to represent 2^k amplitude values, called quanta. For example, a 16-bit system would encode 65,536 amplitude quanta. In the simplest incarnation, binary 0000 0000 0000 0000 would represent decimal 0, and binary 1111 1111 1111 1111 would represent 65,535. In practice, however, that might not be the most efficient mapping of the audio waveform. Thus a different arrangement of the PCM data might be employed. Two examples of alternative binary coding are (*a*) sign and magnitude binary and (*b*) two's complement binary notation. In sign and magnitude notation the absolute value of samples are expressed in binary code; however, their sign is expressed in the leftmost bit. In two's complement notation two ascending binary counts are used, the leftmost bit again representing sign. More specifically, negative numbers are formed by taking the complement of the positive equivalent and adding 1; for example, the two's complement of 01000 is 10111 + 00001 = 11000. Humans appreciate two's complement because the left digit always denotes the sign of the number; digital circuits appreciate it because subtraction can be performed with an addition operation, thus simplifying calculations.

With methods such as PWM, PPM, and PAM, only one pulse is needed to represent the amplitude value whereas in PCM several pulses per sample are required. As a result, PCM may require a higher-bandwidth channel. However, PCM forms a very robust signal in that only the presence or absence of a pulse is necessary to read the signal. In addition, a PCM signal can be regenerated without loss. Therefore the quality of a PCM transmission depends on the qual-

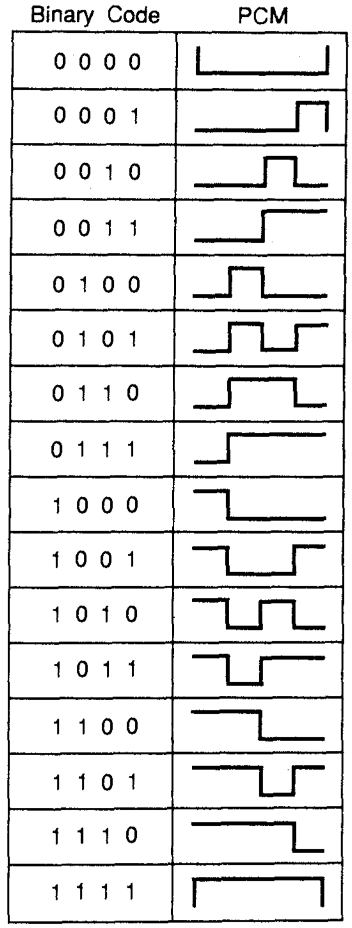


Figure 1. Binary words are coded directly into the PCM waveform.

ity of the sampling, quantizing, and coding processes, not the quality of the channel itself. In addition, depending on the sampling rate and the capacity of the channel, several PCM signals may be simultaneously conveyed by using time division multiplexing. This greatly expedites use of PCM. Several types of modulation are reviewed in Figure 2.

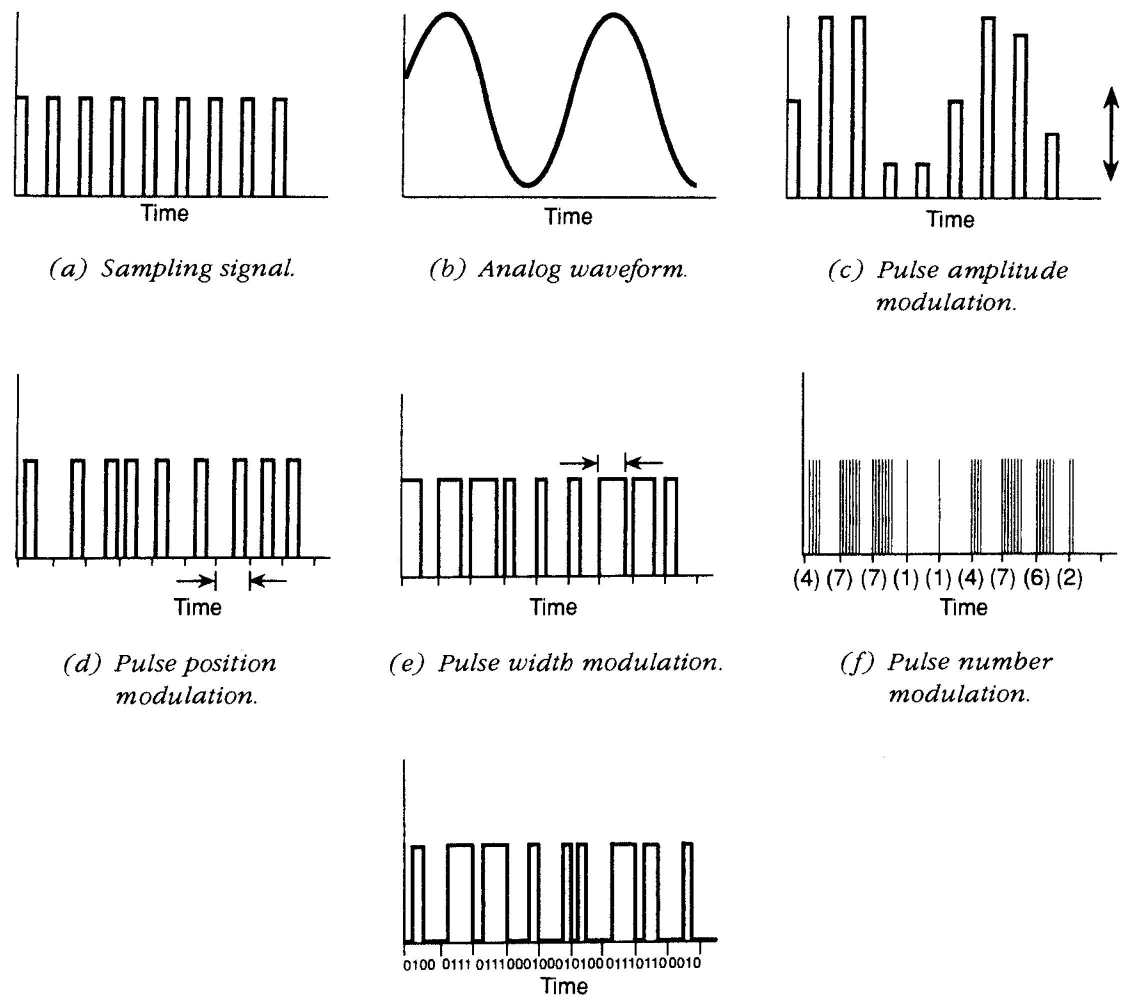
Theoretical Performance of Pulse Code

A pulse code's efficiency can be measured through comparison of the information capacity of the message versus the capacity of the encoded signal itself. As Shannon [1948] has pointed out, the information capacity of the message characterizes the entropy of the signal. The coding or format efficiency may be defined as

$$\eta = C_m/C_c,$$

where C_m is the information capacity of the signal and C_c is the information capacity of the coded signal.

Substituting expressions for C_m and C_c, we obtain



(a) Sampling signal. *(b) Analog waveform.* *(c) Pulse amplitude modulation.*

(d) Pulse position modulation. *(e) Pulse width modulation.* *(f) Pulse number modulation.*

(g) Pulse code modulation.

Figure 2. A variety of modulation methods may be used to represent an analog waveform.

$$\eta = \frac{S_m \sum_{i=1}^{n} P_i \log_2(1/P_i)}{S_c \sum_{i=1}^{n} Q_i \log_2(1/Q_i)},$$

where

S_m is the message symbol rate in symbols per second,
S_c is the coded symbol rate in symbols per second,
P_i is the probability of a logical 1 or 0 message bit occurring,
Q_i is the probability of a coded level occurring,
i is the ith message symbol.

Alternatively, coding efficiency may be characterized by the relationship between timing of the minimum duration and the format's clock period:

$$\eta = t_{\min}/T,$$

where t_{min} is the minimum duration timing and T is the clock period.

In codes in which each symbol represents a bit, efficiency can be defined as the ratio of actual information transmitted to the theoretical maximum transmitted rate. In other words, this compares the number of data bits to the maximum number of symbols. Nominally, the highest efficiency of such codes is 1, with efficiency of 1/2 common in codes devised to exclude DC components. However, codes may be written in which several data bits are expressed by one symbol. Since a clock period can convey a number of bits, the coding or format efficiency must be redefined as the number of data bits contained in one symbol compared to a symbol with a maximum number of transitions:

$$\eta = r/s,$$

where r is the number of data bits in a symbol and s is the maximum number of transitions within a clock period.

The reciprocal of this value is known as the format conversion ratio.

A formula proposed by Hartley and Shannon [1949] may be used to study the information transfer of a pulse modulation system. Given a transmitted signal's power and bandwidth, and power of disturbing white noise, the Hartley-Shannon law determines the maximum information that can be accurately conveyed:

$$I = B \log_2(1 + T/N),$$

where

I is the maximum information capacity in bits per second,
B is the bandwidth of the channel in hertz,
T is the power of the transmitted signal,
N is the power of disturbing Gaussian noise.

For example, ideally a PAM system would always transmit an amount of information equal to I; information lost will be proportional to the noise in the system.

According to a derivation undertaken by Haykin [1978], the maximum information capacity of a PCM signal in bits per second may be shown to be

$$I_{PCM} = B \log_2(1 + 12T/a^2N),$$

where a is a constant typically equal to 10. Pulse code modulation is a quantized noise-limited modulation method, and, in particular, in a PCM system, power and bandwidth are interchanged logarithmically.

Practical Limitations of Pulse Code

An efficient pulse code format must restrict DC components which could disrupt timing synchronization, and it must observe the maximum transmission bandwidth available. Data patterns must be analyzed to determine presence of these two limits: DC components and maximum signal rate. Existence of a DC component or noise can seriously compromise the success of the transmission of a signal. Any amplitude variation upward or downward, referred to as baseline wander, changes the reference by which the receiver distinguishes positive or negative values. In AC-coupled receivers, such as in fiber-optic systems, the baseline is established purely by the demodulated data signal itself.

To compound the degradation, the base line wander and other problems can cause a time-axis variation known as jitter to accumulate through a system, leading to random errors. For example, the effects of jitter on the sampling clock of an A/D converter are quite similar to FM modulation; the input frequency acts as the carrier, and clock jitter acts as the modulation frequency. The jitter acts to reduce the amplitude of the input signal and to add sideband components equally spaced at either side of the input frequency, at a distance equal to multiples of the jitter frequency. The effect of sinusoidal sampling clock time jitter on an A/D converter may be described as

$$v(t) = A \cos[\omega_i(t + J \sin \omega_j)],$$

where

A is the amplitude of the input signal,
ω_i is the frequency of the input signal,
J is the peak amplitude of the jitter,
ω_j is the jitter frequency.

The effect of jitter thus increases as the input signal frequency increases; specifically, jitter amplitude error increases with input signal slew rate. In practice, the first sideband is potentially audible and can thus decrease the dynamic range of the system. Peak clock jitter of less than 200 ps is often mandated.

Use of crystal-controlled phase-locked loops and other measures can limit jitter to an acceptable level. Master clocks must not be distributed through a studio. Rather, each piece of sampling equipment must contains its own clock, frequency locked to a distributed master clock. A signal can be successfully recovered provided the jitter margin does not exceed

$$\Delta t = \pm \tau/2,$$

where $\tau = 1/T$ and T is the sampling period.

Signal integrity and problems such as DC offset, jitter, peak shift, and noise, which affect signal integrity can be observed by an eye pattern, as shown in Figure 3. This display uses an oscilloscope to overlay a superimposed collection of regenerated data sequences. When the oscilloscope is triggered at the data rate, the dynamic changes in the signal are visible. Amplitude and temporal variations which deteriorate the signal will reduce the eye opening and hence the receiver's performance to the point where pulse shaping can no longer retrieve the signal. The amount of deterioration can be gauged by measuring the extent of amplitude variations and forming an eye opening ratio,

$$E = \frac{a_2}{a_1} = \frac{a_1 - 2\Delta a}{a_1},$$

where

a_1 is the outside amplitude,
a_2 is the inside amplitude,
Δa is the amplitude variation.

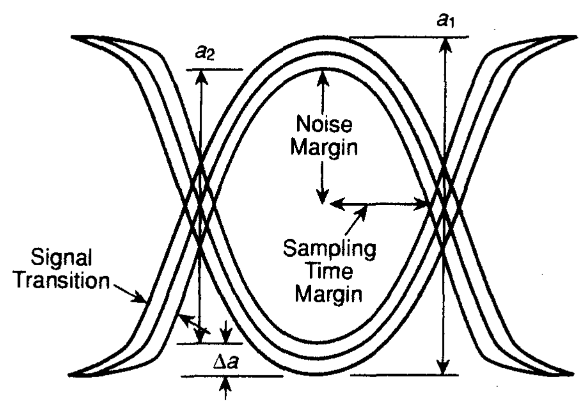


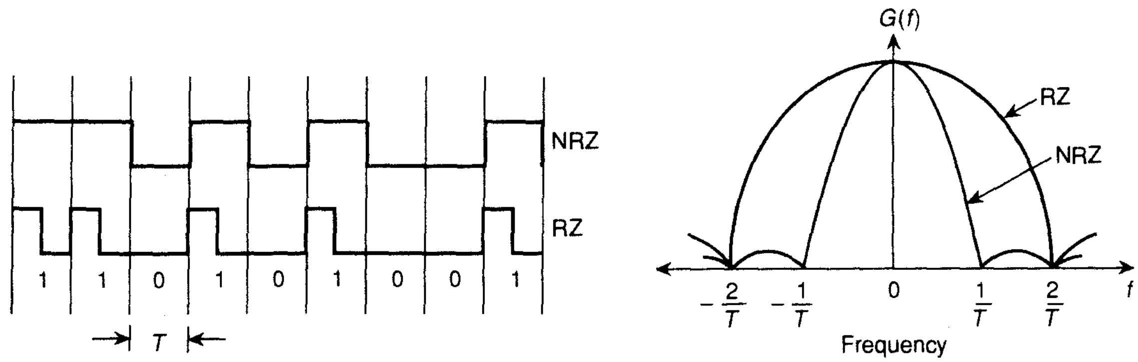
Figure 3. Signal integrity can be evaluated through the eye pattern. The eye opening ratio is taken from the ratio of the amplitudes a_1 and a_2.

Variations are measured from the center of the opening in the eye pattern. The width of the eye gives the percentage of the data period available to ascertain its logical value, and the height shows the maximum difference between these levels during the available time.

Bit Coding

As we shall see, the analysis of a sampling pulse is relatively easy because of its periodic nature in the time domain; Fourier analysis clearly shows its spectrum. However, a data stream differs in that the data pulses occur aperiodically and in fact can be considered to be random. Analysis may require evaluation of the power spectrum at each frequency, a difficult task. The resulting power spectral density, or power spectrum, $G(f)$, statistically shows the response of the data stream $f(t)$. For example, Figure 4a shows two data streams with identical data content but different pulse types: full binary NRZ and half binary RZ. Figure 4b shows the power spectrum curves for the two pulse types. Signal peak amplitudes are normalized, and the pulses have equal periods. Absolute magnitude is used; thus all frequency components are plotted positively. Viewing the power spectrum, we see that a transmission waveform ideally should have minimal energy at low frequencies to expedite equalization, and minimal energy at high frequencies to reduce crosstalk. A code is optimized through manipulation of pulse shape, number of logical levels, and repetition rate.

Using statistical methods, the worst-case (widest) spectrum of a signal can be determined from its probability distribution. For NRZ, the widest spectrum takes place when data alternates between 1 and 0; for RZ, the worse case is



(a) Two data streams of different type but of identical data.

(b) The power spectrum curves for the data streams.

Figure 4. The power spectrum statistically shows the response of the data stream. *(From Morris)*

incurred with continuous 1s. In both cases, when this occurs, the signal level spends its time exactly split, with equal probability, between the positive and negative levels. The average power is thus $f(t)^2$ for both NRZ and RZ. They differ in their pulse durations, however, and because of this NRZ is more efficient in terms of spectrum. The NRZ signal uses the entire bandwidth to convey its information whereas the RZ signal with a pulse duration of $T/2$ uses only half of its bandwidth, wasting half.

Performance may also be analyzed in terms of bit error probability, that is, the receiver's net ability to distinguish between a 1 and 0 value. The error probability is determined as a function of the signal-to-noise (S/N) ratio:

$$P_e = f(S_v/N_{\text{rms}}),$$

where S_v is the magnitude of the signal voltage, varying from $+V$ to $-V$, and N_{rms} is the rms noise voltage in the received signal.

In the case of a NRZ signal, a bit error occurs when the noise has a higher (and opposite) magnitude than the signal. The error probability can be written

$$P_e = \text{Prob}(N_{\text{rms}} > S_v/2).$$

And the bit error probability is written

$$P_e(\text{NRZ}) = {}^1\!/_2\, \text{erfc}(S_v/N_{\text{rms}})^{1/2},$$

where erfc is the complementary error function defined as

$$\text{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^{\infty} \exp(-z^2)\, dz.$$

To further minimize decoding errors, formats can be developed in which data is conveyed with data patterns that are as individually unique as possible. For example, in the EFM code devised for the compact disc format, 8-bit symbols are translated into 14-bit symbols, carefully selected for maximum difference between symbols. In this way, invalid data can be more easily recognized. Similarly, a data symbol could be created based on previous adjacent symbols, and the receiver could recognize the symbol and its past history as a unique state. A state pattern diagram is used in which all transitions are defined, based on all possible adjacent symbols.

Sampling Spectra

Discrete time sampling is founded on the concept of a rectangular impulse of infinitesimal width. Because of its infinite frequency response, and the requirement that the channel have infinite bandwidth, an ideal impulse is physically impossible to create and eludes everyone but mathematicians. In practice we consider a rectangular pulse of fixed amplitude A and finite width τ. Such an impulse is represented in Figure 5.

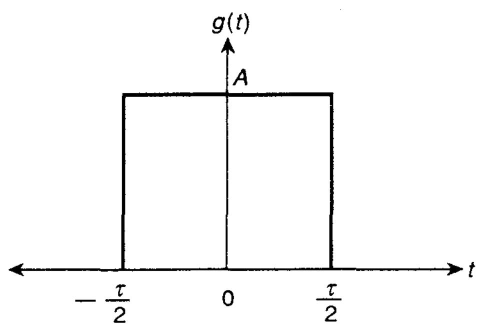


Figure 5. A rectangular pulse of fixed amplitude and width: an impulse.

Although data signals are often characterized by their time domain properties, the performance of a transmission channel is usually best described by its frequency domain properties. Specifically, it is important to know the bandwidth required for successful transmission of a single pulse. The Fourier transform may be used to describe the time domain function in the frequency domain. The Fourier transform of a time function $T(t)$ can be found:

$$F(f) = \int_{-\tau/2}^{\tau/2} T(t)\, e^{-j\omega t}\, dt,$$

where $\omega = 2\pi/\tau$.

Given a single rectangular pulse of duration τ and amplitude A, centered at $t = 0$ (Figure 6),

$$F(f) = \int_{-\tau/2}^{\tau/2} Ae^{-j\omega t}\, dt = A\tau \frac{\sin(\omega\tau/2)}{(\omega\tau/2)}.$$

If we define $x = \omega\tau/2$, then the transformation of the pulse is shown to be

$$F(f) = A\tau \frac{\sin x}{x}.$$

This function is shown in Figure 6. It can be seen that it is composed of a fundamental cosinewave and its harmonics, that its maximum value occurs at $x = 0$, and that it approaches zero as x approaches $\pm\infty$. The width of the center lobe is exactly $2/\tau$, and the frequency response passes through zero at multiples of $1/\tau$. Importantly, it demonstrates the fundamental nature of sampling as a modulation process; the frequency pattern of the function shows that the rectangular time pulse has modulated the amplitude of a carrier frequency. The center frequency may be shifted without altering the shape of the envelope itself. Clearly, this spectrum extends to infinity, and thus ideal transmission of the pulse would require a system with infinite bandwidth. As we shall see, however, only the central lobe is required: A finite bandwidth will suffice.

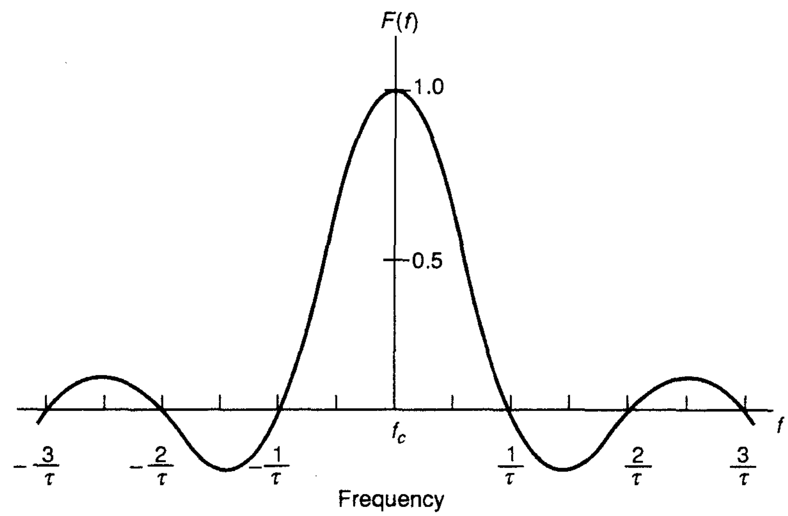


Figure 6. A single-pulse impulse response composed of a fundamental cosinewave and its harmonics.

Given an understanding of the properties of a single pulse, it is useful to examine a series of such pulses with a periodic repetition of T. This leads to the creation of a practical sampling signal as a periodic series of pulses of fixed amplitude and finite width. It is defined as follows:

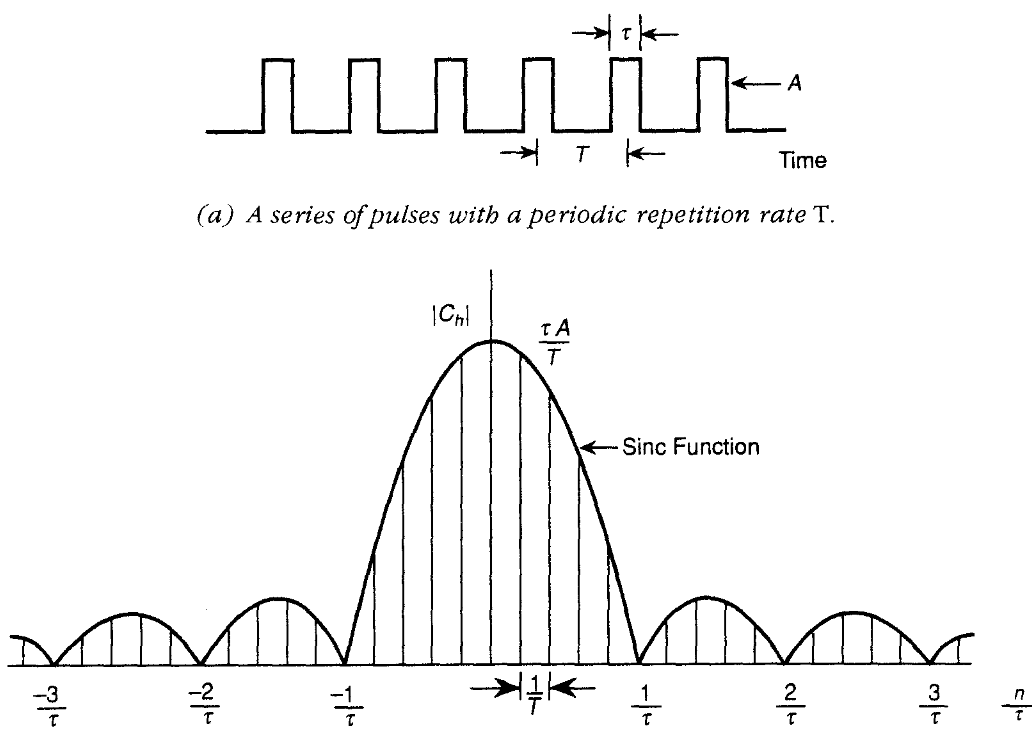
$$S(t) = \begin{cases} A & \text{when } -\tau/2 < t < \tau/2, \\ 0 & \text{for the remainder of the period.} \end{cases}$$

where τ is the width of the pulse, and it is periodically repeated with a period of T seconds. This sampling signal is shown in Figure 7a. Evaluating the complex Fourier coefficient C_n, we observe

$$C_n = \frac{1}{T}\int_{-\tau/2}^{\tau/2} A\, e^{-j2\pi nt/T}\, dt = \frac{A}{n\pi}\sin(n\pi\tau/T).$$

To simplify the results, we can employ the sinc function defined as

$$\operatorname{sinc} x = \frac{\sin \pi x}{\pi x}.$$



(a) A series of pulses with a periodic repetition rate T.

*(b) The frequency spectrum of (a) is defined at discrete values of 1/*T *along the contour of a sinc function.*

Figure 7. Illustration of the properties of a series of pulses—a sampling signal.

The sinc function has a maximum value at $x = 0$, and approaches zero as x approaches infinity. It oscillates through positive and negative values passing though zero at integer multiples. Using the sinc function, we may rewrite the expression for C_n:

$$C_n = \frac{\tau A}{T} \operatorname{sinc}(n\tau/T).$$

The frequency spectrum of this function is defined at discrete values of n; that is, as equally spaced spectral lines with heights corresponding to the discrete frequency components. This spectrum is shown in Figure 7b. The spectral lines are spaced according to the period T, and the envelope is determined by the pulse duration τ, duty ratio τ/T, and pulse amplitude A with zero crossings at frequencies that are multiples of $1/\tau$.

The spectral response of a series of sampling pulses thus creates spectral lines with amplitudes that follow the same contour as that of a single pulse. The spectrum bandwidth is not affected by the pulse repetition frequency; rather, the bandwidth is determined by the pulse width τ. The shorter the duration of the pulse, the greater is the frequency spread of the bandwidth. It is thus clear that transmission of narrower pulses requires a channel with higher bandwidth. From a frequency domain standpoint, wider pulses might appear advantageous. However, as viewed in the time domain, narrower pulses permit a greater repetition rate and, for example, permit time multiplexing of channels. In any case, it is not a higher repetition rate that necessitates higher bandwidth but the narrow width of the pulses. Similarly, as we shall see, a condition known as aperture error can be minimized by decreasing the duration of the pulse width.

Using the sampling signal $S(t)$, it is possible to define the nature of the sampled signal $f_s(t)$:

$$f_s(t) = f(t)S(t),$$

where

$f_s(t)$ is the sampled signal,
$f(t)$ is the message signal,
$S(t)$ is the sampling signal.

Moreover, we may obtain an expression for the frequency spectrum of the sampled signal. The multiplication of these two time domain functions may be represented as the convolution of their spectra:

$$F_s(\omega) = \frac{1}{2\pi} F(\omega) * S(\omega),$$

where * denotes the convolution of these frequency domain signals.

Mathematical derivation yields the following equation:

$$F_s(\omega) = Ad \sum_{n=-\infty}^{\infty} \frac{\sin n\pi d}{n\pi d} F(\omega - n\omega_0).$$

. More obviously, a graphic representation of this expression is shown in Figure 8. It shows that the spectrum of the message signal contains both positive and negative sidebands centered at the impulses defined by the sampling

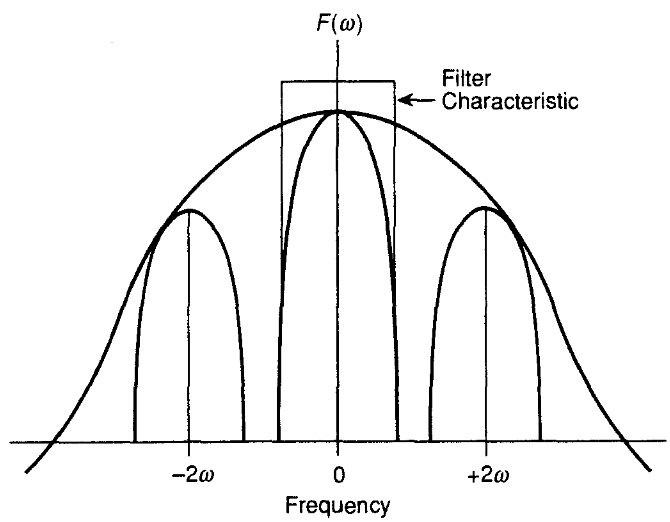


Figure 8. The spectrum of the message signal contains positive and negative sidebands centered at the impulse defined by the sampling function.

function, and placed at multiples of the sampling frequency. In addition, their amplitude again clearly follows the $(\sin x)/x$ contour predicted by the Fourier transform of the sampling signal. It is important to note that complete information of the message signal is held in each sideband. Filtering of the sidebands can thus retrieve complete information. Clearly, the bandwidth is finite and has insignificant energy at high frequencies.

A POSTSCRIPT: Before leaving the topic (for the moment) we should note that the $(\sin x)/x$ function occurs repeatedly in sampling theory and in fact is often called the sampling function. In addition, the function is found in other fields; for example, the $(\sin x)/x$ function occurs in optics, where Fraunhoffer lines are analyzed.

ANOTHER POSTSCRIPT: In the case of ideal sampling with a pulse of infinitesimal width and infinite bandwidth, its spectral lines are placed at multiples of the sampling rate, as in natural sampling. The amplitudes of the lines, however, remain constant across the spectrum.

The Sampling Theorem

Given the fact that any practical channel is bandlimited, it is important to know the maximum transmission rate afforded by a channel. This question was answered by Harry Nyquist, an engineer working in Bell Telephone Laboratories, in his classic paper, "Certain Topics in Telegraph Transmission Theory," published in 1928. He demonstrated that a message of S Hz may be completely characterized by samples taken at less than $S/2$ seconds apart. This sampling condition is now well known as the Nyquist sampling theorem, which, in other words, states that the periodic sampling signal must be at least twice the highest information signal frequency. Looked at in another way, the message signal must be bandlimited to a frequency half that of the sampling frequency. Because messages do not always follow this criterion, it is necessary to low-pass filter them at half the sampling frequency, sometimes known as the Nyquist frequency. Although Nyquist proposed his result in terms of telegraph transmission, the result is equally valid for any kind of digital data transmission including, of course, digital audio.

The basis of sampling theory can be demonstrated with a simple example. Given an arbitrary signal $g(t)$, it may be sampled at a uniform rate T_s, yielding a pulse amplitude signal $g_s(t)$ defined at sample times. Sampling theory determines the conditions under which the sampled signal can uniquely define the original signal. The Fourier transform of the sampled signal consists of a series of delta functions spaced $1/T_s$ Hz apart. In particular, the Fourier transform may be expressed as

$$G_s(f) = \frac{1}{T_s} \sum_{n=-\infty}^{\infty} G(f - n/T_s).$$

Thus as a result of time sampling the original signal in the time domain, we obtain a periodic spectrum in the frequency domain with a period equal to the sampling rate. Suppose now that the signal is bandlimited with no components above S Hz, or in other words, the Fourier transform of the signal has the property that $G(f) = 0$ for $|f| > S$. In addition, let the sampling period be $T_s = 1/2S$. Then we may write the spectrum as

$$G(f) = \frac{1}{2S} \sum_{n=-\infty}^{\infty} g(n/2S) e^{-jn\pi f/S}, \qquad -S \leq f \leq S.$$

Because of the relationship of $g(t)$ to $G(f)$ via the inverse Fourier transform, the signal $g(t)$ is uniquely defined by the sample values $g(n/2S)$ for $-\infty \geq n \geq \infty$. Thus the sample values $g(n/2S)$ contain all the information of $g(t)$. Furthermore, this information may be extracted from the samples using reconstruction methods. Using the inverse Fourier transform, we obtain

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

This may be expressed as

$$g(t) = \sum_{n=-\infty}^{\infty} g(n/2S) \frac{1}{2S} \int_{-S}^{S} e^{j2\pi f(t - n/2S)}\, df.$$

Evaluating, we obtain

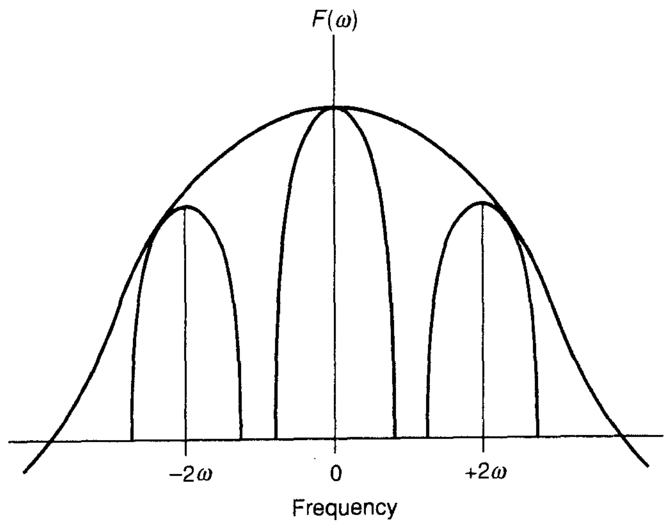
$$g(t) = \sum_{n=-\infty}^{\infty} g(n/2S)\, \mathrm{sinc}(2St - n).$$

Thus the sinc function may be used as an interpolation function to reconstruct the original signal $g(t)$ from the sample values $g(n/2S)$. Each sample is multiplied by its interpolation function, and added to the functions of all other samples to obtain the signal waveform $g(t)$. Importantly, the sinc function obtained above also represents the response of an ideal low-pass filter of bandwidth S, given an input signal of samples $g(n/2S)$. In other words, the original signal may be reconstructed exactly by passing the representing samples through a low-pass filter with a bandwidth S. Thus as Nyquist stated, a signal bandlimited to S Hz may be completely represented by specifying values of the signal at a period of $1/2S$ seconds. Further, a signal bandlimited to S Hz may be completely reconstructed from samples taken at a rate of $2S$ Hz. In other words, for a signal bandwidth of S, a sampling rate of $2S$ is required. This, of course, is the key component which permits the transformation of analog signals and digital sequences.

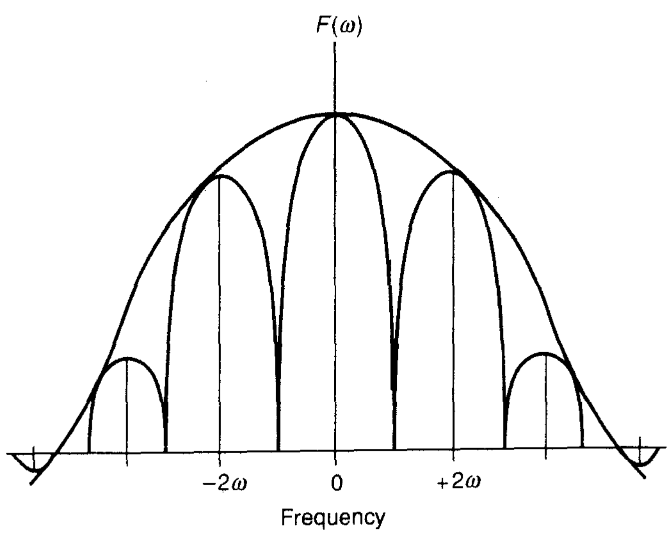
Summarizing, the spectrum of the message signal itself is repeated at multiples of the sampling frequency within the envelope of the sampled signal. When proper bandlimiting is provided, the spectrum repeats itself without overlap; if the signal's bandwidth is less than the Nyquist half-sampling frequency, the image sidebands are separated by a guard band, as shown in Figure

9a. The signal can easily be retrieved by filtering out higher-frequency sideband spectra, leaving only the first sideband. As noted, complete information is contained in each sideband, and thus, for example, a negative sideband could theoretically be used.

If the signal's bandwidth is exactly equal to the Nyquist frequency, a condition called critical sampling, it is still possible to retrieve the original signal by filtering. This, however, is the limiting case; it is shown in Figure 9b. If the signal's bandwidth exceeds the Nyquist sampling frequency, spectra will overlap, as shown in Figure 9c. In this case, aliasing occurs, and it is impossible to completely retrieve the original signal. Degradation increases as overlap increases, that is, as the new bandwidth increases. At best, filtering at half the sampling frequency will lose information in the spectrum above half the

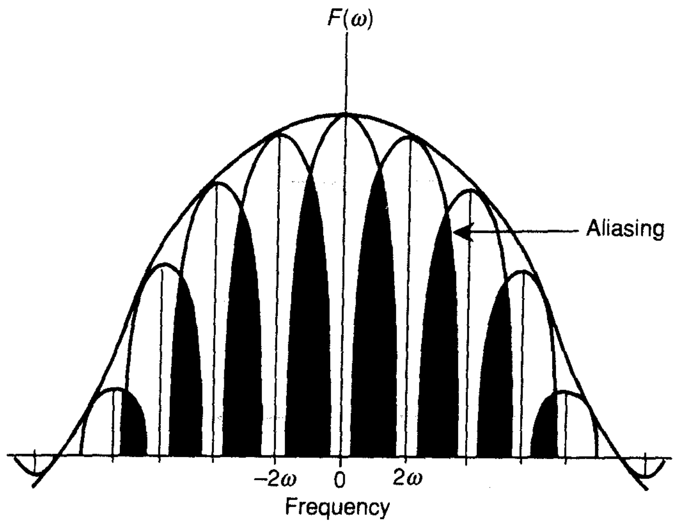


(a) The signal's bandwidth is less than half the sampling frequency.



(b) The signal's bandwidth is equal to half the sampling frequency.

Figure 9. Illustration showing the relationship between signal bandwidth and sampling frequency.



(c) The signal's bandwidth is greater than half the sampling frequency; aliasing results.

Figure 9. (cont.)

sampling frequency, and yet pass information from the inverted tail of the higher-frequency spectrum.

Information data, sent as pulses of duration T, can utilize a signaling rate of $1/T$ through a channel of bandwidth $f = 1/2T$, as shown in Figure 10. Each pulse conveyed through the bandlimited channel will exhibit a pulse response of $(\sin x)/x$. It is important to note that for any given pulse response at a sampling time, the pulse response of all other received signals theoretically passes through zero at that time, and at exact multiples of T seconds. This is because the channel acts as an ideal low-pass filter with ideal cutoff. The condition in which the response of side lobes does not pass through zero at sampling time decreases the system's ability to distinguish and accurately regenerate data; this interference condition is known as intersymbol interference, and can be caused by amplitude distortion and group delay. As noted, an eye pattern composed of regenerated data sequences can be used to evaluate a channel's performance.

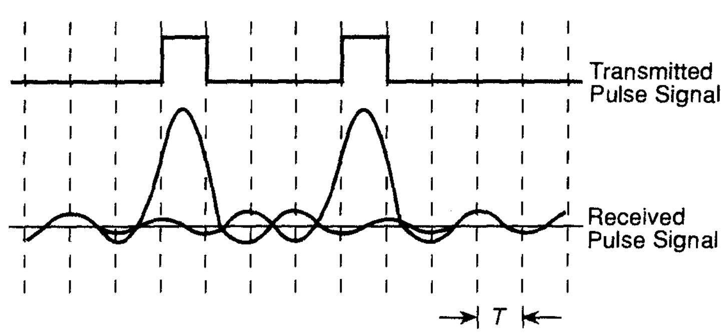


Figure 10. Message data sent as pulses of duration T can utilize a signaling rate of $1/T$ through a channel of bandwidth $f = 1/2T$.

Low-Pass Filtering

As noted, a low-pass filter must precede and follow every digitization system to bandlimit the signal and thus prevent aliasing, and to remove image spectra. In either case, an analog low-pass filter presents the prototypical circuit to accomplish these tasks. Its transfer function may be written as

$$f(s) = \frac{V_{\text{out}}}{V_{\text{in}}} = \frac{1}{1+S} = \frac{1}{1+j(\omega/\omega_0)},$$

where the angular frequency is $\omega = 2\pi f$ and ω_0 is the angular frequency at the filter's cutoff frequency of $f_c = 1/RC$.

The filter's cutoff frequency can be calculated:

$$f_c = 1/2\pi RC.$$

An ideal low-pass filter perfectly transmits all frequencies inside the passband and infinitely attenuates all frequencies in the stopband. An ideal filter is noncausal; its impulse response shows response prior to the time the unit impulse is applied. More importantly, the impulse response exhibits the sinc x function.

In practice, either analog or digital filters may be used. Although cascaded analog filter stages, resulting in a so-called brickwall characteristic, would bandlimit the signal, other artifacts such as ringing and phase nonlinearity encourage use of digital filters.

Such digital filters accept an input data sequence and output a low-pass filtered data sequence. Generally, given an input $x(n)$, output $y(n)$, and a and b, which are constant coefficients describing the filter's response,

$$y(n) = \sum_{i=0}^{M} a_i x(n-i) - \sum_{i=1}^{N} b_i y(n-i).$$

It can be seen that past output samples are used to form present output samples. Such a filter is said to be *recursive*; an example of such a digital filter is shown in Figure 11. When only past and present input samples are used, the equation is simplified to

$$y(n) = \sum_{i=0}^{M} a_i x(n-i).$$

The filter is said to be nonrecursive; an example of a nonrecursive filter is shown in Figure 12. A nonrecursive filter is especially attractive in many digital audio applications because of its phase linearity. The theory underlying the design of digital filters is explored in detail in Chapter 10. Oversampling, a method employed in these filters, is discussed below.

In any case, whatever the type of low-pass filter employed, the output waveform is comprised of the summed response of the individual impulse responses, as shown in Figure 13. Thanks to the miracle of superposition, together these individual responses form the reconstructed waveform.

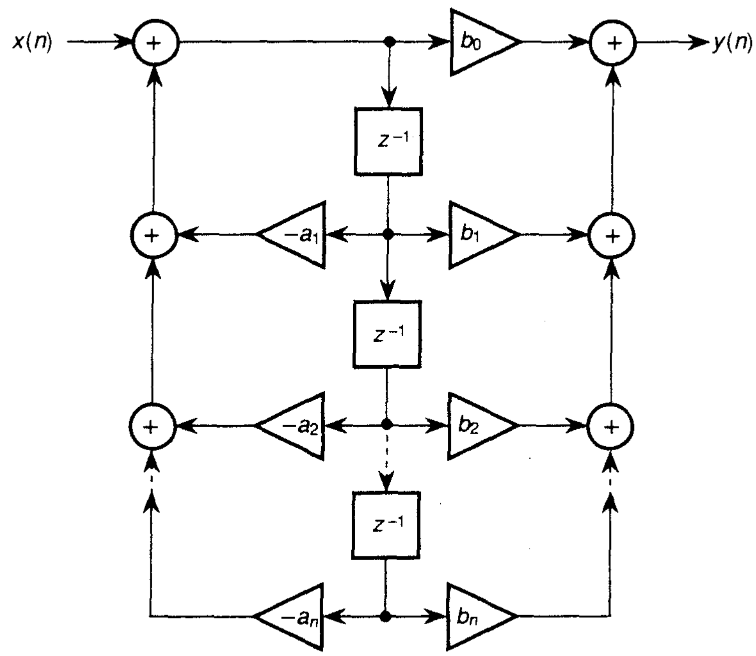


Figure 11. An example of an IIR filter using recursive structure.

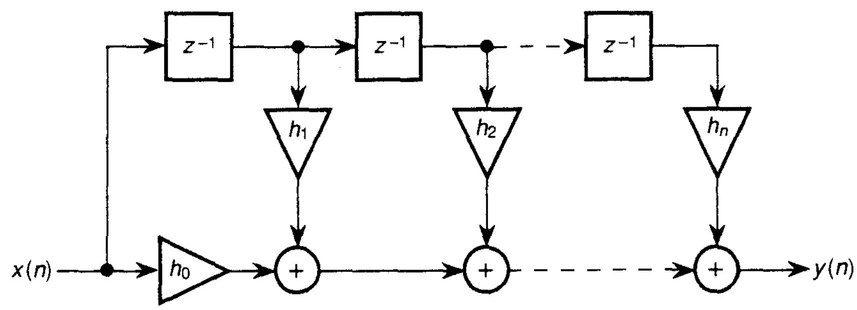


Figure 12. An example of an FIR filter using nonrecursive structure.

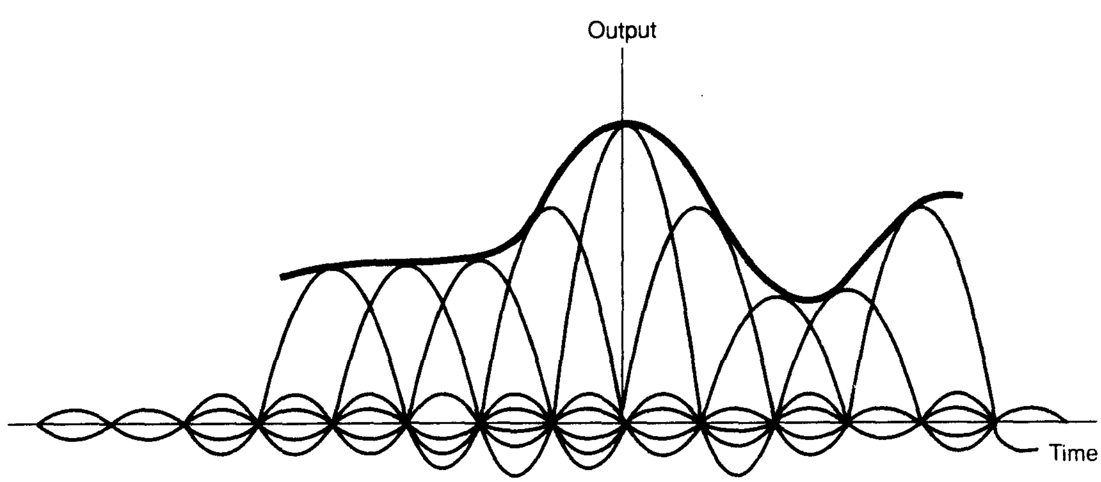


Figure 13. The summed response of the individual (low-pass filtered) impulse responses yields the reconstructed analog waveform.

Moreover, this waveform passes exactly through the same points as the original filtered input waveform. On the other hand, we are dealing with systems of finite bandwidth; in other words, in practice the impulse response must be truncated. Whereas the Fourier transform of an infinite (ideal) impulse response would create a rectangular pulse, a finite (real-world) impulse response would create a function exhibiting the Gibbs phenomenon, as shown in Figure 14. This is certainly not ringing as in analog systems but the mark of a finite-bandwidth digital system.

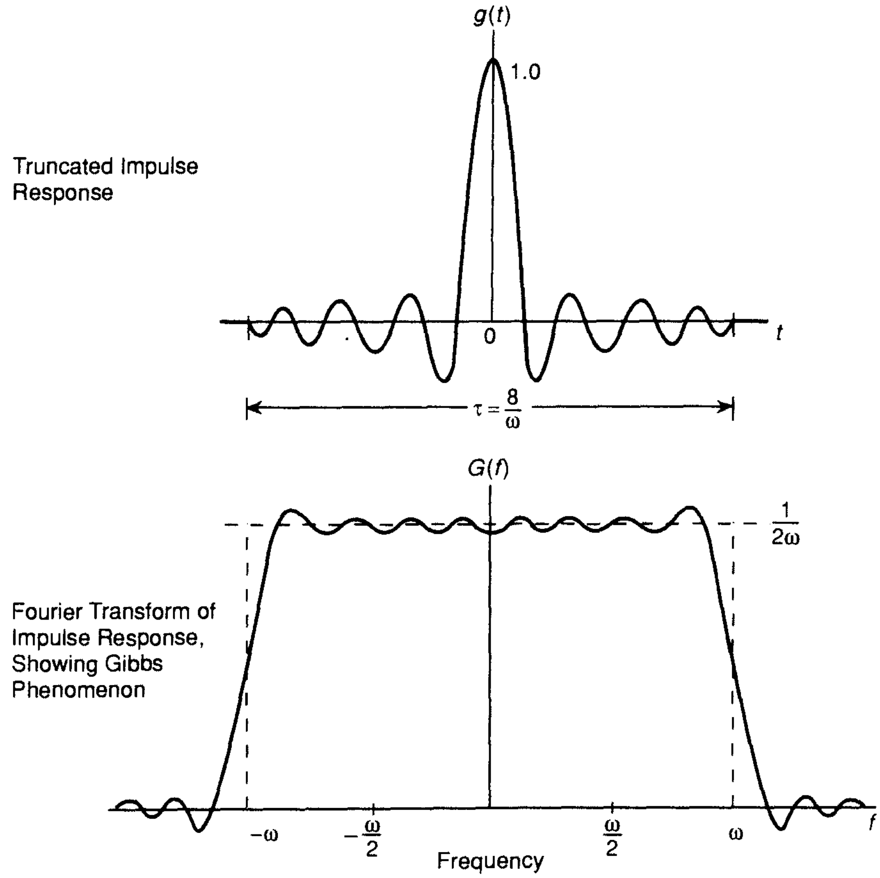


Figure 14. Illustration of the Gibbs phenomenon. *(From Haykin)*

Sample-and-Hold Circuits

Clearly, a sampling system must incorporate input filtering to prevent aliasing, as well as output filtering to remove spurious spectra. A sampling system may also employ two sample-and-hold circuits. An input sample-and-hold circuit may be used to periodically select sample points from the input signal at the sampling period T, and hold those analog values for a duration τ. Ideally, the amplitude of the pulse does not change. During this time, an input encoder or quantizer forms a representation of the sample. An output sample-and-hold circuit may also be required, for entirely different reasons; a slight frequency error must be corrected at the system output.

A sample-and-hold circuit, in lengthening pulses from duration τ to the sampling period T, performs the same transform as a low-pass filter. Given an input pulse of duration τ, its transform is equal to

$$H(\omega) = \frac{\tau}{T}\frac{\sin[(\tau/2)\omega]}{(\tau/2)\omega},$$

where ω is the angular frequency of the input signal.

We observe the $(\sin x)/x$ function once again, which accounts for the filtering that takes place. The contour of this function is the same as that present in defining the sampling spectra; however, the zero points differ. Specifically, the contour of a sample-and-hold circuit is zero at the sampling frequency. At the output of the sample-and-hold circuit, $\tau = T$. Thus

$$H(\omega) = \frac{\sin[(\tau/2)\omega]}{(\tau/2)\omega}.$$

When compared to the input signal bandwidth, as shown in Figure 15, the attenuation of the sample-and-hold contour represents high-frequency loss in the signal. At the maximum input frequency (half the sampling frequency) we observe that

$$\omega = \pi/T.$$

Thus the high-frequency value at half the sampling rate is

$$H(\pi/T) = \frac{\sin(\pi/2)}{\pi/2} \cong 0.64.$$

This is equivalent to a high-frequency attenuation of approximately 4 dB at half the sampling frequency. One solution is equalization; if an equalizer with an inverse response were placed in series, an overall flat response would result. Alternatively, the width of the output pulse can be manipulated using an output sample-and-hold circuit. As noted earlier in our discussion of sampling spectra, the bandwidth of the response is determined by the pulse width τ. The shorter the duration of the pulse, the greater is the frequency spread of the bandwidth. Using the same principle, we observe that aperture error can be minimized by decreasing the duration of the pulse width, as shown in Figure 16. In other words, by decreasing the pulse duration, an output sample-and-hold circuit can be used to more closely approximate the output of an impulse train of an ideal D/A converter. For example, whereas the frequency error at half the sampling frequency is roughly 4 dB when $\tau = T$, it decreases to 0.2 dB when τ is decreased to $\frac{1}{4}T$:

$$H(\pi/T) = \frac{T}{4} = \frac{\sin(\pi/8)}{\pi/8} \cong 0.97.$$

Equalization could easily be used to compensate for this small error.

Further shortening of pulse duration would overly decrease average voltage level and hence diminish the S/N ratio.

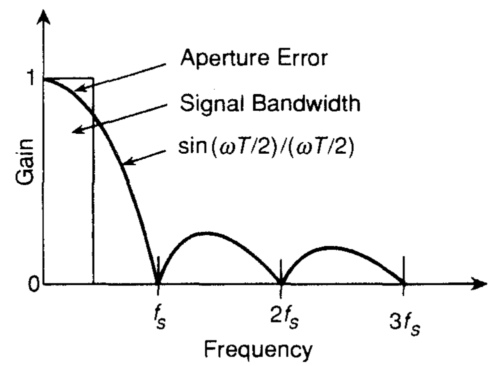
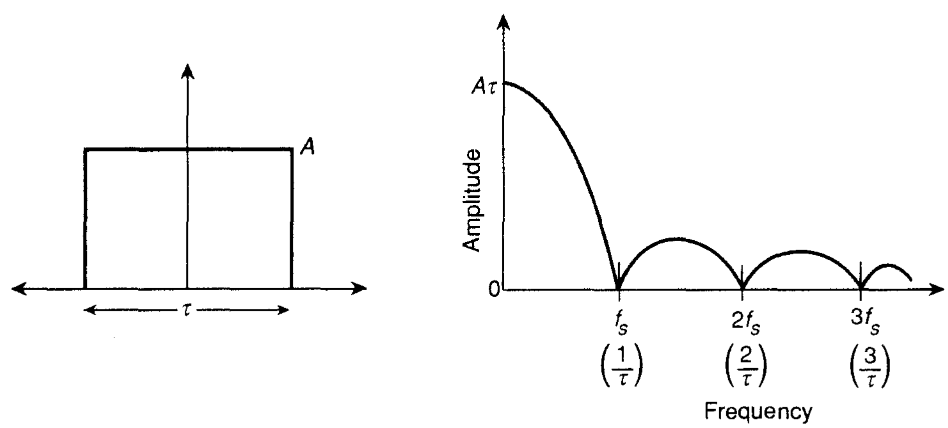
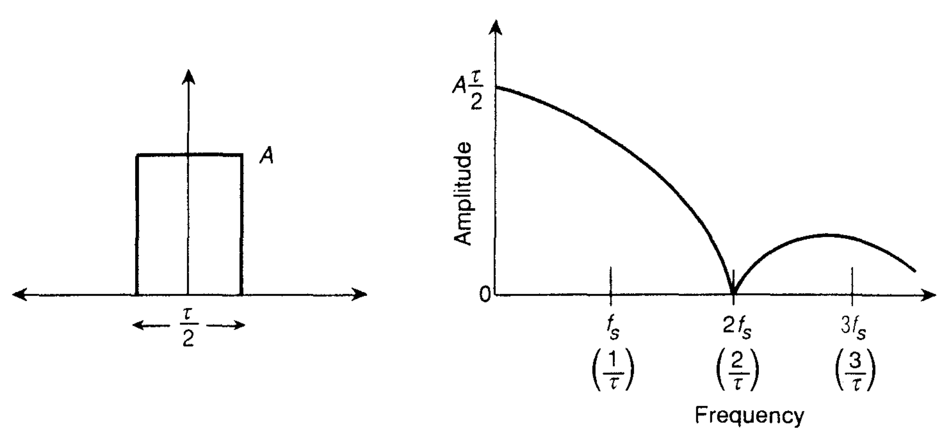


Figure 15. The sample-and-hold function creates a high-frequency loss in the signal when compared with the input bandwidth; this is aperture error.



(a) A longer duration pulse yields rapid high-frequency rolloff.



(b) A shorter duration pulse yields an extended high-frequency response.

Figure 16. Aperture error can be minimized by decreasing the duration of the output pulse width.

Oversampling

Most contemporary digital audio systems employ a sampling rate of 44.1 or 48 kHz. Some critics have suggested that these sampling rates be doubled. These higher sampling rates would extend frequency response from 20 kHz to 50 kHz. Unfortunately, this extended frequency response is beyond the range of human hearing. The only practical advantage of a higher sampling rate is the decrease in demands on the low-pass filters which precede and follow a digital audio system. The need to sharply limit audio energy at frequencies higher than half the sampling frequency encouraged the use of analog brickwall filters; such filters introduce phase nonlinearities. This problem, however, can be avoided in both A/D and D/A conversion by using oversampling techniques; with these techniques, phase nonlinearities are negligible.

As we have seen, a look at the spectrum of the reconstructed signal reveals the identity of the high-frequency signals, which are images of the original audio signal's spectrum, repeated at multiples of the original sampling frequency. In the case of a 44.1 kHz sampling format, this generates image spectra at 88.2 kHz, 132.3 kHz, 176.4 kHz, 220.5 kHz, etc. As we observed, this is a natural consequence of using sampled waveforms. Although these frequencies are well above the highest frequency audible to humans, they could interfere with downstream electrical equipment. For example, they could affect an analog tape recording by beating against a bias oscillator, or they could affect FM stereo radio transmission. A high-quality digital audio product must suppress these high frequencies. Although this high-frequency information can be removed from the analog waveform using an analog filter, there is a more efficient way. Using a digital filter, the data itself may be processed before it is reconstructed as an analog waveform. With proper processing, the problem of filtering may be largely solved while the data is still in digital form.

In an oversampling filter, audio samples are input and subjected to computation which implements digital filtering of the audio signal. Additional audio samples are generated between the original samples through interpolation, hence the output sampling rate is increased. Intermediate samples are multiplied by fixed coefficients corresponding to their contribution to the overall response of the filter. The output filtered sample is produced by summing together the multiplication products. The spectrum of the signal is changed, with the images appearing at multiples of the new (oversampled) sampling rate. This additional data creates a more linear waveform and shifts unwanted modulation noise to an extreme frequency, where it can be removed without audible effect.

For example, in an eight-times oversampling filter, seven new audio samples are computed for each input sample; an input data rate of 44.1 kHz would be raised to an output rate of 352.8 kHz. Modulation artifacts are shifted to a band centered at 352.8 kHz, where they are easily removed with an analog low-pass filter. The accuracy of such digital filters is precise, yielding passband ripple on the order of 0.00001 dB. In addition, the stopband suppression is greater than 120 dB.

We may define an oversampling ratio as

$$R = f_a/f_s,$$

where f_a is the oversampling frequency and f_s is the input Nyquist sampling frequency.

Oversampling initially requires insertion of $(R-1)$ zero samples per input sample. The samples created by oversampling must be placed symmetrically between the Nyquist input samples. A low-pass filter is used to bandlimit the input data to $f_s/2$, with spectral images at integer multiples of $(R \times f_s)$. Moreover, low-pass filtering creates the intermediate sample values, providing interpolation. Rather than perform multiplications on zero samples, redundancy may be observed to design a transposed FIR filter in which the number of multiplications is reduced by $(R-1)/R$.

Oversampling benefits a digital system's dynamic range by reducing requantization noise. Requantization noise adds $Q^2/12$ noise power to the signal, where Q is 1 LSB; this noise is uniformly distributed over the entire baseband. With oversampling, the spectrum is increased so the noise in the audio band is relatively reduced. For example, a four-times oversampling filter, without any noise shaping, reduces in-band requantization noise by 6 dB. Noise shaping is considered in more detail in Chapter 12.

Consider the eight-times oversampling digital filter in Figure 17 in which seven intermediate samples are generated for each input sample. This filter consists of a shift register of 24 delay elements, each delaying a sample for one sampling period. Each 16-bit sample is multiplied by a 12-bit coefficient. This generates words of 28 bits. The multiplication products are summed eight times during each period and are then output from the filter. Thus eight times as many samples are present after oversampling, with new intermediate values

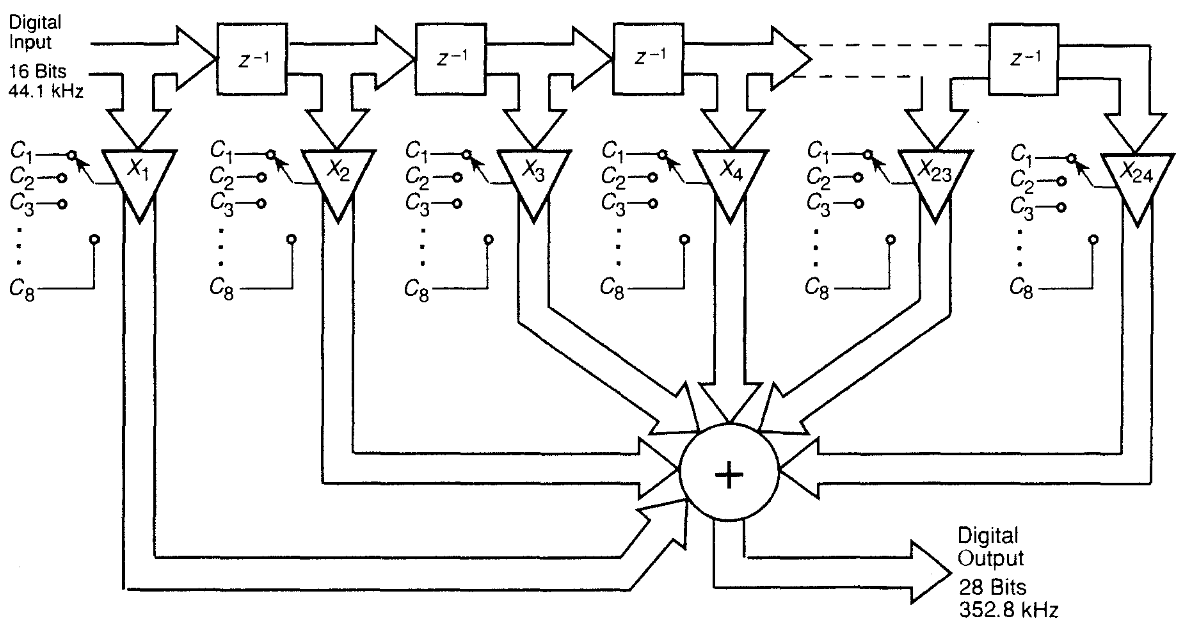


Figure 17. An example of an eight-times oversampling filter using a twenty-four element transversal architecture.

calculated by the filter. The sampling frequency is increased to 352.8 kHz. The filter's coefficients produce a transition region between 20 and 24.3 kHz and again around 330 kHz. Figure 18 shows the effect of the filtering. Note that when the effect of the output sample-and-hold circuit is accounted for, the amplitudes of the output spectra are attenuated.

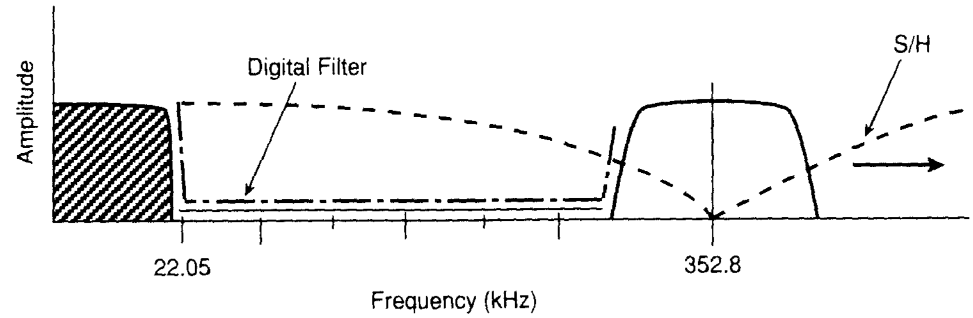


Figure 18. Digital filtering extends the placement of the image spectra, suppressing intermediate components. The sample-and-hold function further suppresses high-frequency image spectra.

Following digital filtering, the data is converted to the analog domain with a D/A converter. Because the multiplication in the filter generates words of 28 bits, designers are free to choose words of 16, 18, or 20 bits or more, to apply to 16-, 18-, or 20-bit D/A converters. However, any remaining bands must be completely suppressed by an anti-imaging analog filter following the converter. The oversampling rate determines where the images are placed and hence the kind of post filter needed to remove them. Generally, since the remaining band is so high in frequency, a filter with a gentle response can be used. It is a noncritical design, and its low order guarantees good phase linearity; a simple *RC* low-pass filter may be employed.

A digital filter thus efficiently accomplishes the task of anti-image filtering without resorting to analog brickwall filters. In addition, oversampling may be used to decrease the wordlength required to represent a digital audio signal. Information theory assures us that when the output sampling rate is greater than the input sampling rate, the output wordlength may be less than the input wordlength without loss of source data. This opportunity leads to the concept of low-bit D/A converters, as described in Chapter 12.

Quantization

With PCM systems, degradation occurs when a sample is measured and assigned an approximate amplitude value. In other words, quantization results in a quantization error, manifested as noise that can be correlated to the signal. This error may be reduced by increasing the number of quantization quanta (ie. longer wordlength); this in turn increases the bandwidth required. The signal-to-quantization noise power ratio increases exponentially with bandwidth. This is an efficient relationship which approaches the theoretical maximum, and is a hallmark of coded systems, such as PCM.

With uniform quantization, an analog signal is mapped into a number of quanta of equal height and placement. The infinite number of points on the analog waveform must be quantized to the finite number of quanta; clearly, this introduces an error. No matter what the medium, the quality of the representation increases as the number of quanta is increased; for example, a high-quality color picture may require 512 color levels, and high-quality music may require 65,536 amplitude levels. On the other hand, only a few levels can still carry reasonable information content; for example, two amplitude levels can (barely) convey intelligible speech.

An analytical means of expressing this quality, in terms of degradation, is the mean square quantizing error $\overline{E^2(t)}$. If Q is the amplitude of a quantum, and E is the difference between the analog signal and the quantized signal, then E must lie between $-Q/2$ and $+Q/2$. Assuming that E possesses equal probability in the range from $-Q/2$ to $+Q/2$,

$$\overline{E^2} = \int_{-\infty}^{\infty} E^2 P(E)\, dE,$$

where $P(E)$ is the probability density function of E such that

$$P(E) = \begin{cases} 1/Q & \text{when } -Q/2 \le E \le Q/2, \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$\overline{E^2} = \int_{-Q/2}^{Q/2} E^2 \frac{1}{Q}\, dE = \frac{1}{Q} \int_{-Q/2}^{Q/2} E^2\, dE = \frac{Q^2}{12}.$$

Furthermore, if k is the number of bits used to represent the signal,

$$Q = 1/2^{k-1)}.$$

Thus,

$$\frac{Q^2}{12} = \frac{2^{-2k}}{3}.$$

This assumes that the transmitted signal is bandlimited and sampling adheres to the Nyquist theorem. Given a noise spectrum of $\pm f_n$, and sampling rate f_s, it can be shown that the total quantization noise power is equal to

$$N_Q = \int_{-f_n}^{f_n} N(f)\; df = \frac{Q^2}{12}.$$

The noise power is flat, and the level of the noise power spectral density is

$$N(f) = \frac{Q^2}{12} \cdot \frac{1}{f_s} = \frac{2^{-2k}}{3 f_s}.$$

Additional analysis would show that the spectral content of the quantization error signal lies in the bandwidth of the message signal. The result is only valid given the assumption that the error probability is equal for any amplitude. Furthermore, as noted above, oversampling decreases in-band noise power as a function of the oversampling rate. Specifically, the in-band noise power is equal to

$$N_B = \int_{-f_n}^{f_n} N(f)\ df = \frac{Q^2}{12}\frac{2f_n}{f_s}.$$

Given an expression for noise power, it is relatively simple to derive an expression for S/N ratio for a PCM system. Assuming a sinusoidal message signal with peak amplitude A normalized between ± 1, the average signal power S is

$$S = A^2/2.$$

The S/N ratio power is thus

$$\mathrm{S/N} = \frac{A^2}{2}\frac{12}{Q^2} = \frac{6A^2}{Q^2}.$$

If there are n quantization steps, and the wordlength is k bits in length, then

$$Q = 2/n,$$

and

$$Q = \frac{2}{2^k - 1}.$$

Thus

$$\mathrm{S/N} = \frac{6A^2}{4}(2^k - 1)^2 = \frac{3}{2}A^2(2^k - 1)^2.$$

Written as a decibel expression,

$$(\mathrm{S/N})_{\mathrm{dB}} = 10\log_{10}(3/2) + 20\log_{10} A + 20\log_{10}(2^k - 1).$$

Thus, for a sinusoidal signal of maximum allowed input amplitude, this expression can be simplified to

$$(\mathrm{S/N})_{\mathrm{dB}} = 6.02k + 1.76.$$

This result, however, assumes that the quantization error is uniformly distributed, and quantization is accurate enough to prevent signal correlation in the error waveform. This is generally true for high-amplitude audio signals but certainly is not the case for low-amplitude signals.

To reduce quantization error, as noted, the quantum size can be made smaller; however, this requires an increase in transmission bandwidth. In some PCM systems, quantization error is minimized by use of nonuniform quantization quantum sizes. Such systems attempt to tailor quantum sizes to best suit the statistical properties of the signal. This is logical because the probability distribution of most message signals is not uniform, as assumed above. For example, speech signals would be best served by an exponential-type quantization distribution; this assumes that small-amplitude signals are more prevalent than large signals. Many quantization levels at low amplitudes, and fewer at high amplitudes, should result in decreased error. Companding, with compression prior to uniform quantization, and expansion following quantization, can be used to achieve this result. Quantization and requantization conversion may be performed with a variety of techniques. Multibit circuits are described in Chapter 3; low-bit techniques are described in Chapter 12.

Dither

The effects of quantization error may be reduced by increasing wordlength, or by using methods such as companding. A far more efficient technique, however, is the use of dither. Fundamentally, dither is a low-level noise signal added to the message signal prior to quantization. Dither decorrelates and minimizes the effect of quantization and requantization error to the point of elimination, causing a digital system to act more like an analog system in this respect; however, a digital system, properly dithered, far exceeds the signal-to-noise performance of an analog system. On the other hand, an undithered digital system can be inferior to an analog system, particularly under low-level signal conditions. A high-quality digital audio system demands use of dither prior to quantization; in addition, digital computation may require use of dither. For example, gain changing, oversampling interpolation, equalization, and sample rate conversion all require dithering to reduce the effects of quantization and noise modulation as well as to limit cycle oscillations possible in IIR and noise shaping filters. There are many possible dither signals, nonsubtractive as well as subtractive; alternatively, as noted, digital dither may be employed.

Analog dither, applied prior to a linear A/D converter, causes the A/D converter to make additional level transitions which preserve low-level signals through duty cycle modulation—a kind of pulse width modulation. This linearizes the quantization process and helps eliminates error because the average value of the output reflects the input signal. Harmonic distortion and intermodulation products are converted to wide-band noise. A price, however, is paid in a slightly raised noise floor; noise power is $Q^2/4$, where Q is 1 LSB. Because of the ease with which Gaussian noise may be generated in the analog domain, it is often selected for this application. The mathematical expression for a probability density function (PDF) of a zero-mean random noise with rms value of v_{rms} is given by

$$P(v) = \frac{1}{v_{\text{rms}}\sqrt{2\pi}} \exp(-v^2/2v_{\text{rms}}{}^2).$$

A Gaussian dither eliminates distortion and reduces noise floor modulation. However, it results in a relatively higher noise floor compared with other kinds of dithering. Other types of dither generally produce better results than Gaussian dither. In some cases, the input analog noise floor itself acts as sufficient dither. For example, a pseudo random noise generator may be used to produce the required dither signal. Lipshitz and Vanderkooy [1989] have shown that dither with a rectangular probability density function (RPDF) eliminates distortion products caused by quantization but does not eliminate noise floor modulation. RPDF dither of $\pm Q/2$ adds $Q^2/12$ noise power. RPDF dither is a uniformly distributed random voltage between $\pm 1/2$ LSB:

$$P(v) = \begin{cases} 1/Q & \text{when } -Q/2 \leq v \leq Q/2, \\ 0 & \text{otherwise,} \end{cases}$$

where Q is one quantizing quantum or LSB step size.

Alternatively, dither with a triangular probability density function (TPDF) can be employed; TPDF can be obtained by summing two random numbers with uniform distribution, that is, by summing two RPDF signals. The total width of the signal is $2Q$. TPDF dither minimizes both distortion and noise floor modulation; the noise floor, however, is somewhat higher than in RPDF dither; TPDF dither adds $Q^2/6$ noise power. RPDF and TPDF dither are easily generated in the digital domain and are preferable to Gaussian dither. Alternatively, a high-frequency dither placed at the Nyquist frequency can be used to alter the spectrum of quantization error to minimize its audibility. An example of the effects of high-frequency triangular dither is shown in Figure 19; a 22 kHz triangular dither of 1/2 LSB peak amplitude is added to a −90 dB, 330 Hz sinewave quantized with 16 bits. The dither creates a pulse width modulation signal in which the average value approximates the sinewave input signal. The frequency response plots show that harmonic distortion components are clearly reduced in the dithered signal. In addition, this high-pass dither does not add as much noise to the signal as broadband dither. In this case, however, use of a discrete dither frequency leads to intermodulation products. A wideband dither signal would help alleviate this artifact.

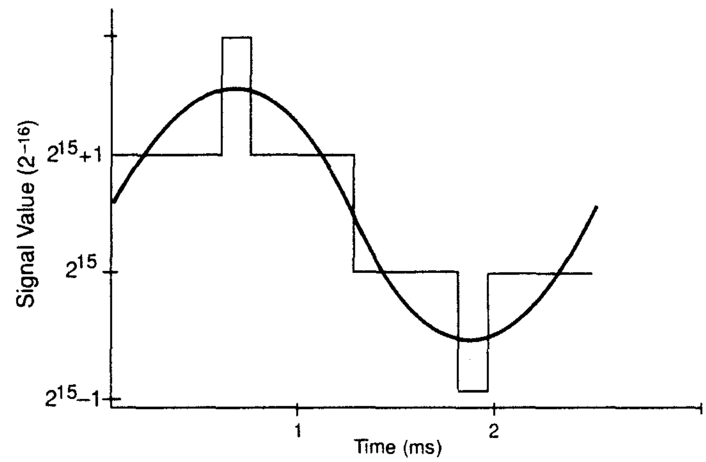
As noted, digital dither is advantageous where signal manipulation takes place. For example, the truncation associated with multiplication can cause objectionable error, especially as it accumulates through multiple calculations. In effect, such rounding adds the same distortion and noise modulation errors as requantization, even if the input digital signal is properly dithered. Infinite impulse response noise shaping filters with digital feedback can exhibit limit cycle oscillations with low-level signals if gain reduction or certain equalization processing (resulting in gain changes) is performed; digital dither can be used to randomize these cycles.

As Vanderkooy and Lipshitz [1987] have shown, truncated or rounded digital words can be redithered with RPDF or TPDF fractional numbers, as shown in Figure 20. An RPDF dither word D_1 is added to a digital audio word with integer part P_i and fractional part P_f. The carry bit acts as dither for the rounding process in the same way that 1 LSB analog RPDF dither would affect an A/D converter. When a statistically independent RPDF dither D_2 is added, TPDF dither results. This TPDF dither noise power is $Q/6$, and rounding noise power is $Q^2/12$, so total noise power is $Q^2/4$. The final sum has integer part S_i and fractional part S_f, which become S upon rounding.

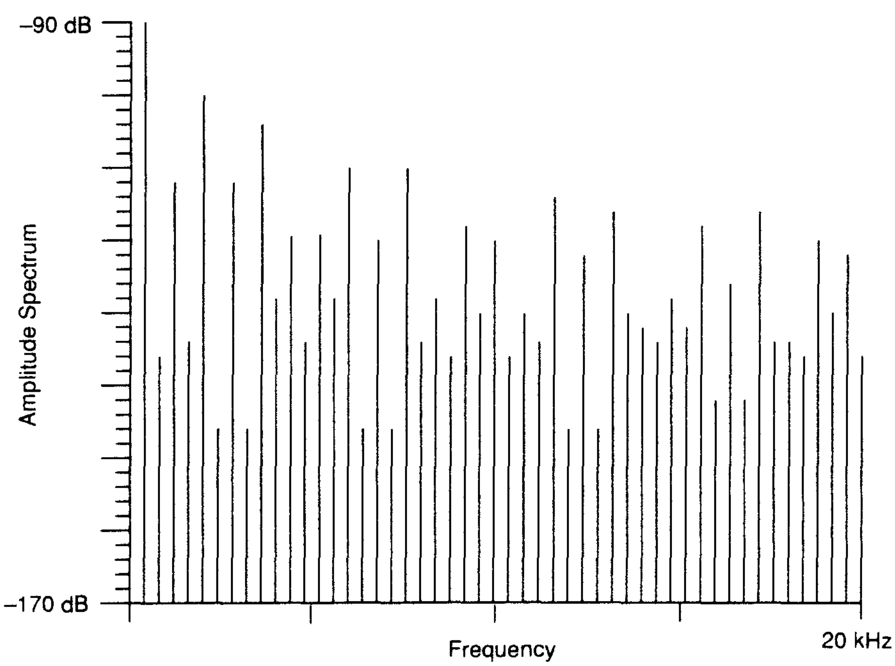
In cases of gain fading, TPDF appears to be a better choice than RPDF because of its elimination of noise modulation as well as distortion, at the expense of a slightly higher noise floor. To minimize audibility of this noise penalty, high-pass TPDF dither may be most appropriate in rounding. The TPDF statistics are not changed; however, dither samples are correlated. The power spectrum of the dither signal, normalized for total noise power $Q^2/6$, showing the autocorrelation is

$$f(V) = \int_{-\infty}^{\infty} f(V + v)p(v)\ dV.$$

Average dither noise power is $Q^2/6$, with no noise at 0 Hz and double the average value at the Nyquist frequency—hence the term high-pass dither. This



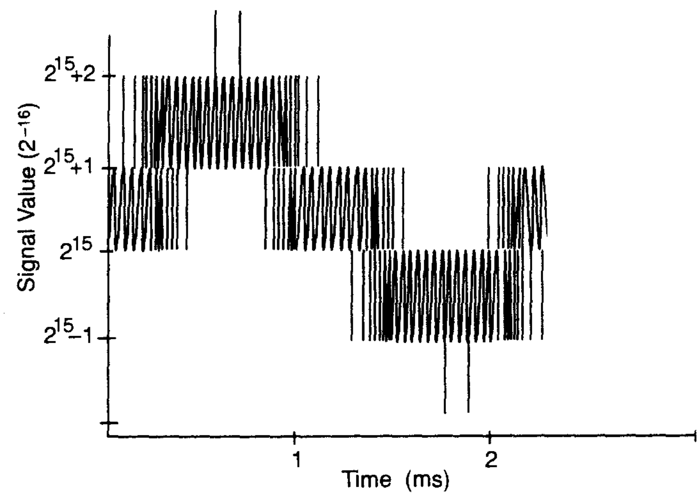
(a) A −90 dB, 330 Hz sinewave quantized to 16 bits, without dither.



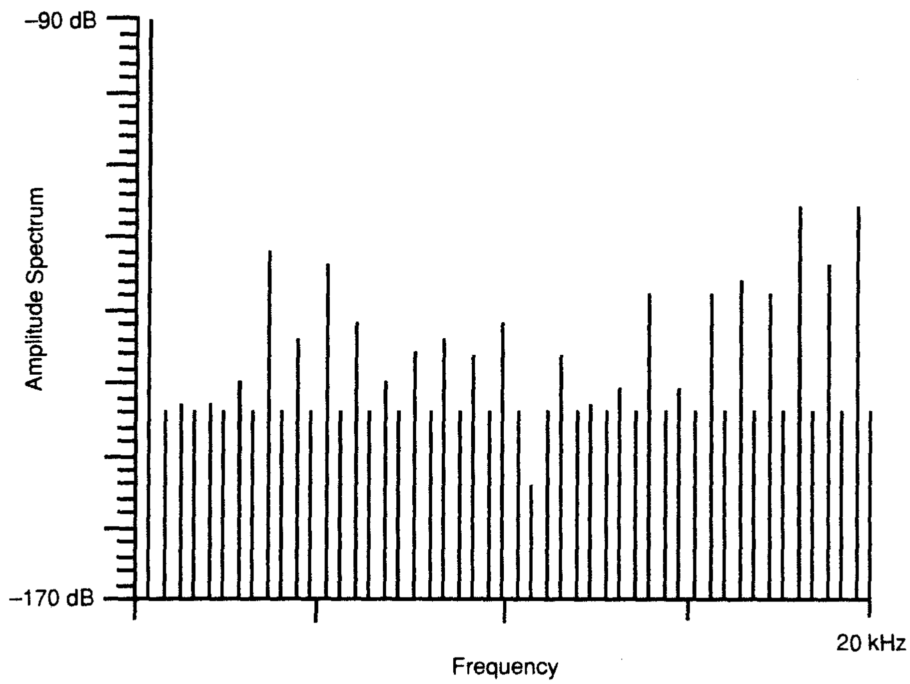
(b) Frequency spectrum of the undithered quantized sinewave showing harmonic distortion components.

Figure 19. Illustration of the effects of high-frequency triangular dither.
(From Dijkmans and Naus)

shaping becomes more pronounced with noise increasingly shifted outside the audio band, as the oversampling ratio is increased, for example, by two or four times. These and related topics are discussed in more detail in Chapter 12. As already noted, the audible effect of a noise penalty is lessened when oversampling is employed because in-band noise is relatively decreased proportional to the oversampling rate. Further reduction of the $Q^2/12$ requantization noise power can be achieved through noise shaping circuits, as described in Chapter 12.



(c) A 22 kHz triangular dither of 1/2 LSB peak amplitude is added to the −90 dB, 330 Hz sinewave quantized to 16 bits.



(d) Frequency spectrum of the dithered quantized sinewave showing reduced distortion components.

Figure 19. (cont.)

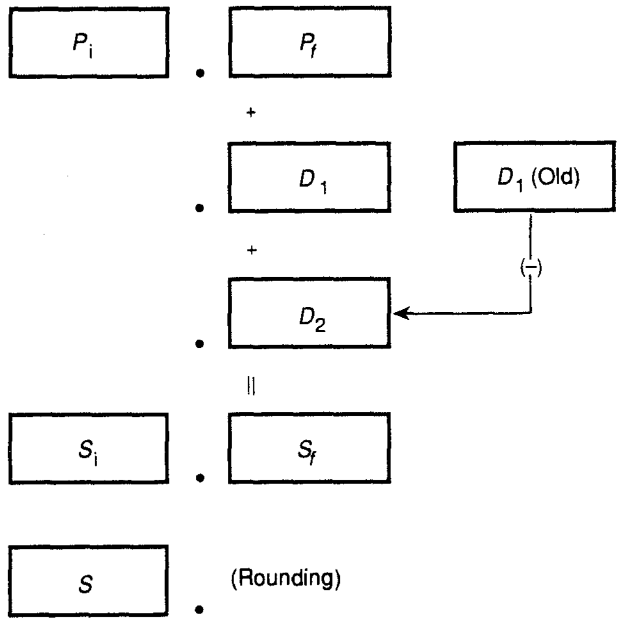


Figure 20. An example of the use of digital redithering during truncation or rounding computation. *(From Vanderkooy and Lipshitz)*

Conclusion

Although complex, the world of digital audio modulation and sampling is based on well-understood principles. Different forms of modulation, although none are perfect, offer a variety of advantages for virtually any application. In practice, a digital audio system may variously employ a PAM signal during encoding, PCM during transmission, and PWM during decoding—each providing a unique asset. Throughout the process, the cornerstone of discrete time sampling yields completely predictable results, thanks to properties of the impulse response, summation, and other fortunate happenstances. Unfortunately, bandwidth demands, quantization, and other negative attributes can affect the success of a system. However, engineers can make use of clever methods, such as dither, to minimize their harm and help ensure the successful operation of the audio digitization system.

References

AES Recommended Practice for Professional Digital Audio Applications Employing Pulse Code Modulation: Preferred Sampling Frequencies. AES5-1984 (ANSI S4.28-1984), *Journal of the Audio Engineering Society*, vol. 32, no. 10, October, 1984.

Blesser, B. A., and B. N Locanthi, "The Application of Narrow-Band Dither Operating at the Nyquist Frequency in Digital Systems to Provide Improved

Signal-to-Noise Ratio over Conventional Dithering," *Journal of the Audio Engineering Society*, vol. 35, no. 6. June, 1987.

Finger, R. A., "On the Use of Computer Generated Dithered Test Signals," *Journal of the Audio Engineering Society*, vol. 35, no. 6, June, 1987.

Gersho, A., "Quantization," *IEEE Communication Society Magazine*, vol. 15, September, 1977.

Harris, S., "The Effects of Sampling Clock Jitter On Nyquist Sampling Analog-to-Digital Converters, and on Oversampling Delta-Sigma ADCs," *Journal of the Audio Engineering Society*, vol. 38, no. 7/8, July/August, 1990.

Hawksford, M. O. J., "Chaos, Oversampling, and Noise Shaping in Digital-to-Analog Conversion," *Journal of the Audio Engineering Society*, vol. 37, no. 12, December, 1989.

Haykin, S., *Communication Systems*, John Wiley & Sons, New York, 1983.

Jayant, N., and L. Rabiner, "The Application of Dither to the Quantization of Speech Signals," *Bell System Technical Journal*, vol. 51, 1972.

Morris, D. J., *Pulse Code Formats for Fiber Optical Data Communication*, Marcel Dekker, 1983.

Muraoka, T., Yamada, Y., and M. Yamazaki, "Sampling Frequency Considerations in Digital Audio Standards," *Journal of the Audio Engineering Society*, vol. 26, no. 4, April, 1978., Erratum, *ibid.*, vol. 26, no. 7/8, July/August, 1978.

Owen, F. F. E., *PCM and Digital Transmission Systems*, McGraw-Hill Book Co., 1982.

Pohlmann, K. C., *Principles of Digital Audio*, 2nd edition, Howard W. Sams & Co., 1989.

Rossum, D., "Digital Dither in Music and Sound Synthesis," *Audio in Digital Times: 7th AES Conference Proceedings*, May, 1989.

Shannon, C. E., "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, 1948.

———, "Communication in the Presence of Noise," *Proceedings of the Institute of Radio Engineers*, vol. 37, 1949.

Trischitta, P. R., and E. L. Varma, *Jitter in Digital Transmission Systems*, Artech House, 1989.

Vanderkooy, J., and S. P. Lipshitz, "Digital Dither," *AES Preprint 2412*, December, 1986.

———, "Digital Dither: Signal Processing with Resolution Far Below the Least Significant Bit," *Audio in Digital Times: 7th AES Conference Proceedings*, May, 1989.

———, "Dither in Digital Audio," *Journal of the Audio Engineering Society*, vol. 35, no. 12, December, 1987.

———, "Resolution Below the Least Significant Bit in Digital Audio Systems with Dither," *Journal of the Audio Engineering Society*, vol. 32, no. 3, March, 1984, Erratum, *ibid.*, vol. 32, no. 11, November, 1984.

Widrow, B., "A Study of Rough Amplitude Quantization by Means of Nyquist Sampling Theory," *IRE Transactions on Circuit Theory*, vol. CT-3, December, 1956.

———, "Statistical Analysis of Amplitude-Quantized Sampled-Data Systems," *Transactions of the AIEE (Applications and Industry)*, vol. 79, January, 1961.

Wong, P. W., "Quantization Noise, Fixed-Point Multiplicative Roundoff Noise, and Dithering," *IEEE Transactions*, ASSP, vol. 38, no. 2, February, 1990.