1 Subscriber Loop Transmission Concepts and Signal Conversion

Chapter 1 Acronyms

A/D	analog-to-digital
ADM	adaptive delta modulation
ADPCM	adaptive differential pulse code modulation
AM	amplitude modulation
ANSI	American National Standards Institute
APC	adaptive predictive coding
APK	amplitude-phase keying
ASBC	adaptive subband coding
ATM	asynchronous transfer mode
BER	bit error rate or bit error ratio
BETRS	basic exchange telecommunications radio service
CCITT	The International Telegraph and Telephone Consultative Committee
codec	COder-DECoder
CVSDM	continuously variable slope delta modulation
D/A	digital-to-analog
dc	direct current
DCME	digital compression multiplex equipment
DLC	digital loop carrier
DM	delta modulation
DPCM	differential pulse code modulation
DSI	digital speech interpolation
DTE	data terminal equipment
DTMF	dual tone multifrequency
FCC	Federal Communications Commission
FDM	frequency division multiplexing
FEC	forward error correction
FM	frequency modulation
FSK	frequency-shift keying
FTTC	fiber-to-the-curb
FTTH	fiber-to-the-home
HDSL	high bit-rate digital subscriber line
HDTV	high-definition television
ISDN	integrated services digital network
LD-CELP	low delay code-excited linear prediction
LPC	linear predictive coding
MOS	mean opinion score
MP-LPC	multipulse linear predictive coding
PAM	pulse-amplitude modulation
PCM	pulse code modulation
PM	phase modulation
PSK	phase-shift keying
PSTN	public switched telephone network
QDU	quantizing distortion unit
RF	radio frequency
rms	root mean square

continued

SDR	signal-to-quantizing distortion ratio	TASI	time assignment speech interpolation
SE-LPC	stochastically excited linear predictive coding	TDM	time division multiplexing
SNR	signal-to-noise ratio	TLP	transmission level point
SONET	synchronous optical network	VNL	via net loss
SPE	synchronous payload envelope	VQL	variable quantizing level
		VT	virtual tributary

1.1 Subscriber Loops

A subscriber loop is that part of a telecommunication transmission system between a subscriber's premises and the serving central office. Subscribers can be individuals or businesses, and in this book the term *subscriber* is used interchangeably with end-user, or just user. Loops consist of twisted metallic cable pairs, radio links, or optical fibers and serve as end-links in the communication channel between users. The loop usually carries low-traffic volumes compared to interoffice facilities.

Loops can take on many forms, from a simple twisted cable pair carrying an analog voicegrade signal between two points to a complex network of optical fibers and associated interfaces carrying combinations of voicegrade and video signals and high-speed digital data. This book concentrates on landline digital loops. A detailed treatment of analog loops is given in a companion volume [1]. Some digital loop configurations are shown in Figure 1-1.

The twisted pair loop of Figure 1-1(a) is ubiquitous—it appears almost everywhere in the public network. In digital applications, it carries digital signals covering the range from near dc to over 3 Mbps, with bandwidths extending upwards of 6 MHz. When multiplexers are used with digital loops, a number of individual analog and digital channels can be combined into an aggregate data stream. Digital loop carrier (DLC), conceptually shown in Figure 1-1(b), is a typical system that uses this technique.

Where it is uneconomical to install twisted pair or fiber optic cable, such as in rural areas, digital or analog radio systems are used, as shown in Figure 1-1(c). With digital radio, analog voice frequency signals at remote terminals (subscriber stations) are digitally encoded and then used to modulate a radio frequency (RF) carrier for transmission to a base station (central office station) that serves one subscriber in point-to-point applications or many subscribers in point-to-multipoint applications. Examples of digital radios used in subscriber loop applications are those licensed by the Federal Communications Commission (FCC) as Basic Exchange Telecommunications Radio Service (BETRS) under the Rural Radio Service rules [2].

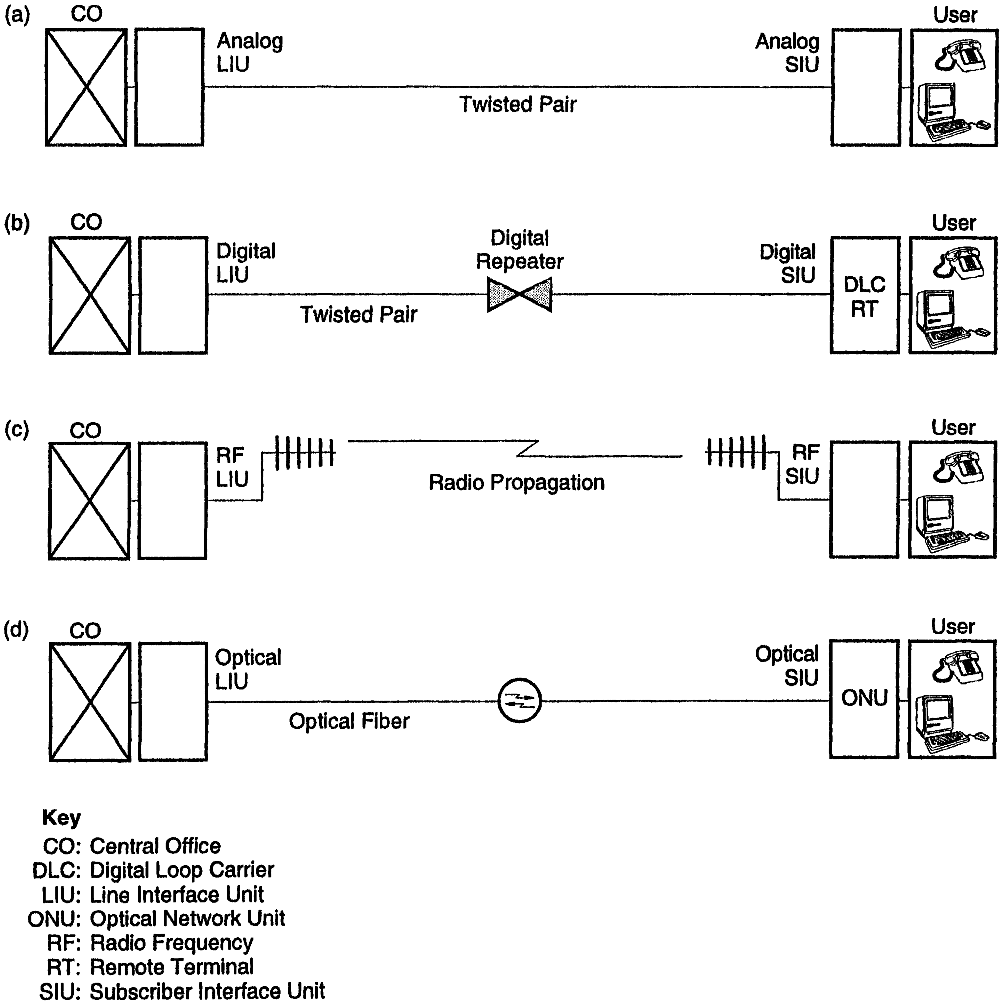


Figure 1-1 Subscriber loop configurations: (a) Analog twisted pair loop; (b) Digital loop carrier twisted pair loop; (c) Radio-derived loop; (d) Optical fiber loop

The most widely used and most mature technology used in loops is the metallic twisted pair. The newest and fastest growing segment of the loop plant uses fiber optic cables. A typical configuration is shown in Figure 1-1(d). The transmission rates presently used in optical fiber loop applications are nominally 1.5 Mbps, 6.3 Mbps, and 44.7 Mbps, although fiber-to-the-home (FTTH) and fiber-to-the-curb (FTTC) systems may use much higher rates and possibly wideband analog signals, as well (for example, television). By proper selection of high-speed line interface cards, many DLC systems may be connected directly to optical fibers. Some systems interface with the Synchronous Optical Network (SONET), which has tributary applications in the digital loop environment.

1.2 Analog and Digital Transmission

Subscriber loops carry two basic types of telecommunications traffic: voice and data. Analog voltages representing voice signals are produced by an electro-acoustical device such as a microphone in a telephone set. When such signals are transmitted on an analog telecommunications channel, only temporary frequency translations (for example, by frequency division multiplexing equipment) are made to them while they are in-transit. To be usable to the listener, voice signals must be retranslated to their original form. In a digital transmission environment, voice signals are first converted to digitally encoded signals and then transmitted over the loop as baseband digital voltage pulses. Various voice encoding methods are described later.

Data signals (for example, from a computer port or other data terminal equipment, or DTE) are inherently digital. Data signals can be transmitted over the loop in two ways: as baseband digital pulses or as modulated voice frequency tones. Analog modems convert the source digital signals to voice frequency tones. Different tone frequencies, amplitudes, or phases, or a combination, are used to indicate the binary value of each bit or group of bits.

When a single frequency, amplitude, or phase symbol is used to represent more than one bit (for example, two, three, or four bits), the modem uses multilevel modulation, giving a symbol (baud) rate that is lower than the bit rate. For example, four bits represented by one symbol is the same as four bits per baud encoding.

Virtually all present digital loop systems are bidirectional; that is, transmission and reception can be made simultaneously at each terminal. A twisted pair digital loop requires either two cable pairs (one transmit and one receive, also called 4-wire) or one cable pair (also called 2-wire), depending on the technology used. The familiar T1-carrier and subrate digital data systems, which were designed in the 1960s and 1970s, respectively, require 4-wire loops, while the integrated services digital network (ISDN) loop, designed in the 1980s, requires 2-wire loops. The high bit-rate digital subscriber line (HDSL), designed in the 1990s, requires 4-wire loops or, optionally for reduced bit rate, 2-wire loops. These systems are described in detail in later chapters.

1.3 Baseband and Passband Signals

A baseband signal is a signal that has frequency components approaching zero frequency, or *direct current* (dc). A baseband signal is not changed by frequency or phase translation; that is, it is does not consist of a modulated carrier.

A passband, or modulated, signal, on the other hand, is a carrier signal whose amplitude, frequency, or phase (or some combination of all three) is altered in response to a baseband signal. This causes the carrier or its sideband components to convey the information contained in the baseband signal. The modulation process translates the baseband signal frequency to a new frequency centered on the carrier but displaced from it. Figure 1-2 shows baseband and modulated signals.

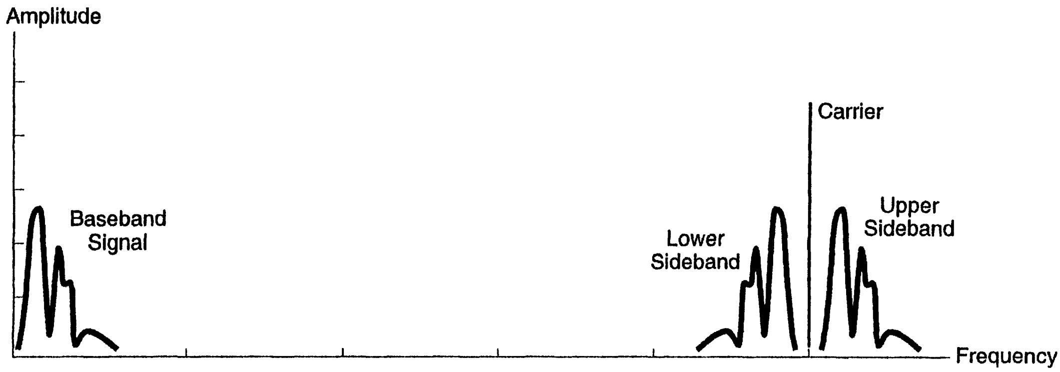


Figure 1-2 Baseband and modulated signals

1.4 Analog and Digital Modulation Methods

Analog carrier signals can be modulated by other analog signals (analog modulation) or by digital signals (digital modulation). Some analog modulation methods are AM (amplitude modulation), FM (frequency modulation), and PM (phase modulation). A commercial radio broadcast transmitter is an example of a device that uses an analog signal (voice or music) to modulate another analog signal (carrier) using either AM or FM.

A data set or modem is an example of a device that uses a digital signal (digital source data) to modulate an analog signal (carrier). When a digital signal is used in this application, the process is called keying. Modems typically use phase-shift keying (PSK), frequency-shift keying (FSK), or amplitude-phase keying (APK), or a combination, to convert the digital signals to analog. The foregoing analog and digital modulation methods are not the subject of this book; there is an abundance of literature on the subject. For example, see [3] for theoretical treatment and [4] for the application of modulation in modem design.

Modulation as used above is a specific form of signal encoding (frequency or phase translation) for the purposes of transmitting analog or digital information on an analog telecommunication channel. The modulation concept is not restricted to analog channels. Digital channels also can be used to transmit both analog and digital source information, and, for many applications, digital telecommunication channels are preferred.

The initial choice of a modulation method (analog or digital) is determined by the noise and bandwidth characteristics of the transmission channel to be used. The ultimate choice is one of economics. In some systems, particularly existing networks, this choice is a matter of convenience or policy. In many applications, digital transmission channels have significant advantages over analog channels. These are:*

*Adapted from [5] with additions.

- Ease of multiplexing
- Ease of signaling
- Use of modern technology
- Integration of transmission and switching
- Uniformity of transmission format
- Reduction of residual transmission loss variation of analog source signals
- Complete signal regeneration
- Operation at low signal-to-noise/interference ratios
- Better error performance
- Performance monitoring capability
- Accommodation of a number of diverse telecommunication services on a single circuit
- Ease of encryption

Although these advantages are significant, the disadvantages of digital transmission cannot be overlooked. Some disadvantages are:

- Imprecise analog-to-digital conversion of analog source signals (voice and modem) to digital format, which introduces nonlinear distortion
- Requirement for analog interfaces, particularly for voiceband services
- Increased bandwidth requirements
- Increased echo problems due to increased delay
- Need for precise synchronization of interconnected digital circuits

When analog source signals are transmitted on a digital channel, a *channel coding* technique called pulse code modulation (PCM) frequently is used. A theoretical development of PCM can be found in [5] and [6].

PCM refers to the use of a code to represent the instantaneous value of a sampled analog waveform. In other words, the analog waveform and the information contained in it are converted to a digital waveform carrying the same information. The conversion from analog to digital is called encoding; the conversion from digital to analog is called decoding. A device that performs these functions is called a codec (COder-DECoder). Modern digital transmission systems use different forms of PCM including delta modulation (DM), adaptive predictive coding (APC), differential PCM (DPCM), adaptive differential PCM (ADPCM), or linear predictive coding (LPC). These techniques differ in the specific method used to encode or decode the analog signal.

1.5 Digital Hierarchy

The digital hierarchy defines the different levels of digital multiplexing. The hierarchy used in North America, shown in Table 1-1, is based on the characteristics

Table 1-1 Digital Hierarchy

Digital Signal Level	Bit Rate	Equivalent 4 kHz Voice Channels[a]	Typical Transmission Media[b]
DS-0	64.00 kbps	1	TP
DS-1	1.544 Mbps	24	TP
DS-1C	3.152 Mbps	48	TP
DS-2	6.312 Mbps	96	FO, RD, CX
DS-3	44.736 Mbps	672	FO, RD, CX
DS-4	274.176 Mbps	4,032	FO, CX

[a]Using 64 kbps encoding.
[b]TP = twisted pair cable, FO = fiber optic cable, RD = radio, CX = coaxial cable.

and optimum information rates of the available transmission media at the time the hierarchy was conceived. The hierarchy segregates the various levels of information processing rates into well-defined blocks. The blocks at each level, shown conceptually in Figure 1-3, consist of:

- Analog-to-digital and digital-to-analog signal conversion terminals (channel banks)
- Multiplexers (source and subaggregate)
- Digital processing terminals
- Cross-connects
- Transmission facilities

At present, twisted pair digital loops carry traffic at the DS-1C, DS-1, and DS-0 rates or less (subrate). Many optical fiber transmission systems operate at the DS-2 rate, and many DS-3 rate systems are deployed in the network between central offices and user premises. The number of these higher-speed applications is expected to grow, and the DS-3 rate soon will be considered a common digital loop transmission rate. Not part of the North American digital hierarchy is the CEPT-1 (also called E1) rate of 2.048 Mbps, common in Europe. This rate has found some application in North America, particularly in private networks that cross international boundaries.

It is important to distinguish between correct terminology and colloquial usage. The terms "DS-1" and "T1" often are used interchangeably, but this is incorrect. The term DS-1 refers to a particular digital speed in the digital hierarchy, while the term T1, or T1-carrier, refers to a digital transmission system that happens to operate at the DS-1 rate. T1 is called "repeatered T1-carrier" in this book to clearly distinguish it from other digital loop transmission systems operating at the DS-1 rate.

Another hierarchy exists for SONET. Although SONET is beyond the scope of this book, a table of rates is given in Table 1-2 for reference and comparison. In SONET, the payload signals are carried in the synchronous payload envelope (SPE). Various quantities of DS-1 and other rate signals can be mapped as virtual tributaries (VTs) into the SPE. These are given in Table 1-3 for VT1 through VT6.

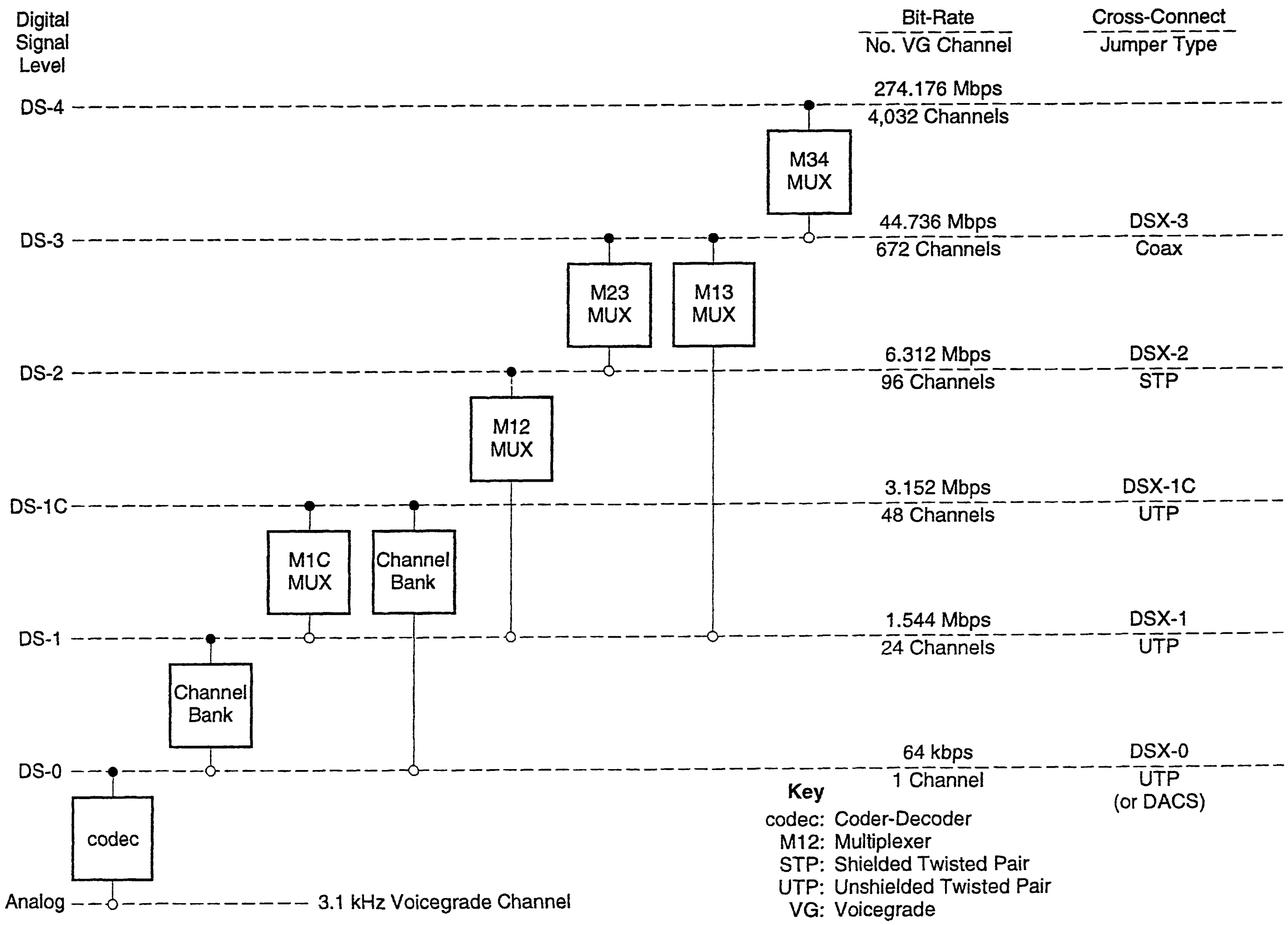


Figure 1-3 Digital hierarchy building blocks

Table 1-2 SONET Hierarchy

Optical Carrier Level	Optical Line Rate	Synchronous Transport Signal Level
OC-1	51.840 Mbps	STS-1
OC-3	155.520 Mbps	STS-3
OC-9	466.560 Mbps	STS-9
OC-12	622.080 Mbps	STS-12
OC-18	933.120 Mbps	STS-18
OC-24	1,244.160 Mbps	STS-24
OC-36	1,866.240 Mbps	STS-36
OC-48	2,488.320 Mbps	STS-48

Table 1-3 SONET Synchronous Payload Envelope

Virtual Tributary	VT Rate	Payload Rate	Designation	Number Mapped into SPE
VT1.5	1.728 Mbps	1.544 Mbps	DS-1	28
VT2	2.304 Mbps	2.048 Mbps	CEPT-1	21
VT3	3.456 Mbps	3.152 Mbps	DS-1C	14
VT6	6.912 Mbps	6.312 Mbps	DS-2	7

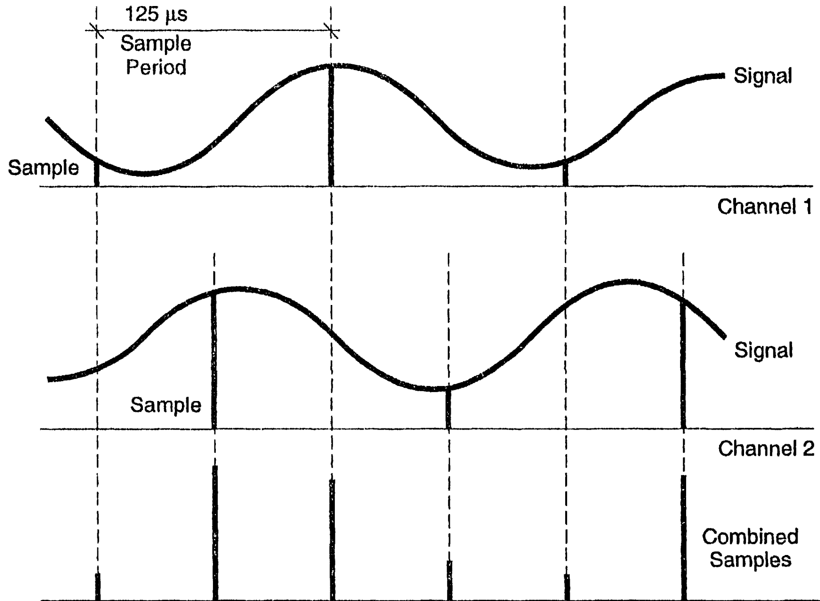
1.5.1 Multiplexing

Multiplexing is the process of combining multiple-user channels of information into a single, larger network channel. With analog multiplexing, several small bandwidth channels can be combined using frequency division multiplexing (FDM) into a larger bandwidth channel. Here, the bandwidth is measured in hertz (Hz). With digital multiplexing, the individual user, or source, channels are combined into a single, larger digital stream (aggregate) using time division multiplexing (TDM). TDM is shown conceptually in Figure 1-4.

In TDM the bandwidth is measured in terms of bit rate, or bits per second. The correlation between the bit rate and the actual transmission bandwidth in Hz depends on the digital coding scheme used. When a nominal 4 kHz analog channel is encoded at an 8 kHz rate into eight bits, the resulting digital bandwidth is 8 kHz $\times$ 8 bits = 64 kbps, or DS-0 rate. If transmitted as a baseband signal, the DS-0 signal has significant spectrum components beyond 64 kHz. Line coding and spectrum requirements are discussed in Chapter 3.

Multiplexers can be classified into two basic categories for the purposes of the following discussion:

- Multiplexers that have as inputs source information (for example, user data or voice). These will be called *source multiplexers* in the following discussion.
- Multiplexers that have as inputs the low aggregate rates from source multiplexers. These will be called *subaggregate multiplexers*.



0010	1110	1101	0011	0001	1101	Quantized Binary Values
00101110		11010011		00011101		Frames
001011101101001100011101						Combined Binary Bit Stream

Figure 1-4 Time division multiplexing concepts

With source multiplexers, the source signals can be analog or digital. Where the source signals are analog, an analog-to-digital (A/D) converter is required to convert the signals to the required digital format for transmission. A digital-to-analog (D/A) converter is required at the other end of the circuit to restore the signal to its analog form. A/D and D/A conversion is discussed later. Early source multiplexers were called channel banks, and source signals were exclusively analog. Modern source multiplexers have the capability of accepting both analog and digital source signals with the proper choice of plug-in channel units.

Subaggregate multiplexers are classified by a simple nomenclature, M*xy*, where *x* is the low-speed multiplex input level, typically from a source multiplexer, and *y* is the high-speed output level. For example, an M13 multiplexer will accept as inputs 28 DS-1 rate signals and as output, one DS-3 rate signal. An M23 multiplexer will accept as inputs seven DS-2 rate signals and as output, one DS-3 rate signal. Finally, an MX3 multiplexer will accept as inputs a combination of DS-1 and DS-2 rate signals and as output, one DS-3 signal. Other combinations are possible, limited only by one's (or a vendor's) imagination.

A typical channel bank will encode each of 24 voice channels at a 64 kbps rate (DS-0) and multiplex these into a single aggregate digital stream at a 1.544

Mbps rate (DS-1). The DS-1 rate includes the payload at 1.536 Mbps (24 × 64 kbps) plus framing overhead at 8 kbps. Early multiplexers had a filter and sampler for each channel. The pulse-amplitude-modulated (PAM) signals from each sampler were placed on a PAM bus and combined through TDM into a quantizer (codec), which was shared among all 24 channels.

Modern channel banks and intelligent multiplexers have a codec per channel, and the combining of all channels is done on a PCM bus. This arrangement allows more flexible signaling protocols and accommodation of different types of user-side digital channel units. Additional signal processing is required of the PCM signal at the output of the codec to condition it according to the specific requirements of the transmission medium to be used.

Voice encoding does not have to be at the DS-0 rate. Encoders operating at lower bit rates are common. A typical speed is 32 kbps. An encoder operating at 32 kbps will effectively double the number of voice channels on a DS-1 rate digital loop carrier from 24 to 48. Voice encoding is discussed later.

Several independent source data channels can be combined into a single digital stream, but the aggregate rate does not necessarily have to be the sum of the individual rates; the aggregate can be higher or lower. In many multiplexers, the aggregate rate is higher than the sum of the source data rates because of signaling and framing overhead. In many situations, there is some diversity in the time any given channel is active due to the "bursty" nature of digital data traffic. A statistical multiplexer may be used to fill the gaps in one channel with data from another channel to obtain more efficient use of the aggregate channel. Statistical multiplexing uses temporary storage registers, or buffers, which introduce delay.

The opposite of multiplexing is inverse multiplexing, which is used to pool relatively inexpensive, low-speed switched and dedicated digital circuits on the network-side to support high-speed aggregate signals on the user-side. Such a multiplexer spreads the higher-speed user's data stream across multiple lower-speed network circuits. For example, a 384 kbps signal on the source (user-) side of the inverse multiplexer can be allocated to two 56 kbps dedicated circuits and five 56 kbps switched digital circuits on the network-side. This is illustrated in Figure 1-5.

As traffic demand on the user-side of the inverse multiplexer increases, the multiplexer automatically dials up additional 56 kbps switched circuits until enough network capacity is connected end-to-end. This might occur when a video conference is scheduled, as shown in Figure 1-6. During low-traffic periods, the switched 56 kbps circuits are idle. This dynamic bandwidth allocation is sometimes called "rubber bandwidth" [8].

At the sending end, the inverse multiplexer segments the user data and sends these data over the individual dedicated and switched 56 kbps channels. At the receiving end, the inverse multiplexer accepts the data from the individual channels and reorders the segments. The inverse multiplexer buffers the bit streams at the receiving end to compensate for the slight variations in transit times across the individual network channels. Depending on its sophistication, the inverse multiplexer may monitor the integrity of the connections and take action to replace failed or failing channels with new dialed channels.

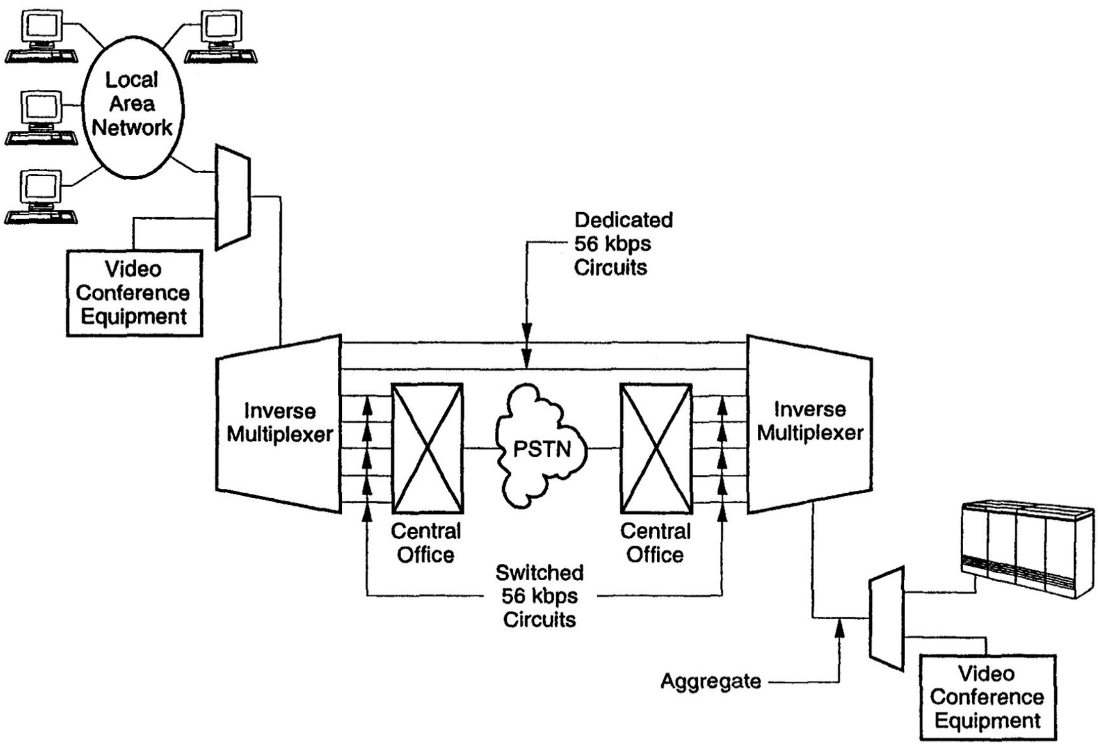


Figure 1-5 Inverse multiplexing

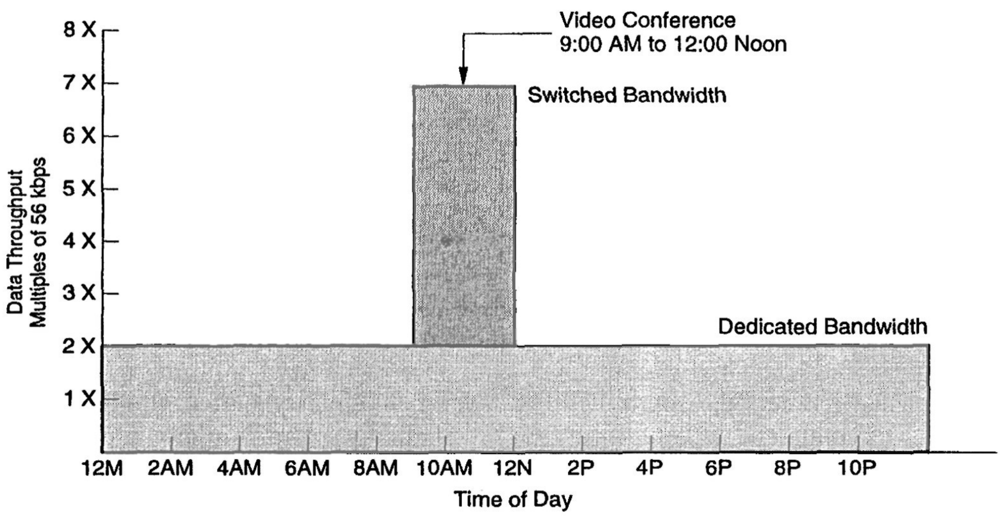


Figure 1-6 Dynamic bandwidth allocation

1.5.2 Pulse Stuffing

The input digital signals to subaggregate multiplexers are not always synchronized, nor need to be. Multiplexers use pulse stuffing (also called justification) techniques to synchronize several asynchronous or independent data streams. In this case, the aggregate digital stream has a higher rate than the sum of the individual channel rates to accommodate the extra pulses.

Extra pulses are inserted in the individual incoming streams until the channel rates are equal to that of a locally generated clock in the multiplexer, as shown in Figure 1-7. Stuffed pulses are inserted at fixed locations of each frame so they may be identified and removed at the far-end multiplexer. Pulse stuffing requires an elastic store or buffer of at least one frame in the multiplexer, although buffers with the capacity of two frames frequently are used. Pulse stuffing causes variations in the bit stream timing as bits are added. Pulse de-stuffing leaves gaps in the recovered bit stream, which must be smoothed out.

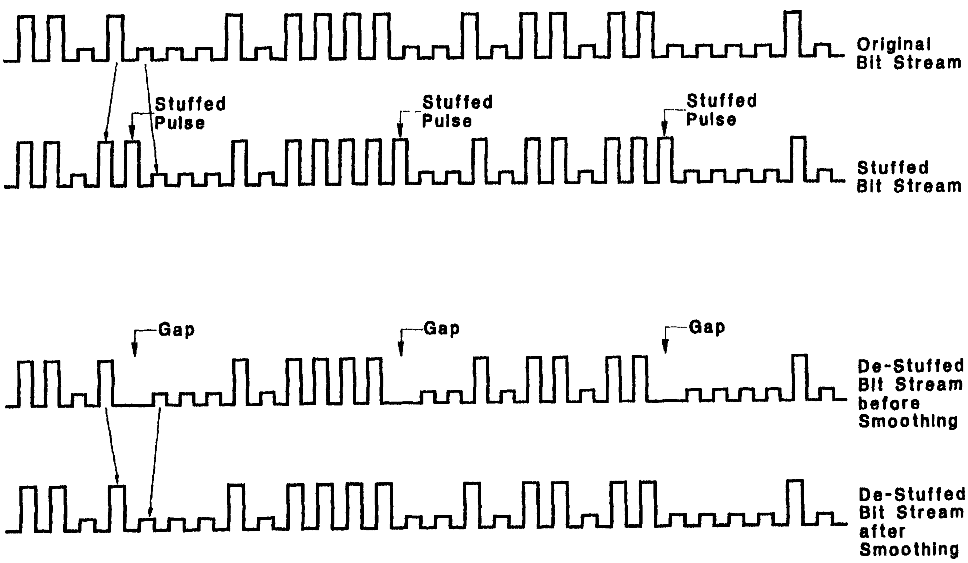


Figure 1-7 Pulse stuffing

The digital hierarchy is based on multiplexers with bit stuffing capabilities at the DS-2, DS-3, and DS-4 rates. For example, the DS-2 rate consists of four DS-1 signals with a total rate of 6.176 Mbps (4 × 1.544 Mbps). Bit stuffing is used to raise the rate of each DS-1 to 1,545,796 bps. In addition to the DS-1 signals, the DS-2 rate includes 128,816 bps of overhead for alignment and bit stuffing control. This brings the aggregate to 6.312 Mbps (4 × 1,545,796 + 128,816). The aggregate is an even multiple of the 8 kHz sampling rate. An M12 multiplexer, which multiplexes four DS-1 signals to one DS-2 signal, is illustrated in Figure 1-8(a).

Similarly, the aggregate rate for a DS-3 signal is 44.736 Mbps, which can consist of a multiplexed combination of seven DS-2 (from an M23 multiplexer) or

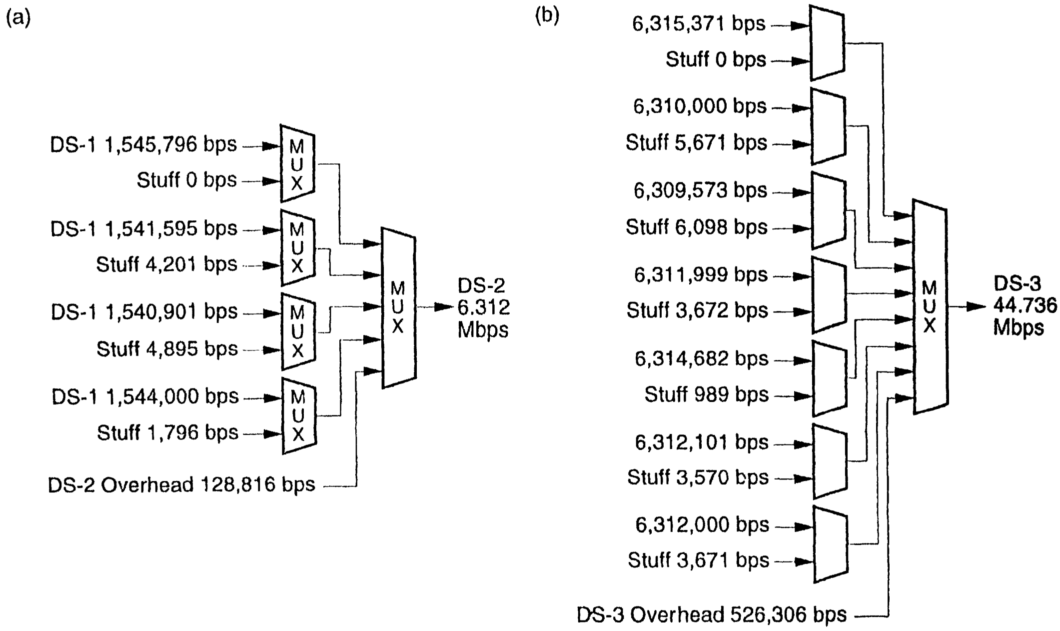


Figure 1-8 Multiplexer with bit stuffing: (a) M12 multiplexer; (b) M23 multiplexer

28 DS-1 (from an M13 multiplexer) signals. Multiplexers with a DS-3 aggregate usually are capable of mixing any appropriate combination of DS-1 and DS-2 signals (for example, five DS-2 and eight DS-1).

To arrive at the DS-3 rate, groups of four DS-1 signals are first multiplexed into a DS-2 signal as previously described. The seven DS-2 signals may be asynchronous because they are formed in different, unsynchronized multiplex terminals, so additional stuffing is required to bring them to the DS-3 rate. Bit stuffing is used to raise the rate of each DS-2 signal from nominal 6,312,000 bps to 6,315,671 bps. In addition to the stuffed DS-2 signals, the DS-3 includes 526,306 bps of overhead for alignment, error checking, multiplex control, and bit stuffing control. This brings the aggregate to 44.736 Mbps (7 × 6,315,671 + 526,306), which also is an even multiple of the 8 kHz sampling rate.* An M23 multiplexer is illustrated in Figure 1-8(b).

1.5.3 Scrambling

Scrambling is used on digital transmission systems to:

- Prevent the transmission of repetitive data patterns, which can cause discrete spectral components and crosstalk interference

*The totals do not add exactly due to rounding errors of fractional rates.

- Ensure sufficient timing energy is transmitted and available in the received signal even though the source bit string may contain a large number of consecutive binary zeros (no pulses)
- Ensure the data signals in the two directions of a 2-wire digital loop are uncorrelated and thereby do not bias the echo canceller adaptation
- Encrypt voice and data signals in secure applications

Completely randomized (scrambled) data provide an almost continuous spectrum and continuous spread of energy at the output of multiplexers. On the other hand, repetitive patterns, such as repeating binary-ones sequences or alternating binary 1-0 sequences, generate line spectra (or concentration of energy at a particular frequency), which can lead to crosstalk or interference problems.

Data scrambling ensures that relatively short repetition patterns are transformed into randomized signals. Scrambling does not guarantee that a long string of zeros will always be scrambled into anything other than another long string of zeros. Similarly, there is a finite probability that a random sequence of ones and zeros could be scrambled into a long string of zeros. Nevertheless, scrambling is used in systems that depend on a minimum-ones density for timing extraction. The ones density requirements of digital loop systems are discussed in Chapter 3.

1.6 Analog Signal Conversion and Channel Coding

1.6.1 A/D Conversion

Channel banks and intelligent multiplexers belong to a general class called primary multiplexers. Primary multiplexers convert analog signals (voice signals or tones from a modem) to digital and back again to analog using PCM. The conversion process is called A/D and D/A conversion.

Early channel banks performed little more than voice encoding and decoding, but later versions added digital interfaces. Some primary multiplexers have a large range of features, including direct digital interfaces, digital bandwidth allocation, and digital cross-connect capabilities, and are popularly called intelligent multiplexers.

PCM is a method of digitizing or quantizing an analog signal waveform. As in any A/D conversion, the quantization process produces an approximation of the signal sample. Such an approximation can introduce an error into the digital representation because of the finite number of bits available to represent the value. This error can be made arbitrarily small by representing the approximation by a large number of bits, thereby providing a high level of precision.

In practice, there is a tradeoff between the amount of error and the amount of data representing each signal sample. The goal of the A/D process is to quantize the signal sample in the smallest number of bits that result in a tolerable error. A linear quantization with 12 to 14 bits is the minimum required to produce an accurate digital representation of speech signals over their full range.

The number of bits required is reduced to eight by exploiting the nonlinear characteristics of human hearing. The ear is more sensitive to quantization noise in low-level signals than to noise in large signals. This nonlinearity can be matched by a logarithmic quantization function, which adjusts the data size in proportion to the input signal. Two quantization functions, or encoding laws, are used in telephony applications: the μ-law and A-law. The μ-law is used in North America and Japan, whereas the A-law is used in Europe.

A/D conversion using PCM, shown in the upper part of Figure 1-9, requires three discrete steps, in order:

1. Filtering
2. Sampling
3. Quantization

A device that provides the A/D conversion functions is called an encoder. The encoder filter is used to limit the bandwidth of the source analog signal. In voice telephony, sampling is done at 8,000 Hz to comply with the Nyquist sampling theorem, which requires sampling at twice the highest frequency of interest.* Ideal filters would allow capturing a full 4 kHz of bandwidth at the 8 kHz sampling rate.

Since ideal filters are not available, practical filters are built to start attenuating the upper frequency at approximately 3,400 Hz. This effectively provides a guardband. If proper filtering is not used, the frequency components higher than 4 kHz

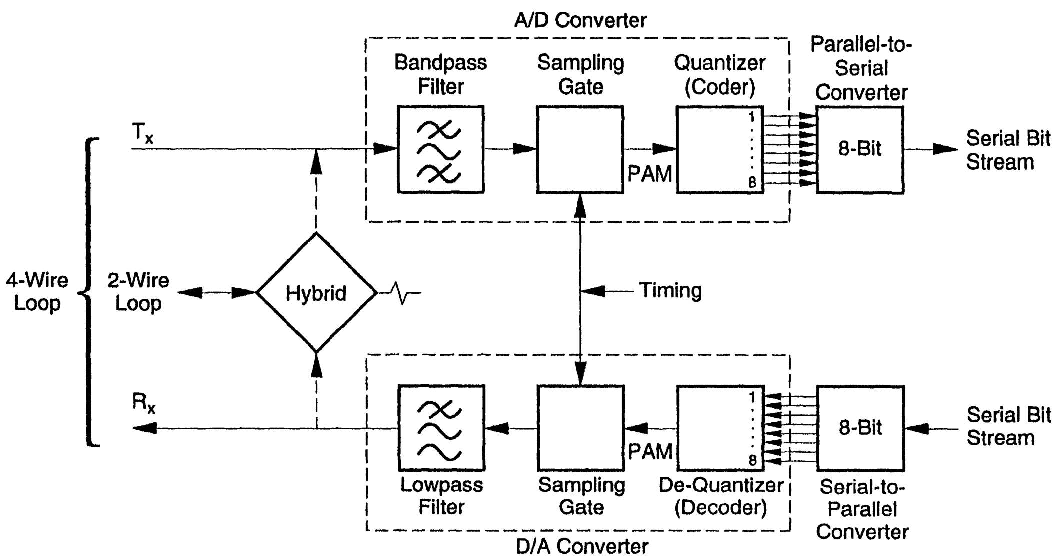


Figure 1-9 A/D converter

*The Nyquist sampling theorem is considered the basis for information theory. Further information can be found in [9].

will lead to aliasing, or foldover of the higher-frequency components into the lower components at the codec output.

Aliasing is illustrated in Figure 1-10. Assume the sampling rate is 8 kHz (8,000 samples per second). A 2 kHz input signal frequency is shown in line (a). The 8 kHz sampling rate provides four samples during each cycle of this waveform. The critical signal frequency of 4 kHz is shown in line (b). Such a waveform frequency is sampled, at most, twice during each cycle. Two samples are sufficient to recover accurately the original waveform at the decoder.

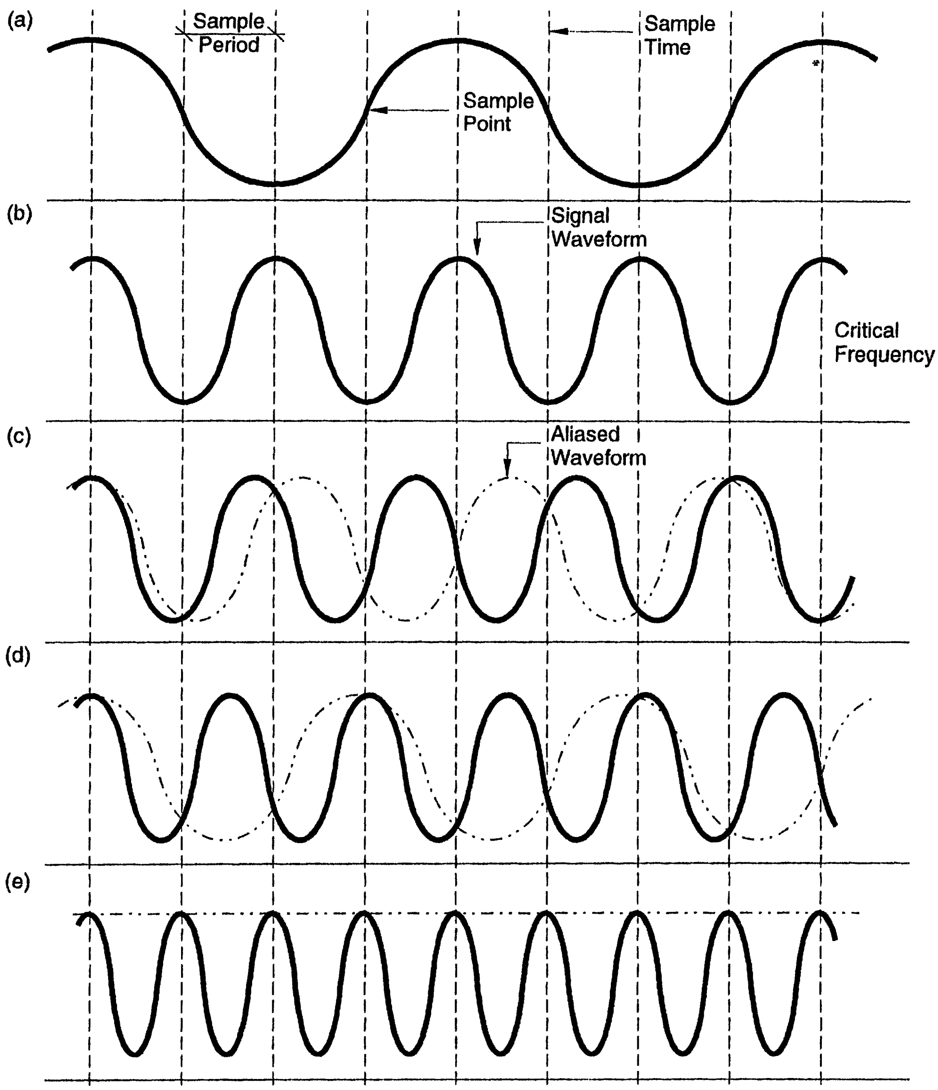


Figure 1-10 Aliasing

In line (c), the signal frequency, shown by a solid line, is raised to approximately 4.5 kHz. When decoded, the sampled points will produce not only the original 4.5 kHz signal waveform, but also the aliased waveform at (8 kHz − 4.5 kHz =) 3.5 kHz, which is shown by the dotted line. The input signal is further increased to approximately 5.2 kHz, shown by the solid line in line (d). The original waveform and (8 kHz − 5.2 kHz =) 2.8 kHz waveform will be produced in the decoder. In line (e), the input signal is 8 kHz (solid line), which results in zero frequency, or dc (dashed line) at the decoder output.

By properly filtering all frequency components greater than one-half the sampling frequency (in this case, greater than 4 kHz), the amplitude of aliased waveforms will be reduced to acceptably low levels.

In voicegrade applications, the source signal is limited to the range of 300 to 3,400 Hz or a bandwidth of 3,100 Hz (usually called a nominal 4 kHz channel). The lower frequency cutoff is used to reduce powerline hum. The performances of typical 2-wire and 4-wire voice channel filters are shown in Figure 1-11.

The periodic samplings are said to capture the time and amplitude information of the original signal. The output of the sampler is a string of pulses whose ampli-

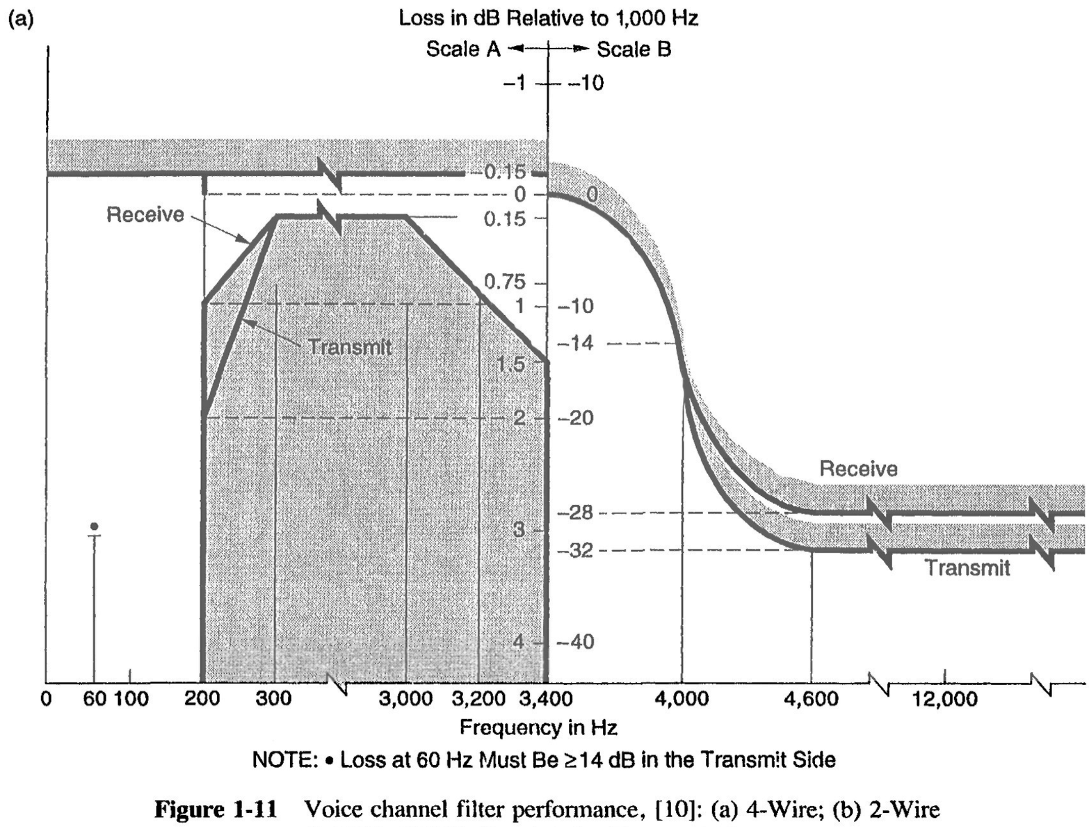


Figure 1-11 Voice channel filter performance, [10]: (a) 4-Wire; (b) 2-Wire
© 1982 AT&T. Reprinted with permission

(b)

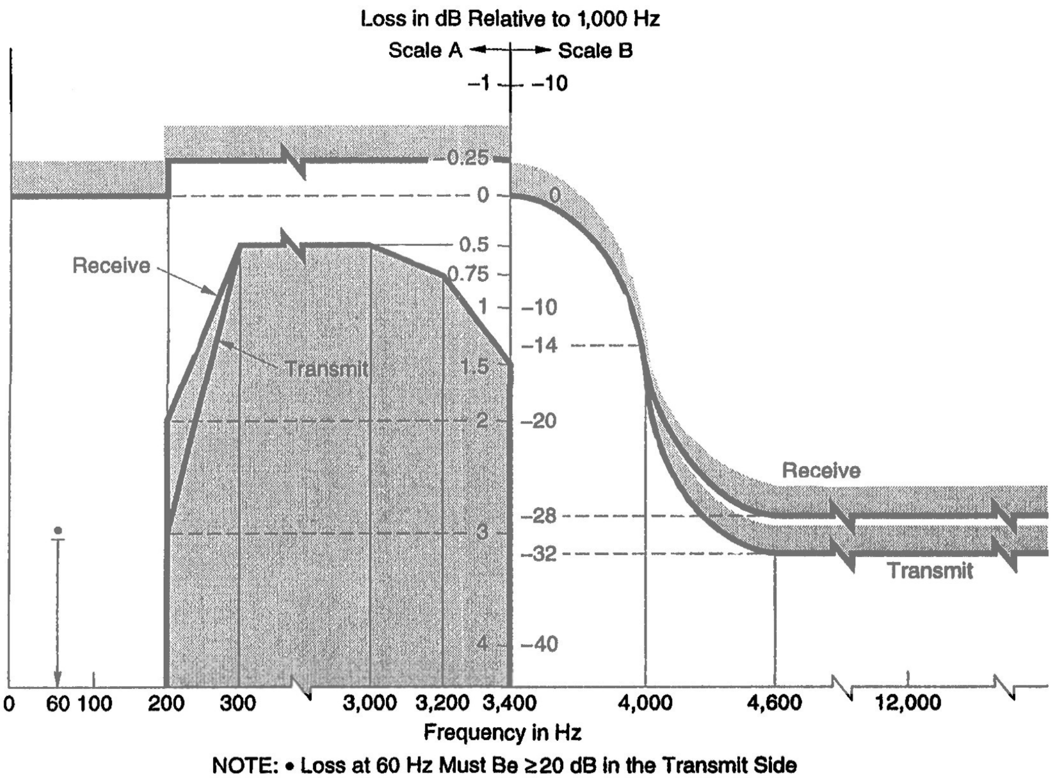


Figure 1-11 Continued

tudes correspond to the amplitude of the analog source signal at sampling time. This pulse string is called a PAM signal, as shown in Figure 1-12.

The output of the sampler is fed into a quantizer. The quantizer converts each PAM signal to an 8-bit binary code word. The code word is a digital representation of the amplitude information at sample time. Since an analog source signal is continuous, quantization naturally limits the resolution of the signal's value at any particular sample time, as shown in Figure 1-13. The decision values shown in this illustration represent the quantizing interval boundaries. Therefore, an input signal falling anywhere between two decision values will be encoded to the quantized value, which is halfway between the two corresponding decision values. The distance (or voltage) between each decision value is the same for a linear encoder but different for a nonlinear encoder, as will be seen.

The typical digital transmission system uses a code word length of eight bits, which limits the resolution to 2^8, or 256 individual binary values (±0 through 127). This resolution defines the error of the sampled signal. Such errors are manifested in the form of distortion (noise) when the digital signal is converted back to analog.

Signal-to-quantizing distortion ratio (SDR) defines the performance of an encoder. When the analog signal peak meets but does not exceed the coding range,

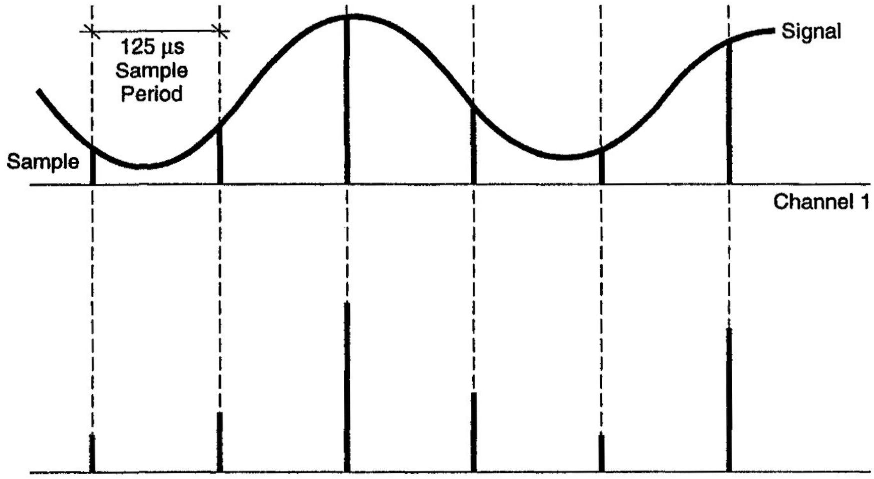


Figure 1-12 PAM signal

the encoder is said to be fully loaded. For a sine wave, the SDR is given as*

$$SDR_{dB} = 6n + 1.8 \text{ dB}$$

where SDR_{dB} = the signal-to-quantizing distortion ration in dB
n = the number of quantizing bits

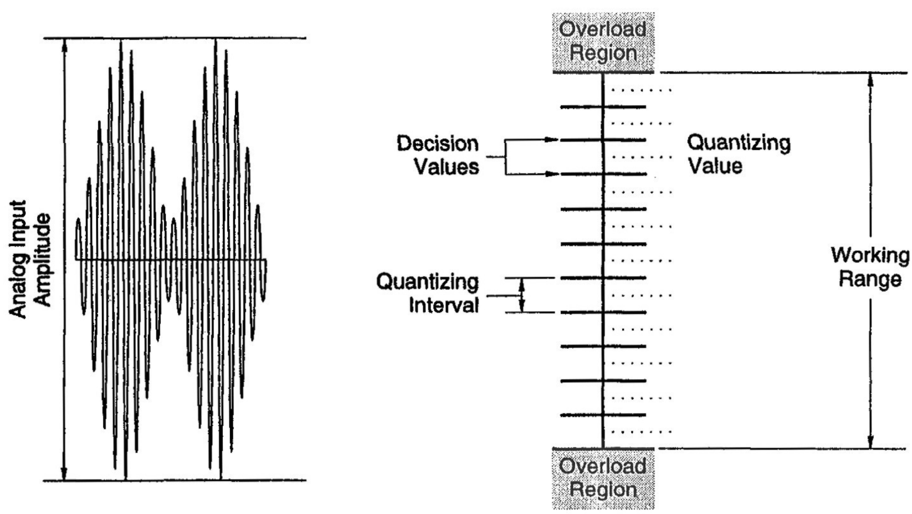


Figure 1-13 Quantizer resolution

*For a derivation of this equation, see [12].

A fully loaded 8-bit encoder will provide a best-case SDR of almost 50 dB, as seen from this expression. With a linear encoder, however, the SDR suffers considerably with lower-amplitude input signals because the noise power from linear encoding error is the same for all input amplitudes (the noise stays the same as the signal is reduced).

Different generations of channel banks have been called D1, D2, D3, D4, and (the most modern) D5. The D1 and D2 channel banks initially used a 7-bit word length for the encoded analog (voice) signal and one bit for signaling. The D3, D4, and D5 channel banks assigned all eight bits to the encoded voice signal in all but the sixth and 12th signaling frames (giving an effective 7-5/6 bits throughout the 12 frames).* The change from 7-bit word to 8-bit word provided about 5 dB improvement in SDR. If the full eight bits are used for voice encoding, a 6 dB improvement would be gained over a 7-bit encoder, as seen from the previous equation.

The typical range of voice level for any particular talker is about 25 dB. The range among a number of talkers is about 40 dB [13]. To cover the very large majority of talkers (99%), encoders are designed to have a 50 dB dynamic range. As previously mentioned, the average distortion noise power of a linear encoder is independent of input signal amplitude. Therefore, in voice applications, this encoder would give a 40 to 50 dB difference in the SDR for a range of talkers and talker volume levels (amplitudes).

EXAMPLE 1-1

Assume a fully loaded 8-bit linear encoder has a 50 dB SDR. Find the SDR for the lowest expected talker volume (level 40 dB down from fully loaded). For purposes of this problem, assume the talker power is equivalent to sine-wave power.

The full-load signal-to-quantizing distortion ratio, SDR_f, can be defined as 10 log (Psf/Pd) where Psf is the full-load signal power and Pd is the quantizing distortion power. If the signal power drops by 40 dB, it is the same as subtracting 40 dB from this expression. Quantizing distortion power remains the same. Therefore, the new SDR = SDR_f − 40 dB = 50 dB − 40 dB = 10 dB.

The amplitude distribution of voice signals is not uniform; that is, some amplitudes are more probable than others. Figure 1-14 shows the statistical distribution of instantaneous speech amplitudes relative to the nominal root mean square (rms) amplitude for a single talker. The graph shows that low amplitudes occur with a greater probability than high amplitudes. For example, the probability that the instantaneous amplitude will exceed the rms value (ratio of 1.0) is less than 15%, and more than 50% of the time the instantaneous amplitude will be less than one-fourth the rms value (ratio of 0.25).

It is possible to optimize the average SDR by using a nonlinear encoder that favors the weak signal amplitudes at the expense of the higher amplitudes. Such an encoder provides a higher SDR for more probable low amplitudes or, equivalently,

---

*In every sixth and 12th frame, only seven bits are used for source signal encoding; the eighth bit is used for supervisory signaling. See Chapter 3 for further discussion of channel bank types and framing.

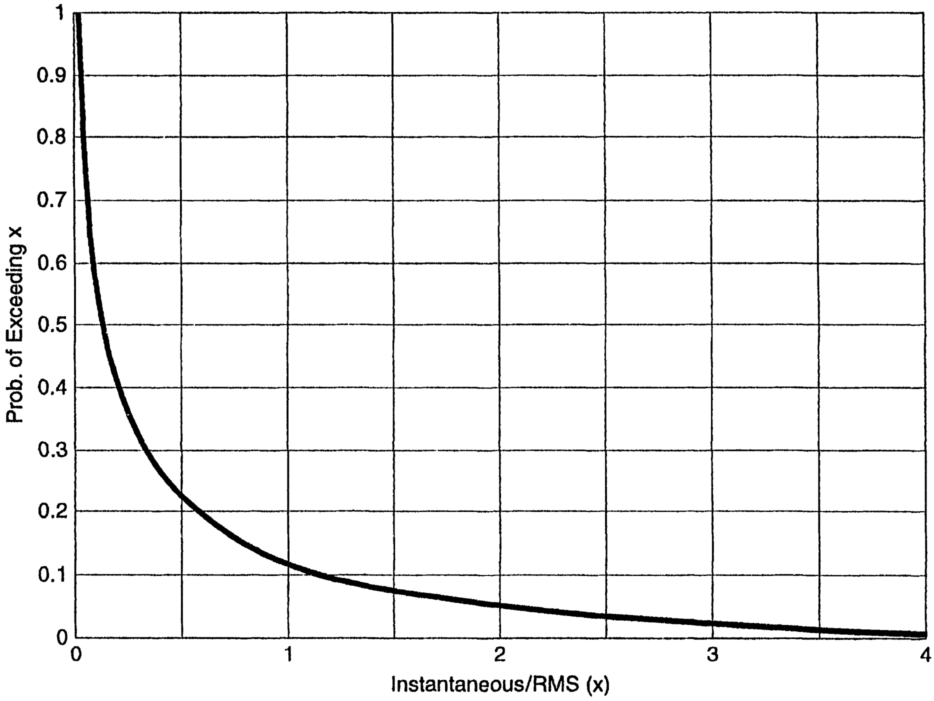


Figure 1-14 Speech amplitude distribution, adapted from [14]

lower distortion at low amplitudes. Similarly, this encoder provides a lower SDR for less probable high amplitudes or, equivalently, higher distortion at higher amplitudes. Therefore, more of the speech energy, which is at the lower amplitudes, is encoded using "fine-grained" quantizing to reduce the overall noise.

In effect, the nonlinear encoder *com*presses the signal and the associated decoder ex*pand*s the signal, so the process is called *companding*. There are a number of ways to implement the compression/expansion process, including mapping, predistortion, and digital processing [13]. The transfer and error characteristics of a linear encoder are shown in Figure 1-15(a). The characteristics of a nonlinear encoder typically used in telecommunications are shown in Figure 1-15(b).

The nonlinear encoders used in North America have a modified logarithmic compression characteristic, called μ-law, such that the normalized output as a function of the input is given by

$$Output_{\mu}(\mathrm{i}) = sign(\mathrm{i})\left[\frac{\ln(1 + \mu|\mathrm{i}|)}{\ln(1 + \mu)}\right]$$

where $\mu = 255$ and i is the input signal.*

*In early 7-bit encoders, $\mu = 100$ was used.

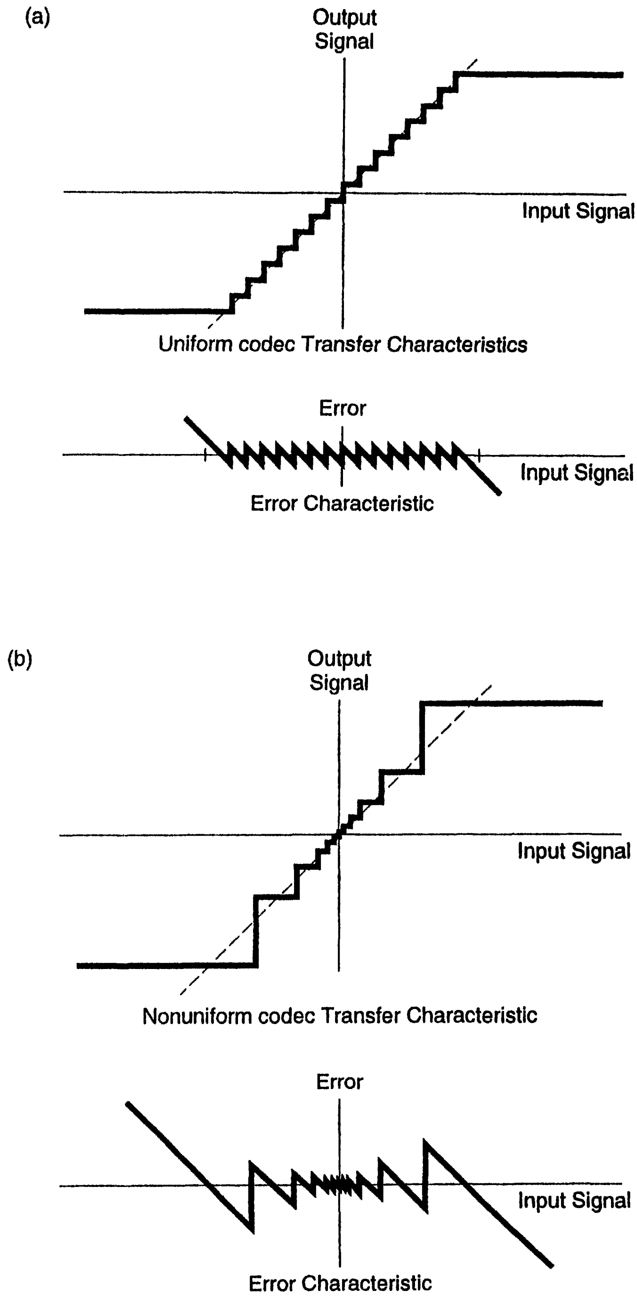


Figure 1-15 Encoder transfer characteristics, [13]: (a) Uniform encoder; (b) Nonuniform encoder
© 1982 AT&T. Reprinted with permission

With small signal levels, around −40 dBm0, this encoding function provides about 25 dB improvement in SDR over a linear encoder, as seen in Figure 1-16.* Referring back to the expression for SDR, this is equivalent to about four additional quantizing bits in the pulse code. Therefore, for an 8-bit code, the μ-law encoding provides an equivalent to a 12-bit code for low signal levels. At an input level of −10 dBm0, the SDR performance of the μ-law encoder is the same as a linear encoder. Figure 1-16 also shows the SDR for a nonsinusoidal signal that has a speech-like amplitude distribution.

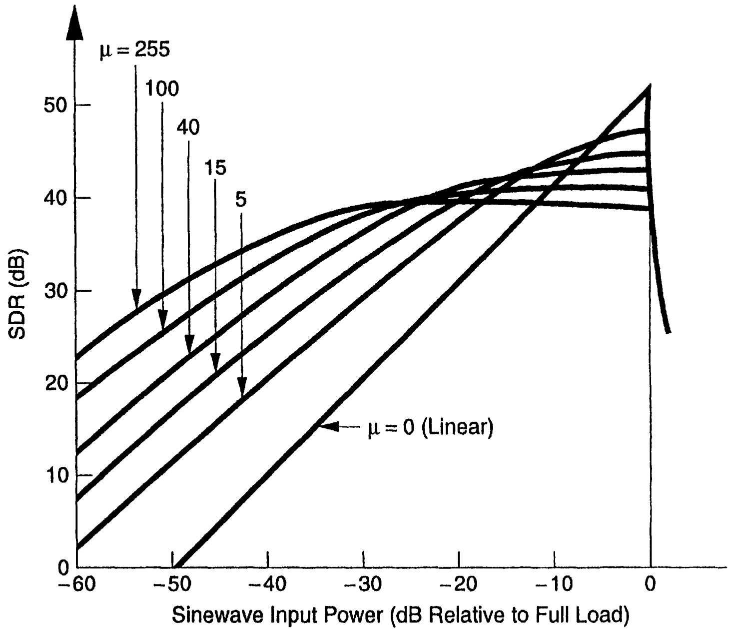


Figure 1-16 8-Bit μ-law codec SDR performance, [13]
© 1982 BELLCORE. Reprinted with permission

The maximum average signal input power to a μ-law encoder is approximately +3 dBm0 (3.17 dBm theoretical maximum). Signals with greater power are simply encoded at a level equal to +3 dBm and, in effect, are clipped at this level.

The idealized μ-law function defined above is approximated in real encoders by piecewise, linear segments. Both the ideal function and a 15-segment piecewise linear approximation are shown in Figure 1-17. The 15-segment μ-law approximation is used in North America. The length of successive segments increases by a factor of two, with eight segments of each polarity. The segments about zero are collinear, so these are counted as one. See [11] for the recommended characteristics of a μ-law encoder.

*The nomenclature dBm0 stands for decibels with respect to 1 milliwatt referred to the zero transmission level point (0 TLP). The 0 TLP normally is taken as the (theoretical) center of central office switching systems. See [1] for a more thorough discussion of TLP.

In early channel banks, the piecewise linear approximations were required because of technology limitations. Such approximations, however, are still used in modern equipment for compatibility and because it is possible to convert from a linear code to a compressed code or from a compressed code to a linear code without any additional impairments. This simplifies mixing ADPCM (or other types of PCM) signals on PCM transmission systems and converting PCM values derived from other encoding laws (for example, the A-law used in Europe) to the μ-law. The A-law is defined in [11]. Also, the linear code allows for precise digital pads and level control in digital loop terminal equipment.

Analog samples are encoded using an 8-bit folded (signed) binary code, as shown below. Table 1-4 gives the accompanying code set.

D1 D2 D3 D4 D5 D6 D7 D8

where D1 = sign (polarity) bit (binary zero for negative, binary one for positive)

(a)

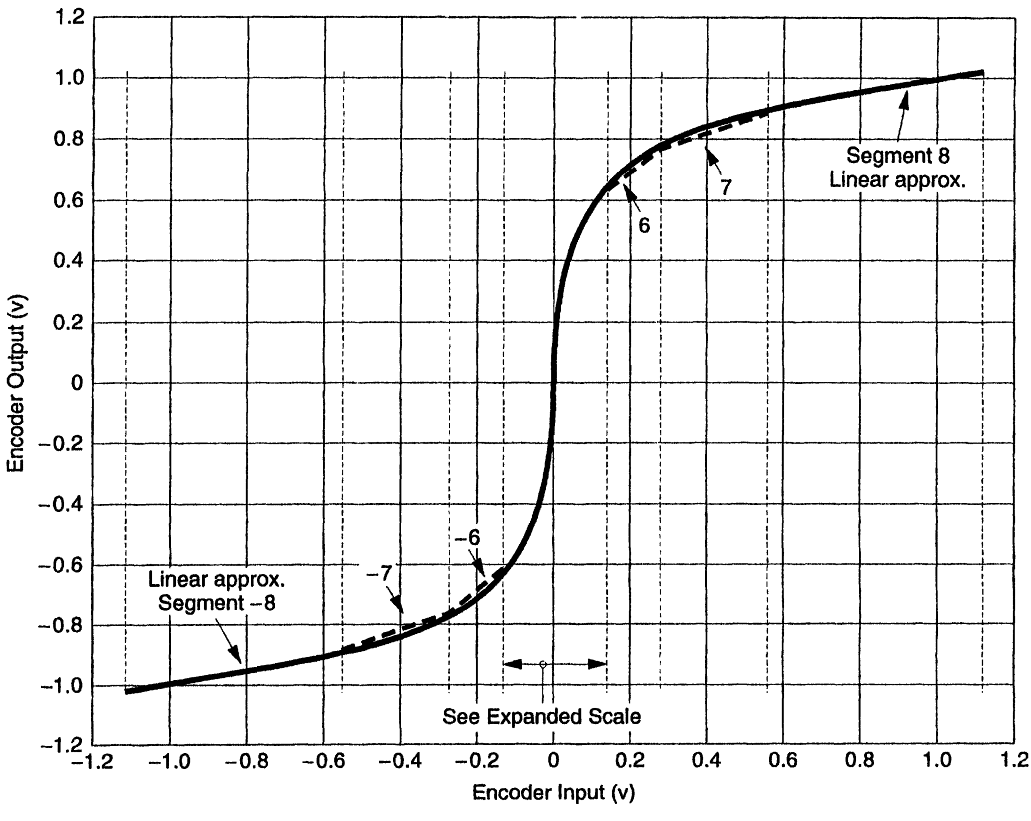


Figure 1-17 Idealized and 15-segment approximated compression curves: (a) Idealized; (b) 15-Segment

(b)

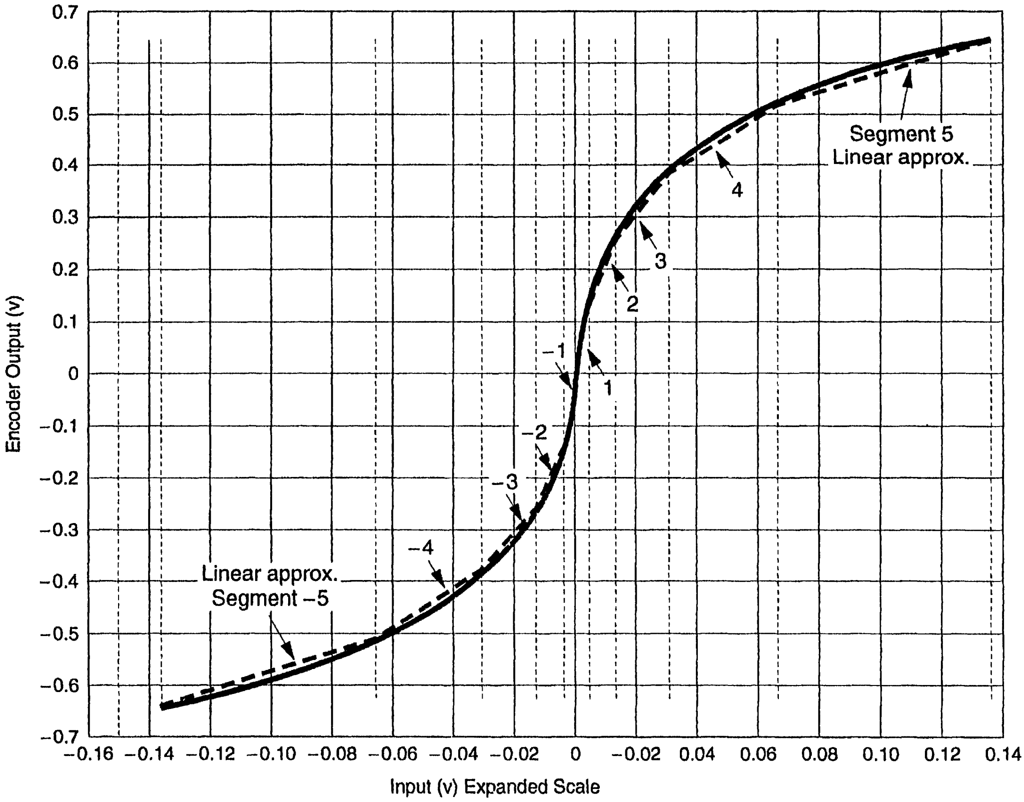


Figure 1-17 Continued

D2 D3 D4 = 3-bit binary code representing one segment of an 8-segment approximation to the magnitude (numbered segments 1 through 8), where D2 is the most significant bit of the segment code

D5 D6 D7 D8 = 4-bit binary code representing a 16-step approximation to the magnitude within the segment (numbered steps 0 through 15), where D5 is the most significant bit of the step code

As previously seen, the amplitude distribution of speech signals favors low signal levels. Columns (6) and (7) of Table 1-4 indicate that these low amplitudes are encoded with a high density of binary zeros (no pulses). Ordinarily, this would result in poor performance of repeatered T1-carrier span lines because the clocking performance of the repeaters requires a high density of binary ones (pulses). Therefore, prior to transmission, the encoded values of the segment and step codes are complemented as shown in column (8) to create a high density of binary ones. The sign bit is left unchanged. The decoder output shown in column (9) is the average of the normalized input values. This table is valid for both positive and negative values.

Table 1-4 Coding for $\mu = 255$, 15-Segment Companding

Segment (1)	Segment Endpoints (2)	Step Size (3)	Normalized Input Range (4)	Decision Value (5)	Segment Code D2 D3 D4 (6)	Step Code D5 D6 D7 D8 (7)	Transmitted Code Word (8)	Decoder Output (9)
	0		0–1	0		0 0 0 0	1 1 1 1 1 1 1	0
			1–3	1		0 0 0 1	1 1 1 1 1 1 0	2
1		2	.	.	0 0 0	.	.	.
			.	.		.	.	.
			27–29	14		1 1 1 0	1 1 1 0 0 0 1	28
	31		29–31	15		1 1 1 1	1 1 1 0 0 0 0	30
	31		31–35	16		0 0 0 0	1 1 0 1 1 1 1	33
			35–39	17		0 0 0 1	1 1 0 1 1 1 0	37
2		4	.	.	0 0 1	.	.	.
			.	.		.	.	.
			87–91	30		1 1 1 0	1 1 0 0 0 0 1	89
	95		91–95	31		1 1 1 1	1 1 0 0 0 0 0	93
	95		95–103	32		0 0 0 0	1 0 1 1 1 1 1	99
			103–111	33		0 0 0 1	1 0 1 1 1 1 0	107
3		8	.	.	0 1 0	.	.	.
			.	.		.	.	.
			207–215	46		1 1 1 0	1 0 1 0 0 0 1	211
	223		215–223	47		1 1 1 1	1 0 1 0 0 0 0	219
	223		223–239	48		0 0 0 0	1 0 0 1 1 1 1	231
			239–255	49		0 0 0 1	1 0 0 1 1 1 0	247
4		16	.	.	0 1 1	.	.	.
			.	.		.	.	.
			447–463	62		1 1 1 0	1 0 0 0 0 0 1	455
	479		463–479	63		1 1 1 1	1 0 0 0 0 0 0	471

Table 1-4 Continued

Segment (1)	Segment Endpoints (2)	Step Size (3)	Normalized Input Range (4)	Decision Value (5)	Segment Code D2 D3 D4 (6)	Step Code D5 D6 D7 D8 (7)	Transmitted Code Word (8)	Decoder Output (9)
	479		479–511	64		0 0 0 0	0 1 1 1 1 1 1 1	495
			511–543	65		0 0 0 1	0 1 1 1 1 1 1 0	527
5		32	.	.	1 0 0	.	.	.
			.	.		.	.	.
			927–959	78		1 1 1 0	0 1 1 0 0 0 0 1	943
	991		959–991	79		1 1 1 1	0 1 1 0 0 0 0 0	975
	991		991–1055	80		0 0 0 0	0 1 0 1 1 1 1 1	1023
			1055–1119	81		0 0 0 1	0 1 0 1 1 1 1 0	1087
6		64	.	.	1 0 1	.	.	.
			.	.		.	.	.
			1887–1951	94		1 1 1 0	0 1 0 0 0 0 0 1	1919
	2015		1951–2015	95		1 1 1 1	0 1 0 0 0 0 0 0	1983
	2015		2015–2143	96		0 0 0 0	0 0 1 1 1 1 1 1	2079
			2143–2271	97		0 0 0 1	0 0 1 1 1 1 1 0	2207
7		128	.	.	1 1 0	.	.	.
			.	.		.	.	.
			3807–3935	110		1 1 1 0	0 0 1 0 0 0 0 1	3871
	4063		3935–4063	111		1 1 1 1	0 0 1 0 0 0 0 1	3999
	4063		4063–4319	112		0 0 0 0	0 0 0 1 1 1 1 1	4191
			4319–4575	113		0 0 0 1	0 0 0 1 1 1 1 0	4447
8		256	.	.	1 1 1	.	.	.
			.	.		.	.	.
			7647–7903	126		1 1 1 0	0 0 0 0 0 0 0 1	7775
	8159		7903–8159	127		1 1 1 1	0 0 0 0 0 0 0 0	8031

Note: D1 = binary 1 for positive inputs and binary 0 for negative inputs

In Table 1-4, the segments are defined in column (1). By inspection of Table 1-4, segment 1 has 16 equal steps, as do the other segments; but the input voltage—represented by the normalized input range in column (4)—between each step is the smallest. The normalized range of ±8,159 is chosen so all magnitudes are represented by an integer. It is easy to visualize the purpose of the normalized units if each unit is considered to be some voltage, say 1 mV. For step 1, the step size is two normalized units and for step 8, 256 units, as shown in column (3). Therefore, an input signal falling in step 1 at sampling time is encoded with a resolution of two normalized units; similarly, an input signal falling in step 8 will be encoded with a considerably coarser resolution of 256 units.

The endpoints of each segment are shown in column (2). The decision value shown in column (5) defines the boundary between adjacent quantizing intervals. Column (6) gives the binary representation of the corresponding segment and column (7) gives the binary representation of the corresponding step (also called quantizing level). The transmitted code word in column (8) is the complement (or inverse) of the D2 through D8 bits. Finally, column (9) shows the output of the decoder at the far-end for each range of normalized inputs. For example, if the input at the encoder falls in the normalized range of 479 to 511, the decoder output always will be 495 normalized units.

A voltage level of 0 volts may be encoded as binary 1 0 0 0 0 0 0 0 or 0 0 0 0 0 0 0 0 and transmitted as 1 1 1 1 1 1 1 1 or 0 1 1 1 1 1 1 1 depending on the background noise at the sampling instant. The first bit is the polarity bit and is not shown in the table. The coarsest quantizing structure is at segment 8 of each polarity. A negative voltage peak at the maximum input power level will be encoded as binary 0 1 1 1 1 1 1 1 (negative polarity, segment 8, step 15) and transmitted as 0 0 0 0 0 0 0 0. If this occurs in an analog channel unit, zero code suppression is used, which changes the all-zeros word to binary 0 0 0 0 0 0 1 0 (negative polarity, segment 8, step 13). This modification to an all-zeros code is not done when a channel is used to transmit digital data in a clear channel format. An example will illustrate the use of Table 1-4.

EXAMPLE 1-2

A −10 dBm0 test tone level is inserted into the channel unit of a T1-carrier channel bank, which uses a $\mu = 255$, 15-segment encoder. Assume the maximum input power level of the encoder is +3.0 dBm0. Find the transmitted code word, output voltage, and voltage quantizing error at the negative peak of the waveform at the far-end decoder.

The maximum input power level of 3.0 dBm0 equates to an absolute power of 1.995 mW. Assuming a 600 Ω impedance, the maximum input voltage is 1.0941 Vrms. The peak value of a sinusoidal waveform is 1.414 times its rms value. For the maximum input voltage, the negative peak is at −1.0941 Vrms × 1.414 = −1.5471 V-peak. This corresponds to the normalized input value of 8,159 units, as shown in column (4). The absolute power of a −10 dBm0 test tone is 0.1 mW and the voltage is 0.2449 Vrms. The negative peak voltage is −0.3464 V-peak. The ratio of the two voltages is 0.3464/1.5471 = 0.2239. Therefore, the test tone is 22.39% of the maximum allowed input voltage and, in normalized units, is represented by 0.2239 × −8,159 units = −1,827 units. The sign bit is binary zero for negative polarity. In columns (2) and (4), the value falls between segment endpoints 991 and 2,015, which corresponds to segment 6 in column (1). The segment code for segment 6 is binary 1 0 1, as found in column (6). By simple extension of the values in columns (4) and (7), the step code corresponding to the input value 1,827 is found to be binary 1 1 0 1, which corresponds to an input range of 1,823 to 1,887. Therefore, the complete 8-bit code word is binary 0 1 0 1 1 1 0 1. The actual transmitted code word is the complement, or 0 0 1 0 0 0 1 0 (the sign bit is preserved). The decoder output is the center of the input range values (found by a simple arithmetic average), or $-(1{,}823 + 1{,}887)/2 = -1{,}855$. The decoder output value corresponds to an output voltage of $(1{,}855/8{,}159) \times -1.5471$ V-peak = 0.3517 V-peak. Therefore, the voltage quantizing error is $0.3517 - 0.3464 = 5.3$ mV.

EXAMPLE 1-3

Figure 1-18 shows approximately 1 ms of a 600 Hz sinusoidal input signal and a graphical representation of a 15-segment, $\mu = 255$ encoder. Three samples are taken in succession at a rate of 8,000 samples/second. Find the transmitted code word corresponding to each sample.

All three samples are negative, so bit D1 = binary zero for each. The first sample is in negative segment 8 (D2 D3 D4 = binary 1 1 1) and falls closest to step 12 (D5 D6 D7 D8 = 1 1 0 0). All bits, except the sign bit, are complemented, giving a transmitted code word of 0 0 0 0 0 0 1 1. The transmitted code word for the other samples can be found by a similar analysis. The three encoded samples are shown in Table 1-5.

Each voicegrade channel encoded at the DS-0 rate requires one of 24 available timeslots in a DS-1 rate system. Using a lower bit-rate encoding, two or more voicegrade channels can be inserted into one timeslot. Similarly, but in an opposite sense, channel units for wideband audio circuits, such as are required between a radio or television studio and a broadcast transmitter site, require more than one timeslot. The equivalent timeslots for various analog circuit types are shown in Table 1-6. The DS-0 encoding rate is the basic building block for any channelized system and not just systems operating with a DS-1 aggregate rate.

The final step in the A/D conversion process is the conversion of the parallel 8-bit words into a serial PCM bit stream by a parallel-to-serial converter. This converter can be as simple as a parallel load shift register in which the eight data bits are bulk loaded from the quantizer and then output one bit at a time. In repeatered T1-carrier systems and other channelized systems operating at the DS-1 rate, the individual bit streams from 24 separate timeslots are mixed to provide a 192-bit data stream. A framing bit is then added in each 24-channel group, as shown in Figure 1-19. An idle voice channel (that is, a channel not equipped with a channel unit) is encoded as all binary ones.

The following summarizes the steps required to encode 24 individual analog signals using PCM and standard (D4 channel bank) framing for transmission at the DS-1 rate:

1. The input signal is filtered to remove all frequency components above about 3.4 kHz to prevent aliasing and below approximately 200 Hz to remove powerline hum.
2. The input signal voltage is sampled at a rate of 8,000 samples/second, resulting in 8,000 pulse-amplitude-modulated voltage samples/second.

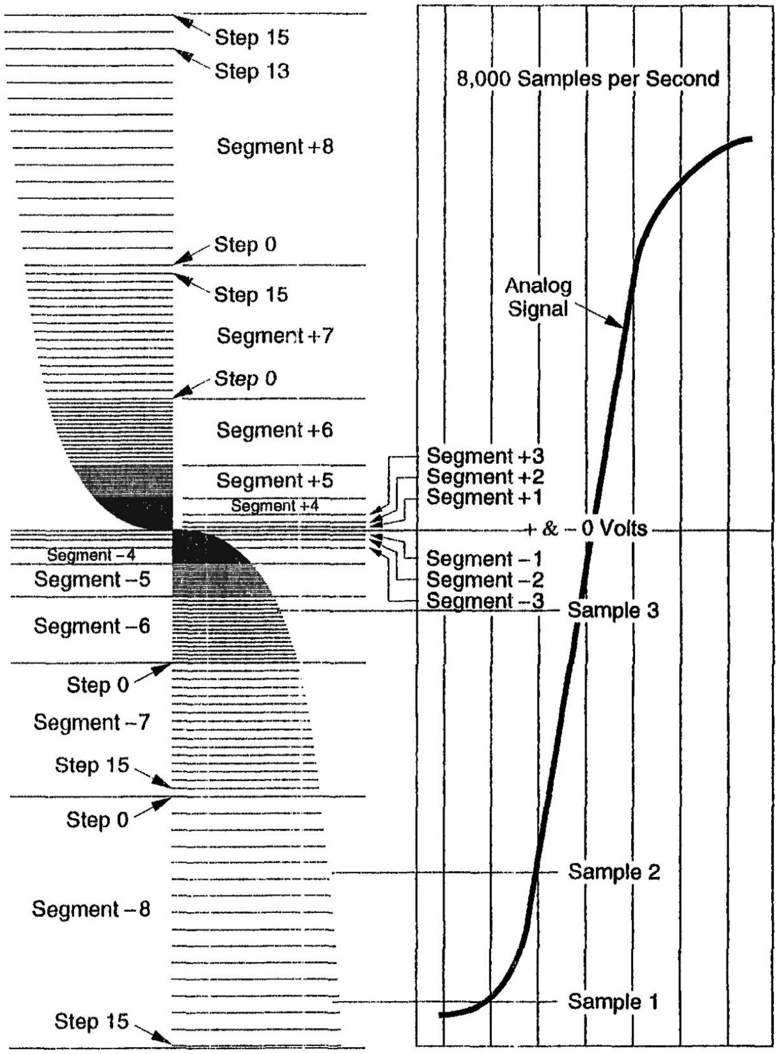


Figure 1-18 Illustration for Example 1-3

Table 1-5 Encoded Values for Example 1-3

Sample	Polarity Bit D1	Segment Bits D2 D3 D4	Step Bits D5 D6 D7 D8	Transmitted Code Word
Sample 1	0	1 1 1	1 1 0 0	0 0 0 0 0 0 1 1
Sample 2	0	1 1 1	0 1 0 1	0 0 0 0 1 0 1 0
Sample 3	0	1 0 1	0 1 0 0	0 0 1 0 1 0 1 1

Table 1-6 Voiceband and Audio Program Wideband Program Channel Characteristics

Circuit Bandwidth	Encoding	Equivalent Timeslots	Channel Sampling Rate	Bit Rate per Channel
3.1 kHz	ADPCM	$\frac{1}{2}$ per direction	8 kHz	32 kbps
3.1 kHz	PCM	1 per direction	8 kHz	64 kbps
5 kHz Studio	PCM	2 per direction	16 kHz	128 kbps
8 kHz Studio	PCM	4 per direction	32 kHz	256 kbps
15 kHz Studio	PCM	6 per direction	48 kHz	384 kbps

3. Each PAM voltage sample is compared to the discrete quantizing levels of the encoder and assigned to the nearest value.
4. The assigned value is encoded as an 8-bit signed binary word composed of a polarity bit, three segment bits, and four step bits.
5. The 8-bit word is combined with samples from the other 23 channels to form a 192-bit frame (24 channels × 8 bits/channel = 192 bits). These 192 bits represent the information payload in each frame.

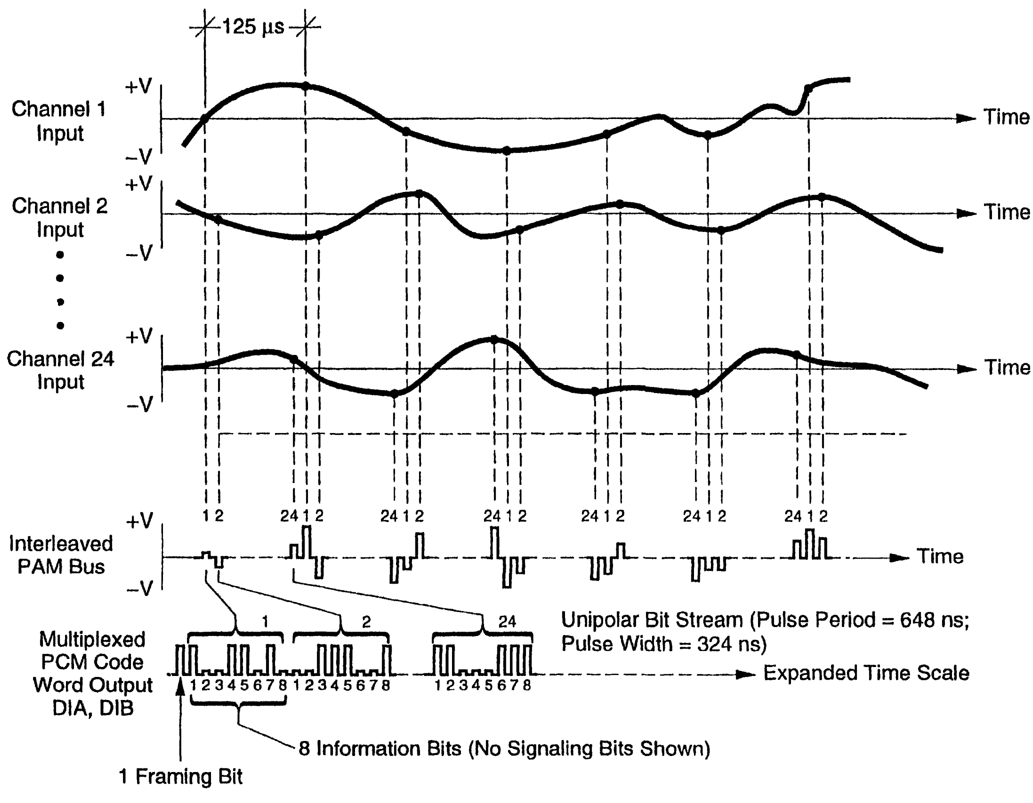


Figure 1-19 DS-1 rate system with 193-bit serial data frame

6. A single framing bit is added to the beginning of the 192 payload to complete the total frame of 193 bits. The frame width is 125 μs.
7. The 193-bit frame is transmitted at a rate of 1.544 Mbps (24 channels/frame $\times$ 8 bits/sample/channel $\times$ 8,000 samples/second + 1 framing bit/frame $\times$ 8,000 frames/second = 193 bits/frame $\times$ 8,000 frames/second = 1.544 Mbps).

1.6.2 D/A Conversion

D/A conversion, or decoding, of a signal is a mirror image of the encoding process. The serial bit stream is converted to parallel 8-bit words and then converted to a serial string of PAM signals of equivalent amplitude. The PAM signals are then filtered to extract the time and amplitude relationships of the original signal and to re-establish its continuity.

An incoming digital code word is converted to a PAM signal with an exact amplitude (within tolerance limits of the electronics) for each code word value. The limitations of the decoder are different than the encoder in that there are no additional quantizing errors introduced at the decoder. In the encoder, the signal was converted to a code word with a discrete step value, even though the signal fell anywhere within the range covered by that step.

1.7 Voice and Voiceband Encoding

1.7.1 Introduction

The largest proportion of *traffic* carried on telecommunication transmission systems is voice-type traffic.* On active circuits during the busiest hour of the transmission system, voice is present about 40% of the time; the rest of the time silence is encoded. As a result, it can be said that the largest proportion of *signals* comprise silence throughout any given day.

Digitally encoding a voice signal increases the bandwidth required for transmission. This is a clear disadvantage not to be overlooked during the design or choice of a transmission system. However, there are a number of advantages to digital encoding that usually far outweigh the bandwidth disadvantage, as previously discussed.

Voice encoding can be accomplished by three basic coder types:

- Waveform coders
- Vocoders
- Hybrid coders

*One inter-exchange carrier that carries traffic between the United States and the Far East reports that the majority of its traffic is facsimile [15].

With waveform coders, the actual voice waveform is encoded. A typical waveform encoder will use PCM, as previously discussed. The recovery of the signal at the output of the decoder provides an explicit approximation of the signal at the input to the encoder. With vocoders, a speech production model is parametrically summarized, and only the parameters are then digitized. Therefore, only a summary of the speech information is transmitted where it is recovered and converted to speech-like sounds at the other end. A hybrid coder uses a combination of a waveform coder and vocoder.

A typical vocoder uses LPC techniques. In the speech model, a time varying digital filter is used. The filter coefficients represent the vocal tract parameters. The filter is driven by a function, which, for voiced speech sounds, is a train of unit pulses at the fundamental, or pitch, frequency of the voice. For unvoiced sounds, the driving function is random noise with a flat spectrum. Voiced sounds are produced by vibration of the vocal chords and unvoiced sounds are produced by the lips and tongue without vocal chord vibration.

Figure 1-20 shows a block diagram of an LPC system. The speech input signal in a 15 to 20 ms window is encoded as three components: (a) a set of filter coefficients; (b) whether the signal is voiced or unvoiced; and (3) pitch period if voiced. These components, in the form of digital code words, are transmitted to the far-end through a digital channel. Several speech channels can be multiplexed into a single DS-0 rate digital channel.

The coding system predicts the value of the speech input signal based on the weighted average of the previous input samples. The number of input samples equals

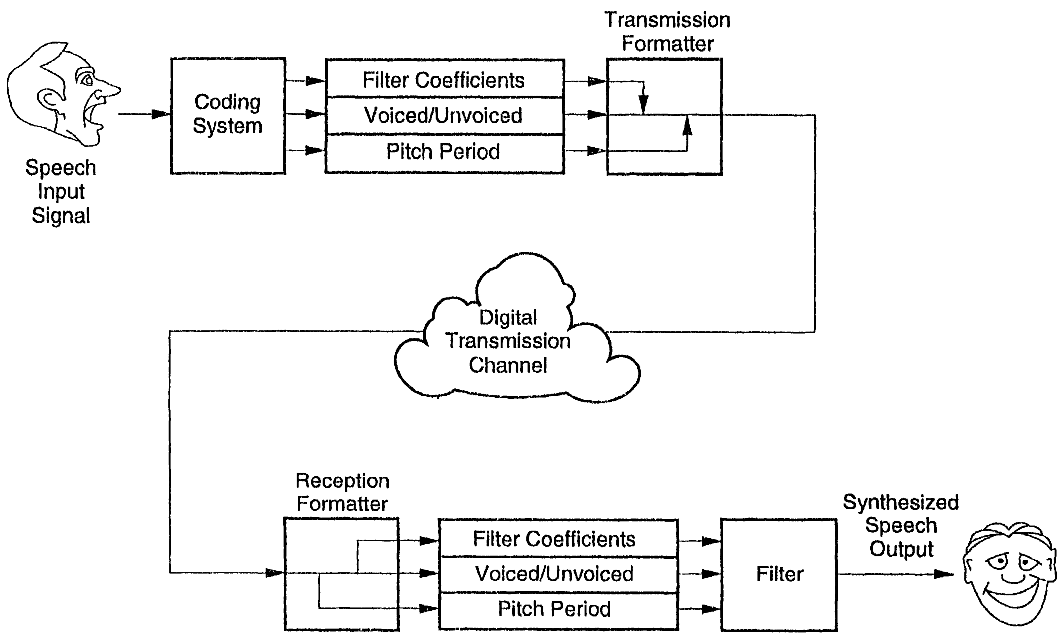


Figure 1-20 Linear predictive coding

the number of filter coefficients. The difference between actual speech input and the predicted value is the prediction error. The problem of linear prediction is to determine a set of filter coefficients in any window that minimizes the mean-squared prediction error.

At the receiving end, the LPC system synthesizes a speech signal by creating a digital filter with the received filter coefficients and driving the filter either with a train of unit impulses at the given pitch (for voiced sounds) or random noise (for unvoiced sounds).

The quality achievable by voice encoding largely depends on the type of encoding as well as the bit rate. As would be expected, the higher bit rates give higher quality for a given encoder type. Speech quality is generally related to digital transmission rate, as shown in Figure 1-21. This figure uses the terms *broadcast quality, toll quality* (or *network quality*), *communications quality,* and *synthetic quality.* Broadcast quality is what one expects from a commercial radio station. Toll quality and network quality are what the general public expects when it uses the public switched telephone network (PSTN). Toll quality can be quantified in terms of a 30 dB or better signal-to-noise ratio (SNR) and 3.1 kHz bandwidth.

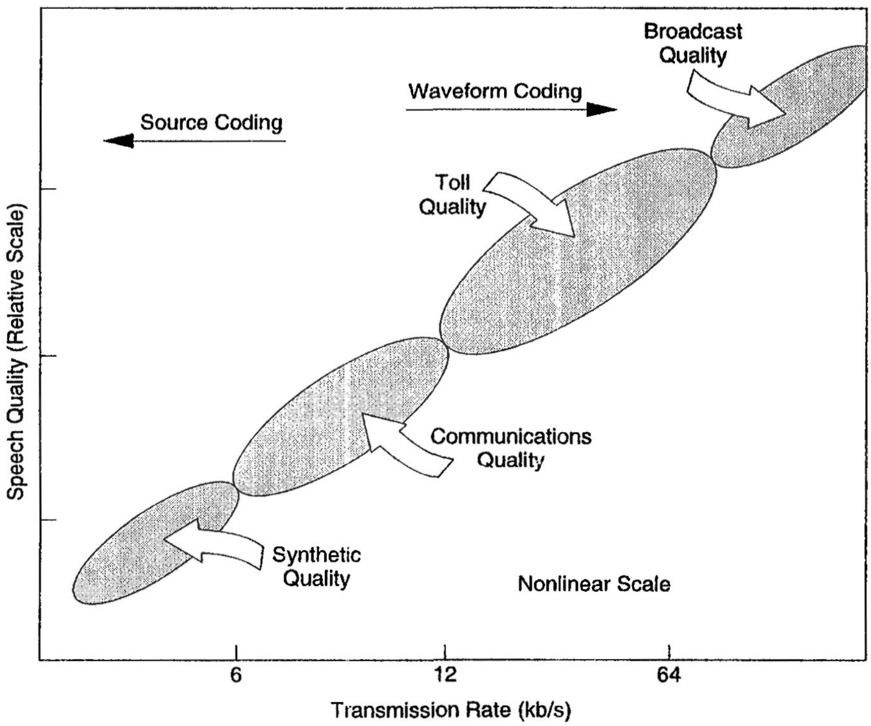


Figure 1-21 Speech quality versus digital transmission rate, [16] Kitawaki and Nagabuchi, © 1988 Institute of Electrical and Electronics Engineers. Reprinted with permission

The term "communications quality" is not as well quantified as toll quality, but it is what would be acceptable on radio communication links for pilots, and military and mobile radio users. A particular voice can be recognized on a communications quality link. Encoding rates for communication quality vary from about 5 kbps to 14 kbps. Synthetic quality, while intelligible, lacks natural voice sounds, and the person talking is not recognizable. The encoding rate for synthetic quality is below approximately 5 kbps.

The subjective aspect at the output of a voice decoder is often categorized by a mean opinion score (MOS), as shown in Table 1-7. Such a score would be given by tests using a large number of speakers, inputs, and listeners. The MOS is a perceptual value that relates the listener's expectations to the distortion in a voice signal.

PCM voice encoding at a 64 kbps rate has become somewhat of a standard against which other encoding methods are judged. A 64 kbps PCM encoder gives an MOS of 4.5 on a single encode/decode link, with an MOS of 4.0 after eight stages [17]. At present, the encoding rate for toll-quality voice encoders falls between 24 kbps and 64 kbps, although 16 kbps encoders have been developed with the required quality. Significantly lower encoding rates will provide the required quality as technologies and the economics of these technologies improve.

Toll-quality voice encoding presently uses waveform coders. Communications quality uses both waveform and hybrid encoders. Synthetic-quality voice encoding currently uses vocoders. Advances in technology and further understanding of the speech process will lead to the use of vocoders with communications, and then toll quality. The quality of the recovered signal at the output of the decoder depends on the SDR (previously discussed), which can be quantified by tests, as well as by subjective evaluations by listeners.

A quality measure used in network planning is the quantizing distortion unit (QDU). The 8-bit PCM codec pair is assigned one QDU to cover quantizing distortion. Under guidelines being developed by The International Telegraph and Telephone Consultative Committee (CCITT), a standards body,14 QDUs of impairment are allowed for an overall international connection [16]. Up to five QDUs are allowed for each national extension and four QDUs, for the international circuit. A 7-bit PCM codec is assigned three QDUs, and a 32 kbps ADPCM codec that meets CCITT recommendation G.721 is assigned 3.5 QDUs. The proposed QDU measurement method assumes a linear additive process, which is inaccurate in the region of larger QDU numbers [16].

Table 1-7 Mean Opinion Score

Mean Opinion Score (MOS)	Descriptive Value	Distortion
1	Bad	Very annoying and objectionable
2	Poor	Annoying but not objectionable
3	Good	Perceptible and slightly annoying
4	Fair	Just perceptible but not annoying
5	Excellent	Imperceptible

1.7.2 Compression

The encoding of any analog source signal is more efficient the lower the bit rate. A linear encoding scheme (for example, linear PCM at 64 kbps) gives a benchmark against which other encoding schemes can be measured. A compact representation of a source signal that results in a lower bit rate can be considered to be compressed. Compression is achieved by the removal of redundancy and reduction of irrelevancy [18].

All encoding methods strive to minimize the bit rate through some measure of compression. There are several dimensions to defining the performance of these encoders, including complexity, delay, signal quality, and bit rate. Ignoring complexity and delay, the performance of an encoder can be demonstrated by measuring the resulting signal quality for a given bit rate or by giving a specified quality at a lower bit rate.

As will be shown in the next section, there are a number of voice encoding methods. This also is true of encoding methods used on relatively wideband nonvoice signals, such as video, where compression is crucial to economical signal transmission. For example, the uncompressed bit rate for a single channel of high-definition television (HDTV) is over a gigabit per second. This single television signal is equivalent to over 20,000 uncompressed voice channels.

1.7.3 Voiceband Encoding Tradeoffs

Low bit-rate encoders suffer from disadvantages that restrict their use or require other special network design considerations. Some low bit-rate encoders perform badly (become very noisy or completely unusable) when the transmission system has relatively high error rates that otherwise would be acceptable with PCM or ADPCM. For example, a bit error rate (BER) of 1E-3 or 1E-2 will cause degraded transmission with PCM or ADPCM, but some low bit-rate encoders will not work at all.

Systems that use low bit-rate encoding are vulnerable to the rapid buildup of distortion through successive (tandem) encode/decode stages. Encoders with these characteristics are said to lack the robustness of 64 kbps PCM encoders. The low bit rates are achieved by optimizing the encoder for voice signals, which can lead to inadequate performance with nonvoice signals such as multifrequency signaling tones, modems, and facsimile signals.

Almost any encoding scheme has some delay associated with it due to the inherent characteristics of filters, samplers, quantizers, and other coding devices such as digital signal processors intermediate to and at each end of a digital path. The delay often is necessary to buffer enough speech to exploit its redundancy and allow predictive coding to take place accurately.

Delay and its effects on echo are closely related. Echo with delay greater than 5 ms can be annoying on voice calls and can cause talker and listener confusion. Delayed echo traditionally has been made less annoying by inserting loss in the receiver end of the connection. The more delay, the more inserted loss is necessary for echo at a given level to be tolerable. For example, a 10 ms delay requires approximately 10 dB of echo path loss on a typical connection for the connection to be considered reasonably good. A delay of 265 ms requires approximately 40 dB.

In the public analog network, the via net loss (VNL) method is used to determine the amount of required loss. The VNL method is described in [1]. In a lossless digital network, to which all telecommunication systems are evolving, echo is tightly controlled by echo cancellers and by minimizing the encoding delays.

Delay also can cause improper echo canceller operation, and excessively delayed signals between modems can lead to modem and DTE operation problems and reduced data throughput. Modern modem standards include echo cancellation methods (for example, CCITT V.32 and V.32bis; see [19]).

Echo can be split into two components: near echo caused by the near local loop, and far echo caused by the remote local loop. The possible echo span is in the order of tens of milliseconds for each loop, but the difference between the near and far echoes may be in the order of hundreds of milliseconds. To account for this, echo cancellers typically have two sections with a bulk delay processor between them. Low bit-rate voiceband encoders, which typically are located in the local loops, introduce additional delay and widen the near and far echo spans. Faulty operation will result if the echo canceller is not designed for the larger local echo spans (also called end-link delay).

Present low bit-rate encoders (< 16 kbps) have a one-way encode/decode delay of 20 to 40 ms, or more. CCITT has targeted a maximum delay of 5 ms (objective < 2 ms) for the next generation of 16 kbps toll-quality encoders [20].

There is a high amount of complexity associated with any low bit-rate encoding scheme. This complexity and its associated cost, in conjunction with the limitations mentioned above, lead to compromises in system design. Table 1-8 summarizes some of the tradeoffs and characteristics.

Table 1-8 Tradeoffs in Digital Voice Encoding[a]

Encoder	Type	Bit Rate	Quality	MOS	Complexity	Delay
PCM	Waveform	64 kbps	Toll	Good-excellent	0.0	0 ms
ADPCM	Waveform	32 kbps	Toll	Good	0.1	0 ms
LD-CELP	Hybrid	16 kbps	Toll	Good	0.1	2 ms
ASBC	Waveform	16 kbps	Comm.-Toll	Good	1.0	25 ms
MP-LPC	Hybrid	8 kbps	Comm.	Fair-good	10.0	35 ms
SE-LPC	Hybrid	4 kbps	Syn.-Comm.	Fair	100.0	35 ms
LPC	Vocoder	2 kbps	Syn.	Bad-poor	1.0	35 ms

[a]Source: [17]

1.7.4 ADPCM

As defined earlier, an extension of the PCM method is called ADPCM. ADPCM uses the predictable and redundant nature of speech to lower the overall encoding requirements. With this method, an algorithm is used to predict the source voice signal. The error between the actual signal and prediction is then coded. Full-amplitude, 8-bit encoding is not needed. In 32 kbps ADPCM, only four bits are used to encode the error. Detailed information on the 32 kbps ADPCM algorithms can be found in [11,21–23].

Typically, a PCM-to-ADPCM translator-encoder (transcoder) is inserted into a PCM system to increase its voice channel capacity. The ADPCM encoder accepts

64 kbps PCM values as input, and the ADPCM decoder outputs 64 kbps PCM values.

ADPCM works as follows: after an analog signal has been encoded according to 8-bit μ-law PCM requirements, it is converted to 14-bit uniform (linear) PCM by simple code word conversion. At this point the uniform PCM signal is subtracted from an estimate of the signal to give a difference signal. An adaptive 15-level quantizer assigns a 4-bit code word to the value of the difference signal. The 4-bit code word is transmitted to the far-end multiplexer at the standard 8 kHz sampling rate. An inverse quantizer at the far-end produces a quantized difference signal from the code word. A signal estimate is obtained from an adaptive predictor, which operates on both the reconstructed signal and the quantized difference signal, thus completing a feedback loop. The building blocks of an ADPCM encoder and decoder are shown in Figure 1-22.

The decoder uses a structure similar to the feedback portion of the encoder together with a uniform PCM to μ-law PCM converter. Both the encoder and decoder update their internal variables based only on the generated ADPCM value, which

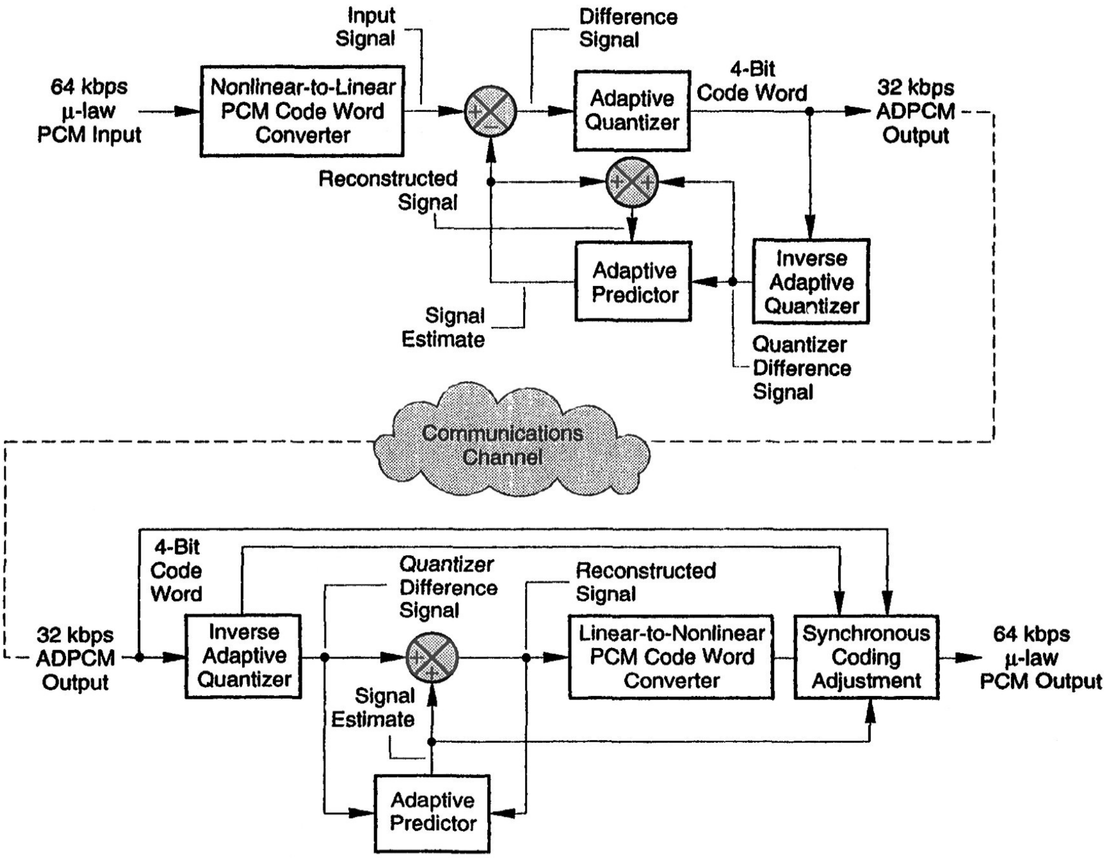


Figure 1-22 ADPCM encoder and decoder, [11]: (a) ADPCM encoder; (b) ADPCM decoder

ensures the encoder and decoder operate in synchrony without the need to send additional data. This is the function of the synchronous adjustment block. The synchronous coding adjustment reduces cumulative distortion on tandem codings (for example, ADPCM pair to PCM pair to ADPCM pair, which would be counted as three tandem connections). In order for this adjustment to work properly, there must be no intermediate digital processes such as digital pads, echo cancellers, digital speech interpolation, or μ-law to A-law converters [11].

A full ADPCM decoder is embedded within the encoder to ensure that all variables are updated based on the same data. In the receiving decoder as well as the decoder embedded in the transmitting encoder, the transmitted ADPCM value is used to update the inverse adaptive quantizer. This, in turn, produces a dequantized version of the difference signal. The dequantized value is added to the value generated by the adaptive predictor to produce the reconstructed speech sample at the decoder output.

The adaptive predictor computes a weighted average of the last six dequantized difference values and the last two predicted values. The ADPCM algorithm also includes a tone and transition detector. If a tone is detected (for example, modem or dual tone multifrequency (DTMF) tones), the predictor is changed to allow rapid adaptation and improved performance.

Since the encoder and decoder operate at an 8 kHz sampling rate and four bits are used, the overall encoding rate is 8 kHz × 4 bits = 32 kbps. Two such signals can be multiplexed into a single 64 kbps timeslot of a DS-1 rate channel bank or multiplexer, as shown in Figure 1-23. Various encoding levels can be used in the adaptive quantizers to achieve other encoding rates, such as 3-bit for 24 kbps and 5-bit for 40 kbps.

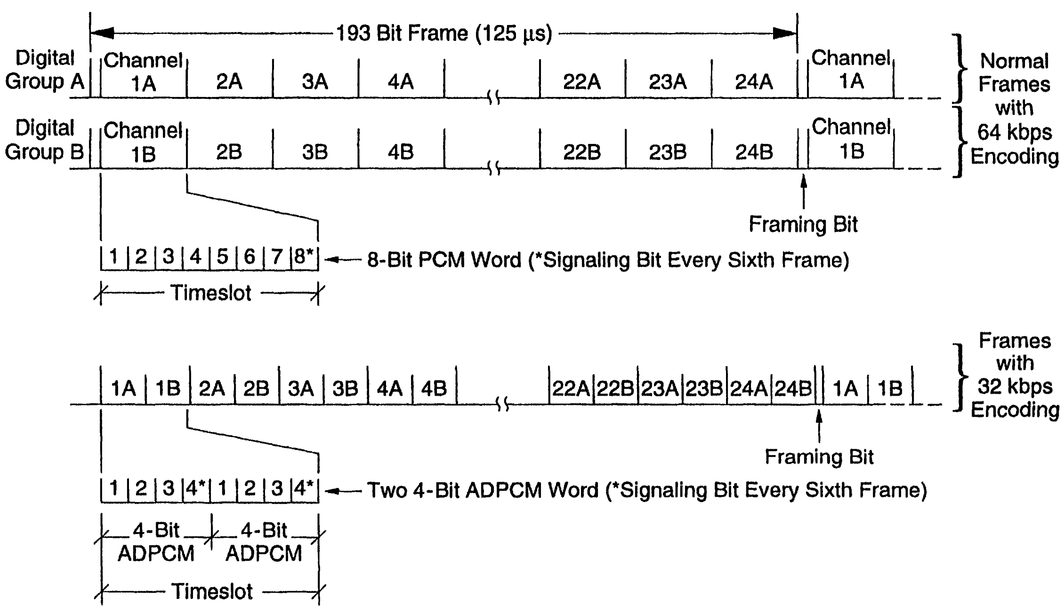


Figure 1-23 32 kbps voice channel multiplexing

The 32 kbps ADPCM was originally designed for voice applications. The voice quality of 32 kbps ADPCM is indistinguishable from 64 kbps PCM provided the system BER is better than 1E-4. The performance of 32 kbps ADPCM is better than PCM at higher BER, as shown in Figure 1-24. Both CCITT and the American National Standards Institute (ANSI) recommended that algorithms for 32 kbps ADPCM are robust enough to operate at 1E-3 BER [11,22]. However, nonvoice source signals and applications require special considerations, as will be discussed in Chapter 2.

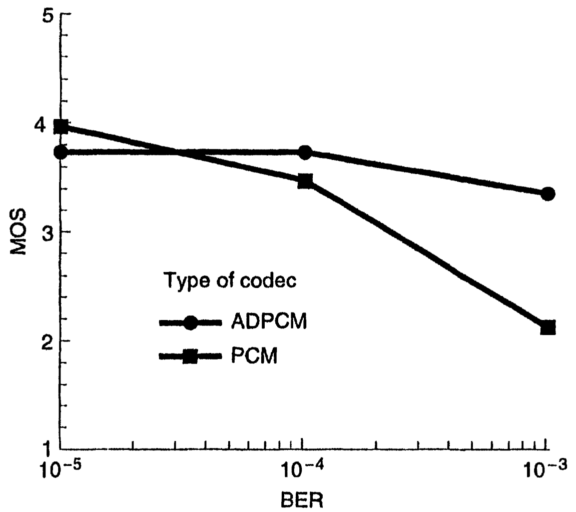


Figure 1-24 Performance of 32 kbps ADPCM and 64 kbps PCM, [24]
© 1992 Artech House. Reprinted with permission

1.7.5 Delta Modulation (DM)

DM is a special case of differential PCM. DM is used to encode the difference between the analog source signal and a prediction of it, but only one bit per sample is used to indicate the error. This bit specifies the sign (positive or negative) of the error and so indicates whether the source signal has increased or decreased since the last sample. As with all modulation schemes that encode a difference signal, DM is subject to slope overload, as shown in Figure 1-25. Its main advantage is in the reduction in encoding rate and subsequent more efficient use of a digital transmission facility.

A variety of delta coding methods are used, including adaptive delta modulation (ADM) and continuously variable slope delta modulation (CVSDM). ADM uses a variable step size, whereas CVSDM encodes the slope differences rather than amplitude differences.

DM systems are found in older multiplexers used with digital loop systems. The exact protocols and encoding algorithms usually are proprietary. This leads to compatibility problems when trying to mix equipment types and manufacturers. DM was used in a number of digital loop systems in the 1970s and 1980s but has given way to ADPCM in many systems. ADPCM has been standardized by ANSI and

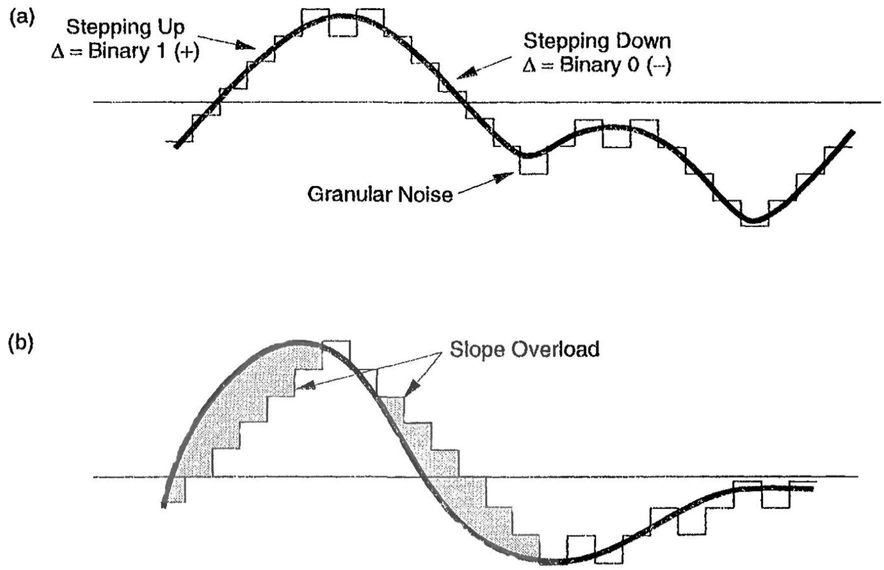


Figure 1-25 Delta Modulation, [5]: (a) DM encoding; (b) DM slope overload

CCITT. The chances of end-to-end equipment compatibility are greater with the ADPCM coding scheme, but specific implementations by different manufacturers still may lead to compatibility problems.

1.7.6 Digital Speech Interpolation (DSI)

The capacity of a digital transmission system that is used exclusively for voice can be doubled by taking advantage of the fact that either direction of transmission is used only about 40% of the time. That is, while one party is talking, the other is listening; also, normal speech contains many natural pauses [25]. A technique called time assignment speech interpolation (TASI) allows the speech power from more than one conversation to be interwoven. The technique was originally applied to intercontinental analog transmission systems using submarine cables. When it is applied to digitized voice, it is called digital speech interpolation (DSI). In this regard, DSI can be considered compression or multiplexing, and the equipment that performs this function is called digital compression multiplex equipment, or DCME.

With DSI, a channel in the transmission system is only connected when the talker is active. Early systems using DSI were not usable for analog data communications except at low speeds, if at all. Modern DCME includes circuitry to detect modem and facsimile tones, demodulate the data signals, and multiplex the data channels into one or more DS-0 rate signals. The far-end reverses the process to regenerate the original signals and deliver them unchanged. Also, the character of voice traffic after it has been compressed using DSI is no longer deterministic but, instead, is statistical and bursty. This has no direct effect on digital loop systems currently in use but can affect the delay and packet loss in asynchronous transfer mode (ATM) systems presently under development [26].

1.7.7 Variable Quantizing Level (VQL)

Variable quantizing level (VQL) is a technique used to encode voice at a 32 kbps rate [27]. It provides a measure of redundancy in the transmitted code word, which makes the technique relatively robust in the face of line errors. However, VQL introduces delay in the order of 6 ms at each encoding point. The added delay is detrimental to transmission quality when echo is present. Therefore, echo cancellers are required in combination with the VQL equipment and, in some implementations, are built-in. VQL algorithms and techniques are not standardized, which precludes mixing different manufacturers' equipment at opposite ends of a circuit. However, the general principles are similar and will be explained here.

In order to achieve a variable quantizing level, the nominal 300 to 3,400 Hz speech signal is first filtered to reduce its high-end response to 3,000 Hz. The speech signal is then sampled at 6,666.67 samples per second using regular PCM encoding techniques. Each block of 40 samples is processed every 6 ms, giving a block processing rate of 166.67 blocks per second. The amplitude of the largest signal in the 40-sample block is encoded into an 8-bit word and then divided into 11 positive and 11 negative equal levels or steps. Each of the 40 samples is then assigned a new 5-bit word, which corresponds to the closest level in the 11 steps (four bits) plus a sign bit.

Every 6 ms, corresponding to the block sampling rate, a code word is formed from a header word, which is formed from the maximum amplitude word (six bits), signaling bits (two bits) and forward error correction bits (four bits), plus the 5-bit words for each of the 40 samples. The resulting code word is 212 bits long, as follows:

$$\text{Maximum amplitude (6 bits)} + \text{signaling (2 bits)} + \text{forward error correction (4 bits)} \\ + 40 \text{ samples} \times 5 \text{ bits/sample} = 212 \text{ bits per block}$$

The forward error correction (FEC) bits protect the maximum amplitude word and two signaling bits, allowing the receiver to detect and correct any single bit error in the header. Since the block processing rate is 166.67 blocks/second and there are 212 bits per block, the overall rate is 35,333.3 bits/second. This exceeds the standard 32 kbps rate required to put two voice channels in a 64 kbps time slot, so further processing efficiency is required.

As previously noted, each sample can assume one of 22 values (five bits). Two such samples taken together can form one of $22 \times 22 = 484$ values. This is equivalent to nine bits, or 4.5 bits per sample (versus five bits for the single sample). This results in a shortened code word, as follows:

$$\text{Maximum amplitude (6 bits)} + \text{signaling (2 bits)} + \text{FEC (4 bits)} \\ + 40 \text{ samples} \times 4.5 \text{ bits/sample} = 192 \text{ bits per block}$$

Since the block rate is 166.67 blocks/second, the overall code word rate is 166.67 blocks/second × 192 bits/block = 32,000 kbps, which is the required rate.

The VQL encoding technique gives an almost constant signal-to-distortion ratio for varying signal levels. As a result, many voiceband modems operating at 9.6 kbps and below may work on a VQL circuit. However, certain modems that require a circuit frequency response up to 3,400 Hz can be adversely affected. In general, successful modem operation above 4.8 kbps is highly dependent on the modulation techniques used.

References

[1] Reeve, W. *Subscriber Loop Signaling and Transmission Handbook: Analog.* IEEE Press, 1992.

[2] *Public Mobile Service.* 47 CFR Part 22, subpart H. Federal Communications Commission, Oct. 1990. Available from USGPO.

[3] Taub, H., Schilling, D. *Principles of Communication Systems,* 2nd Ed. McGraw-Hill Book Co., 1986.

[4] Bingham, J. *The Theory and Practice of Modem Design.* John Wiley & Sons, 1988.

[5] Bellamy, J. *Digital Telephony,* 2nd Ed. John Wiley & Sons, 1991.

[6] Peebles, P. *Digital Communication Systems.* Prentice-Hall, 1987.

[7] Martin, J. *Future Developments in Telecommunications,* 2nd Ed. Prentice-Hall, 1977.

[8] Duncanson, J. "Inverse Multiplexing." IEEE Communications Magazine, Vol. 32, No. 4, April 1994.

[9] Bylanski, P., Ingram, D. *Digital Transmission Systems,* 2nd Ed. IEE Telecommunications Series 4, Peter Peregrinus Ltd., 1980.

[10] *Digital Channel Bank Requirements and Objectives.* AT&T Technical Reference TR43801, Nov. 1982.

[11] *General Aspects of Digital Transmission Systems: Terminal Equipments.* Recommendations G.700–G.772. CCITT Blue Book, Vol. III.4, 1989. Available from National Technical Information Service, Order No. PB89-143887.

[12] Smith, D. *Digital Transmission Systems.* Van Nostrand Reinhold, 1985.

[13] *Transmission Systems for Communications.* Bell Telephone Laboratories Technical Staff, 1982. Available from AT&T Customer Information Center.

[14] "The Transmission of PCM Over Cable." The Lenkurt Demodulator, Vol. 12, No. 1, Jan. 1963.

[15] Personal communication from Len Thomas of Sprint, May 13, 1992.

[16] Kitawaki, N., Nagabuchi, H. "Quality Assessment of Speech Coding and Speech Synthesis Systems." IEEE Communications Magazine, Oct. 1988.

[17] Bartee, T., Editor. *Digital Communications.* Howard W. Sams & Co., 1986.

[18] Jayant, N. "Signal Compression: Technology Targets and Research Direc-

tions." IEEE Journal on Selected Areas in Communications, Vol. 10, No. 5, June 1992.

[19] *Data Communications Over the Telephone Network.* Series V Recommendations, CCITT Blue Book, Vol. VIII.1, 1989. Available from National Technical Information Service.

[20] Chen, J., Atal, B., et al. "A Robust Low-Delay CELP Speech Coder at 16 KB/S." Advances in Speech Coding. Kluwer Academic Publishers, 1991.

[21] American National Standard for Telecommunications. *Digital Processing of Voice-Band Signals—Line Format for 32-kbit/s Adaptive Differential Pulse-Code Modulation (ADPCM).* ANSI T1.302-1989.

[22] American National Standard for Telecommunications. *Digital Processing of Voice-Band Signals—Algorithms for 24-, 32-, and 40-kbit/s Adaptive Differential Pulse-Code Modulation (ADPCM).* ANSI T1.303-1989.

[23] American National Standard for Telecommunications. *Network Performance—Tandem Encoding Limits for 32 kbit/s Adaptive Differential Pulse-Code Modulation (ADPCM).* ANSI T1.501-1988.

[24] Gruber, J., Williams, G. *Transmission Performance of Evolving Telecommunications Networks.* Artech House, 1992.

[25] Keiser, B., Strange, E. *Digital Telephony and Network Integration.* Van Nostrand Reinhold, 1985.

[26] Sriram, K., et al. "Voice Packetization and Compression in Broadband ATM Networks." IEEE Journal on Selected Areas in Communications, Vol. 9, No. 3, April 1991.

[27] Held, G. *Digital Networking and T-Carrier Multiplexing.* John Wiley & Sons, 1990.