Handbook of Computer Networks: Key Concepts, Data Transmission, and Digital and Optical Networks
by Hossein Bidgoli

Line Coding

Asim Loan, *University of Engineering and Technology, Lahore, Pakistan*

INTRODUCTION

Information is sent from one point in the network to another and can be converted to either a digital signal or an analog signal. This chapter deals with the shaping of digital data so that it is suitable for transmission over a channel. The digital messages can be stored in the memory of the computer and include data, text, numbers, graphical images, audio, and video. Passband systems have a relatively constant attenuation over the bandwidth because the carrier frequency is much larger compared to the signal bandwidth and the cross-talk coupling loss is also relatively frequency-independent. In baseband modulation, however, data are transmitted directly without resorting to frequency translation. A variety of waveforms have been studied, and the ones that have good power, spectral efficiency, and adequate timing information are used for converting the data, a sequence of bits, to a digital signal. These baseband modulation waveforms are called *line codes*, *baseband waveforms*, or *pulse-coded modulation* (PCM) codes. Line coding is an issue in baseband systems (compared to the passband systems) that operate over cable because the channel exhibits a large variation in attenuation over the bandwidth of interest and also has a large variation in cross-talk coupling loss. However, improvement in performance of these systems is possible by controlling the power spectrum of the transmitted signal.

In the next section, we give an overview of the fundamentals of optimum detection of binary (and multiple-level) signals in *additive white Gaussian noise* (AWGN) and briefly describe how to calculate the *power spectral density* (PSD) of random digital signals. For more details, the interested reader is referred to Simon, Hinedi, and Lindsey (1995).

The rest of the chapter is broadly divided into two categories of line codes: those that are suitable for transmission over (1) wireline links and (2) optical fiber. Line codes that are appropriate over wireline systems are discussed in the third section and can be broadly divided into four subclasses: *non-return to zero* (NRZ), *return-to-zero* (RZ), *pseudoternary* (PT) and *biphase* (see Figure 1). These line codes will be compared based on their spectral characteristics, bandwidth, error performance, error-detection

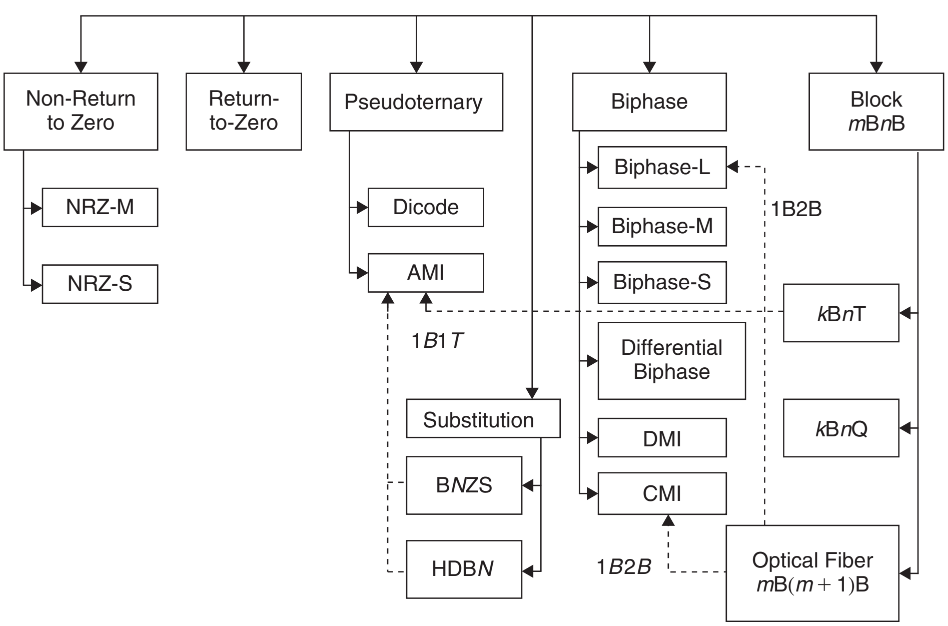

Figure 1: Line code tree

capability, self-synchronization ability, and bit sequence independence or transparency. Substitution codes overcome the shortcomings of the pseudoternary line codes—the most popular wireline line codes—by exchanging the consecutive zeros with various control signals that improve the timing information of the original codes. Block codes are not limited to binary and introduce redundancy through which the characteristics of the line code can be controlled. The nonbinary block codes (ternary and quaternary) are used in wireline systems where bandwidth is a premium. However, because the optical sources and detectors operate in the nonlinear mode, binary block codes are preferred in optical fiber transmission systems where abundant bandwidth is available. The block codes used for optical fiber are discussed in the last major section.

PROBABILITY OF BIT ERROR AND POWER SPECTRAL DENSITY

This section gives an overview of fundamentals of optimum detection of binary (and multilevel) signals in channels disturbed by AWGN and the basics of PSD calculations for digital random signals.

Probability of Bit Error

The probability of making a bit error after transmission through an AWGN channel is discussed in this section. The AWGN channel model implies that channel frequency response is flat and has infinite bandwidth. The only distortion introduced is by the additive white Gaussian noise. This is not an accurate model for practical channels except probably the satellite channel, but it is reasonably accurate as long as signal bandwidth is much narrower than that of the channel.

From the signal-detection viewpoint, binary line code (i.e., code that uses two signals to transmit information) presents one of two hypotheses:

$$\begin{array}{lll} H_1 : s_1(t) & 0 \le t \le T_b & \textit{apriori probability } p_1 \\ H_2 : s_2(t) & 0 \le t \le T_b & \textit{apriori probability } p_2 \end{array} \qquad (1)$$

Signal energies are $E_1 = \int_0^{T_b} s_1^2(t)dt$, $E_2 = \int_0^{T_b} s_2^2(t)dt$, and the correlation coefficient between the two signals is $\rho_{12} = \frac{1}{\sqrt{E_1E_2}} \int_0^{T_b} s_1(t)\, s_2(t)\, dt$ where $|\rho_{12}| \le 1$. When a matched filter receiver is used, the bit error probability, P_b, of equally likely binary signals, in AWGN with zero mean and variance $N_0/2$, is given by Simon, Hinedi, and Lindsey (1995):

$$P_b = Q\left(\sqrt{\frac{E_1 + E_2 - 2\rho_{12}\sqrt{E_1E_2}}{2N_0}}\right) \qquad (2)$$

where $Q(x)$ represents the area under the tail of a unit normal variate—that is, the normal random variable has a zero mean and unit variance. The bit error probability is minimum for antipodal signals—in other words, when $\rho_{12} = -1$ and 3 dB worse for equal energy binary orthogonal signals, or when $\rho_{12} = 0$.

For M-ary signals, the minimum error probability receiver computes (Simon, Hinedi, and Lindsey 1995)

$$l_j = \ln(p_j) - \frac{1}{N_0}\sum_{i=1}^{N}(r_i - s_{ij})^2 \quad j = 1,2,\cdots\cdots,M \qquad (3)$$

and chooses the largest l_j. For equally likely signals, the decision rule reduces to calculating the distance between the signals and selects the signal that is at a minimum distance. The received signal is given by

$$r(t) = s_j(t) + n(t) \quad j = 1,2,\cdots\cdots,M \qquad (4)$$

where $n(t)$ is AWGN with zero mean and variance $N_0/2$. Note also that r_i are statistically independent Gaussian random variables with variance $N_0/2$:

$$r_i = \int_0^{T_b} r(t)\, \Phi_i(t)\, dt \qquad i = 1,2,\cdots\cdots,N \qquad (5)$$

$\Phi_i(t)$ are N orthonormal basis functions. The projection of $s_j(t)$ onto $\Phi_i(t)$ is s_{ij}:

$$s_{ij} = \int_0^{T_b} s_j(t)\, \Phi_i(t)\, dt \qquad \begin{array}{l} i = 1,2,\cdots,N \\ j = 1,2,\cdots,M \end{array} \qquad (6)$$

Power Spectral Density

A general formula for calculating the power spectrum of digitally modulated baseband waveforms is presented below. It can be used for most of the binary line codes. However, for other codes, different methods to calculate PSD will have to be employed.

A digital random signal $s(t)$ can be represented by

$$s(t) = \sum_k a_k g(t - kT_S) \qquad (7)$$

where $\{a_k\}$ represents the binary or multilevel random data sequence and $g(t)$ is a filter that shapes the spectrum. For binary signaling, $T_s = T_b$ where T_b is the time it takes to send one bit and T_s is the symbol duration. The power spectral density of $s(t)$ is (Simon, Hinedi, and Lindsey 1995)

$$S(f) = \frac{|G(f)|^2}{T_S}\sum_n R[n]\exp(-j\omega nT_S) \qquad (8)$$

where $R[n] = E(a_k a_{k+n})$ is the autocorrelation function of the data sequence and, in general, $E(XY) \stackrel{\Delta}{=} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{XY}(x,y)dxdy$. Thus, the PSD of a line code is determined by its (1) pulse shape and (2) the statistical properties of the data sequence.

For uncorrelated $\{a_k\}$:

$$R[n] = \begin{cases} \sigma_a^2 + \eta_a^2 & n = 0 \\ n_a^2 & n \neq 0 \end{cases} \tag{9}$$

where σ_a^2 is the variance of $\{a_k\}$ and η_a^2 is the square of the mean of $\{a_k\}$. Using the Poisson sum formula for binary signals, we have (Simon, Hinedi, and Lindsey 1995):

$$S(f) = \frac{|G(f)|^2}{T_b}\sigma_a^2 + \frac{\eta_a^2}{T_b^2}\sum_n |G(nT_b^{-1})|^2\,\delta(f - nT_b^{-1}) \tag{10}$$

The first term in Equation 10 represents the continuous part of the power spectral density and depends only on the pulse shape used. Pulse shape should be selected such that the desired properties of line codes (see the following section) are met. The second term in the above equation represents discrete components, weighted by the pulse shape, $1/T_b$ a part in frequency that aids in timing recovery. This term can be reduced to zero by (1) a zero mean sequence—that is, one that consists of equally likely and symmetrically placed symbols in the complex plane; or (2) a pulse shape whose spectrum is zero at all multiples of $1/T_b$. Later we will show that unipolar RZ codes, with rectangular pulses, have this discrete component at zero frequency. However, note that polar versions can be made unipolar by rectification, from which the timing can be easily recovered.

In the following, we will assume that 1's and 0's are equally likely in the data sequence. Moreover, $\{a_k\}$ may represent the original data sequence or its differentially encoded form. Later in the chapter, we will see that the NRZ (NRZ-L, NRZ-M, and NRZ-S), RZ (polar or unipolar) family, and pseudoternary (AMI-RZ, AMI-NRZ, dicode RZ, and dicode NRZ) families as well as biphase-L can be written as Equation 7. Thus, the above expression, for PSD—Equation 10—represents their spectral densities. However, biphase-M, biphase-S, and substitution line codes cannot be represented as Equation 7, so other means must by used to calculate their spectral densities (Xiong 2000).

To find the spectral density of a *wide sense stationary* (WSS) signal, we need to find the autocorrelation function $R(\tau)$ first and then, according to the Wiener-Khintchine theorem, the Fourier transform of the function will give the spectral density (Simon, Hinedi, and Lindsey 1995).

BASIC LINE CODES FOR WIRELINE SYSTEMS

This section compares various basic wireline line codes. The comparison is based on their spectral characteristics, bandwidth, error performance, error-detection capability, self-synchronization ability, redundancy, and transparency. The basic line codes considered are:

- the non-return-to-zero family, including *NRZ-level* (NRZ-L), *NRZ-mark* (NRZ-M), and *NRZ-space* (NRZ-S) (including unipolar and polar subclasses);
- the return-to-zero family, including unipolar and polar subclasses;
- the pseudoternary family, including alternate mark inversion, dicode NRZ, and dicode RZ;
- the substitution codes that include binary *N* zero substitution and high-density bipolar *N;*
- the biphase family, which includes biphase-level (or Manchester), biphase-mark, and biphase-space; and
- the multilevel block code family, which includes kBnT and kBnQ codes.

There are other codes that do not belong to any of the classes while some may belong to more than one family.

The reason for a large selection of line codes is because of differences in performance that lead to different applications. The features to look for in choosing a line code are described below (Xiong 2000; Bellamy 2000; Barry, Lee, and Messerschmitt 2003).

Self-synchronization: A digital signal is preferred that includes timing information (self-synchronizing as opposed to sending a pilot signal that wastes bandwidth or synchronizes to a master clock) in the data being transmitted. This can be achieved if there are a suitable number of transitions in the signal that alert the receiver to the beginning, middle, or end of the pulse. Thus, formats with higher transition density are desired.

Spectrum suitable for channel: For lossless transmission, the spectra of the line codes should match the channels over which information is being transmitted. For example, for alternating current (AC) coupled channels, a line code whose spectrum has negligible energy near DC (zero frequency) should be used. Furthermore, line code should have a bandwidth that is small compared to the channel bandwidth so that channel-induced *intersymbol interference* (ISI) is not a problem.

Transmission bandwidth: The bandwidth should be as narrow as possible and may be reduced by using multilevel transmission techniques.

Low error probability: Receivers should be able to recover the information from the corrupted signal with a low probability of bit error.

Error detection capability: Line codes should preferably have a built-in error-detection capability.

Bit sequence independence (transparency): The bit pattern in a line code should not affect the ability to accurately recover the timing information.

Differential coding: If 1's and 0's are encoded such that they are negative of each other and then are transmitted over a medium on which it is impossible to determine an absolute phase reference, then 1's will be decoded as 0's and vice versa. To overcome this polarity inversion problem, either the source is differentially encoded or line codes that are inherently differential are used. In the latter case, a 1 or a 0 causes a transition in the code. Thus, decoding reduces to detecting whether the current state is the same or different from the previous.

Redundancy: By allowing redundancy (i.e., having an information bit rate less than the information-carrying capacity), we can make the transmitted data symbols statistically dependent, regardless of the statistics of the

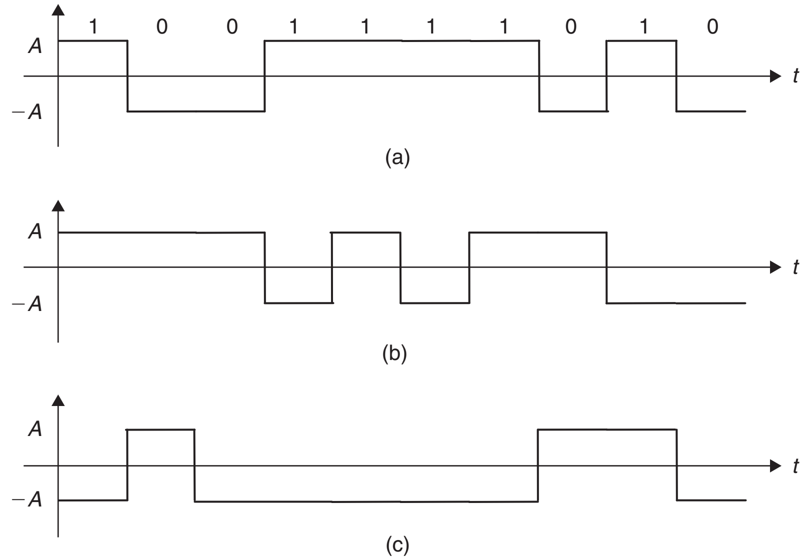

Figure 2: Polar NRZ waveforms—(a) NRZ–L, (b) NRZ—M, (c) NRZ–S

information, and hence exercise control over the power spectrum of the transmitted signal.

Non-Return-to-Zero Codes

Digital transmission systems send pulses along a link, usually a wire or fiber line within a cable. Two levels of pulse amplitude are used to distinguish between a 1 and a 0 in polar NRZ-L format (see Figure 2a). For an equally likely binary data sequence, the waveform has no DC component. The line code is relatively easy to generate, but when the data contain a long string of 0's or 1's, the received signal may not carry enough timing-recovery information and therefore the receiver may lose synchronization.

In NRZ-M [or *NRZ-invert* (NRZ-I)] format, the level changes whenever a *mark* (or 1) occurs in the sequence and no change takes place for a space (or 0) (see Figure 2b). The NRZ-S waveform is similar but encoded in the opposite sense (see Figure 2c). Thus, NRZ-M and NRZ-S are differentially encoded versions of NRZ-L. The main advantage of NRZ-M and NRZ-S over NRZ-L is their immunity to polarity reversals.

The NRZ family can be made unipolar (i.e., only one polarity—positive or negative—is used) by changing the lower level ($-A$ volts) to zero. This line code is referred to as *on-off signaling* and is widely used in optical fiber systems where the sources and detectors operate in the nonlinear region (Bellamy 2000). For equally likely 1's and 0's, unipolar waveforms have a nonzero DC level of $A/2$ volts, whereas their polar counterparts do not. The disadvantage of on-off signaling is waste of power because of the transmitted DC level; thus, the power spectrum is nonzero at DC (see Equation 19).

A constant level in the NRZ-encoded waveforms contains no transitions and hence a lack of timing content. This shortcoming can be overcome by (1) scrambling (or randomizing), which is a process that makes data look more random by eliminating long strings of 1's and 0's; or (2) transmitting a separate synchronization sequence.

NRZ-L is used in digital logic. NRZ-M is used in magnetic tape recording. In telecommunication applications, the NRZ format is limited to short-haul links because of its timing characteristic.

Bit Error Rate of NRZ Codes NRZ-L is polar with

$$\begin{aligned} s_1(t) &= +A \qquad 0 \le t \le T_b \\ s_2(t) &= -A \qquad 0 \le t \le T_b \end{aligned} \tag{11}$$

$\rho_{12} = -1, E_1 = E_2 = A^2T_b$, therefore polar NRZ-L is antipodal with the best error performance—that is, $P_b = Q\left(\sqrt{2A^2T_b/N_0}\right) = Q\left(\sqrt{2E_b/N_0}\right)$, where the average bit energy

$$E_b = (E_1 + E_2)/2 = A^2T_b$$

For unipolar NRZ-L:

$$\begin{aligned} s_1(t) &= +A \qquad 0 \le t \le T_b \\ s_2(t) &= 0 \qquad 0 \le t \le T_b \end{aligned} \tag{12}$$

$\rho_{12} = 0, \quad E_1 = A^2T_b, \quad E_2 = 0,$ or $P_b = Q\left(\sqrt{A^2T_b/2N_0}\right) = Q\left(\sqrt{E_b/N_0}\right)$: 3 dB worse than polar NRZ-L where $E_b = A^2T_b/2$ (see Figure 3).

NRZ-M and NRZ-S are modulated by a differentially encoded data sequence. The coded sequence is then differentially decoded back to the original data sequence. The current and previous bits of coded sequence are used to produce the current bit of the original sequence. Thus, the error probability, P_b', is (Xiong 2000)

$$\begin{aligned} P_b' &= (1-P_b)P_b + P_b(1-P_b) = 2(1-P_b)P_b \\ &\approx 2P_b = 2Q\left(\sqrt{2E_b/N_0}\right) \quad \text{for small } P_b \end{aligned} \tag{13}$$

Note that if the current bit is in error, then it and the next bit (because the current bit is used as a reference) will be erroneously decoded. Thus, in differential decoding, one

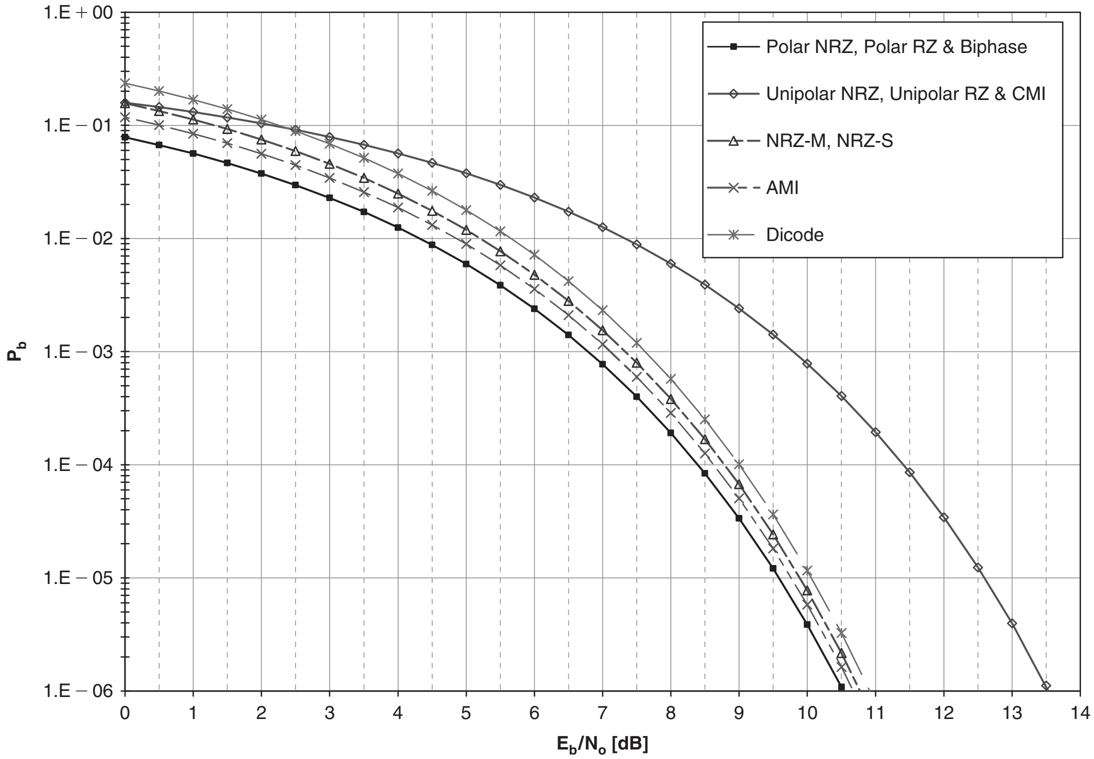

Figure 3: Probability of bit error

bit error affects two decoded bits (see Figure 3). However, the ambiguity in the reference bit will only affect the first decoded bit and, if desired, can be removed by agreeing on the reference bit in advance or by calculating the phase of the reference bit from a known transmitted sequence.

Power Spectral Density of NRZ Codes. NRZ-M and NRZ-S are generated by modulating the differentially encoded data sequences using the NRZ-L format. If the original binary data are equally likely, then the differentially encoded data are also equally likely—that is, statistical properties of the sequences used for modulation are the same for NRZ-L, NRZ-M, and NRZ-S. From the section on the probability of bit error and power spectral density, we know that if their pulse-shaping function is also the same, then the spectral densities are also the same because the power spectrum depends on the pulse shape used and the statistical properties of the data sequence.

NRZ formats' pulse-shaping function is a square pulse in $[0,T_b]$ seconds—that is,

$$g(t) = rect\left(\frac{1}{T_b}\right) = \begin{cases} 1 & 0 \le t \le T_b \\ 0 & else \end{cases} \tag{14}$$

$$\begin{aligned} G(f) &= T_b\left[\frac{\sin(\pi f T_b)}{\pi f T_b}\right]\exp(-j\pi f T_b) \\ &= T_b \mathrm{sinc}(\pi f T_b)\exp(-j\pi f T_b) \end{aligned} \tag{15}$$

Because the data sequence is equally likely, the correlation function in Equation 9 becomes

$$R[n] = \begin{cases} (A)^2\frac{1}{2} + (-A)^2\frac{1}{2} = A^2 & n = 0 \\ (A)(A)\frac{1}{2}\frac{1}{2} + (-A)(-A)\frac{1}{2}\frac{1}{2} + 2(-A)(A)\frac{1}{2}\frac{1}{2} = 0 & n \ne 0 \end{cases} \tag{16}$$

Thus, the power spectral density is

$$S(f) = A^2 T_b\left[\frac{\sin(\pi f T_b)}{\pi f T_b}\right]^2 \tag{17}$$

From Equation 17, we can see that the spectral density is a squared sinc function with first null at $fT_b = 1$. The signal energy is concentrated around zero frequency, and null bandwidth, B_{null}, is R_b. Note that the power spectral density is a non-bandlimited function. There are various definitions of bandwidth, and the interested reader is referred to Benedetto and Biglieri (1999). In this chapter, the bandwidth referred to is always null to null.

In Figure 4, A is selected such that the normalized average power of the polar NRZ-L signal is one. To calculate A, assume that the transmitted sequence is periodic—that is, $\{1\ 0\ 1\ 0\ 1\ 0 \cdots\}$. It can then be easily shown that $A = 1$ satisfies the unity normalized average power constraint.

A DC component of $A/2$ volts is present in unipolar NRZ and appears as an impulse function, with a strength

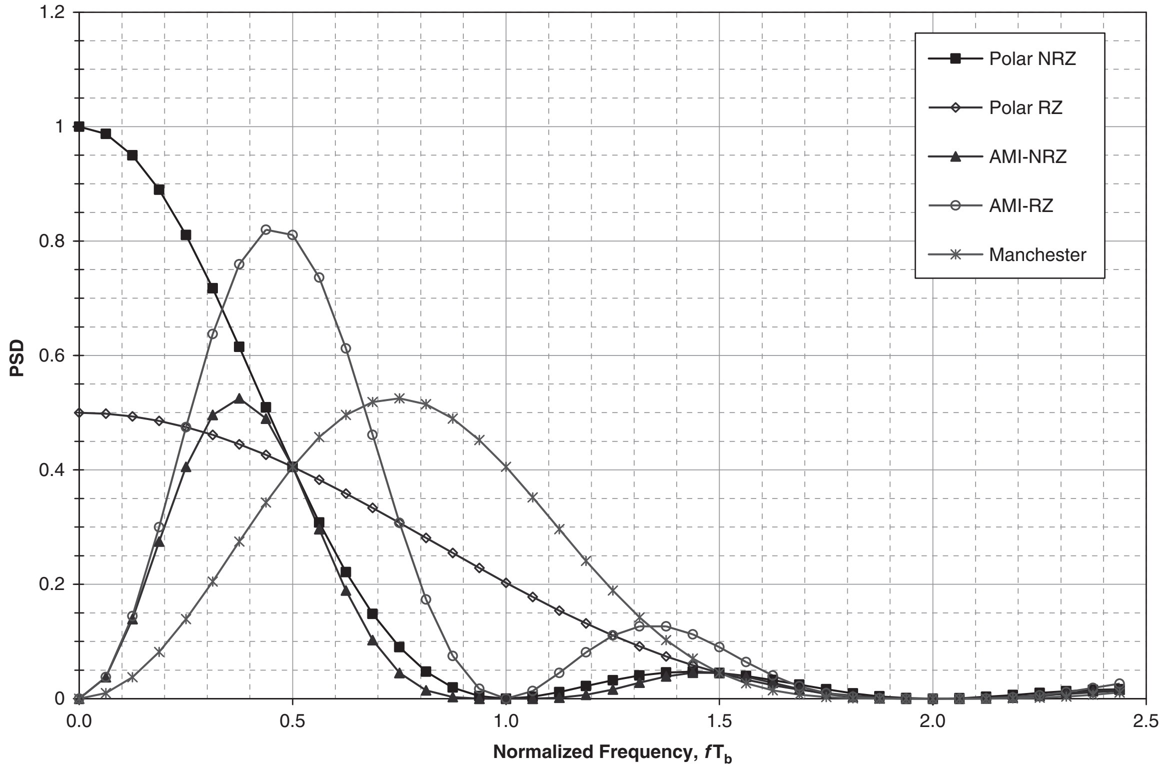

Figure 4: Power spectral density

$A^2/4$ watts at DC in the power spectrum. For equally likely data sequence the values of mean and variance are

$$\eta_a = E(a_k) = (A)\frac{1}{2} + (0)\frac{1}{2} = \frac{A}{2}$$
$$\sigma_a^2 = E[(a_k - \eta_a)^2] = (A - A/2)^2\frac{1}{2} + (0 - A/2)^2\frac{1}{2} = \frac{A^2}{4} \tag{18}$$

Thus, the power spectral density, from Equation 10, is

$$S(f) = \frac{A^2 T_b}{4}\left[\frac{\sin(\pi f T_b)}{\pi f T_b}\right]^2 + \frac{A^2}{4}\delta(f) \tag{19}$$

because the sinc function is zero at multiples of nR_b and exists only for $n = 0$. For $A = \sqrt{2}$, the normalized average power of the unipolar NRZ signal equals 1. Thus, the power spectra of polar NRZ and unipolar NRZ are the same with the only difference that the unipolar NRZ contains an impulse at DC corresponding to the nonzero mean of the signal.

Multiline Transmission, Three-Level Codes. The scheme of *multiline transmission, three-level* (MLT-3) codes is similar to NRZ-I, but it cycles through the three levels (+1, 0, and −1) to transmit a 1—that is, signal transitions from one level (say positive or zero or negative) to the next (say, zero, negative, or positive) at the beginning of 1 and stays at the same level for 0 (see Figure 5). This line coding technique is used in 100BaseT Ethernet ("Line code" undated).

Return-to-Zero Codes

The shortcoming of the NRZ family—lack of timing information—can be overcome by introducing more transitions in the waveform or trading self-synchronization capability for increased bandwidth. The receiver can use these transitions to update and synchronize its clock. To change with every bit, we need more than just two values. One solution is return-to-zero encoding, which uses three values: positive, negative, and zero. However, its disadvantage is that it uses a wider bandwidth than the NRZ waveform.

In polar RZ format, 1's and 0's are represented by positive and negative half-period pulses, respectively

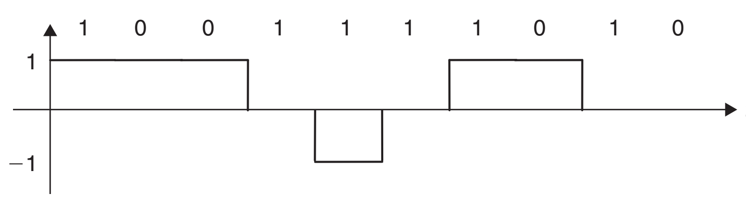

Figure 5: MLT-3

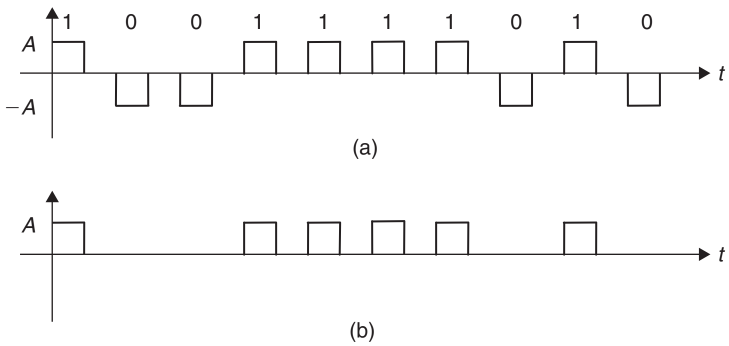

Figure 6: RZ waveforms—(a) polar, (b) unipolar

(see Figure 6a). This waveform ensures two transitions per bit and consequently has no DC component.

In the unipolar RZ format, a 1 is represented by a positive pulse for a half-bit period, which then returns to zero level for the next half period, resulting in a transition in the middle of the bit. A 0 is represented by the zero level for the entire bit period (see Figure 6b). A delta function (i.e., nonzero DC level) exists at $f = 0, \pm R_b$ in the power spectrum of the transmitted signal, which can be used for bit timing recovery at the receiver. However, the disadvantage of this code is that it requires 3 dB more power than polar RZ signaling for the same probability of symbol error. By making the sequence look more random (via scrambling), the long stream of 0's can be eliminated in the unipolar RZ scheme.

Bit Error Rate of RZ Codes. Polar RZ signals are:

$$s_1(t) = \begin{cases} +A & 0 \le t \le T_b/2 \\ 0 & else \end{cases} \qquad (20)$$
$$s_2(t) = -s_1(t)$$

$\rho_{12} = -1$, $E_1 = E_2 = A^2T_b/2$, therefore polar RZ is antipodal and has the best performance with $P_b = Q\left(\sqrt{2E_b/N_0}\right)$. The average bit energy is

$$E_b = (E_1 + E_2)/2 = A^2T_b/2$$

For unipolar RZ:

$$s_1(t) = \begin{cases} +A & 0 \le t \le T_b/2 \\ 0 & else \end{cases} \qquad (21)$$
$$s_2(t) = 0$$

$\rho_{12} = 0$, $E_1 = A^2T_b/2$, $E_2 = 0$ or $P_b = Q\left(\sqrt{A^2T_b/4N_0}\right) = Q\left(\sqrt{E_b/N_0}\right)$: 3 dB worse performance than polar RZ with $E_b = A^2T_b$ (see Figure 3).

When the bit energy of unipolar RZ is the same as that of unipolar NRZ, the amplitude of the former must be $\sqrt{2}$ times that of the latter. However, if amplitudes are fixed, then unipolar NRZ will have twice the energy of unipolar; thus, its error probability is lower.

Power Spectral Density of RZ Codes. For RZ formats, the pulse-shaping function is a square pulse with half-bit duration—that is,

$$g(t) = rect\left(\frac{t}{T_b/2}\right) = \begin{cases} 1 & 0 \le t \le T_b/2 \\ 0 & else \end{cases} \qquad (22)$$

$$G(f) = \frac{T_b}{2}\left[\frac{\sin(\pi f T_b/2)}{(\pi f T_b/2)}\right]\exp(-j\pi f T_b/2) \qquad (23)$$

For an equally likely data sequence, the correlation function of the polar RZ signaling is the same as that for polar NRZ—that is,

$$R[n] = \begin{cases} A^2 & n = 0 \\ 0 & n \ne 0 \end{cases} \qquad (24)$$

The power spectral density is

$$S(f) = \frac{A^2T_b}{4}\left[\frac{\sin(\pi f T_b/2)}{(\pi f T_b/2)}\right]^2 \qquad (25)$$

As can be seen from Equation 25, the spectral density is a stretched version with frequency axis scaled up twice—that is, all bandwidths are double that of NRZ (i.e., the null bandwidth, B_{null}, is $2R_b$). The value of $A = \sqrt{2}$ makes the normalized average power of the polar RZ signal equal to 1 and has been used to plot the PSD in Figure 4.

For the unipolar RZ format, the pulse shape is the same as above in Equation 22. The data sequence, its mean, and its variance are same as those for unipolar NRZ (Equation 18). Thus, the power spectral density, after substituting Equation 23 into Equation 10, is

$$S(f) = \frac{A^2T_b}{16}\left[\frac{\sin(\pi f T_b/2)}{(\pi f T_b/2)}\right]^2 \times \left[1 + R_b\sum_{n=-\infty}^{\infty}\delta(f - nR_b)\right] \qquad (26)$$

The value of $A = 2$ satisfies the constraint of unit normalized average power of the unipolar RZ signal. It can be easily seen from Equation 26 that the bandwidths are double those of the corresponding NRZ signal.

Pseudoternary Codes

The baseline or DC wander is caused by a slow decay in amplitude of a long string of 1's or 0's or whenever there is an imbalance in the number of 1's and 0's (Bellamy 2000). The impairment makes the receiver lose its reference to distinguish between the two levels and is a problem that can be solved through coding of transmitted data symbols—that is, a zero at DC is introduced in the spectrum by forcing the data symbols to be correlated. Moreover, the only way redundancy can be introduced without increasing the baud rate is by increasing the number of levels (Barry, Lee, and Messerschmitt 2003). In a pseudoternary line code, we use three-level data symbols to transmit one bit of information, which also helps reduce baseline wander. However, with these codes we also suffer a reduction in noise immunity because for the same peak power level, a smaller noise level causes an error (compared to biphase codes).

Three levels ($\pm A$, 0 volts) are used. The most popular, *alternate mark inversion* (AMI) codes, belong to this group. They are often called *bipolar codes* (Bellamy 2000).

In AMI-RZ format, a 1 is represented by an RZ pulse and consecutive 1's are represented with alternating polarities. A 0 is represented by the zero-level, hence an average voltage level of zero is maintained (see Figure 7a). In AMI-NRZ, coding rule is the same as AMI-RZ except that the symbol pulse has a full length of T_b — 100 percent duty cycle (see Figure 7b). The average level is still zero, but like unipolar RZ, the lack of transitions in a string of 0's causes synchronization problems.

Dicode NRZ and dicode RZ also belong to this group. These formats are also called *twinned binary* and encode the changes in a data sequence—that is, in dicode NRZ, a 1 to 0 transition or a 0 to 1 transition changes signal polarity. If the data remain constant, then a zero-voltage level is output (see Figure 7c). In dicode RZ, the same coding rule applies except that the pulse is only half-bit wide (see Figure 7d). Dicodes and AMI codes are related through differential coding (Xiong 2000).

These formats are used in baseband data transmission and magnetic recording. AMI-RZ formats are used in telemetry systems and are also being used by AT&T for T1 carrier systems.

Bit Error Rate of Pseudoternary Codes. The AMI-NRZ code consists of three types of signals:

$$\begin{array}{llll} H_1 : s_1(t) = A & 0 \le t \le T_b & p_1 = 1/4 & \textit{Transmit}\ 1 \\ H_2 : s_2(t) = -A & 0 \le t \le T_b & p_2 = 1/4 & \textit{Transmit}\ 1 \\ H_3 : s_3(t) = 0 & 0 \le t \le T_b & p_3 = 1/2 & \textit{Transmit}\ 0 \end{array} \qquad (27)$$

The probability of bit error in AWGN is given by the following expression:

$$\begin{aligned} P_b &= p_1 \Pr(error|s_1) + p_2 \Pr(error|s_2) + p_3 \Pr(error|s_3) \\ &= p_1 \int_{-\infty}^{-A/2} f_N(n|s_1)dn + p_2 \int_{A/2}^{\infty} f_N(n|s_2)dn + 2p_3 \int_{A/2}^{\infty} f_N(n|s_3)dn \\ &= 2p_1 \int_{-\infty}^{-A/2} f(n|s_1)dn + 2p_3 \int_{A/2}^{\infty} f(n|s_3)dn \\ &= \frac{1}{2} Q\left[\frac{A}{2\sigma_n}\right] + Q\left[\frac{A}{2\sigma_n}\right] = \frac{3}{2} Q\left[\frac{A}{2\sigma_n}\right] \end{aligned} \qquad (28)$$

For the matched filter receiver, the above probability of error becomes (Couch 2001)

$$P_b = \frac{3}{2} Q\left[\sqrt{\frac{2E_b}{N_0}}\right] \qquad (29)$$

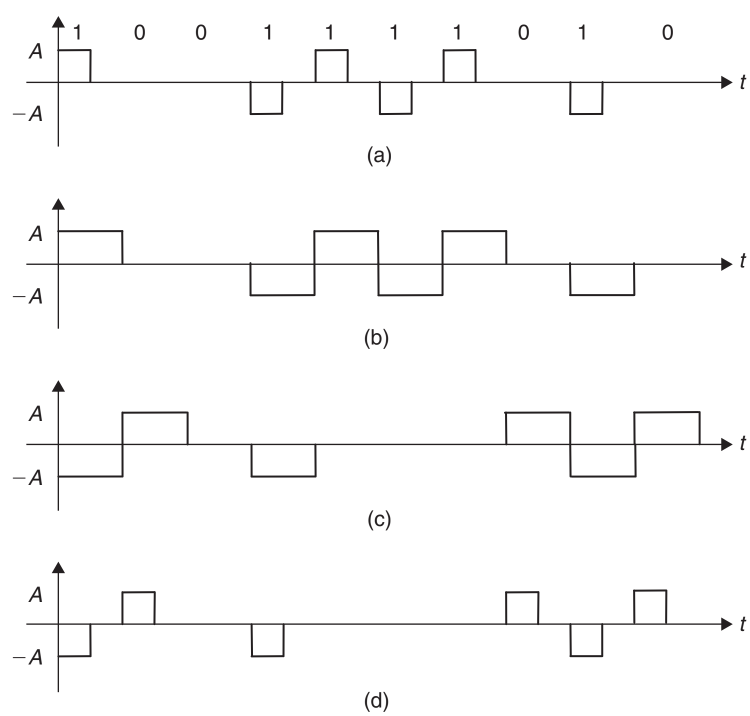

Figure 7: Pseudoternary waveforms—(a) AMI-RZ, (b) AMI-NRZ, (c) dicode NRZ, (d) dicode RZ

For AMI-RZ, the same error performance, as in Equation 29, is obtained (Couch 2001) (see Figure 3). Dicodes and AMI are related via differential encoding. Thus, similar to NRZ-M and NRZ-S, the probability of bit error for the dicode is twice the bit error probability of its AMI counterparts (Xiong 2000).

Power Spectral Density of Pseudoternary Codes. For AMI codes, the data sequence takes on three values:

$$a_k = \begin{cases} +A & binary\,1 \quad p_A = 1/4 \\ -A & binary\,1 \quad p_{-A} = 1/4 \\ 0 & binary\,0 \quad p_0 = 1/2 \end{cases} \tag{30}$$

$$R[0] = E(a_k^2) = \frac{1}{4}(A)^2 + \frac{1}{4}(-A)^2 + \frac{1}{4}(0)^2 = \frac{A^2}{2} \tag{31}$$

Adjacent bits are correlated because of alternate mark inversion. An adjacent bit pattern in the original binary sequence must be one of these: (1,1), (1,0), (0,1), and (0,0). The possible values of product $a_k\, a_{k+1}$ are: $-A^2$, 0, 0, and 0, and each has a probability of 0.25. Thus,

$$R[1] = E(a_k a_{k+1}) = \frac{1}{4}(A)(-A) + \frac{1}{4}(0) + \frac{1}{4}(0) + \frac{1}{4}(0) = -\frac{A^2}{4} \tag{32}$$

When $n > 1$, $a_k\, a_{k+n}$ are uncorrelated. Possible values of product $a_k\, a_{k+n}$ are $\pm A^2$, 0, 0, and 0, and each occurs with a probability of 0.25 (Couch 2001). Thus,

$$R[n > 1] = E(a_k a_{k+n}) = \frac{1}{8}(+A)(+A) + \frac{1}{8}(+A)(-A) = 0$$
$$R[n] = \begin{cases} A^2/2 & n = 0 \\ -A^2/4 & |n| = 1 \\ 0 & |n| > 1 \end{cases} \tag{33}$$

Thus, the power spectral density for AMI-RZ is

$$\begin{aligned} S(f) &= \frac{1}{T_b}|G(f)|^2 \left[\frac{1}{2} - \frac{1}{4}\exp(j\omega T_b) - \frac{1}{4}\exp(-j\omega T_b)\right] \\ &= \frac{1}{T_b}|G(f)|^2 \left[\frac{1}{2} - \frac{1}{2}\cos(\omega T_b)\right] \\ &= \frac{A^2 T_b}{4}\left[\frac{\sin(\pi f T_b/2)}{(\pi f T_b/2)}\right]^2 \sin^2(\pi f T_b) \end{aligned} \tag{34}$$

For $A = 2$, the normalized average power of the AMI-RZ signal is 1. The null bandwidth, B_{null}, is R_b (see Figure 4).

The power spectral density of AMI-NRZ can be obtained by replacing $T_b/2$ with T_b in $G(f)$ of AMI-RZ because both of them have the same coding rules; the only difference is in the pulse width. Thus,

$$S(f) = \frac{A^2 T_b}{2}\left[\frac{\sin(\pi f T_b)}{\pi f T_b}\right]^2 \sin^2(\pi f T_b) \tag{35}$$

a value of $A = \sqrt{2}$ satisfies the normalized average power constraint of unity. The null bandwidth, B_{null}, is R_b. Note that the bandwidths of AMI-RZ and AMI-NRZ are the same.

Dicodes can be constructed using AMI rules and a differentially encoded sequence—that is,

$$s(t) = \sum_{k=-\infty}^{\infty} d_k g(t - kT) \tag{36}$$

where $\{d_k\}$ is the pseudoternary sequence derived from the original data sequence $\{a_k\}$ and is

$$d_k = a_{k-1} - a_k = \begin{cases} +1 & binary\,1 \quad p_1 = 1/4 \\ -1 & binary\,1 \quad p_{-1} = 1/4 \\ 0 & binary\,0 \quad p_0 = 1/2 \end{cases} \tag{37}$$

Thus, the power spectral density of dicodes is the same as those of AMI; therefore, the bandwidth is also the same (Xiong 2000).

Substitution Codes

The AMI code is a preferred choice because of its (1) narrow bandwidth, (2) lack of DC component, (3) error-detection capability (resulting from alternate mark inversion), (4) ease in synchronization (because of transitions in each binary 1 bit), and (5) absence of DC wander (because of inefficient use of the ternary code space). However, a string of 0's will result in a long period of zero level that will cause loss of synchronization. This problem results from the linearity of the AMI code because linearity implies that an all-zero bit sequence is translated into a zero signal. The solution to this problem is therefore to modify the line code and make it nonlinear. One way to alleviate this problem is to substitute the long string of zeros with a special sequence with intentional bipolar violations that can be readily detected and consequently replaced at the receiver. Two popular zero substitution families are considered in this section: (1) *binary* N-*zero substitution* (B*N*ZS) and (2) *high-density bipolar* N (HDB*N*). These codes have been used in T1 carrier systems.

Binary N-Zero Substitution Codes. The B*N*ZS code modifies AMI by performing a substitution for a block of N consecutive zeros. The substituted block, which contains one or more positive or negative pulses to ensure timing recovery, takes advantage of the fact that only 2^{N+1} patterns of N transmitted symbols are allowed by AMI and therefore substitutes one of the nonallowed blocks for the all-zeros block. At the receiver, this nonallowed block is readily detected and replaced with the all-zero bit block (Barry, Lee, and Messerschmitt 2003).

All B*N*ZS formats have no DC component and retain the balanced feature of AMI. There are two kinds of B*N*ZS codes: (1) nonmodal and (2) modal.

In nonmodal code, two substitution sequences are allowed, and the choice among them is based solely on the polarity of the pulse immediately preceding the zeros to be replaced. Substitution sequences must contain an equal number of positive and negative pulses to maintain

DC balance. The substitution sequences may also contain zeros. For nonmodal codes, N must be at least 4 (Xiong 2000). Some practical balanced nonmodal codes are B6ZS and B8ZS. For example, in the B6ZS case, if the polarity of the pulse preceding the six zeros is negative (−), then the sequence to be substituted is 0 − + 0 + − whereas if the polarity of the pulse preceding the six zeros is positive (+), then the suggested substitution sequence is 0 + − 0 − + (Bellamy 2000). Thus, the polarity of the pulse immediately preceding the six zeros and the polarity of the last pulse in the sequence to be substituted is the same. For B8ZS, the two substitution sequences corresponding to a negative and positive pulse, preceding the eight zeros, are 0 0 0 − + 0 + − and 0 0 0 + − 0 −+, respectively (Bellamy 2000).

In the modal case, more than two substitution sequences are provided, and the choice of the sequences is based on the polarity of pulse immediately preceding the zeros to be replaced as well as the previous substitution sequence used. For modal codes, N is 2 or 3 (Xiong 2000). Modal code substitution sequences need not be balanced, and balance is achieved by properly alternating the sequences. A block of 0 0 0 is replaced by B 0 V or 0 0 V, where B represents a normal bipolar alternation that conforms to AMI rule, V represents bipolar violation, and 0 represents no pulse. If the polarity of the pulse preceding the three zeros, is negative (positive) and the number of bipolar pulses since last substitution is odd, then replace 0 0 0 with 0 0 V. However, if the preceding pulse is negative (positive) and if the number of 1's since last substitution is even, then replace 0 0 0 by B 0 V (see Figure 8). In a long-time average, the polarity of these extra pulses will cancel each other so that there will be no DC component.

B3ZS and B6ZS are specified for DS-3 and DS-2, respectively, and B8ZS is specified as an alternative to AMI for DS-1.

No results of bit error probability of BNZS codes are available in the literature. They are conditioned AMI codes, therefore their bit error probabilities must be quite close to those of AMI codes (Xiong 2000).

The power spectrum calculation for BNZS codes is based on a flow graph of the pulse states and is quite involved. Moreover, the spectrum depends on the substitution sequence used and the statistical property of the data sequence. The interested reader is referred to (Xiong 2000) and (Bellamy 2000).

High-Density Bipolar *N* Codes. High-density bipolar N codes limit the number of consecutive 0's to N by replacing the $(N + 1)$th zero by a bipolar violation. Moreover, to avoid a DC component, the code is made modal. Substitution sequences are (Xiong 2000)

$$HDB : B \ 0 \ 0 \ \cdots \ 0 \ \mathrm{V} \,\mathrm{or}\, 0 \ 0 \ 0 \ \cdots \ 0 \ \mathrm{V} \qquad (38)$$

In HDB3 line coding, strings of greater than three zeros are excluded. The coding algorithm is similar to the B3ZS algorithm described in the previous section. If the polarity of the pulse preceding the three zeros is negative (positive) and the number of bipolar pulses since last substitution is odd, then replace 0 0 0 0 by 0 0 0 V. However, if the preceding pulse is negative (positive) and if the number of 1's since last substitution is even, then replace 0 0 0 0 with B 0 0 V (see Figure 9).

Two commonly used HDBN codes are HDB2 and HDB3. HDB2 is identical to B3ZS. HDB3 is used for coding of 2.048, 8.448, and 34.368 Mbps streams within the European digital hierarchy.

Biphase Codes

With the help of substitution codes, the pseudoternary family of codes gains self-synchronizing capability. However, this comes at a price: not the bandwidth (NRZ-L has the same bandwidth) but the increased code space (from binary to ternary). The pseudoternary family does not use the code space efficiently. It will be seen in the section on multilevel block line codes that the code efficiency of AMI is only 63 percent; to achieve higher code efficiencies, the input block size will have to be increased. The desirable characteristics (error detection, self-synchronizing, and lack of DC wander capabilities) of pseudoternary codes can be achieved with binary codes but at the expense of bandwidth. The biphase family of codes uses one cycle of square wave to represent 1's and its opposite phase to represent 0's. The most famous code belonging to this family is the biphase level, which is better known as *Manchester* (or *diphase* or *split-phase*). The Manchester encoding signal is shown in Figure 10a. Note that the pulse shapes for 0 and 1 are arbitrary and can be exchanged. For Manchester encoding, the signal contains strong timing components because a transition is present in the center of each bit interval. Moreover, the positive and negative polarities of 0 and 1 are equal, therefore there is no DC wander. The only drawback—besides twice the bandwidth, which makes it vulnerable to near-end cross-talk and ISI (Barry, Lee, and Messerschmitt 2003)—is that it has no error-detection capability.

In the *biphase-mark code*, a transition is always present at the beginning of each bit. For a 1, there is a second transition in the middle of the bit, while there is no second transition for a 0—that is, it is encoded as a level

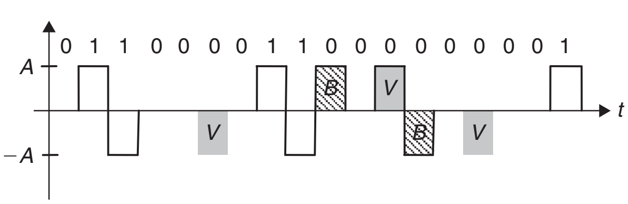

Figure 8: B3ZS-coded waveform

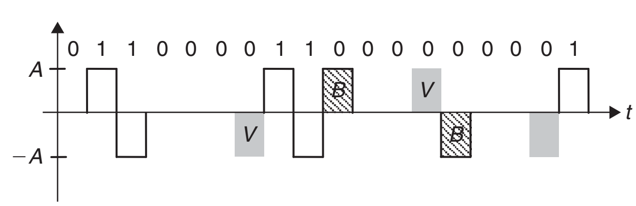

Figure 9: HDB3-coded waveform

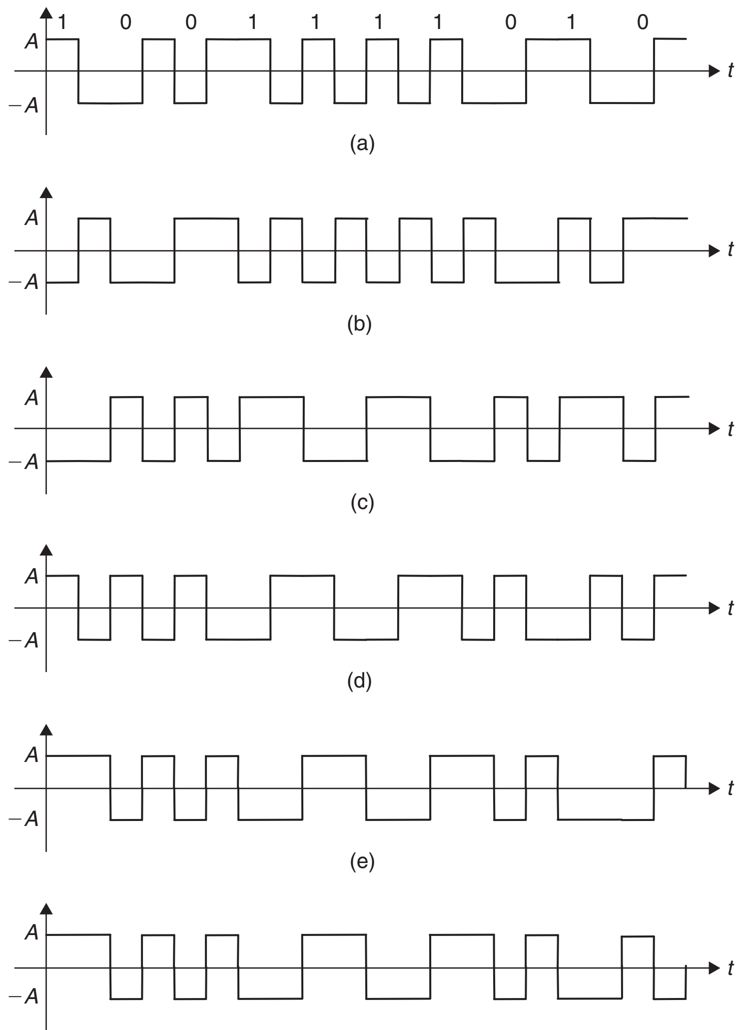

Figure 10: Biphase waveforms—(a) biphase-L, (b) biphase-M, (c) biphase-S, (d) differential biphase, (e) CMI, (f) DMI

(see Figure 10b). In the *biphase-space code*, the opposite coding rule applies (see Figure 10c). Because there is at least a transition in each bit interval, the codes contain adequate timing information.

Conditioned biphase level is differentially encoded biphase-L and is also referred to as *differential Manchester*. Like all differentially encoded signals, the information resides in signal transitions—that is, a change in level—at the beginning of the bit interval, from the preceding one, which indicates that a 0 was transmitted; whereas absence of a transition means a 1 was sent (see Figure 10d). Because data are differentially encoded, this format is also immune from polarity reversals. The biphase-S is identical to Differential Manchester if the zero reference of the former is changed by half a bit.

The *code mark inversion* (CMI) code is a combination of the AMI-NRZ and Manchester. It is similar to AMI-NRZ because 1's are alternately encoded as A or $-A$ volts. It is related to Manchester because a 0 is represented by a half-cycle square wave of a particular phase, such as level A volts for first half bit and $-A$ volts for second half bit, or vice versa (see Figure 10e). The signal has no energy at DC, and the transition density over AMI-NRZ is improved significantly. Error detection is possible by monitoring occurrence of alternating 1's. There is no ambiguity between 1's and 0's. The technique is insensitive to polarity reversals because decoding is possible by simply comparing the second half of the bit with the first half.

The *differential mode inversion* (DMI) code is again a combination of AMI-NRZ and Manchester. Its coding rule for 1's is the same as that of CMI. However, its coding rule for 0's is different: $0 \rightarrow -A, A$ or $0 \rightarrow A, -A$ so that no pulses in a sequence have pulse widths wider than the bit duration (see Figure 10f).

To be used for digital subscriber loops, the power spectrum of the transmitted signal should (1) be zero at zero frequency because no DC transmission passes through a hybrid transformer and (2) be low at high frequencies because transmission attenuation in a twisted pair is most severe at high frequencies and between adjacent twisted pairs increases dramatically at high frequencies because of increased capacitive coupling (Haykin 2004).

The biphase code is a good choice where implementation simplicity is desirable and the distance between transmitter and receiver is modest, as in a *local area network* (LAN). The biphase family has been used in magnetic recording, optical communications, and in some satellite telemetry links. The Ethernet or IEEE 802.3 standard for LAN uses Manchester coding. Differential Manchester has been used for IEEE 802.5 standard for token ring, using either baseband coaxial cable or twisted-pair. Because it uses differential coding and is immune to polarity inversions, differential Manchester is preferred for a twisted-pair channel. CMI has been chosen for coding of 139.246 Mbps multiplex within the European digital hierarchy.

Bit Error Rate of Biphase Codes. Biphase-L signals are antipodal with

$$s_1(t) = \begin{cases} +A & 0 \le t \le T_b/2 \\ -A & T_b/2 \le t \le T_b \end{cases} \qquad (39)$$
$$s_2(t) = -s_1(t)$$

$\rho_{12} = -1$, $E_1 = E_2 = A^2T_b = E_b$, and therefore the probability of error is $P_b = Q\left(\sqrt{2A^2T_b/N_0}\right) = Q\left(\sqrt{2E_b/N_0}\right)$, which is the same as NRZ-L because both have the same average bit energy and are antipodal (see Figure 3).

Conditioned biphase-L has a bit error probability approximately two times that of biphase-L because it is just differentially encoded biphase-L.

The error performance of CMI is 3 dB worse than Manchester signaling when bit-by-bit detection is used (Bellamy 2000):

$$P_b \approx Q\left(\sqrt{\frac{E_b}{N_0}}\right) \qquad (40)$$

This error performance is the same as that of unipolar codes (see Figure 3). For details, see Xiong (2000). Improved performance can be obtained by using maximum likelihood Viterbi decoding (Bellamy 2000).

Power Spectral Density of Biphase Codes. For biphase-L, the pulse shape is half-positive and half-negative—that is,

$$g(t) = \begin{cases} 1 & 0 \le t \le T_b/2 \\ -1 & T_b/2 \le t \le T_b \\ 0 & else \end{cases} \qquad (41)$$

$$G(f) = T_b\left[\frac{\sin(\pi f T_b/2)}{(\pi f T_b/2)}\right]\sin(\pi f T_b/2)\exp(-j\pi f T_b) \qquad (42)$$

The correlation of the data sequence is

$$R[n] = \begin{cases} (A)^2\frac{1}{2} + (-A)^2\frac{1}{2} = A^2 & n = 0 \\ (A)(A)\frac{1}{2}\frac{1}{2} + (-A)(-A)\frac{1}{2}\frac{1}{2} + 2(A)(-A)\frac{1}{2}\frac{1}{2} = 0 & n \ne 0 \end{cases} \qquad (43)$$

Thus, the power spectral density is

$$S(f) = A^2T_b\left[\frac{\sin(\pi f T_b/2)}{\pi f T_b/2}\right]^2 \sin^2(\pi f T_b/2) \qquad (44)$$

For $A = 1$, the normalized average power constraint of unity is satisfied (see Figure 4). The power spectrum of conditioned biphase-L is the same as that of the biphase-L because it is merely a differentially encoded biphase-L, and differential encoding does not change the probability density function of equally likely data.

Although biphase-M and biphase-S use different waveforms to encode data, they have the same power spectrum if marks and spaces are equally likely in a data sequence. Note that their waveforms are close to that of biphase-L in terms of pulse shapes and number of transitions (biphase-S is differentially encoded biphase-L if the zero reference is shifted by half a bit for the former), therefore their power spectrum is the same as that of biphase-L.

Multilevel Block Line Codes

In a block code, k bits are mapped into n data symbols drawn from an alphabet of size L with the constraint

$$2^k \le L^n \qquad (45)$$

When equality is not met, there is available redundancy that can be used to accomplish desirable goals such as minimizing baseline wander or providing energy for timing.

Two basic techniques used in block coding are (1) translation of a block of input bits to a block of output symbols that uses more than two levels per symbol or (2) insertion of additional binary pulses to create a block of n binary symbols that is longer than the number of information bits m. The first technique applies to cases in which bandwidth is limited but multilevel transmission is possible, such as metallic wires used for digital subscriber loops. The second technique is mainly used in optical transmission where modulation is limited to two levels (on-off, optical sources and detectors operate in a nonlinear mode) but can withstand a small increase in transmission rate because a wide bandwidth is available. All basic line codes can be viewed as special cases of block codes.

Ternary *k*B*n*T Codes

This is a class of codes that maps k binary bits into n ternary symbols, where $n < k$. AMI can be considered as 1B1T code.

The efficiency of a block code is the ratio of the actual information rate to theoretical maximum information rate. NRZ encodes one bit into a 1-binary encoded symbol, therefore its efficiency is $\log_2(2)/\log_2(2) = 1$. AMI encodes one bit into a 1-ternary symbol, therefore its efficiency is $\log_2(2)/\log_2(3) = 0.63$. Manchester encodes one bit into two binary symbols, therefore its efficiency is $\log_2(2)/[2 \times \log_2(2)] = 0.5$.

AMI and its derivative pseudoternary codes (dicode RZ and dicode NRZ) transmit only one bit per symbol whereas the capacity of a ternary symbol is $\log_2(3) = 1.58$ bits.

Table 1: Efficiency of *k*B*n*T Codes

k	**1**	**3**	**4**	**6**	**7**
n	1	2	3	4	5
$\eta = \frac{k}{n}\frac{\log_2(2)}{\log_2(3)}$	0.63	0.95	0.84	0.95	0.89
	1B1T	3B2T	4B3T	6B4T	7B5T

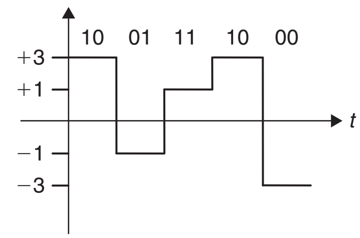

Figure 11: The 2B1Q waveform

Moreover, little control over the power spectrum is provided. A much broader class of pseudoternary codes—the kBnT codes, where k is the number of information bits and n is the number of ternary symbols per block—addresses these shortcomings. If we choose the largest k possible for each n, we get a table of possible codes (up to $k = 7$) (see Table 1).

As the block size increases, we generally achieve greater efficiency. Greater efficiency implies better noise immunity on many channels, because it translates into a lower symbol (baud) rate for a given bit rate and hence a reduced noise bandwidth. However, greater efficiency also implies reduced redundancy, and hence less control over the statistics of transmitted signal (power spectrum, timing recovery, density of 1's, etc.). The 4B3T code seems to be a reasonable compromise between these competing goals and hence has been used widely in digital subscriber loop applications.

The 4B3T Code. This code maps a combination of four bits to twenty-seven possible combinations of three ternary digits. The ternary sequence 000 is not used, but all other combinations are used. The code uses three codebooks; a codebook defines a mapping between bit patterns of input and output for an encoder. The three codebooks contain words that are biased toward positive polarity, negatively biased, or neutral. The algebraic sum of recently transmitted symbols, also known as the *digital sum* or *disparity*, determines which codebook is to be used. If disparity is positive, then code words from the negatively biased codebook are selected and vice versa. If disparity is neutral, then code words from the neutral codebook are selected. This ensures that the transmitted code has zero DC content.

For the 4B3T line code, the power spectral density depends on the selection of a particular codebook. Multiple choices of codebooks exist because sixteen information bits can be mapped to twenty-seven different ternary codes in a variety of ways.

The 4B3T coding is used on T148 span lines developed by ITT Telecommunications (Bellamy 2000).

The 8B6TCode. The 8B6T code substitutes an eight-bit group with a six-symbol ternary code. Eight bits can represent 256 possibilities; a six-symbol ternary signal can represent 729 possibilities—that is, some codes are not used. Thus, a code can be designed that can maintain synchronization and also has error checking capability. This code is used in 100 Mbps Ethernet (Moloney 2005).

Quaternary Code: 2B1Q. In this code, binary data are mapped into one of four levels. The disadvantage with multilevel signaling is that it requires greater SNR for a given error rate. Thus, the 2B1Q code represents a four-level pulse amplitude modulated signal (see Figure 11). Assuming 1 and 0 are equiprobable, the 2B1Q code has zero DC on the average. The signaling baud rate is half of the bit rate. It offers the greatest baud reduction and the best performance with respect to near-end cross talk and ISI. It is because of this desirable property that the 2B1Q code has been adopted as the North American standard for digital subscriber loops ("Line code" undated).

The power spectrum of the 2B1Q code is given by Chen (1998):

$$S(f) = \frac{5}{9} T_{2B1Q,baud\,rate} \left[\frac{\sin(\pi f T_b)}{\pi f T_b} \right]^2 \tag{46}$$

CODES FOR OPTICAL FIBER SYSTEMS

Line codes used in optical fiber systems should also satisfy the requirements that are desired for basic line codes (see the previous "Basic Line Codes for Wireline Systems"). Because the optical sources and detectors operate in nonlinear mode, the best suited line code is the one that has two levels: unipolar scrambled (or randomized) NRZ-L or on-off keying. Because there is plenty of bandwidth available (no advantage is gained by using multilevel signaling) in optical fiber systems, the inherent disadvantage of a lack of timing information in the line code can be overcome by including extra timing transitions for a small bandwidth penalty. Moreover, there is no polarity ambiguity in direct detection optical receivers, therefore resorting to differential encoding is unnecessary. The DC level of the line code affects the gain of some photodiodes. However, it is easy to control variations in the DC level in an optical system than in a wireline system, so for optical fiber the line codes should be DC-constrained instead of DC-balanced.

In intensity modulated practical optical fiber communication systems, a symbol is represented by the optical source's light intensity, therefore AMI cannot be used in such systems because it uses three levels (-1 for off, 0 for half intensity, and $+1$ for full intensity) and thus suffers from nonlinearity of the source. However, in phase modulated optical fiber communication systems, optical AMI has been implemented by coding the phase of light with π phase difference to achieve the coding of +1 and −1. When high data speed is not a requirement, then Manchester or CMI codes are used in optical fibers. When CMI is used for optical transmission, two levels of $\pm A$ are replaced by A and 0—that is, the code waveform is unipolar. This unipolar CMI has the same spectral shape as polar CMI with the only difference that the spectrum has a DC component. Unipolar DMI is also used in optical transmission systems (Xiong 2000).

Just as with wireline systems, block coding can also be used in optical fiber transmission systems. The m information bits are mapped to a block of n binary symbols where $n > m$. The additional pulses can be selected such that timing content is increased in the block line code. The modulation of choice is binary because the optical sources and devices operate in the nonlinear mode.

The *m*B*n*B Block Codes

Transitions in a data stream can be ensured by scrambling (or randomizing) the long string of zeros. The same goal can be achieved by introducing *redundancy*. Redundancy not only gives flexibility over controlling timing and DC wander but also aids in detecting errors. Block coding to some extent achieves these two goals.

A special case of block codes occurs when $L = 2$, which means we transmit a binary signal. For this case, we have the simpler constraint $k \leq n$—in other words, we must transmit a block of n bits that is larger than the number of bits k at the input to the line coder. These codes are useful for media such as optical fiber that prefer to operate in one of the two states (on or off) and for which the additional bandwidth required is easier to achieve.

One primary motivation in the design of line codes has been the elimination of the DC content of the coded signal because of AC coupling of the medium. It might appear that this problem goes away for media such as optical fiber and magnetic recording because there are no transformers required. However, this is not really the case because it is especially difficult to build DC-coupled high-speed electronics for preamplification and so forth (Barry, Lee, and Messerschmitt 2003).

When $n > m + 1$, because of decoding logic or a small look-up table, it is difficult to implement these codes in high-speed transmission systems. For $n = m + 1$, the mB ($m + 1$)B codes are known as *bit-insertion codes* and are popular in high-speed fiber optical transmission because of its simple codec design. However, to increase code efficiency, a large m (greater than 1) must be chosen.

The 2B3B DC-Constrained Code. This code constrains rather than suppresses the DC component and thus makes it possible to use redundancy for error detection. Data bits 1 and 0 are converted to + and −, respectively. Then a third symbol + or − is added to make combinations of one + and two −'s. The code produces a constant DC component of −1/3 (Xiong 2000).

The 4B5B Code. With a four-bit input block, we can have sixteen different groups. With a five-bit code, we can have thirty-two possible codes. This means we can map some of the five-bit groups to four-bit groups. Some of the five-bit codes are not used. We can employ a strategy to choose only the five-bit codes that help us in synchronization and error detection. To achieve synchronization, we can use the five-bit codes in such a way that we do not have more than three consecutive 0's or 1's—that is, select five-bit codes that contain no more than one leading 0 and no more than two trailing 0's. Because only a subset of five-bit codes are used, if one or more of the bits in the block is changed in such a way that one of the unused codes is received, the receiver can easily detect the error. After substitution, we can use one of the basic wireline line coding techniques discussed above to create a signal. This code is used for Ethernet 100 Mbps and *fiber distributed data interface* over optical fiber (Moloney 2005). The 4B5B code increases the line rate by 25 percent. CMI and biphase (1B2B codes) increase the line rate by 100 percent.

The 7B8B code is being used in a 565-Mbps terrestrial system being developed by British Telecom and in a 280-Mbps NL1 submarine system being developed by STC of Great Britain.

The Carter Code (8B9B). This code was proposed for PCM systems. The eight-digit character is transmitted either unchanged or with the digits inverted (i.e., marks for spaces and spaces for marks), depending on which condition will reduce total disparity (the numerical sum of symbols in the sequence) since the start of transmission. Thus, the DC component will be zero over a long period. The efficiency of the Carter code is 0.9464.

The m B1P Code. This code inserts an odd parity bit after every m bits. Odd parity ensures that there is a 1 in every $m + 1$ bits. A Manchester signal is a 1B1P signal. The 24B1P line code is being used in the trans-Pacific submarine cable (TPC-3) system.

The m B1C Code. In the coding process, a complementary bit (i.e., complement to the last bit of the block) is inserted at the end of every block of m information bits. The maximum number of consecutive identical symbols is $m + 1$, which occurs when the inserted bit and m succeeding bits are the same. The Manchester code is a degenerate case of the 1B1C code (Bellamy 2000). The code has been adopted by Nippon Telegraph and Telephone (NTT) in its F-1.6G optical system in the form of 10B1C.

CONCLUSION

Based on the coding rules, the simplest of all line codes is the NRZ family. The NRZ signal occupies a narrow bandwidth, but it lacks other desired characteristics such as adequate timing content, error-detection capability, and lower signal energy near low frequencies. Timing content is increased in the RZ-coded signal by allowing more transitions at the expense of the signal bandwidth. The RZ-coded signals, however, still have substantial energy near DC, which also make them unsuitable for AC-coupled circuits. The three-level pseudoternary codes provide the same timing information as RZ-coded signals but in a narrower bandwidth. Their spectra have no DC components, and near-DC components are small, also making the codes suitable for transmission on AC-coupled circuits. Their performance, however, is worse than polar NRZ

Table 2: Line Codes and Digital Systems

Code	Digital system
Non-return to zero	Digital logic and short-haul telecommunication links
Differential NRZ	Magnetic tape recording
Multiline transmission-3	100baseT Ethernet
Alternate mark inversion	Telemetry systems, especially AT&T T1 carrier systems
Binary N zero substitution	DS1 (N = 8), DS2 (N = 6), and DS3 (N = 3)
High-density bipolar 3	European digital hierarchy: E2, E3, and E4
Biphase (1B1C or 1B2B)	Local area network
Differential biphase	IEEE 802.5 token ring
Code mark inversion (1B2B)	139.246-Mbps European digital hierarchy
4B3T	T148 span line
8B6T	100-Mbps Ethernet
2B1Q	Digital subscriber loops
Optical Fiber Systems	
4B5B	100-Mbps Ethernet and FDDI over optical fiber
7B8B	565-Mbps system of British Telecom
8B9B	Pulse coded modulation
24B1P	Trans-Pacific submarine cable (TPC-3)
10B1C	F-1.6 G optical by NTT

signals. Substitution codes can be used in conjunction with pseudoternary codes to alleviate the latter's timing recovery problem if a long string of zeros occurs. The biphase codes occupy the same bandwidth as the RZ-coded signals, but their spectra contain no DC or very small near-DC components. Their error performance is no better than the NRZ-coded signals. The diphase code contains ample timing recovery information because it contains at least one transition per symbol; this property makes it widely used. However, the AMI family of codes is more suitable for band-limited channels and is therefore preferred over diphase. The block line codes add redundancy, which aids in error detection and benefits timing recovery. However, the disadvantage is that the transmission rate is increased and should only be used on channels where bandwidth availability is not an issue. Block codes are not limited to only binary: Multilevel block codes increase the efficiency of the line codes, reduce bandwidth requirements, and can be designed to have a zero or constant DC component. Finally, the binary block codes that are easier to encode and decode—for example, bit-insertion block codes—are more suitable for optical transmission systems because the devices operate in the nonlinear mode.

Table 2 lists digital systems where the aforementioned line codes are being used.

GLOSSARY

Alternate mark inversion (AMI): A bipolar line coding technique that uses three levels—positive, negative, and zero—and is the most widely used technique because it has good spectral characteristics, better error-detection capability, and superior timing recovery ability compared to other line codes.

Binary *N* zero substitution (B*N*ZS): A line coding technique that replaces a string of N zeros by another sequence that is readily detected and replaced at the receiver by an all-zeros sequence.

Biphase codes: A line coding technique that uses half-period pulses with different phases. Manchester or diphase or split phase codes belong to this family.

Bipolar: A line coding technique that uses three levels: positive, negative, and zero.

Codebook: A codebook defines mapping between input and output bit patterns of an encoder.

Code mark inversion (CMI): A line coding technique that replaces the zero level in the AMI waveform by two levels: $+A$ for first half bit duration and $-A$ for the second half bit duration or vice versa.

High-density bipolar *N* (HDB *N*): A line coding technique that limits the number of consecutive zeros to N by replacing the $(N + 1)$th zero by a bipolar violation.

***k*B*n*T codes:** A family of line code that maps k binary digits into n ternary symbols, where $n < k$. AMI is a 1B1T code.

Line coding: The process of converting the binary digits to waveforms that are suitable for transmission over the baseband channel.

***m*B*n*B codes:** This family of line codes converts a block of m binary digits into a block of n binary digits, where $m < n$. CMI and DMI are 1B2B codes.

Non-return to zero (NRZ): A line coding technique in which the signal does not return to zero during the duration of the bit.

Pseudoternary (PT): A line coding technique that uses three-level data symbols to transmit one bit of information. AMI codes also belong to this group and these are also called *bipolar codes*.

Return to zero (RZ): A line coding technique in which the signal returns to zero during the duration of the bit.

Unipolar: A line coding technique that uses only one signal level for bit representation in addition to zero.

CROSS REFERENCES

See *Digital Communications Basics*; *Digital Phase Modulation and Demodulation*; *Optical Fiber Communications*.

REFERENCES

Barry, J. R., E. A. Lee, and D. G. Messerschmitt. 2003. *Digital communication*. New York: Springer.

Bellamy, J. C. 2000. *Digital telephony*. New York: John Wiley & Sons.

Benedetto, S., and E. Biglieri. 1999. *Principles of digital transmission with wireless applications*. Norwell, MA: Kluwer Academic.

Chen, W. Y. 1998. *Simulation techniques and standards development for digital subscriber line systems*. New York: Macmillan Technical Publishing.

Couch, L. W. 2001. *Digital and analog communication systems*. Upper Saddle River, NJ: Prentice Hall.

Haykin, S. 2004. *Communication systems*. New York: John Wiley & Sons.

Line code. Undated. (Retrieved from www.answers.com/topic/line-coding.)

Moloney, A. 2005. Line coding (retrieved from www.electronics.dit.ie/staff/amoloney/dt021_4/lecture-9.pdf).

Simon, M. K., S, M. Hinedi, and W. C. Lindsey. 1995. *Digital modulation techniques: Signal design and detection*. Upper Saddle River, NJ: Prentice Hall.

Xiong, F. 2000. *Digital modulation techniques*. Boston: Artech House Publishers.

FURTHER READING

The following reading list includes the original sources for those who are interested in learning more about the history of line codes.

4B3T Code

Waters, D. B. 1983. Line codes for metallic systems. *International Journal on Electronics*, 55: 159–69.

Binary *N* Zero Substitution

Johannes, V. I., A. G. Kaim, and T. Walzman. 1969. Bipolar pulse transmission with zero extraction. *IEEE Transactions on Communications*, 17: 303–10.

Carter Code

Carter, R. O. 1965. Low disparity binary coding system. *Electronic Letters*, 1(3): 67.

CMI and 2B3B DC-Constrained Code

Takasaki, Y., M. Tanaka, N. Maeda, K. Yamashita, and K. Nagano. 1976. Optical pulse formats for fiber optic digital communications. *IEEE Transactions on Communications*, 24: 404–13.

Dicode

Aaron, M. R. 1962. PCM transmission in the exchange plant. *Bell System Technical Journal*, 14(99): 99–141.

Dicode and High-Density Bipolar *N* Codes

Croisier, A. 1970. Introduction to pseudoternary transmission codes. *IBM Journal of Research Development*, July, pp. 354–67.

*m*B1C Code

Yoshikai, N. K., K. Katagiri, and T. Ito. 1984. *m*B1C code and its performance in an optical communication system. *IEEE Transactions on Communications*, 32(2): 163–8.