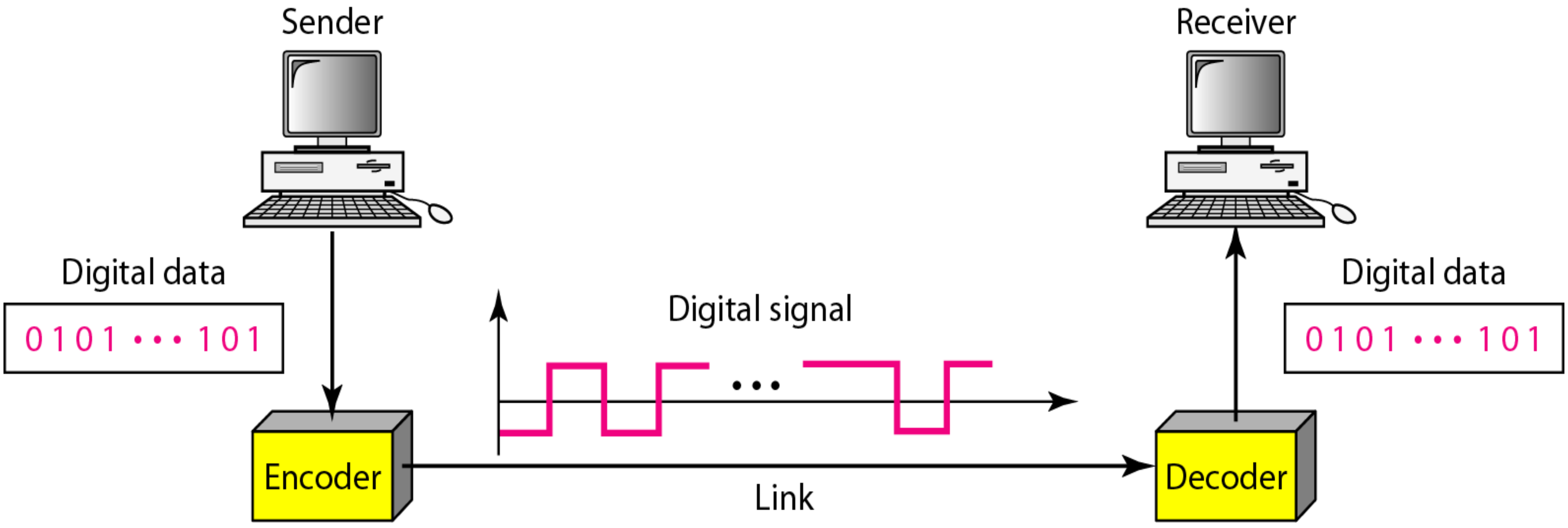
Line Encoding Scheme

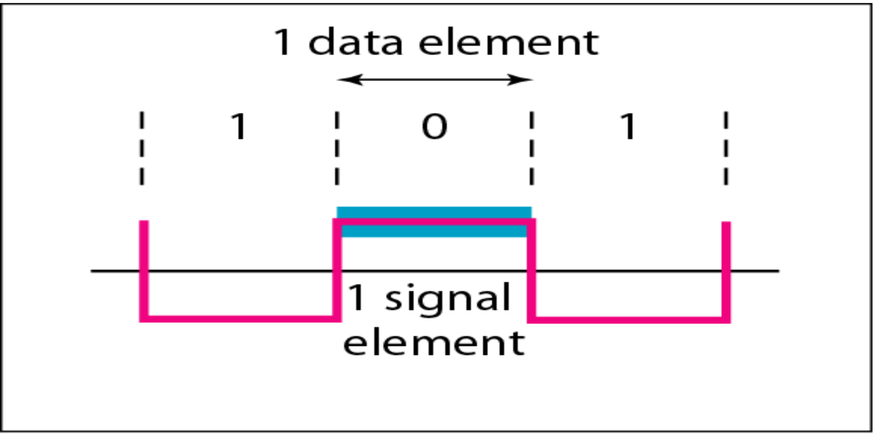
Line Encoding

- Line coding is the process of converting digital data to digital signals.
- We assume that data, in the form of text, numbers, graphical images, audio, or video, are stored in computer memory as sequences of bits.
- Line coding converts a sequence of bits to a digital signal.

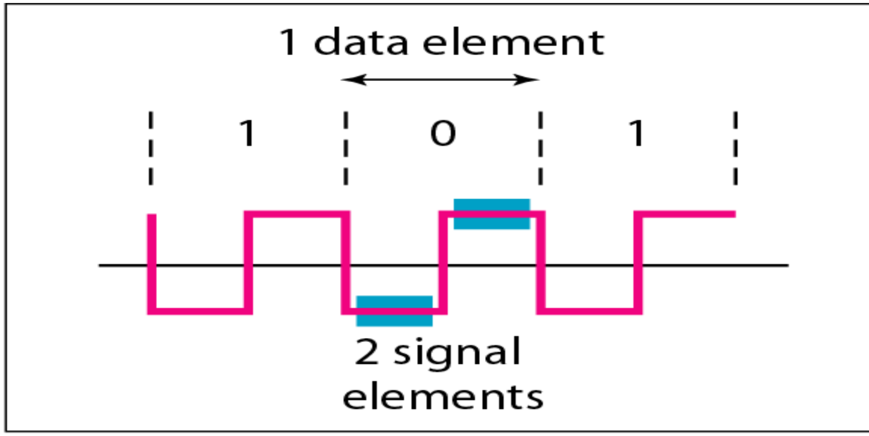
Line Encoding



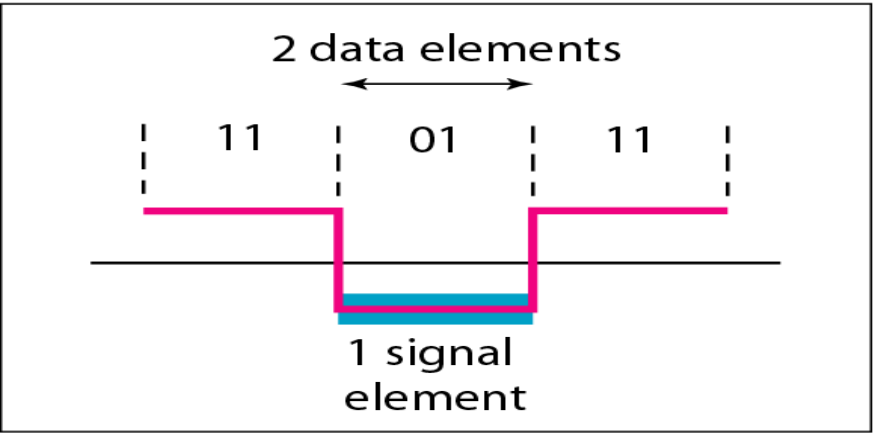
Signal Element Versus Data Element



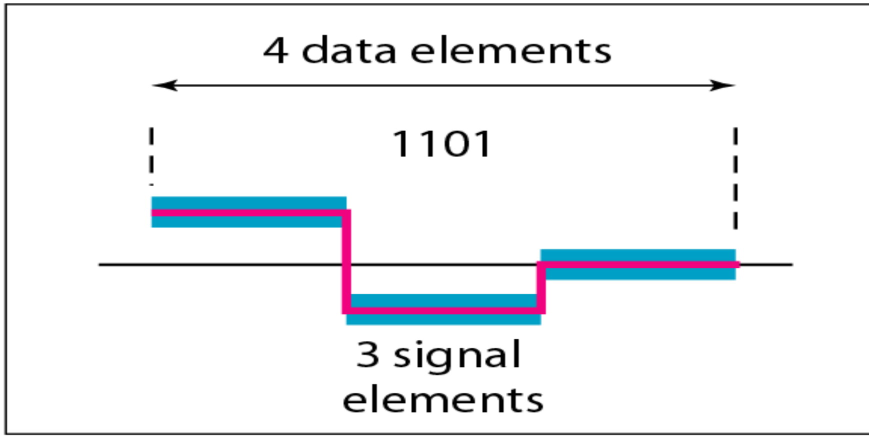
a. One data element per one signal element ($r = 1$)



b. One data element per two signal elements $\left(r = \frac{1}{2}\right)$



c. Two data elements per one signal element ($r = 2$)



d. Four data elements per three signal elements $\left(r = \frac{4}{3}\right)$

Data Rate Versus Signal Rate

- One goal in data communications is to increase the data rate while decreasing the signal rate.
- Increasing the data rate increases the speed of transmission; decreasing the signal rate decreases the bandwidth requirement.
- In our vehicle-people analogy, we need to carry more people in fewer vehicles to prevent traffic jams.
- We have a limited bandwidth in our transportation system.

Example

Question: A signal is carrying data in which one data element is encoded as one signal element (r = 1). If the bit rate is 100 kbps, what is the average value of the baud rate if c is between 0 and 1?

Solution: We assume that the average value of c is 1/2 . The baud rate is then

$$\boxed{S = c \times N \times \frac{1}{r} = \frac{1}{2} \times 100{,}000 \times \frac{1}{1} = 50{,}000 = 50 \text{ kbaud}}$$

Bandwidth

- The bandwidth is theoretically infinite, but many of the components have such a small amplitude that they can be ignored. The effective bandwidth is finite.
- In other words, although the actual bandwidth of a digital signal is infinite, the effective bandwidth is finite.
- We can say that the bandwidth (range of frequencies) is proportional to the signal rate (baud rate). The minimum bandwidth can be given as

$$\mathbf{B_{min} = c \times N \times (1/r)}$$

Example

Question: The maximum data rate of a channel is Nmax = $2 \times B \times \log_2 L$ (defined by the Nyquist formula). Does this agree with the previous formula for N_{max}?

Solution: A signal with L levels actually can carry $\log_2 L$ bits per level. If each level corresponds to one signal element and we assume the average case ($c = 1/2$), then we have

$$\boxed{N_{max} = \frac{1}{c} \times B \times r = 2 \times B \times \log_2 L}$$

Baseline Wandering

- In decoding a digital signal, the receiver calculates a running average of the received signal power. This average is called the baseline.
- The incoming signal power is evaluated against this baseline to determine the value of the data element.
- A long string of 0s or 1s can cause a drift in the baseline (baseline wandering) and make it difficult for the receiver to decode correctly.
- A good line coding scheme needs to prevent baseline wandering.

DC Components

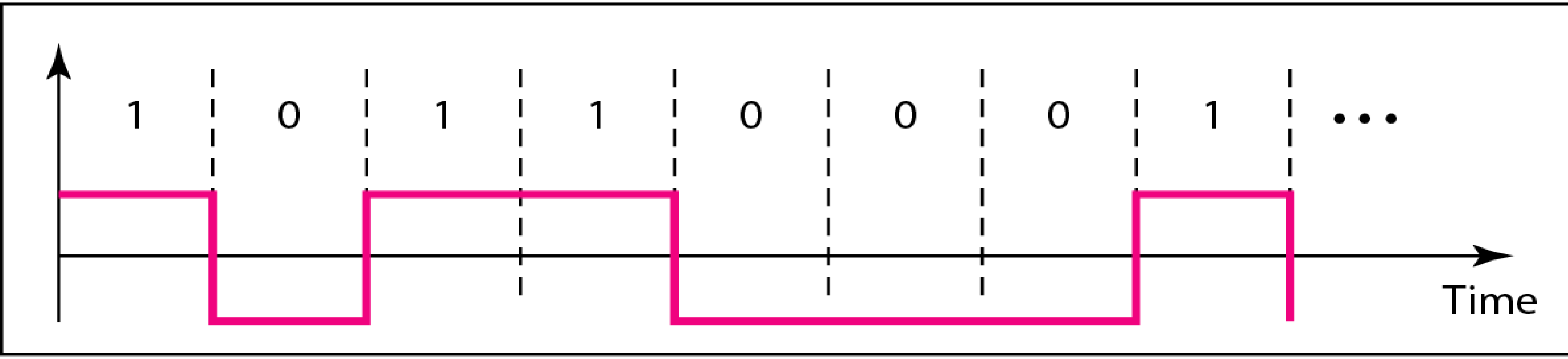
- When the voltage level in a digital signal is constant for a while, the spectrum creates very low frequencies (results of Fourier analysis). These frequencies around zero, called DC (direct-current) components, present problems for a system that cannot pass low frequencies or a system that uses electrical coupling (via a transformer).
- For example, a telephone line cannot pass frequencies below 200 Hz. Also a long-distance link may use one or more transformers to isolate different parts of the line electrically. For these systems, we need a scheme with no DC component.

Self-synchronization

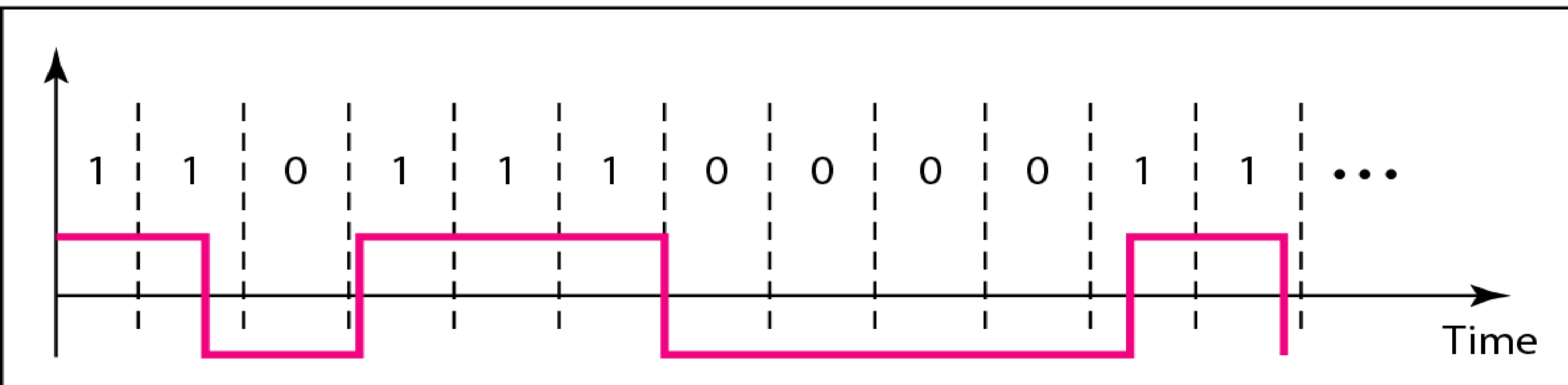
- To correctly interpret the signals received from the sender, the receiver's bit intervals must correspond exactly to the sender's bit intervals.
- If the receiver clock is faster or slower, the bit intervals are not matched and the receiver might misinterpret the signals.

Effect of lack of synchronization

Figure shows a situation in which the receiver has a shorter bit duration. The sender sends 10110001, while the receiver receives 110111000011.



a. Sent



b. Received

Example

Question: In a digital transmission, the receiver clock is 0.1 percent faster than the sender clock. How many extra bits per second does the receiver receive if the data rate is 1 kbps? How many if the data rate is 1 Mbps?

Solution: At 1 kbps, the receiver receives 1001 bps instead of 1000 bps.

1000 bits sent	1001 bits received	1 extra bps

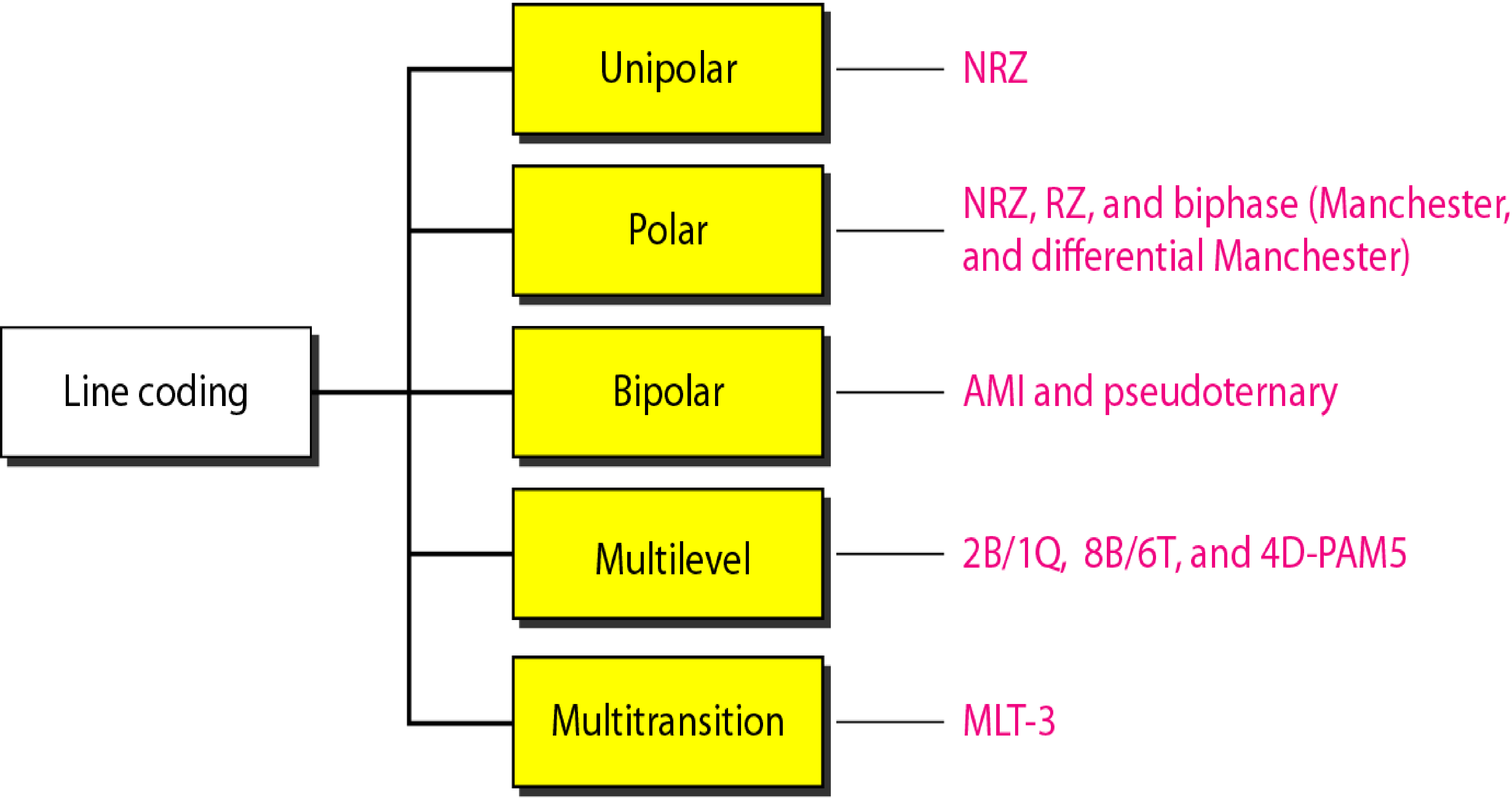
At 1 Mbps, the receiver receives 1,001,000 bps instead of 1,000,000 bps.

1,000,000 bits sent	1,001,000 bits received	1000 extra bps

Line Encoding

- Built-in Error: Detection It is desirable to have a built-in error-detecting capability in the generated code to detect some of or all the errors that occurred during transmission. Some encoding schemes that we will discuss have this capability to some extent.
- Immunity to Noise and Interference: Another desirable code characteristic is a code I that is immune to noise and other interferences. Some encoding schemes that we will discuss have this capability.
- Complexity: A complex scheme is more costly to implement than a simple one. For example, a scheme that uses four signal levels is more difficult to interpret than one that uses only two levels.

Line Encoding Schemes

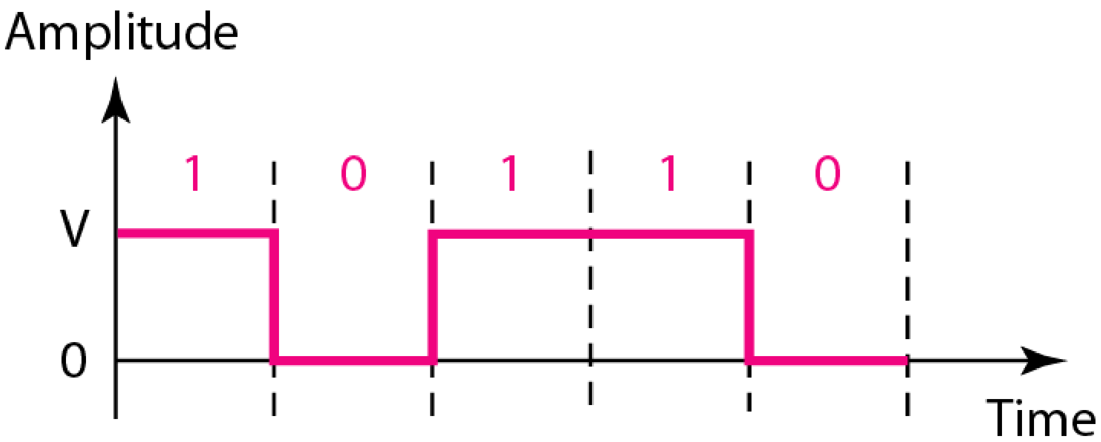


Unipolar Signaling

- Unipolar signaling is also called as **On-Of-Keying** or simply **OOK**.
- The presence of pulse represents a **1** and the absence of pulse represents a **0**.
- There are two variations in Unipolar signaling −

1. Non Return to Zero (NRZ)
2. Return to Zero (RZ)

Unipolar NRZ scheme



$$\frac{1}{2}V^2 + \frac{1}{2}(0)^2 = \frac{1}{2}V^2$$

Normalized power

Unipolar NRZ scheme

- The advantages of Unipolar NRZ are –

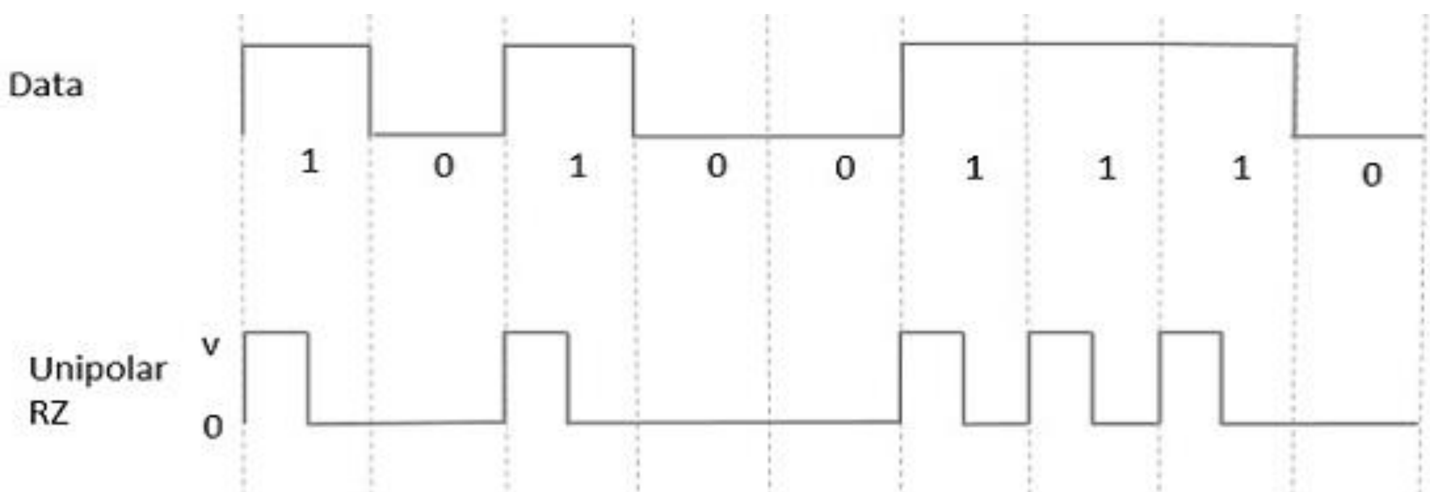
1. It is simple.
2. A lesser bandwidth is required.

- The disadvantages of Unipolar NRZ are –

1. No error correction done.
2. Presence of low frequency components may cause the signal droop.
3. No clock is present.
4. Loss of synchronization is likely to occur (especially for long strings of **1s** and **0s**).

Unipolar RZ scheme

In this type of unipolar signaling, a High in data, though represented by a **Mark pulse**, its duration $\mathbf{T_0}$ is less than the symbol bit duration. Half of the bit duration remains high but it immediately returns to zero and shows the absence of pulse during the remaining half of the bit duration.



Unipolar RZ scheme

- The advantages of Unipolar RZ are −

1. It is simple.
2. The spectral line present at the symbol rate can be used as a clock.

- The disadvantages of Unipolar RZ are −

1. No error correction.
2. Occupies twice the bandwidth as unipolar NRZ.
3. The signal droop is caused at the places where signal is non-zero at 0 Hz.

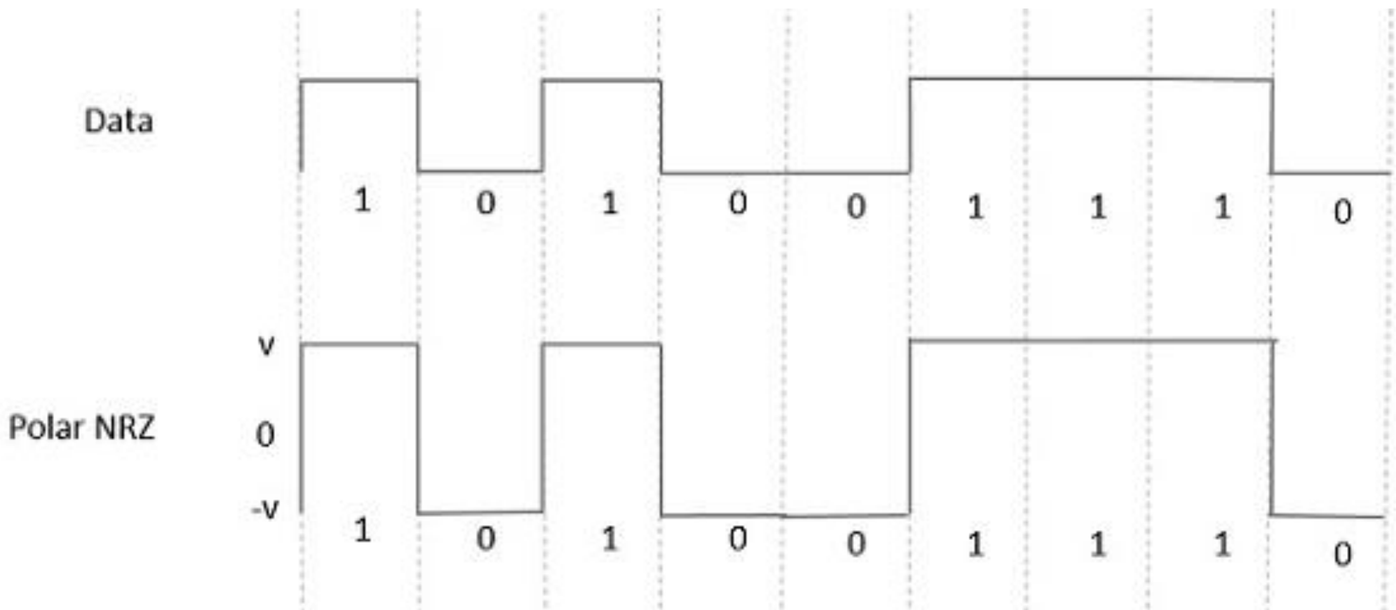
Polar Signaling

- In polar schemes, the voltages are on the both sides of the time axis. For example, the voltage level for 0 can be positive and the voltage level for I can be negative.
- There are two methods of Polar Signaling. They are −

1. Polar NRZ
2. Polar RZ

Polar NRZ

In this type of Polar signaling, a High in data is represented by a positive pulse, while a Low in data is represented by a negative pulse. The following figure depicts this well.



Polar NRZ

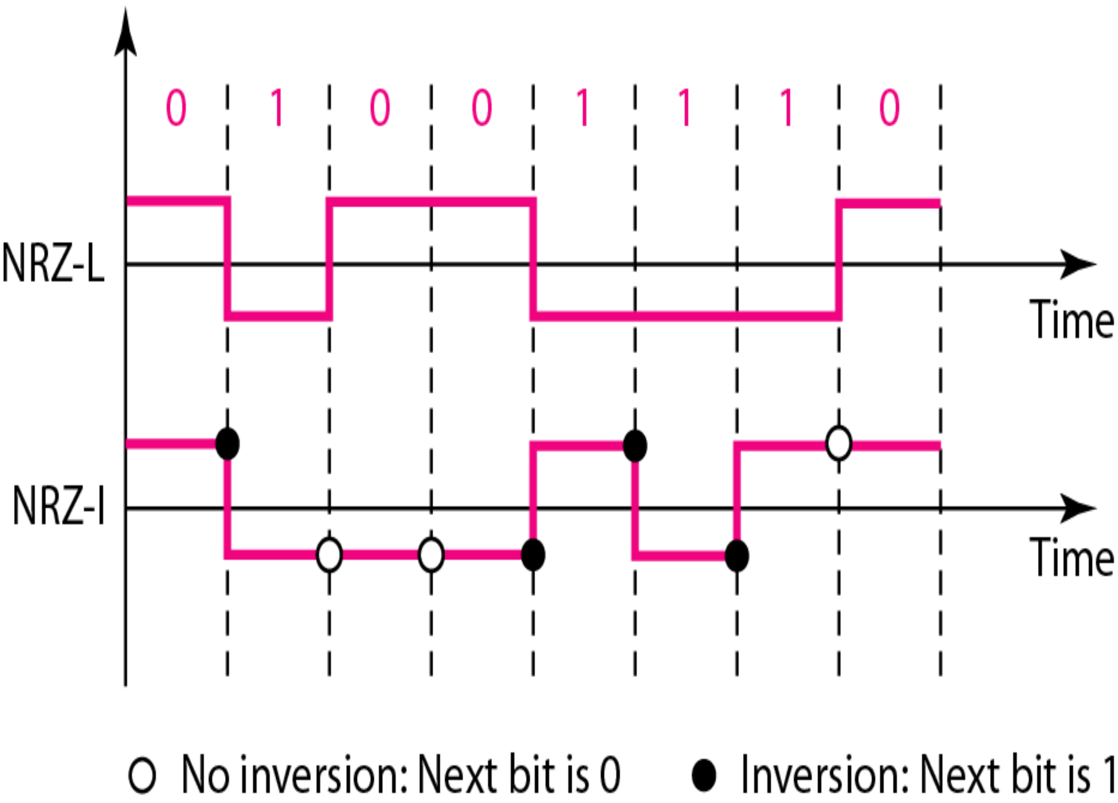
- The advantages of Polar NRZ are –

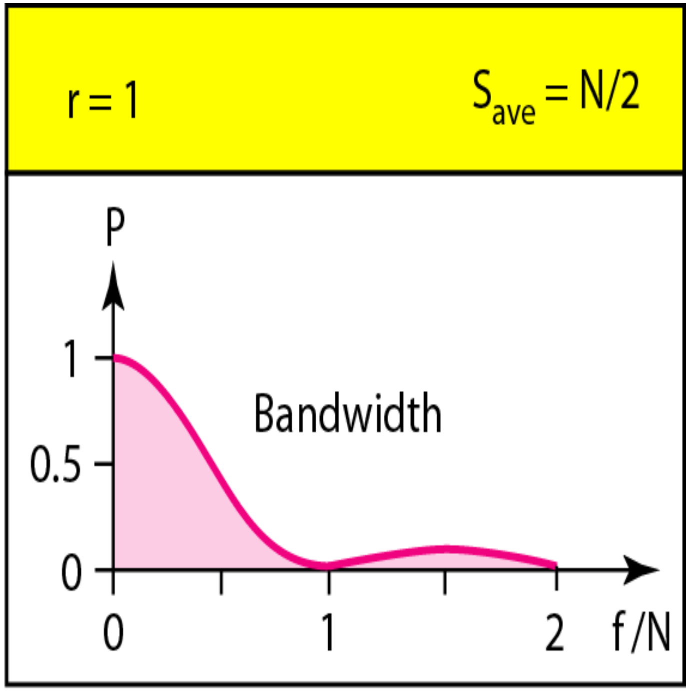
1. It is simple.
2. No low-frequency components are present.

- The disadvantages of Polar NRZ are –

1. No error correction.
2. No clock is present.
3. The signal droop is caused at the places where the signal is non-zero at **0 Hz**.

Polar NRZ-L and NRZ-I schemes

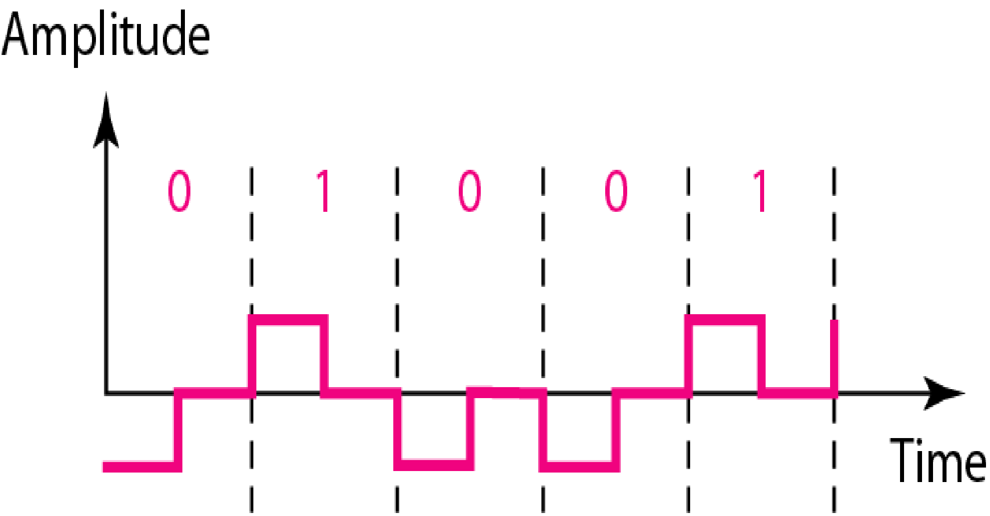


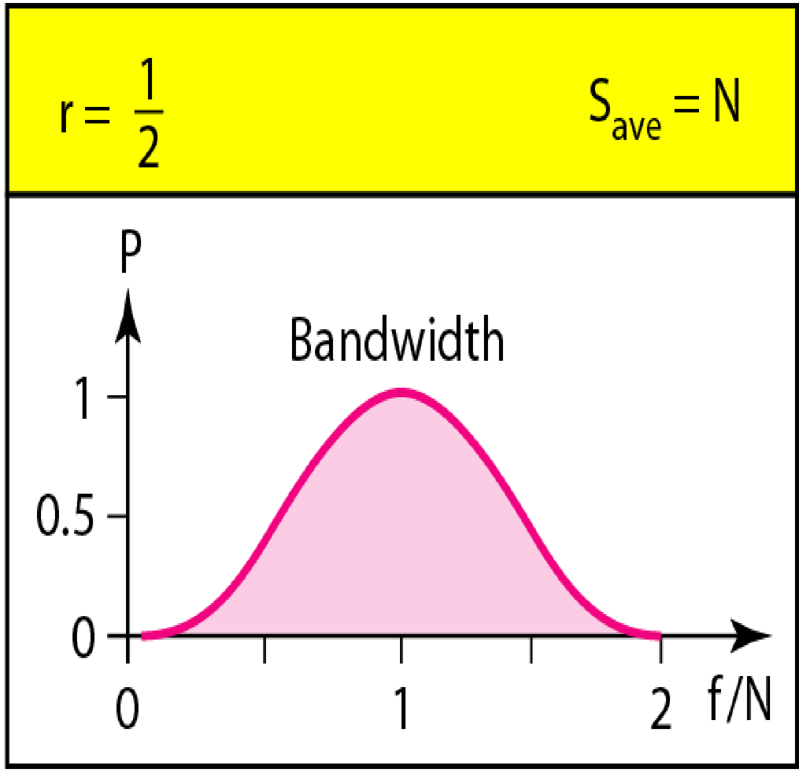


Polar RZ

- In this type of Polar signaling, a High in data, though represented by a Mark pulse, its duration T0 is less than the symbol bit duration. Half of the bit duration remains high but it immediately returns to zero and shows the absence of pulse during the remaining half of the bit duration.
- However, for a Low input, a negative pulse represents the data, and the zero level remains same for the other half of the bit duration. The following figure depicts this clearly.

Polar RZ





Polar RZ

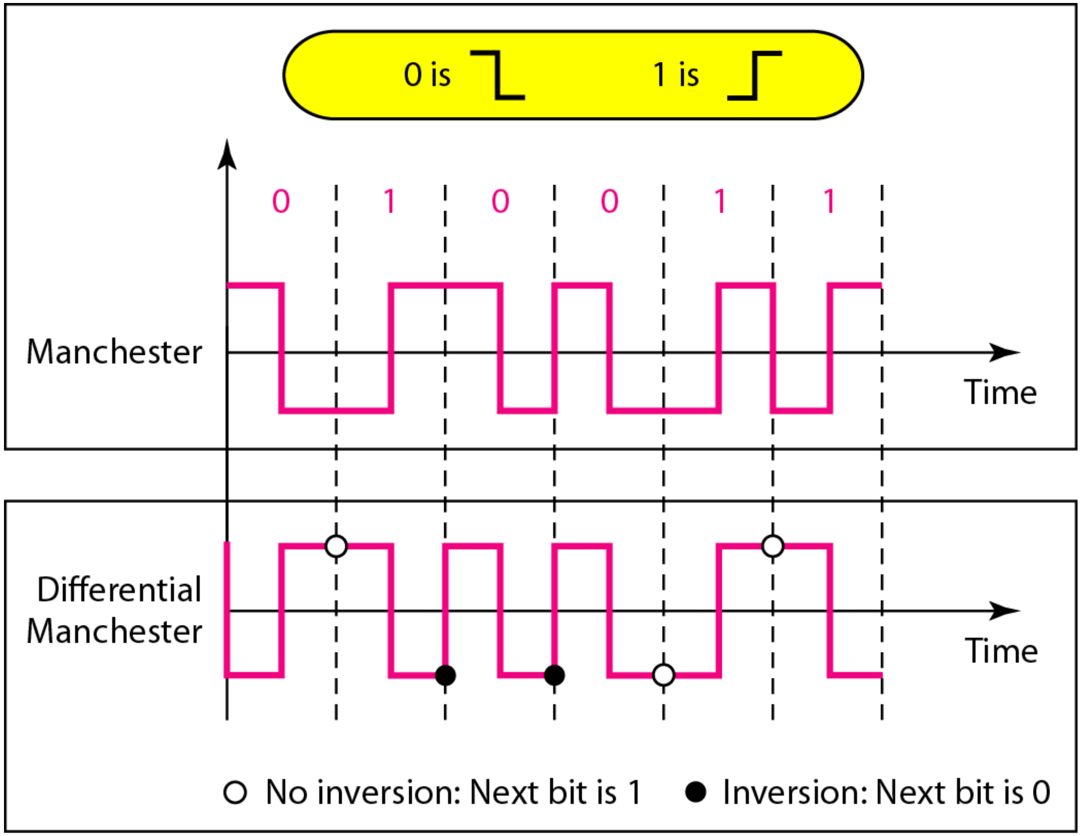
- The advantages of Polar RZ are −

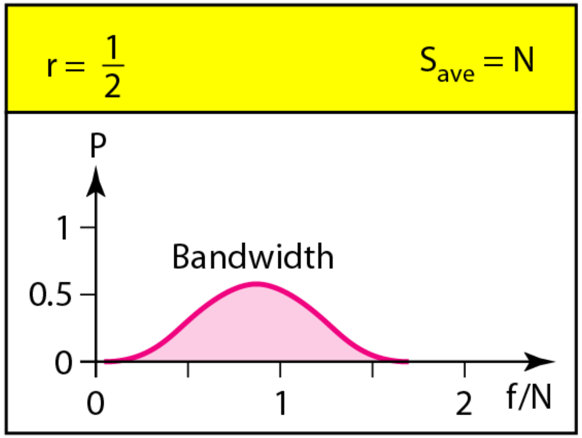
1. It is simple.
2. No low-frequency components are present.

- The disadvantages of Polar RZ are −

1. No error correction.
2. No clock is present.
3. Occupies twice the bandwidth of Polar NRZ.
4. The signal droop is caused at places where the signal is non-zero at 0 Hz.

Polar biphase: Manchester and differential Manchester schemes

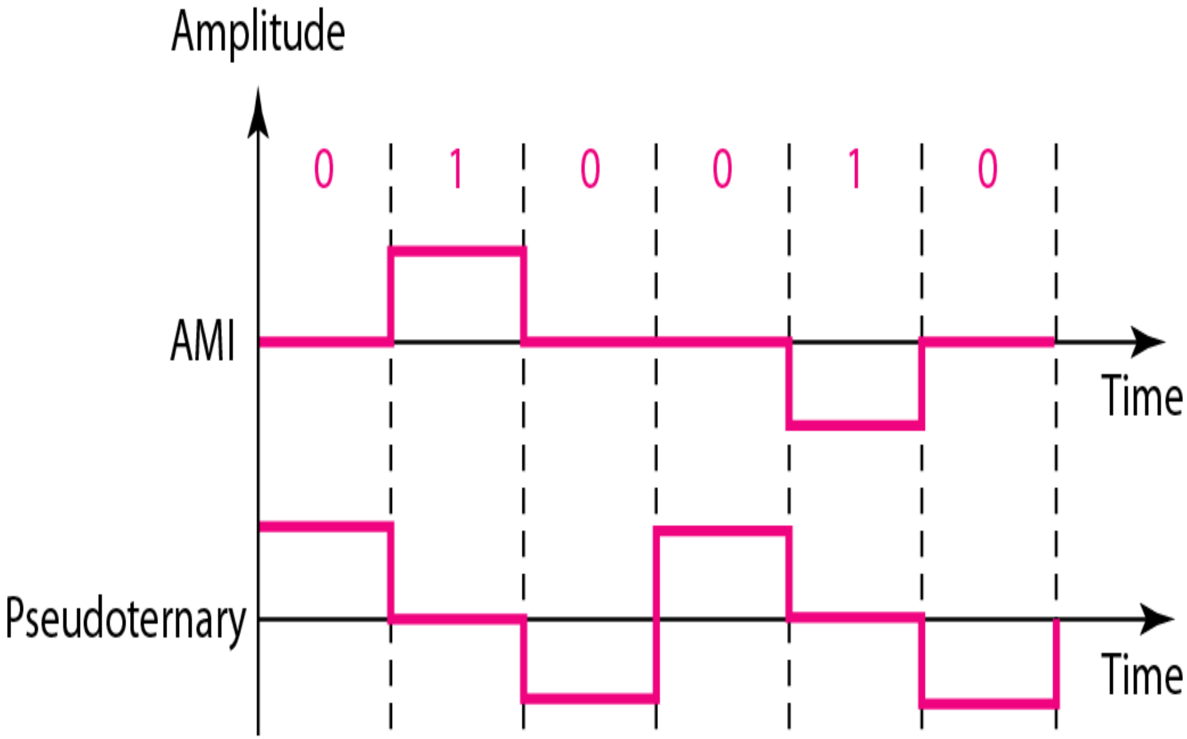


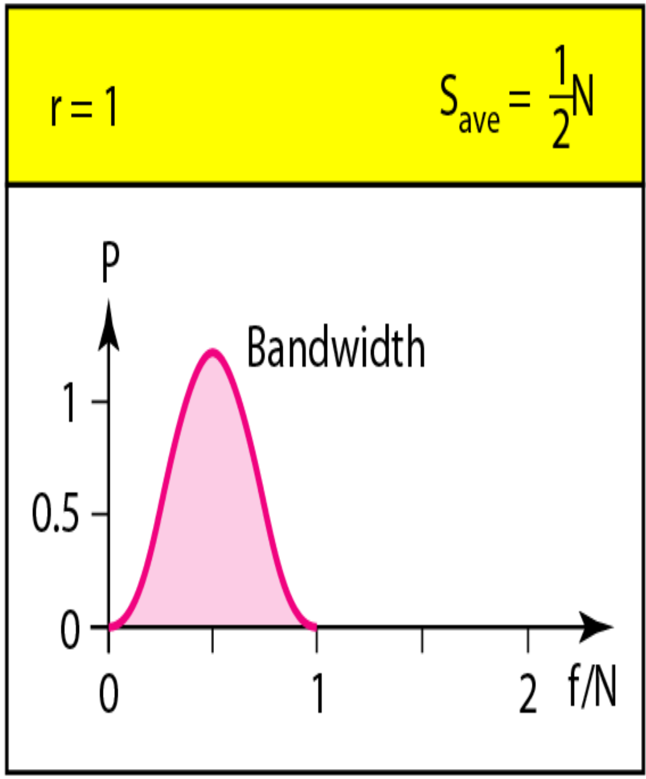


Bipolar Line Encoding

- In bipolar encoding (sometimes called multilevel binary), there are three voltage levels: positive, negative, and zero. The voltage level for one data element is at zero, while the voltage level for the other element alternates between positive and negative.
- An example of this type is Alternate Mark Inversion AMI.
- For a 1, the voltage level gets a transition from + to – or from – to +, having alternate 1s to be of equal polarity. A 0 will have a zero voltage level.

Bipolar schemes: AMI and pseudoternary





Bipolar Line Encoding

- Following are the advantages −

1. It is simple.
2. No low-frequency components are present.
3. Occupies low bandwidth than unipolar and polar NRZ schemes.
4. This technique is suitable for transmission over AC coupled lines, as signal drooping doesn’t occur here.
5. A single error detection capability is present in this.

- Following are the disadvantages −

1. No clock is present.
2. Long strings of data causes loss of synchronization.

Multilevel Schemes

- The desire to increase the data speed or decrease the required bandwidth has resulted in the creation of many schemes.
- The goal is to increase the number of bits per baud by encoding a pattern of m data elements into a pattern of n signal elements.
- We only have two types of data elements (0s and 1s), which means that a group of m data elements can produce a combination of 2^m data patterns.
- We can have different types of signal elements by allowing different signal levels.
- If we have L different levels, then we can produce Ln combinations of signal patterns.

Multilevel Schemes

- If $2^m = L^n$, then each data pattern is encoded into one signal pattern.

- If $2^m < L^n$, data patterns occupy only a subset of signal patterns.

- The subset can be carefully designed to prevent baseline wandering, to provide synchronization, and to detect errors that occurred during data transmission.

- Data encoding is not possible if $2^m > L^n$ because some of the data patterns cannot be encoded.

- The code designers have classified these types of coding as mBnL, where m is the length of the binary pattern, B means binary data, n is the length of the signal pattern, and L is the

Multilevel Schemes

- The code designers have classified these types of coding as mBnL, where m is the length of the binary pattern, B means binary data, n is the length of the signal pattern, and L is the number of levels in the signaling.

- A letter is often used in place of L: B (binary) for L = 2, T (ternary) for L = 3, and Q (quaternary) for L = 4.

- Note that the first two letters define the data pattern, and the second two define the signal pattern.

- In mBnL schemes, a pattern of m data elements is encoded as a pattern of n signal elements in which $2^m \leq L^n$.

2B1Q

- The first mBnL scheme we discuss, two binary, one quaternary (2B1Q), uses data patterns of size 2 and encodes the 2-bit patterns as one signal element belonging to a four-level signal.
- In this type of encoding m =2, n =1, and L =4 (quaternary).

	Previous level: positive	Previous level: negative
Next bits	**Next level**	**Next level**
00	+1	-1
01	+3	-3
10	-1	+1
11	-3	+3

Transition table

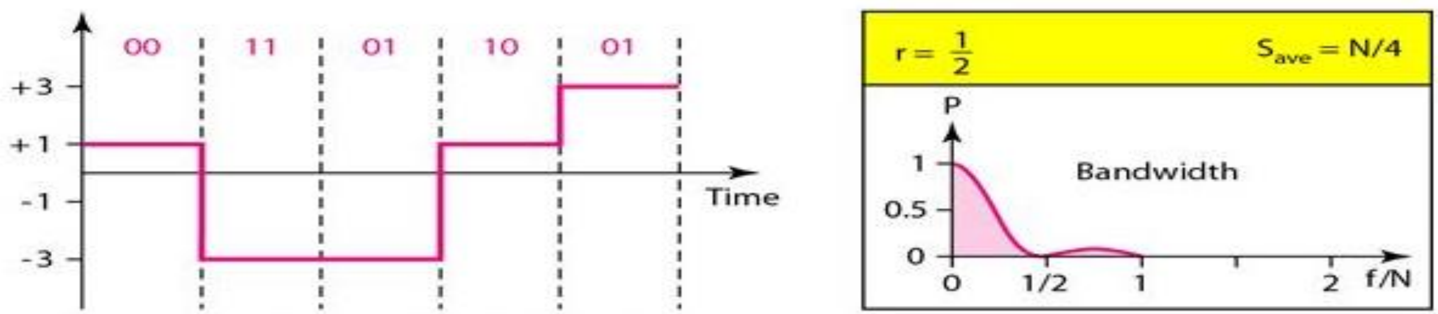


Fig. Multilevel: 2B1Q scheme

2B1Q

- The average signal rate of 2BlQ is S =N/4. This means that using 2B1Q, we can send data 2 times faster than by using NRZ-L.

- However, 2BlQ uses four different signal levels, which means the receiver has to discern four different thresholds. The reduced bandwidth comes with a price.

- There are no redundant signal patterns in this scheme because $2^2 = 4^1$.

8B6T

- A very interesting scheme is eight binary, six ternary (8B6T). This code is used with 100BASE-4T cable.

- The idea is to encode a pattern of 8 bits as a pattern of 6 signal elements, where the signal has three levels (ternary).

- In this type of scheme, we can have 2^8 =256 different data patterns and 3^6 =478 different signal patterns.

- There are 478 - 256 =222 redundant signal elements that provide synchronization and error detection.

- Part of the redundancy is also used to provide **DC balance**. Each signal pattern has a weight of 0 or +1 DC values.

8B6T:

- This means that there is no pattern with the weight -1.
- To make the whole stream DC-balanced, the sender keeps track of the weight.
- If two groups of weight 1 are encountered one after another, the first one is sent as is, while the next one is totally inverted to give

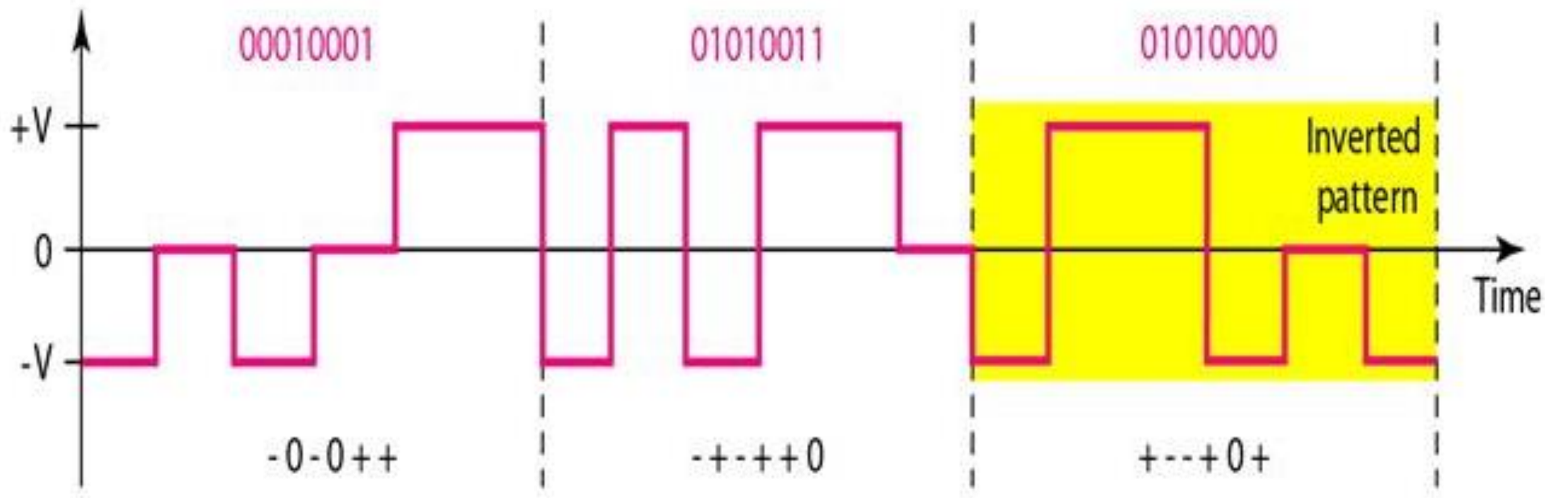


Fig. Multilevel: 8B6T scheme

- Figure shows an example of three data patterns encoded as three signal patterns.
- The three possible signal levels are represented as -,0, and +.
- The first 8-bit pattern 00010001 is encoded as the signal pattern -0-0++ with weight 0; the second 8-bit pattern 010 10011 is encoded as - + - + + 0 with weight +1. The third bit pattern should be encoded as + - - + 0 + with weight +1.
- To create DC balance, the sender inverts the actual signal. The receiver can easily recognize that this is an inverted pattern because the weight is -1. The pattern is inverted before decoding.
- The average signal rate of the scheme is theoretically Save = (1/2) x N x (6/8) ; in practice the minimum bandwidth is very close to 6N18.

4D-PAM5:

- Four dimensional five-level pulse amplitude modulation (4D-PAM5).
- The 4D means that data is sent over four wires at the same time.
- It uses five voltage levels, such as -2, -1, 0, 1, and 2. However, one level, level 0, is used only for forward error detection.
- If we assume that the code is just one-dimensional, the four levels create something similar to 8B4Q. In other words, an 8-bit word is translated to a signal element of four different levels.
- The worst signal rate for this imaginary one-dimensional version is N x 4/8, or N/2.
- The technique is designed to send data over four channels (four

4D-PAMS:

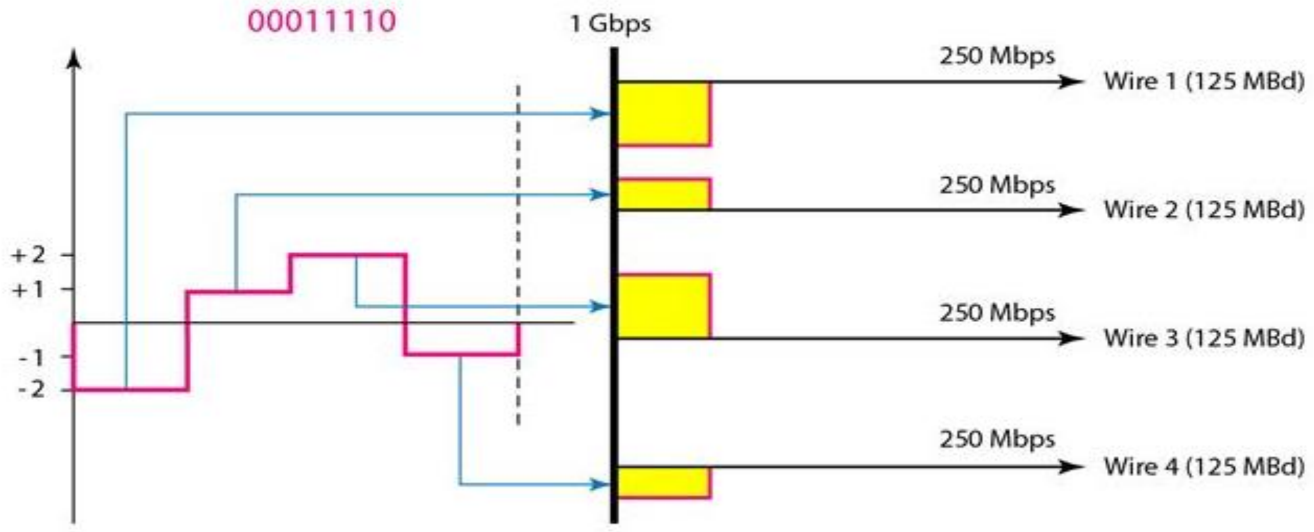


Fig. Multilevel: 4D-PAM5 scheme

- Figure shows the imaginary one-dimensional and the actual four-dimensional implementation. Gigabit LANs use this technique to send 1-Gbps data over four copper cables that can handle 125 Mbaud. This scheme has a lot of redundancy in the signal pattern because 2^8 data patterns are matched to $4^4 = 256$ signal patterns. The extra signal patterns can be used for other purposes such as error detection.

Multiline Transmission: MLT-3

- NRZ-I and differential Manchester are classified as differential encoding but use two transition rules to encode binary data (no inversion, inversion).

- If we have a signal with more than two levels, we can design a differential encoding scheme with more than two transition rules. MLT-3 is one of them.

- The multiline transmission, three level (MLT-3) scheme uses three levels (+v, 0, and - V) and three transition rules to move between the levels.

1. If the next bit is 0, there is no transition.
2. If the next bit is 1 and the current level is not 0, the next level is 0.
3. If the next bit is 1 and the current level is 0, the next level

Multiline Transmission: MLT-3

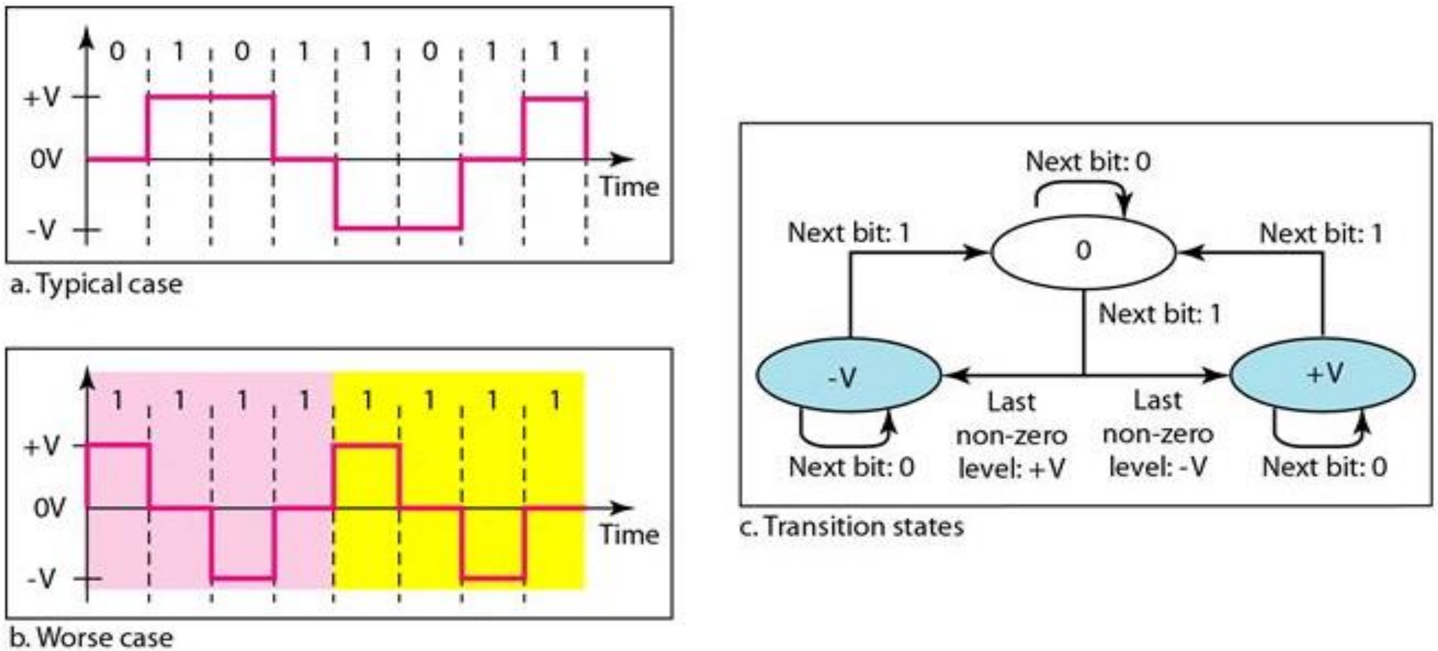


Fig. Multitransition: MLT-3 scheme

- The behavior of MLT-3 can best be described by the state diagram shown in fig.

- The three voltage levels (-V, 0, and +V) are shown by three states (ovals). The transition from one state (level) to another is

Summary of Line Coding Schemes

Category	*Scheme*	*Bandwidth (average)*	*Characteristics*
Unipolar	NRZ	$B=N/2$	Costly, no self-synchronization iflong 0s or Is, DC
Unipolar	NRZ-L	$B=N/2$	No self-synchronization if long 0s or 1s, DC
	NRZ-I	$B=N/2$	No self-synchronization for long aS, DC
	Biphase	$B=N$	Self-synchronization, no DC, high bandwidth
Bipolar	AMI	$B=NI2$	No self-synchronization for long os, DC
Multilevel	2B1Q	$B=N/4$	No self-synchronization for long same double bits
	8B6T	$B=3N/4$	Self-synchronization, no DC
	4D-PAM5	$B=N/8$	Self-synchronization, no DC
Multiline	MLT-3	$B=N/3$	No self-synchronization for long 0s