Unit 8
Physical Layer


Les opératrices d'un standard multiple dans les années 1910.

Bibliography

[Stallings04] Chapters 3.2, 3.3, 5.1 and 5.2

Physical Layer : outline

1. Concepts and Terminology
2. Transmission Impairments
3. Signal Encoding Techniques

Physical Layer : outline

1. Concepts and Terminology
2. Transmission Impairments
3. Signal Encoding Techniques

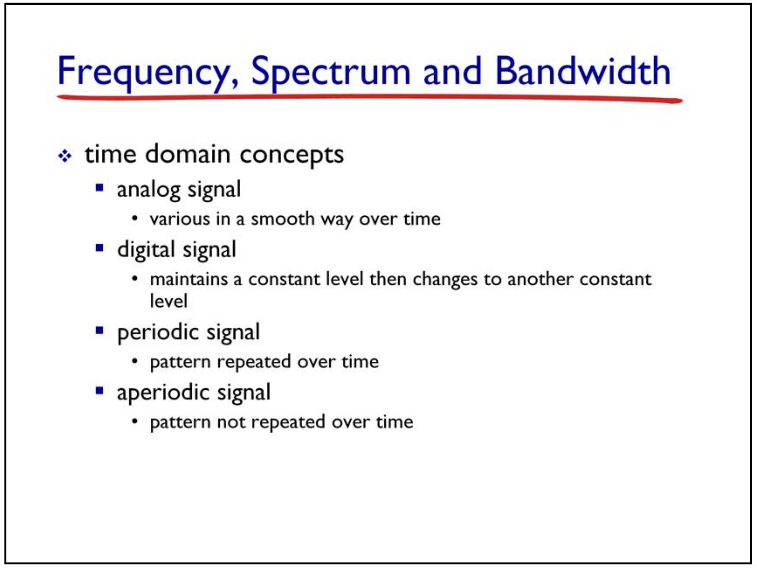
Physical layer

- The physical layer is responsible for transmitting a bit stream over a physical medium.
- It encodes and decodes bits into groups of bits.
- It then transforms a stream of bits into a signal.

Transmission Terminology

- data transmission occurs between a transmitter & receiver via some medium
- guided medium
  - eg. twisted pair, coaxial cable, optical fiber
- unguided / wireless medium
  - eg. air, water, vacuum

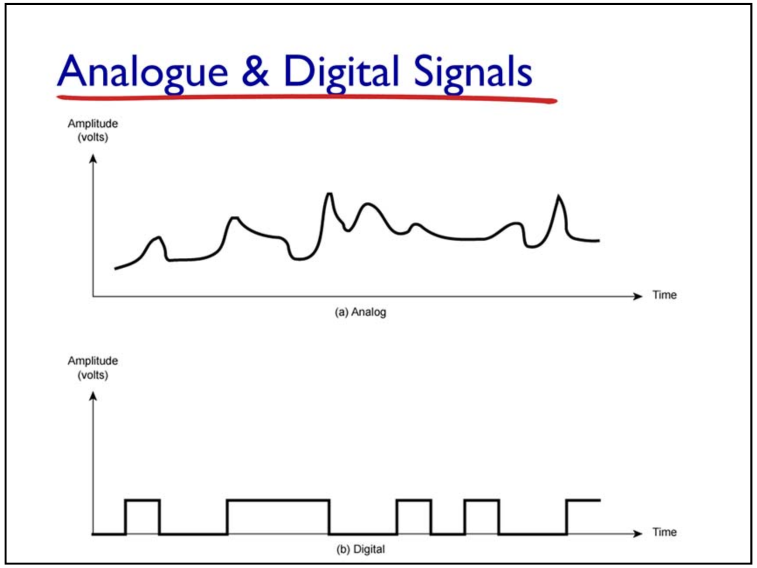
The first section presents some concepts and terms from the field of electrical engineering. This should provide sufficient background to deal with the remainder of the chapter. Data transmission occurs between transmitter and receiver over some transmission medium. Transmission media may be classified as guided or unguided. In both cases, communication is in the form of electromagnetic waves. With **guided media**, the waves are guided along a physical path; examples of guided media are twisted pair, coaxial cable, and optical fiber. **Unguided media**, also called **wireless**, provide a means for transmitting electromagnetic waves but do not guide them; examples are propagation through air, vacuum, and seawater.



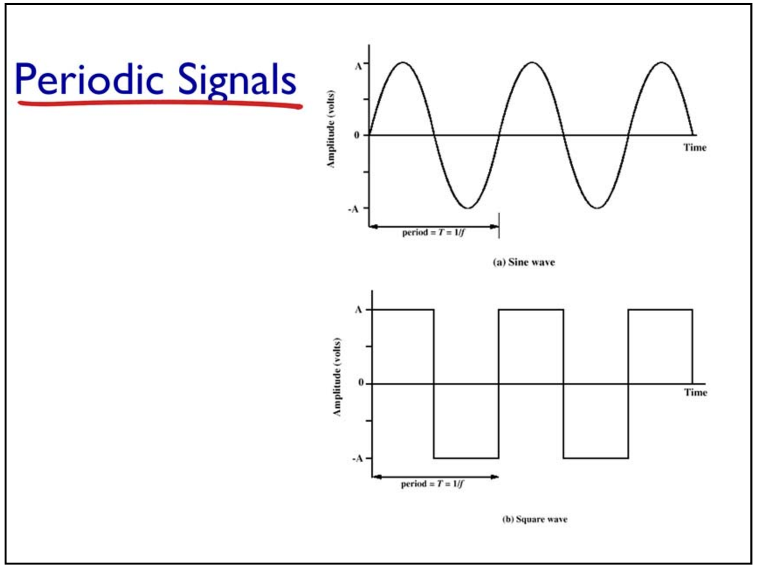
In this book, we are concerned with electromagnetic signals used as a means to transmit data. The signal is a function of time, but it can also be expressed as a function of frequency; that is, the signal consists of components of different frequencies. It turns out that the **frequency domain** view of a signal is more important to an understanding of data transmission than a **time domain** view.

Viewed as a function of time, an electromagnetic signal can be either analog or digital. An **analog signal** is one in which the signal intensity varies in a smooth fashion over time. A **digital signal** is one in which the signal intensity maintains a constant level for some period of time and then abruptly changes to another constant level. This is an idealized definition. In fact, the transition from one voltage level to another will not be instantaneous, but there will be a small transition period.

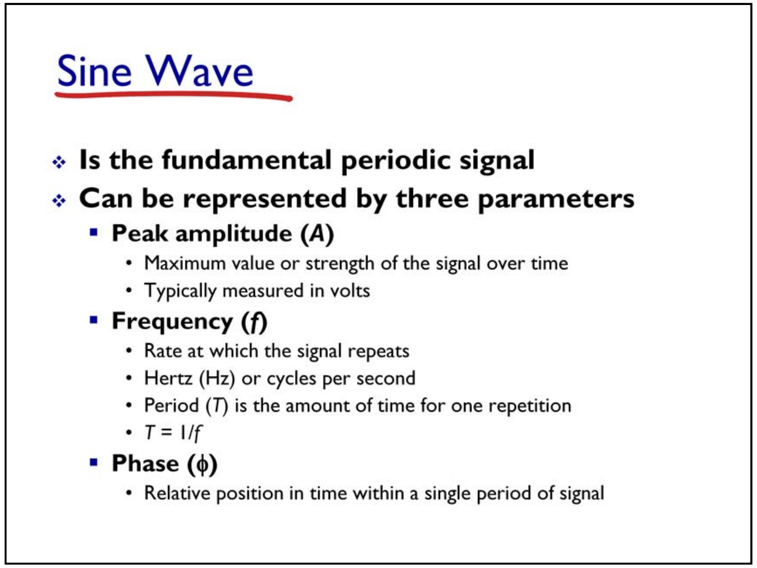
The simplest sort of signal is a **periodic signal**, in which the same signal pattern repeats over time. Otherwise, a signal is **aperiodic**.



Stallings DCC8e Figure 3.1 shows an example of both analog or digital signals. The continuous signal might represent speech, and the discrete signal might represent binary 1s and 0s.



In Stallings DCC8e Figure 3.2 a signal is generated by the transmitter and transmitted over a medium. The signal is a function of time, but it can also be expressed as a function of frequency; that is, the signal consists of components of different frequencies. It turns out that the **frequency domain** view of a signal is more important to an understanding of data transmission than a **time domain** view. Both views are introduced here.



The sine wave is the fundamental periodic signal. A general sine wave can be represented by three parameters: peak amplitude (A), frequency (f), and phase (ϕ). The peak amplitude is the maximum value or strength of the signal over time; typically, this value is measured in volts. The frequency is the rate [in cycles per second, or hertz (Hz)] at which the signal repeats. An equivalent parameter is the period (T) of a signal, which is the amount of time it takes for one repetition; therefore, $T = 1/f$. Phase is a measure of the relative position in time within a single period of a signal, as is illustrated subsequently.

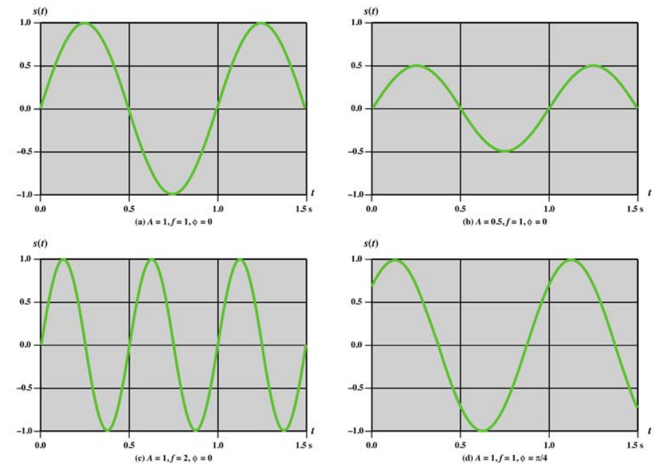


Figure 3.3 $s(t) = A \sin(2\pi ft + \phi)$

The general sine wave can be written $s(t) = A \sin(2\pi ft + f)$

A function with the form of the preceding equation is known as a **sinusoid**. Figure 3.3 shows the effect of varying each of the three parameters. In part (a) of the figure, the frequency is 1 Hz; thus the period is $T = 1$ second. Part (b) has the same frequency and phase but a peak amplitude of 0.5. In part (c) we have $f = 2$, which is equivalent to $T = 0.5$. Finally, part (d) shows the effect of a phase shift of $\pi/4$ radians, which is 45 degrees (2π radians = 360˚ = 1 period).

In Figure 3.3, the horizontal axis is time; the graphs display the value of a signal at a given point in space as a function of time. These same graphs, with a change of scale, can apply with horizontal axes in space. In this case, the graphs display the value of a signal at a given point in time as a function of distance. For example, for a sinusoidal transmission (e.g., an electromagnetic radio wave some distance from a radio antenna, or sound some distance from a loudspeaker), at a particular instant of time, the intensity of the signal varies in a sinusoidal way as a function of distance from the source.

Frequency Domain Concepts

- Signals are made up of many frequencies
- Components are sine waves
- Fourier analysis can show that any signal is made up of components at various frequencies, in which each component is a sinusoid
- Can plot frequency domain functions

In practice, an electromagnetic signal will be made up of many frequencies. For example, the signal

$s(t) = [(4/\pi) \times (\sin(2\pi ft) + (1/3) \sin(2\pi(3f)t)]$

is shown in Figure 3.4c. The components of this signal are just sine waves of frequencies f and $3f$; parts (a) and (b) of the figure show these individual components. There are two interesting points that can be made about this figure:

The second frequency is an integer multiple of the first frequency. When all of the frequency components of a signal are integer multiples of one frequency, the latter frequency is referred to as the **fundamental frequency**. Each multiple of the fundamental frequency is referred to as a **harmonic frequency** of the signal.

The period of the total signal is equal to the period of the fundamental frequency. The period of the component sin $(2\pi ft)$ is $T = 1/f$, and the period of s(t) is also T, as can be seen from Figure 3.4c.

It can be shown, using a discipline known as Fourier analysis, that any signal is made up of components at various frequencies, in which each component is a sinusoid. By adding together enough sinusoidal signals, each with the appropriate amplitude, frequency, and phase, any electromagnetic signal can be constructed. Put another way, any electromagnetic signal can be shown to consist of a collection of periodic analog signals (sine waves) at different amplitudes, frequencies, and phases. The importance of being able to look at a signal from the frequency perspective (frequency domain) rather than a time perspective (time domain) should become clear as the discussion proceeds. For the interested reader, the subject of Fourier analysis is introduced in Appendix A.

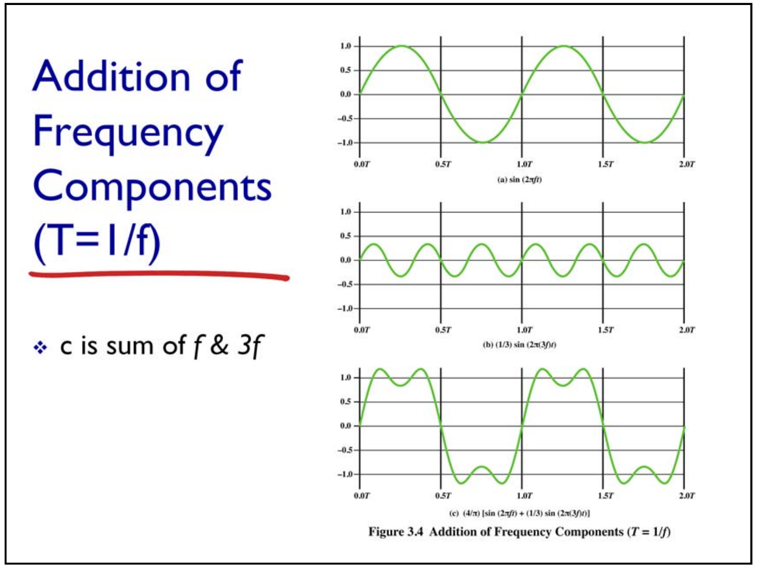
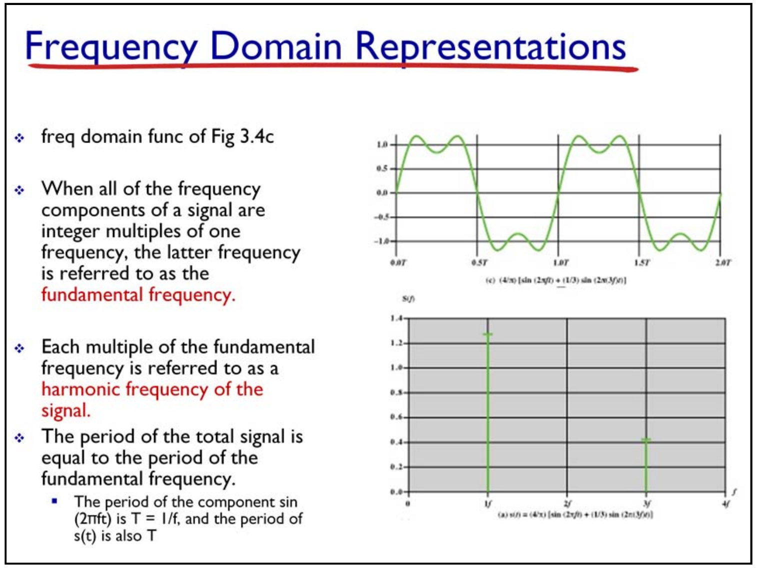


Figure 3.4c, the components of this signal are just sine waves of frequencies *f* and 3f, as shown in parts (a) and (b).



So we can say that for each signal, there is a time domain function $s(t)$ that specifies the amplitude of the signal at each instant in time. Similarly, there is a frequency domain function $S(f)$ that specifies the peak amplitude of the constituent frequencies of the signal. Figure 3.5a shows the frequency domain function for the signal of Figure 3.4c. Note that, in this case, $S(f)$ is discrete.

Spectrum & Bandwidth

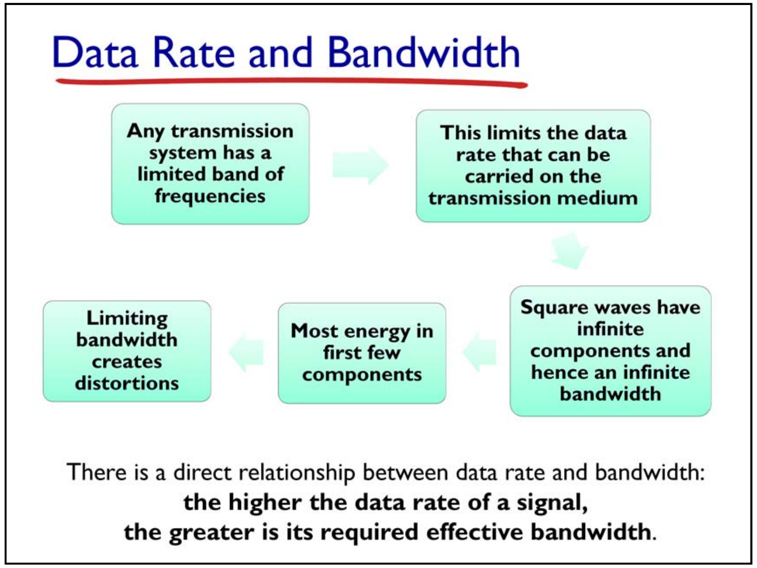
- spectrum
  - range of frequencies contained in signal
- absolute bandwidth
  - width of spectrum
- effective bandwidth
  - often just *bandwidth*
  - narrow band of frequencies containing most energy
- DC (Direct Current) Component
  - component of zero frequency

The **spectrum** of a signal is the range of frequencies that it contains. For Stallings DCC8e Fig 3.4c, it extends from *f* to 3*f*.

The **absolute bandwidth** of a signal is the width of the spectrum (eg is 2f in Fig 3.4c. Many signals, such as that of Figure 3.5b, have an infinite bandwidth.

Most of the energy in the signal is contained in a relatively narrow band of frequencies known as the **effective bandwidth**, or just **bandwidth**.

If a signal includes a component of *zero frequency, it is a direct current (dc) or constant* component.



Although a given waveform may contain frequencies over a very broad range, as a practical matter any transmission system (transmitter plus medium plus receiver) will be able to accommodate only a limited band of frequencies. This, in turn, limits the data rate that can be carried on the transmission medium.

A square wave has an infinite number of frequency components and hence an infinite bandwidth. However, the peak amplitude of the kth frequency component, kf, is only $1/k$, so most of the energy in this waveform is in the first few frequency components.

In general, any digital waveform will have infinite bandwidth. If we attempt to transmit this waveform as a signal over any medium, the transmission system will limit the bandwidth that can be transmitted. For any given medium, the greater the bandwidth transmitted, the greater the cost. The more limited the bandwidth, the greater the distortion, and the greater the potential for error by the receiver.

There is a direct relationship between data rate and bandwidth: the higher the data rate of a signal, the greater is its required effective bandwidth.

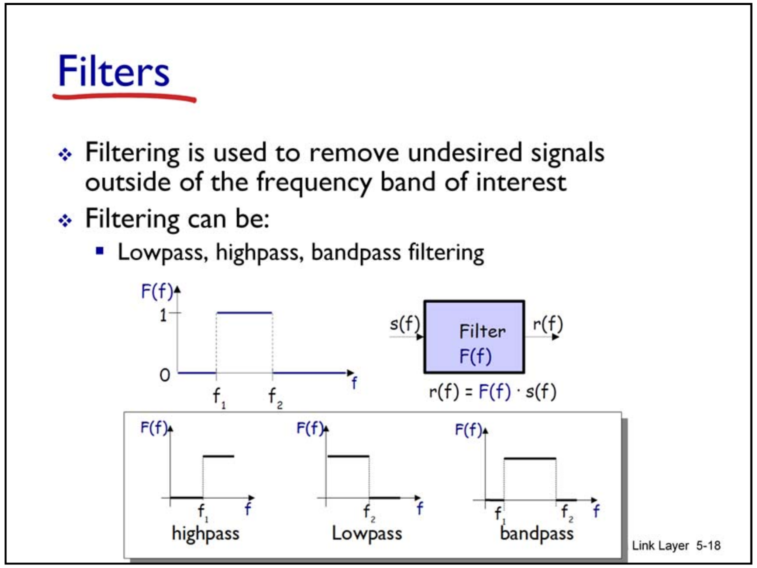
Physical Layer : outline

1. Concepts and Terminology
2. Transmission Impairments
3. Signal Encoding Techniques

Transmission Impairments

- signal received may differ from signal transmitted causing:
  - analog - degradation of signal quality
  - digital - bit errors
- most significant impairments are
  - attenuation and attenuation distortion
  - delay distortion
  - noise

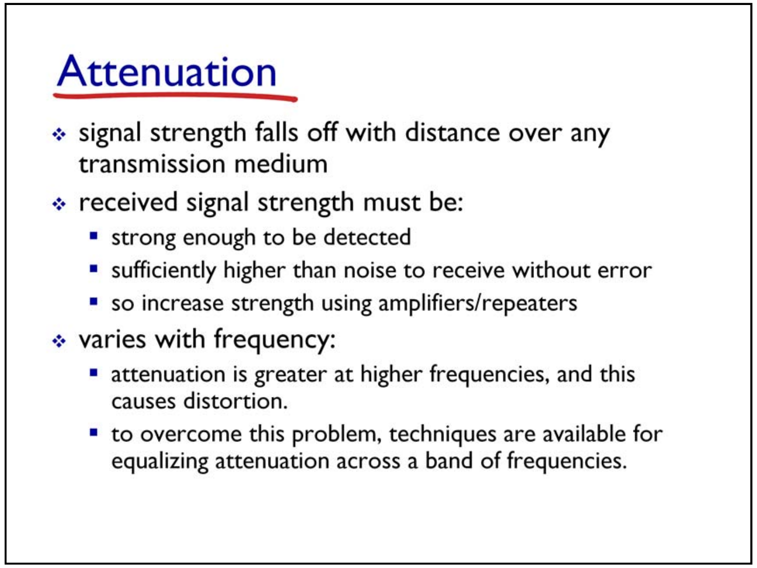
With any communications system, the signal that is received may differ from the signal that is transmitted due to various transmission impairments. For analog signals, these impairments can degrade the signal quality. For digital signals, bit errors may be introduced, such that a binary 1 is transformed into a binary 0 or vice versa.



Lowpass Filter: Only low frequency portion remains

Highpass Filter: Only high frequency (i.e. RF) portion remains

Bandpass Filter: Only desired RF portion remains (> f_1 and <f_2) – Interferer eliminated



Attenuation is where the strength of a signal falls off with distance over any transmission medium. For guided media, this is generally exponential and thus is typically expressed as a constant number of decibels per unit distance. For unguided media, attenuation is a more complex function of distance and the makeup of the atmosphere. See Stallings DCC8e Figure 3.11 on previous slide for illustration of attenuation.

Attenuation introduces three considerations for the transmission engineer. First, a received signal must have sufficient strength so that the electronic circuitry in the receiver can detect the signal. Second, the signal must maintain a level sufficiently higher than noise to be received without error. Third, attenuation varies with frequency. The first and second problems are dealt with by attention to signal strength and the use of amplifiers or repeaters. The third problem is particularly noticeable for analog signals. To overcome this problem, techniques are available for equalizing attenuation across a band of frequencies. This is commonly done for voice-grade telephone lines by using loading coils that change the electrical properties of the line; the result is to smooth out attenuation effects. Another approach is to use amplifiers that amplify high frequencies more than lower frequencies.

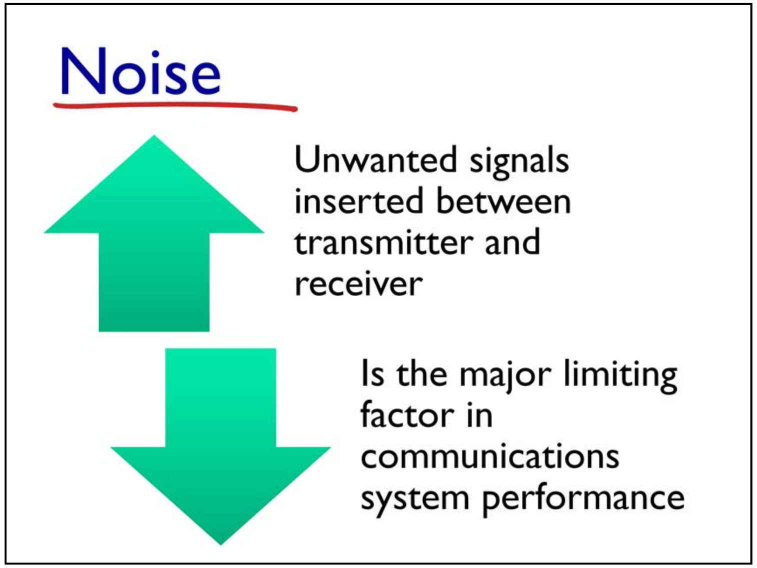
Delay Distortion

- Occurs in transmission cables such as twisted pair, coaxial cable, and optical fiber
  - Does not occur when signals are transmitted through the air by means of antennas
- Occurs because propagation velocity of a signal through a guided medium varies with frequency
  - For a signal with a given bandwidth, the velocity tends to be highest near the center frequency of the signal and to fall off toward the two edges of the band.
  - Thus, various components of a signal will arrive at the receiver at different times resulting in phase shifts between the frequencies
- Particularly critical for digital data since parts of one bit spill over into others causing intersymbol interference, which is a major limitation to maximum bit rate over a transmission channel.
- Equalizing techniques can be used for delay distortion.

Delay distortion is a phenomenon that occurs in transmission cables (such as twisted pair, coaxial cable, and optical fiber); it does not occur when signals are transmitted through the air by means of antennas. Delay distortion is caused by the fact that the velocity of propagation of a signal through a cable is different for different frequencies. For a signal with a given bandwidth, the velocity tends to be highest near the center frequency of the signal and to fall off toward the two edges of the band. Thus, various components of a signal will arrive at the receiver at different times.

This effect is referred to as delay distortion because the received signal is distorted due to varying delays experienced at its constituent frequencies. Delay distortion is particularly critical for digital data. Consider that a sequence of bits is being transmitted, using either analog or digital signals. Because of delay distortion, some of the signal components of 1 bit position will spill over into other bit positions, causing intersymbol interference , which is a major limitation to maximum bit rate over a transmission channel.

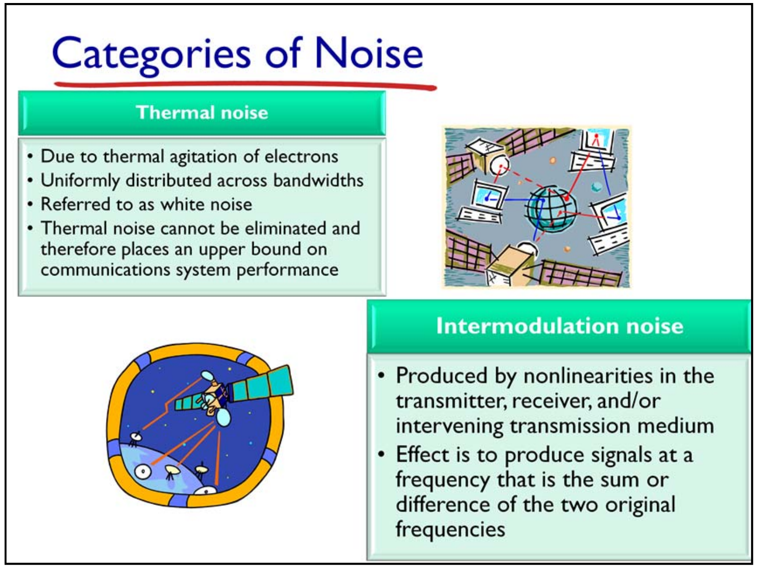
Equalizing techniques can also be used for delay distortion. Again using a leased telephone line as an example, Figure 3.14b shows the effect of equalization on delay as a function of frequency.



For any data transmission event, the received signal will consist of the transmitted signal, modified by the various distortions imposed by the transmission system, plus additional unwanted signals that are inserted somewhere between transmission and reception. The latter, undesired signals are referred to as noise . Noise is the major limiting factor in communications system performance.

Noise may be divided into four categories:

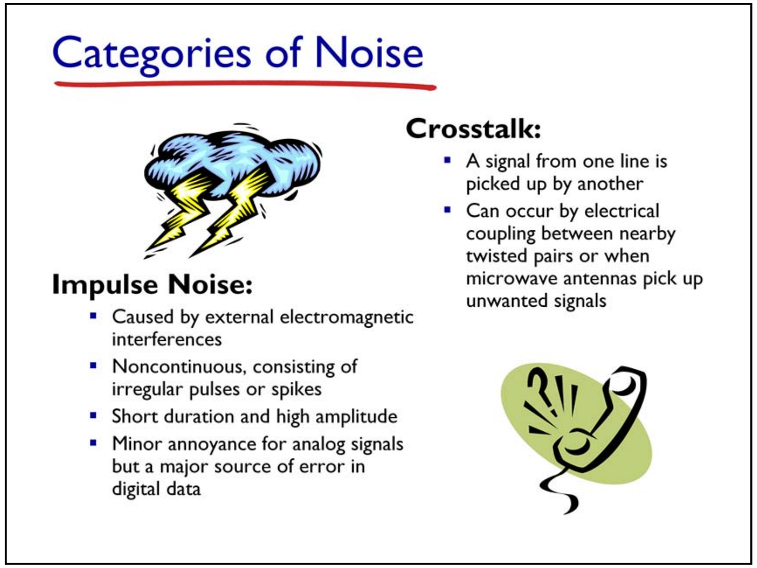
- Thermal noise
- Intermodulation noise
- Crosstalk
- Impulse noise



Thermal noise is due to thermal agitation of electrons. It is present in all electronic devices and transmission media and is a function of temperature. Thermal noise is uniformly distributed across the bandwidths typically used in communications systems and hence is often referred to as white noise. Thermal noise cannot be eliminated and therefore places an upper bound on communications system performance. Because of the weakness of the signal received by satellite earth stations, thermal noise is particularly significant for satellite communication.

When signals at different frequencies share the same transmission medium, the result may be intermodulation noise . The effect of intermodulation noise is to produce signals at a frequency that is the sum or difference of the two original frequencies or multiples of those frequencies. For example, if two signals, one at 4000 Hz and one at 8000 Hz, share the same transmission facility, they might produce energy at 12,000 Hz. This noise could interfere with an intended signal at 12,000 Hz.

Intermodulation noise is produced by nonlinearities in the transmitter, receiver, and/or intervening transmission medium. Ideally, these components behave as linear systems; that is, the output is equal to the input times a constant. However, in any real system, the output is a more complex function of the input. Excessive nonlinearity can be caused by component malfunction or overload from excessive signal strength. It is under these circumstances that the sum and difference frequency terms occur.



Crosstalk has been experienced by anyone who, while using the telephone, has been able to hear another conversation; it is an unwanted coupling between signal paths. It can occur by electrical coupling between nearby twisted pairs or, rarely, coax cable lines carrying multiple signals. Crosstalk can also occur when microwave antennas pick up unwanted signals; although highly directional antennas are used, microwave energy does spread during propagation. Typically, crosstalk is of the same order of magnitude as, or less than, thermal noise.

All of the types of noise discussed so far have reasonably predictable and relatively constant magnitudes. Thus it is possible to engineer a transmission system to cope with them. Impulse noise, however, is noncontinuous, consisting of irregular pulses or noise spikes of short duration and of relatively high amplitude. It is generated from a variety of causes, including external electromagnetic disturbances, such as lightning, and faults and flaws in the communications system.

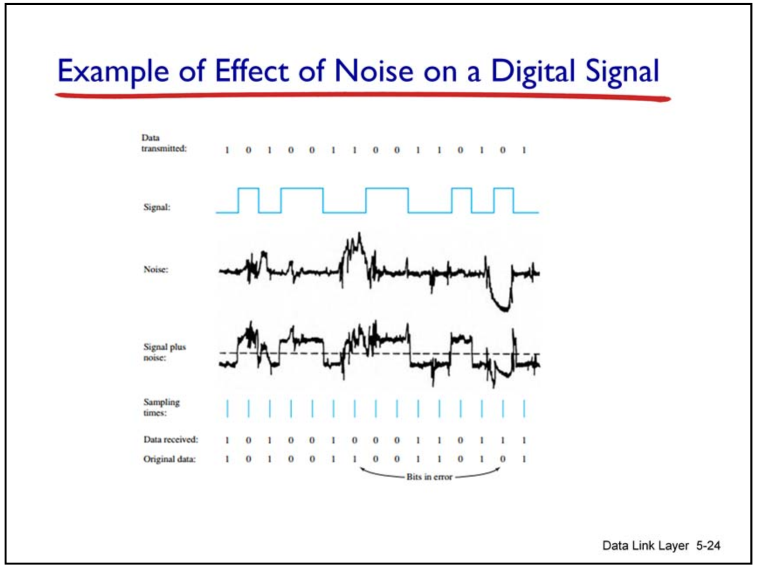


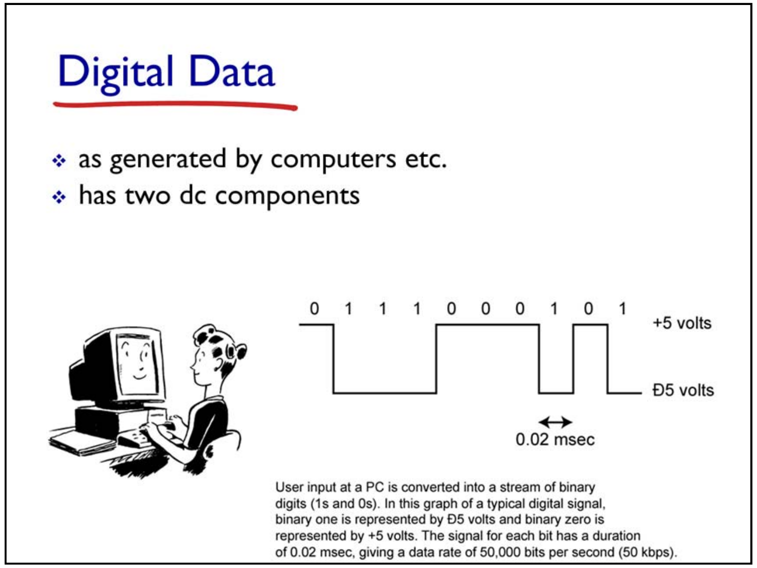
Figure is an example of the effect of noise on a digital signal. Here the noise consists of a relatively modest level of thermal noise plus occasional spikes of impulse noise. The digital data can be recovered from the signal by sampling the received waveform once per bit time. As can be seen, the noise is occasionally sufficient to change a 1 to a 0 or a 0 to a 1.

Physical Layer : outline

1. Concepts and Terminology
2. Transmission Impairments
3. Signal Encoding Techniques

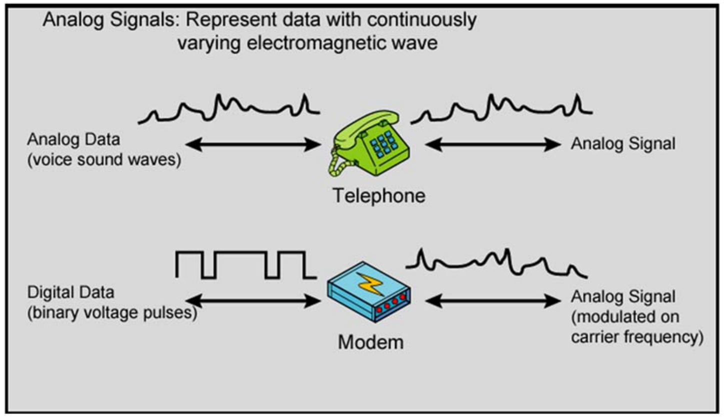
Analog and Digital Transmission

	Analog Signal	Digital Signal
Analog Data	Two alternatives: (1) signal occupies the same spectrum as the analog data; (2) analog data are encoded to occupy a different portion of spectrum.	Analog data are encoded using a codec to produce a digital bit stream.
Digital Data	**Digital data are encoded using a modem to produce analog signal.**	**Two alternatives:** **(1) signal consists of two voltage levels to represent the two binary values;** **(2) digital data are encoded to produce a digital signal with desired properties.**



Lastly consider **binary data**, as generated by terminals, computers, and other data processing equipment and then converted into digital voltage pulses for transmission. This is illustrated in Stallings DCC8e Figure 3.13. A commonly used signal for such data uses two constant (dc) voltage levels, one level for binary 1 and one level for binary 0. Consider the bandwidth of such a signal, which depends on the exact shape of the waveform and the sequence of 1s and 0s. The greater the bandwidth of the signal, the more faithfully it approximates a digital pulse stream.

Analog Signals



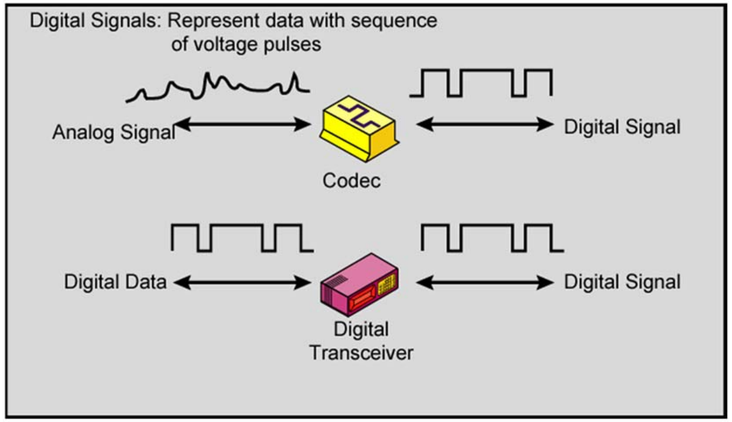
In a communications system, data are propagated from one point to another by means of electromagnetic signals. Both analog and digital signals may be transmitted on suitable transmission media.

An **analog signal** is a continuously varying electromagnetic wave that may be propagated over a variety of media, depending on spectrum; examples are wire media, such as twisted pair and coaxial cable; fiber optic cable; and unguided media, such as atmosphere or space propagation.

As Stallings DCC8e Figure 3.14 illustrates, analog signals can be used to transmit both analog data represented by an electromagnetic signal occupying the same spectrum, and digital data using a modem (modulator/demodulator) to modulate the digital data on some carrier frequency.

However, analog signal will become weaker (attenuate) after a certain distance. To achieve longer distances, the analog transmission system includes amplifiers that boost the energy in the signal. Unfortunately, the amplifier also boosts the noise components. With amplifiers cascaded to achieve long distances, the signal becomes more and more distorted. For analog data, such as voice, quite a bit of distortion can be tolerated and the data remain intelligible. However, for digital data, cascaded amplifiers will introduce errors.

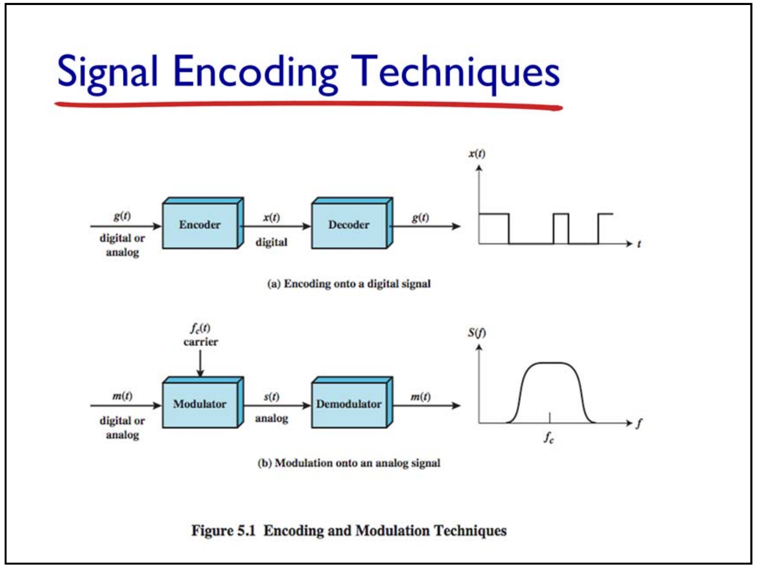
Digital Signals



A **digital signal** is a sequence of voltage pulses that may be transmitted over a wire medium; eg. a constant positive voltage level may represent binary 0 and a constant negative voltage level may represent binary 1.

As Stallings DCC8e Figure 3.14 also illustrates, digital signals can be used to transmit both analog signals and digital data. Analog data can converted to digital using a codec (coder-decoder), which takes an analog signal that directly represents the voice data and approximates that signal by a bit stream. At the receiving end, the bit stream is used to reconstruct the analog data. Digital data can be directly represented by digital signals.

A digital signal can be transmitted only a limited distance before attenuation, noise, and other impairments endanger the integrity of the data. To achieve greater distances, repeaters are used. A repeater receives the digital signal, recovers the pattern of 1s and 0s, and retransmits a new signal. Thus the attenuation is overcome.



Have already noted in Ch 3 that both analog and digital information can be encoded as either analog or digital signals:

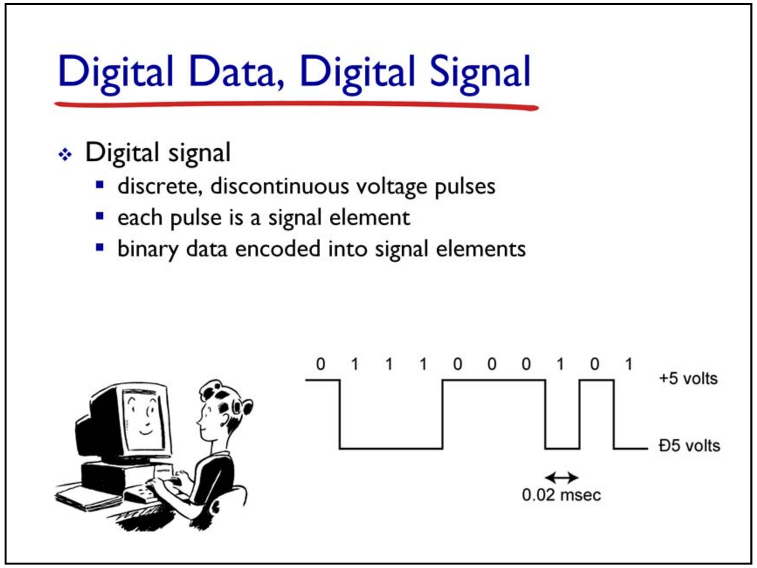
♦**Digital data, digital signals:** simplest form of digital encoding of digital data

♦**Digital data, analog signal:** A modem converts digital data to an analog signal so that it can be transmitted over an analog

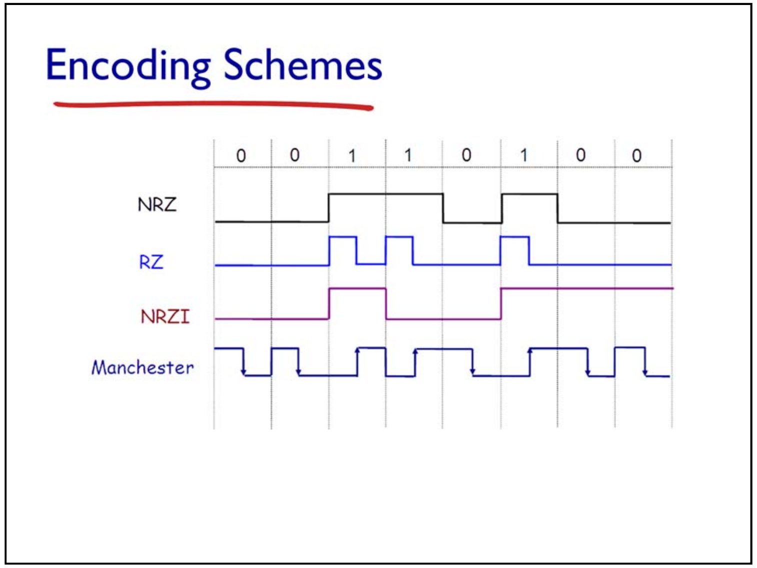
♦**Analog data, digital signals:** Analog data, such as voice and video, are often digitized to be able to use digital transmission facilities

♦**Analog data, analog signals:** Analog data are modulated by a carrier frequency to produce an analog signal in a different frequency band, which can be utilized on an analog transmission system

Stallings DCC8e Fig 5.1 emphasizes the process involved in this. For **digital signaling**, a data source $g(t)$, which may be either digital or analog, is encoded into a digital signal $x(t)$. The basis for **analog signaling** is a continuous constant-frequency f_c signal known as the **carrier signal**. Data may be transmitted using a carrier signal by modulation, which is the process of encoding source data onto the carrier signal. All modulation techniques involve operation on one or more of the three fundamental frequency domain parameters: amplitude, frequency, and phase. The input signal $m(t)$ may be analog or digital and is called the modulating signal, and the result of modulating the carrier signal is called the modulated signal $s(t)$.



Encoding - Digital data to digital signals: A digital signal is a sequence of discrete, discontinuous voltage pulses, as illustrated in Stallings DCC8e Figure 3.13. Each pulse is a signal element. Binary data are transmitted by encoding each data bit into signal elements. In the simplest case, there is a one-to-one correspondence between bits and signal elements. More complex encoding schemes are used to improve performance, by altering the spectrum of the signal and providing synchronization capability. In general, the equipment for encoding digital data into a digital signal is less complex and less expensive than digital-to-analog modulation equipment



We now turn to a discussion of various techniques, which are defined in Stallings DCC8e Table 5.2 and depicted in Figure 5.2 as shown above. **They include:**

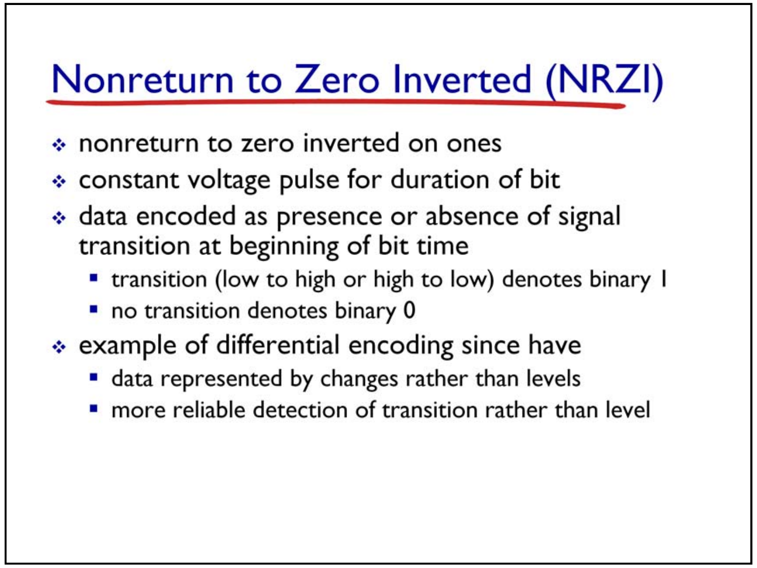
•Nonreturn to Zero (NRZ)

•Return to Zero (RZ)

•Manchester

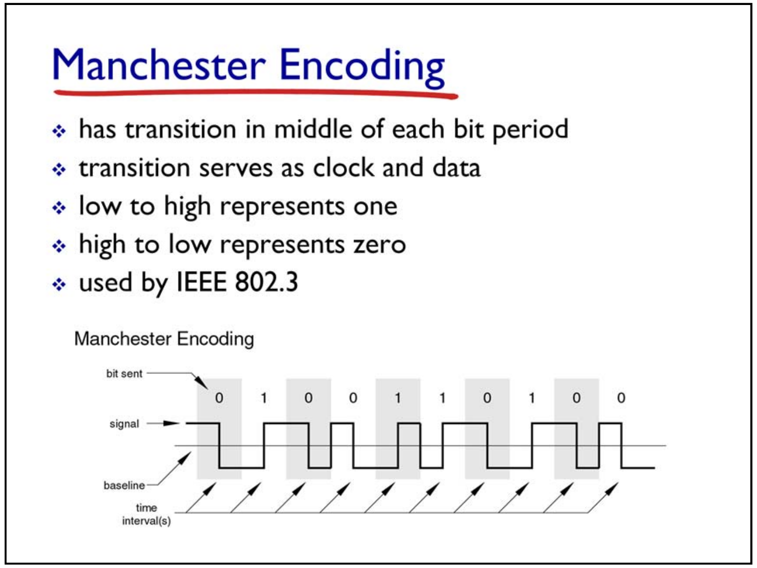
Nonreturn to Zero (NRZ)

- Non-Return to Zero (NRZ)
  - Used by Synchronous Optical Network (SONET)
  - 1=high signal, 0=low signal
  - Long sequence of same bit cause difficulty
    - DC bias hard to detect – low and high detected by difference from average voltage
    - Clock recovery difficult



A variation of NRZ is known as **NRZI** (Nonreturn to Zero, invert on ones). As with NRZ-L, NRZI maintains a constant voltage pulse for the duration of a bit time. The data bits are encoded as the presence or absence of a signal transition at the beginning of the bit time. A transition (low to high or high to low) at the beginning of a bit time denotes a binary 1 for that bit time; no transition indicates a binary 0.

NRZI is an example of **differential encoding**. In differential encoding, the information to be transmitted is represented in terms of the changes between successive signal elements rather than the signal elements themselves. The encoding of the current bit is determined as follows: if the current bit is a binary 0, then the current bit is encoded with the same signal as the preceding bit; if the current bit is a binary 1, then the current bit is encoded with a different signal than the preceding bit. One benefit of differential encoding is that it may be more reliable to detect a transition in the presence of noise than to compare a value to a threshold.



There is another set of coding techniques, grouped under the term *biphase*, that overcomes the limitations of NRZ codes. Two of these techniques, Manchester and differential Manchester, are in common use.

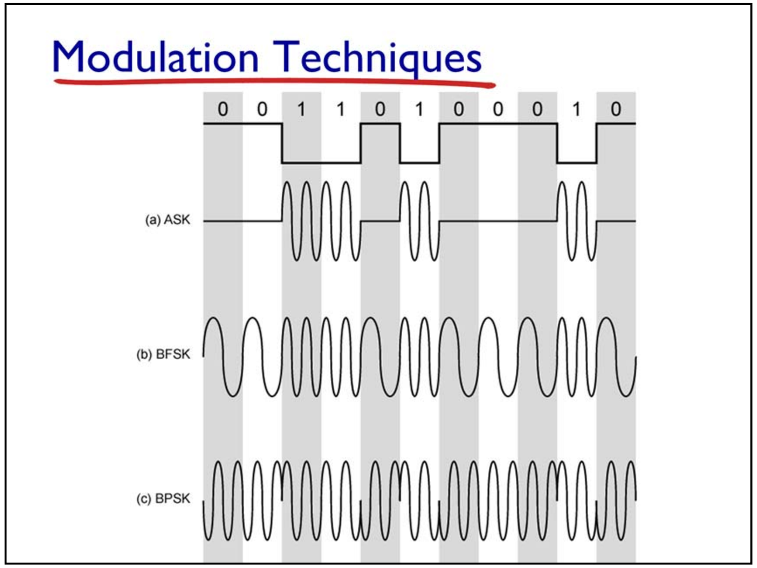
In the **Manchester** code, there is a transition at the middle of each bit period. The midbit transition serves as a clocking mechanism and also as data: a low-to-high transition represents a 1, and a high-to-low transition represents a 0. Biphase codes are popular techniques for data transmission. The more common Manchester code has been specified for the IEEE 802.3 (Ethernet) standard for baseband coaxial cable and twisted-pair bus LANs.

Digital Data, Analog Signal

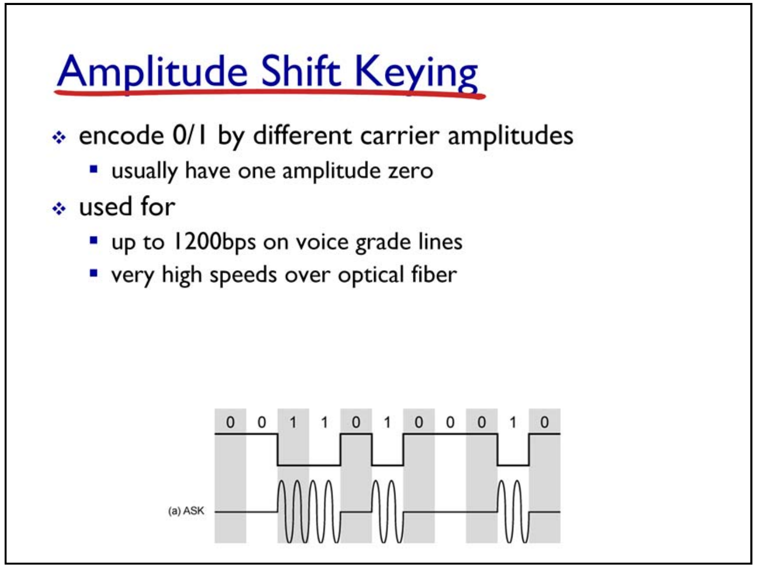
- main use is public telephone system
  - has freq range of 300Hz to 3400Hz
  - use modem (modulator-demodulator)
- encoding techniques
  - Amplitude shift keying (ASK)
  - Frequency shift keying (FSK)
  - Phase shift keying (PSK)

We turn now to the case of transmitting digital data using analog signals. The most familiar use of this transformation is for transmitting digital data through the public telephone network. The telephone network was designed to receive, switch, and transmit analog signals in the voice-frequency range of about 300 to 3400 Hz. It is not at present suitable for handling digital signals from the subscriber locations (although this is beginning to change). Thus digital devices are attached to the network via a modem (modulator-demodulator), which converts digital data to analog signals, and vice versa.

Have stated that modulation involves operation on one or more of the three characteristics of a carrier signal: amplitude, frequency, and phase. Accordingly, there are three basic encoding or modulation techniques for transforming digital data into analog signals, as illustrated in Stallings DCC8e Figure 5.7 (next slide): amplitude shift keying (ASK), frequency shift keying (FSK), and phase shift keying (PSK). In all these cases, the resulting signal occupies a bandwidth centered on the carrier frequency.



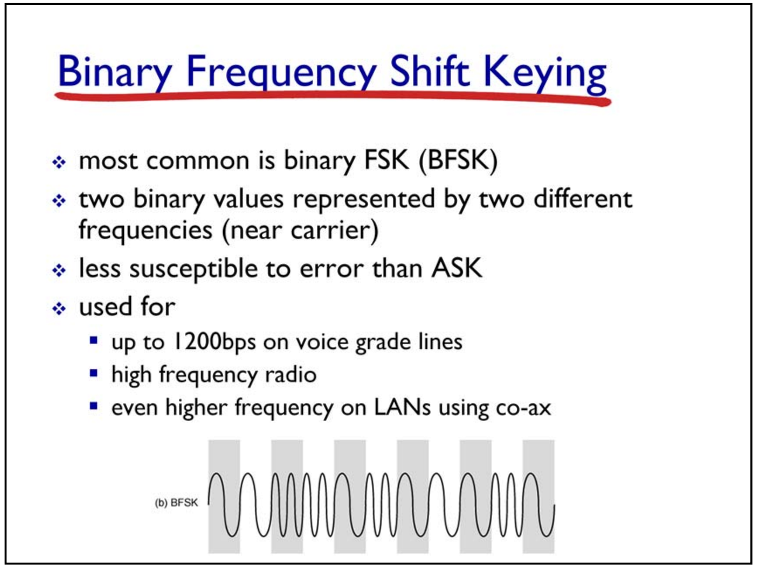
Have stated that modulation involves operation on one or more of the three characteristics of a carrier signal: amplitude, frequency, and phase. Accordingly, there are three basic encoding or modulation techniques for transforming digital data into analog signals, as illustrated in Stallings DCC8e Figure 5.7 (above): amplitude shift keying (ASK), frequency shift keying (FSK), and phase shift keying (PSK). In all these cases, the resulting signal occupies a bandwidth centered on the carrier frequency.



In ASK, the two binary values are represented by two different amplitudes of the carrier frequency. Commonly, one of the amplitudes is zero; that is, one binary digit is represented by the presence, at constant amplitude, of the carrier, the other by the absence of the carrier, as shown in Stallings DCC8e Figure 5.7a.

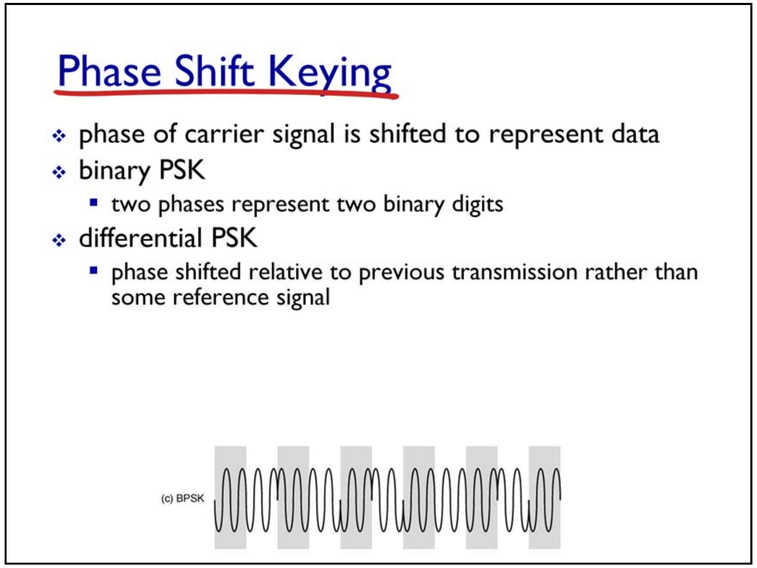
ASK is susceptible to sudden gain changes and is a rather inefficient modulation technique. On voice-grade lines, it is typically used only up to 1200 bps.

The ASK technique is used to transmit digital data over optical fiber, where one signal element is represented by a light pulse while the other signal element is represented by the absence of light.



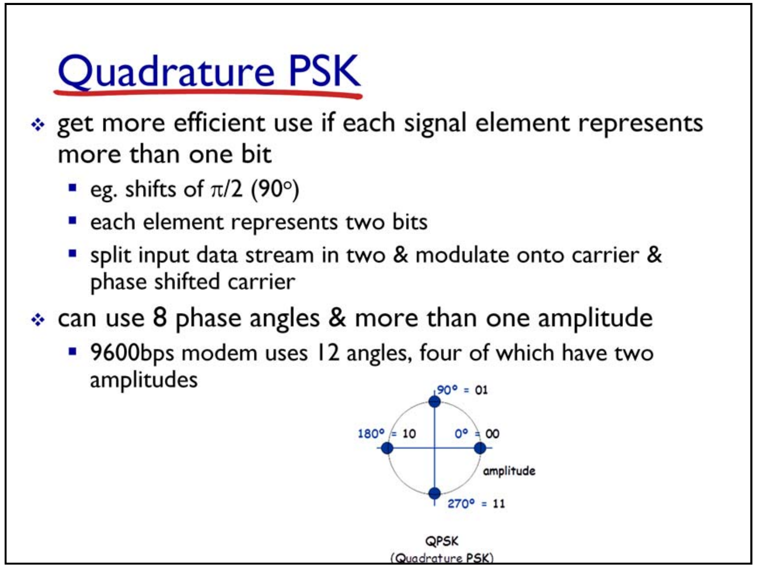
The most common form of FSK is binary FSK (BFSK), in which the two binary values are represented by two different frequencies near the carrier frequency, as shown in Stallings DCC8e Figure 5.7b.

BFSK is less susceptible to error than ASK. On voice-grade lines, it is typically used up to 1200 bps. It is also commonly used for high-frequency (3 to 30 MHz) radio transmission. It can also be used at even higher frequencies on local area networks that use coaxial cable.



In PSK, the phase of the carrier signal is shifted to represent data. The simplest scheme uses two phases to represent the two binary digits (Figure 5.7c) and is known as binary phase shift keying.

An alternative form of two-level PSK is differential PSK (DPSK). In this scheme, a binary 0 is represented by sending a signal burst of the same phase as the previous signal burst sent. A binary 1 is represented by sending a signal burst of opposite phase to the preceding one. This term *differential* refers to the fact that the phase shift is with reference to the previous bit transmitted rather than to some constant reference signal. In differential encoding, the information to be transmitted is represented in terms of the changes between successive data symbols rather than the signal elements themselves. DPSK avoids the requirement for an accurate local oscillator phase at the receiver that is matched with the transmitter. As long as the preceding phase is received correctly, the phase reference is accurate.

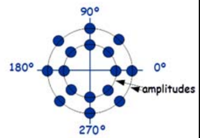


More efficient use of bandwidth can be achieved if each signaling element represents more than one bit. For example, instead of a phase shift of 180˚, as allowed in BPSK, a common encoding technique, known as quadrature phase shift keying (QPSK), uses phase shifts separated by multiples of π/2 (90˚). Thus each signal element represents two bits rather than one.

The use of multiple levels can be extended beyond taking bits two at a time. It is possible to transmit bits three at a time using eight different phase angles. Further, each angle can have more than one amplitude. For example, a standard 9600 bps modem uses 12 phase angles, four of which have two amplitude values, for a total of 16 different signal elements.

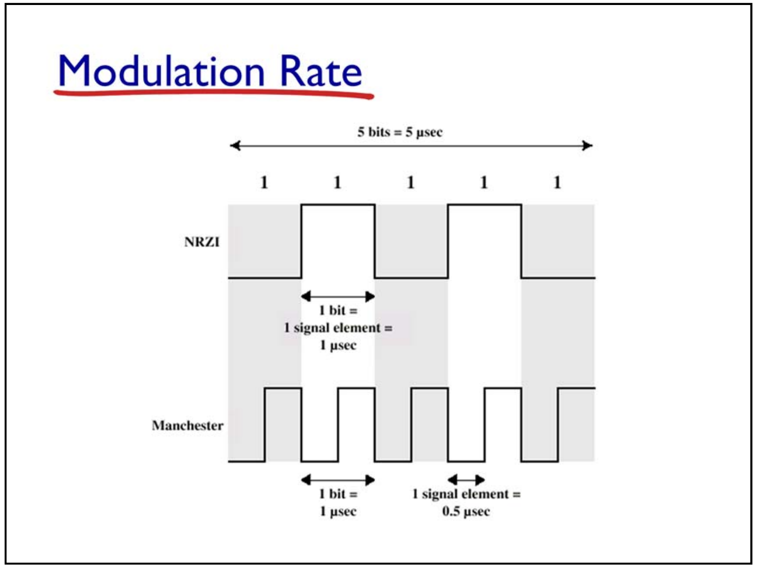
Quadrature Amplitude Modulation

- QAM used on asymmetric digital subscriber line (ADSL) and some wireless
- combination of ASK and PSK
- logical extension of QPSK



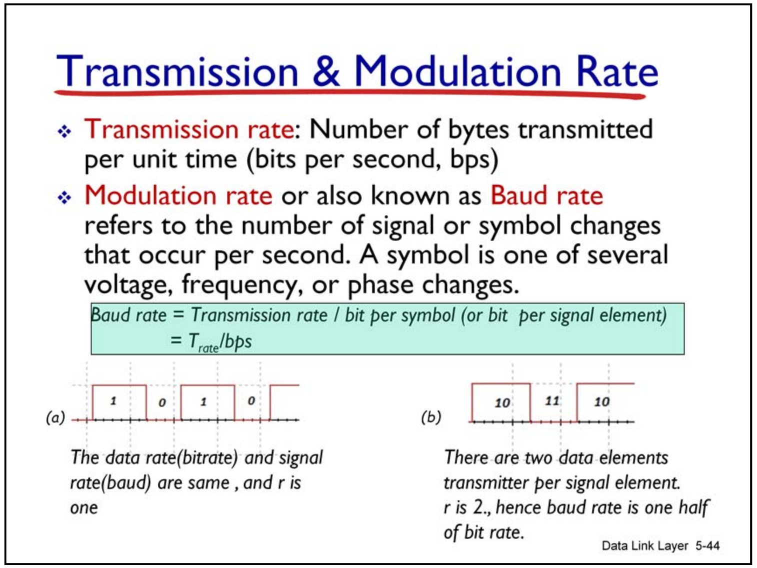
QAM
(Quadrature Amplitude Modulation)

Quadrature amplitude modulation (QAM) is a popular analog signaling technique that is used in the asymmetric digital subscriber line (ADSL), described in Chapter 8, and in some wireless standards. This modulation technique is a combination of ASK and PSK. QAM can also be considered a logical extension of QPSK.



When signal-encoding techniques are used, a distinction needs to be made between data rate (expressed in bits per second) and modulation rate (expressed in baud). The data rate, or bit rate, is $1/T_b$, where T_b = bit duration. The modulation rate is the rate at which signal elements are generated. Consider, for example, Manchester encoding. The minimum size signal element is a pulse of one-half the duration of a bit interval. For a string of all binary zeroes or all binary ones, a continuous stream of such pulses is generated. Hence the maximum modulation rate for Manchester is $2/T_b$. This situation is illustrated in Stallings DCC8e Figure 5.5, which shows the transmission of a stream of binary 1s at a data rate of 1 Mbps using NRZI and Manchester.

One way of characterizing the modulation rate is to determine the average number of transitions that occur per bit time. In general, this will depend on the exact sequence of bits being transmitted. Stallings DCC8e Table 5.3 compares transition rates for various techniques.



Few Examples:

Consider the figures (a) of the slide.

There, the data rate(bitrate) and signal rate(baud) are same and r is one. One data element rides on one signal element (analogous to one person per carriage in a train).

And in image (b) of the slide, you can see that there are two data elements transmitter per signal element. In other words, the bit rate is higher than the baud rate. (Analogous to two passengers per carriage in a train) and here r is two. Hence baud rate is one half of bit rate.
And the next time, when you define the baud rate – its number of signal elements per second and not number of bits per second!

Normally the number of symbols is some power of two. If N is the number of bits per symbol, then the number of required symbols is $S = 2^N$. Thus, the gross bit rate is: R = baud rate x $\log_2 S$ = baud rate x 3.32 $\log_{10} S$ If the baud rate is 4800 and there are two bits per symbol, the number of symbols is $2^2 = 4$. The bit rate is: R = 4800 x 3.32 log(4) = 4800 x 2 = 9600 bits/s

Bit rate Vs Baud rate – the common misconception

The two most common/confused words in digital communication – Bit rate and Baud rate. Generally, communication is concerned with transmission of data. In digital communication, there are two entities that are needed to carry out communication – the data to be transmitted and the signal over which the data is transmitted. Now, we have two entities to be worried about – the data and the signal. The most common misconception is that most people think both travel at the same speed! – NO!

The difference:

Digital data is very different from digital signal. The process of converting digital data to digital signal is called as **line coding**. Now, to discriminate between data and signal, data is what we need to send. But signal is what we can send. So, signal is the carrier which carries data. Also, keep in mind that the smallest entity of the data, that can represent a piece of information is called data element and shortest meaningful unit of a signal is called signal element. Consider this as in the following scenario – Consider a train. Each carriage is a signal element. Each passenger inside the train is a data element. The train as a whole is a signal and all passengers together represent a data.

Data rate and Signal rate:

Data rate – Number of data elements transmitted per second.

Signal rate – Number of signal elements transmitted per second.

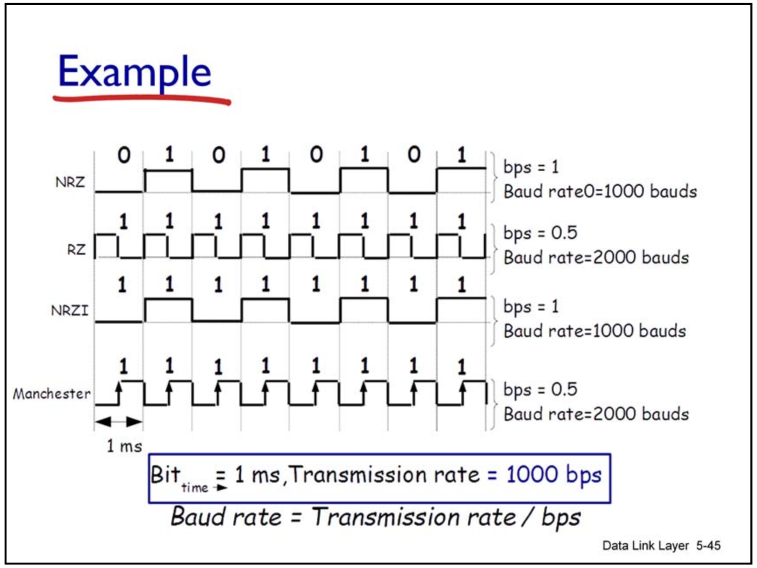
Now, the unit of data rate is bit rate. And the unit of signal rate is pulse rate/ modulation rate/ baud rate or simply baud. From the previous example, we can see that, a carriage in a train can carry more than one person. So, if you consider the number of person is more than one per carriage, you can say that bit rate is greater than baud rate for the signal.

Calculating the baud rate:

Baud rate is calculated using the below formula.
Modulation rate = Transmission rate / bit per signal element = N/ r

here, N is the bit rate and r is the number of data elements carried by each signal element. Here r must be as great as possible for better efficiency. (Stuff more people in a carriage :P)

From the above text, it is clearly inferred that the bit rate must be greater than the baud rate for higher efficiency. The aim is to transmit as many data element as possible in a signal element. There are different methods to do this which are collectively called as line coding schemes.Some of the popular line coding schemes are: Non-return to zero (NRZ), Manchester, Alternate mark inversion (AMI) and also multi level schemes are available.



In Manchester Encoding, a symbol must contain a 1/2 bit. It would be more accurate to say that it takes *two symbols* (clock transitions, in this case) to encode *one* bit.