DATA AND SIGNALS

- All of the forms of information viz. voice, data, image, video) can be represented by electromagnetic signals. Depending on the transmission medium and the communications environment, either **analog or digital signals can be used to convey information**.

- Any electromagnetic signal, analog or digital, is made up of a number of constituent frequencies. A key parameter that characterizes the signal is bandwidth, which is the width of the range of frequencies that comprises the signal. In general, the greater the bandwidth of the signal, the greater its information-carrying capacity.

- A major problem in designing a communications facility is transmission impairment. The most significant impairments are attenuation, attenuation distortion, delay distortion, and the various types of noise. The various forms of noise include thermal noise, intermodulation noise, crosstalk, and impulse noise. For analog signals, transmission impairments introduce random effects that degrade the quality of the received information and may affect intelligibility. For digital signals, transmission impairments may cause bit errors at the receiver.

- The designer of a communications facility must deal with four factors:
    - the bandwidth of the signal,
    - the data rate that is used for digital information,
    - the amount of noise and other impairments, and
    - the level of error rate that is acceptable.

The bandwidth is limited by the transmission medium and the desire to avoid interference with other nearby signals. Because bandwidth is a scarce resource, we would like to maximize the data rate that is achieved in a given bandwidth. The data rate is limited by the bandwidth, the presence of impairments, and the error rate that is acceptable.

The **successful transmission of data depends principally on two factors**:

- **Quality of the signal being transmitted** and
- **Characteristics of the transmission medium**.

Either analog or digital data may be transmitted using either analog or digital signals.

it is common for intermediate processing to be performed between source and destination, and this processing has either an analog or digital character.

The **Chief Impairments** are:

- Attenuation,
- Attenuation Distortion,
- delay distortion, and
- various forms of noise.

Electromagnetic signals are used as a means to transmit data.

The signal is a function of time.

It can also be expressed as a function of frequency; that is, the signal consists of components of different frequencies.

The **frequency domain** view of a signal is more important to an understanding of data transmission than a **time domain** view.

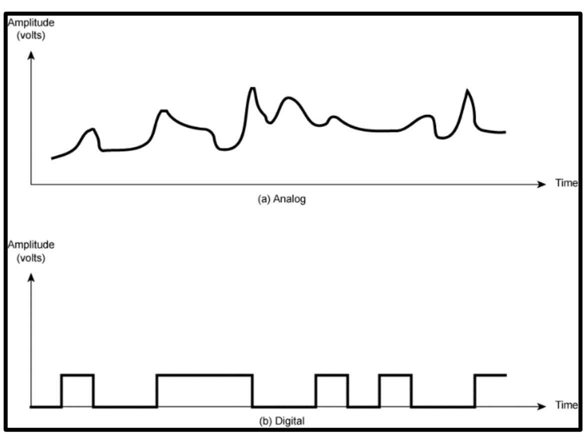
Viewed as a function of time, an electromagnetic signal can be either analog or digital.

Analog signal: The signal intensity varies in a smooth fashion over time.

Digital signal: The signal intensity maintains a constant level for some period of time and then abruptly changes to another constant level.

Digital data take on discrete values.
For example, data are stored in computer memory in the form of 0's and 1's.

They can be converted to a digital signal or modulated into an analog signal for transmission across a medium.



The **transition from one voltage level to another will not be instantaneous**, but there will be a **small transition period**.

Periodic signal: The same signal pattern repeats over time.
Otherwise, a signal is **aperiodic**.

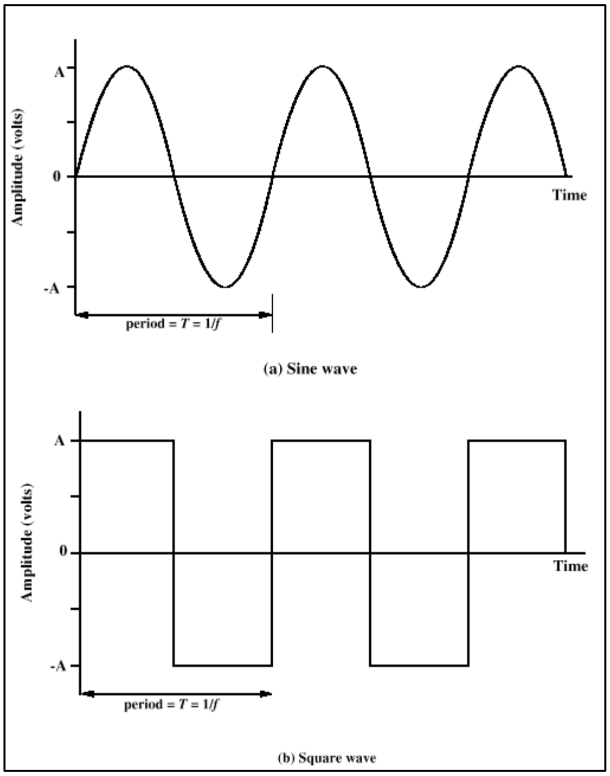
Periodic and Nonperiodic Signals

Both analog and digital signals can take one of two forms: *periodic* or *nonperiodic* (sometimes refer to as *aperiodic,* because the prefix *a* in Greek means "non").

A periodic signal completes a pattern within a measurable time frame, called a period, and repeats that pattern over subsequent identical periods. The completion of one full pattern is called a cycle. A nonperiodic signal changes without exhibiting a pattern or cycle that repeats over time.

In data communications, we commonly use periodic analog signals and nonperiodic digital signals.

Periodic Signal:

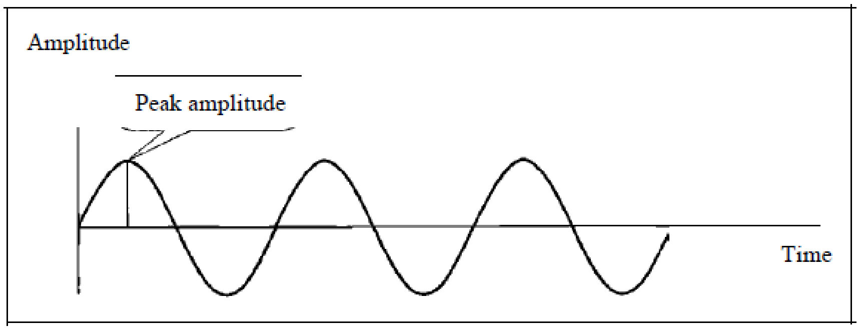


(a) Sine wave

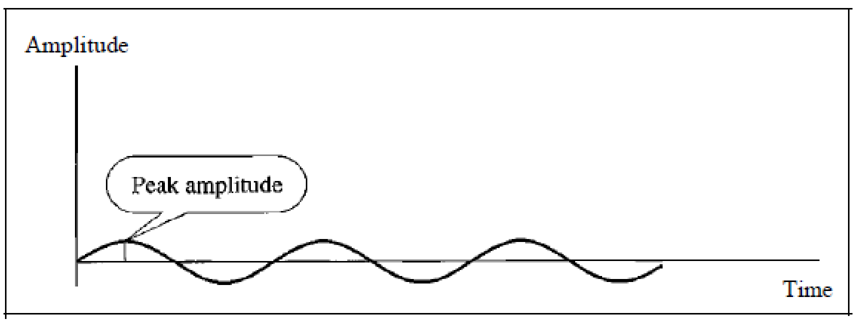
(b) Square wave

The **sine wave is the fundamental periodic signal**.

Two signals with the same phase and frequency, but different amplitudes



a. A signal with high peak amplitude

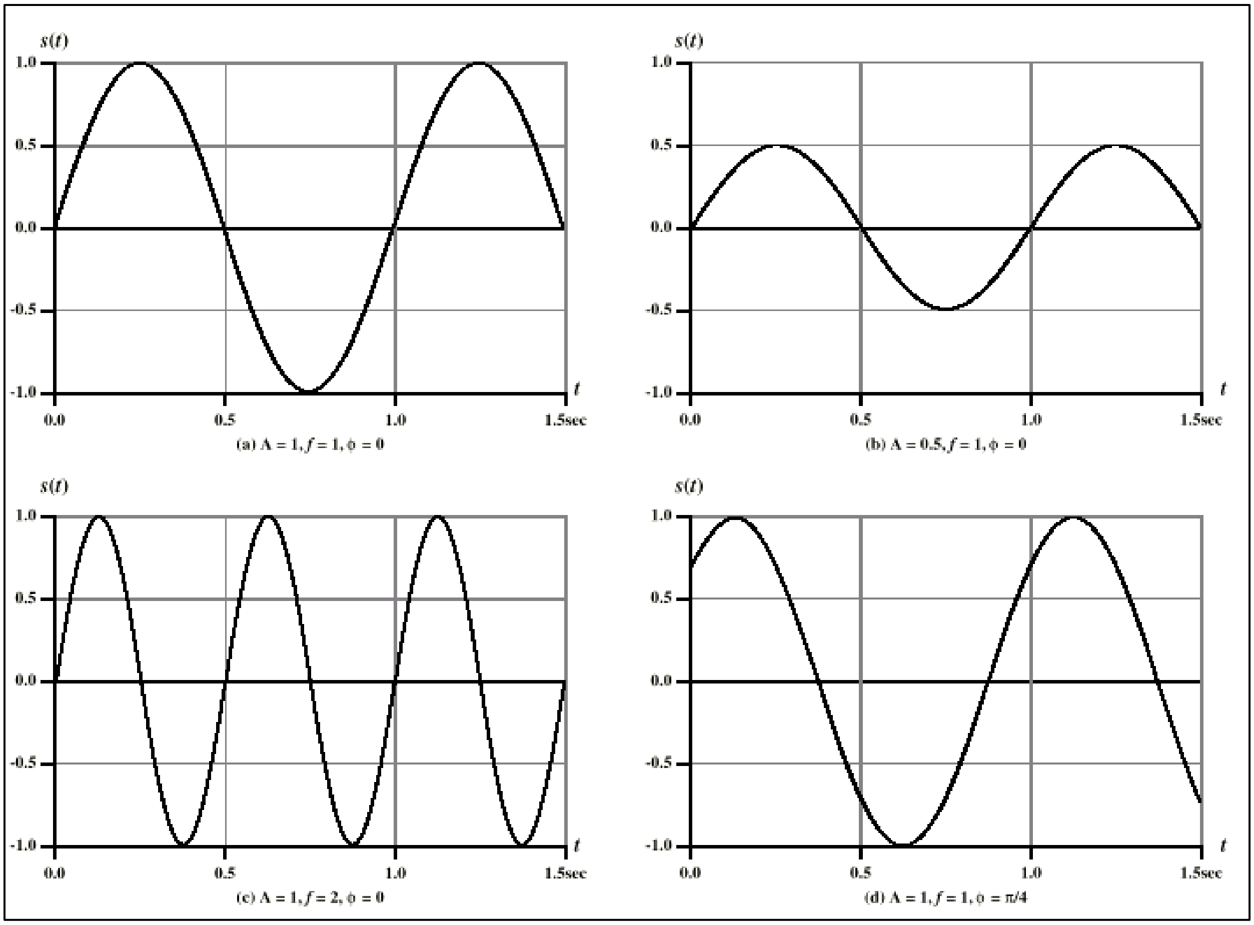


b. A signal with low peak amplitude

A general sine wave can be represented by **three parameters**:

- **peak amplitude (*A*)** - the ***maximum value or strength of the signal over time***; typically measured in volts.
- **frequency (*f*)** - the rate [in cycles per second, or Hertz (Hz)] at which the signal repeats. An equivalent parameter is the **period** (T) of a signal, so $T = 1/f$.
- **phase (ϕ)** - **measure of relative position in time within a single period of a signal**.

Varying Sine Waves - $s(t) = A \sin(2\pi ft + \Phi)$



The general sine wave can be written as: $s(t) = A \sin(2\pi ft + \phi)$, known as a **sinusoid function**.

Figure above shows the effect of varying each of the three parameters, the horizontal axis is time; the graphs display the value of a signal at a given point in space as a function of time.

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).

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.

There is a simple relationship between the two sine waves, one in time and one in space.

The **wavelength** (λ) of a signal is the ***distance occupied by a single cycle***, or, put another way, the ***distance between two points of corresponding phase of two consecutive cycles***.

Assume that the signal is traveling with a ***velocity v***.

Then the wavelength is related to the period as follows: $\lambda = vT$.

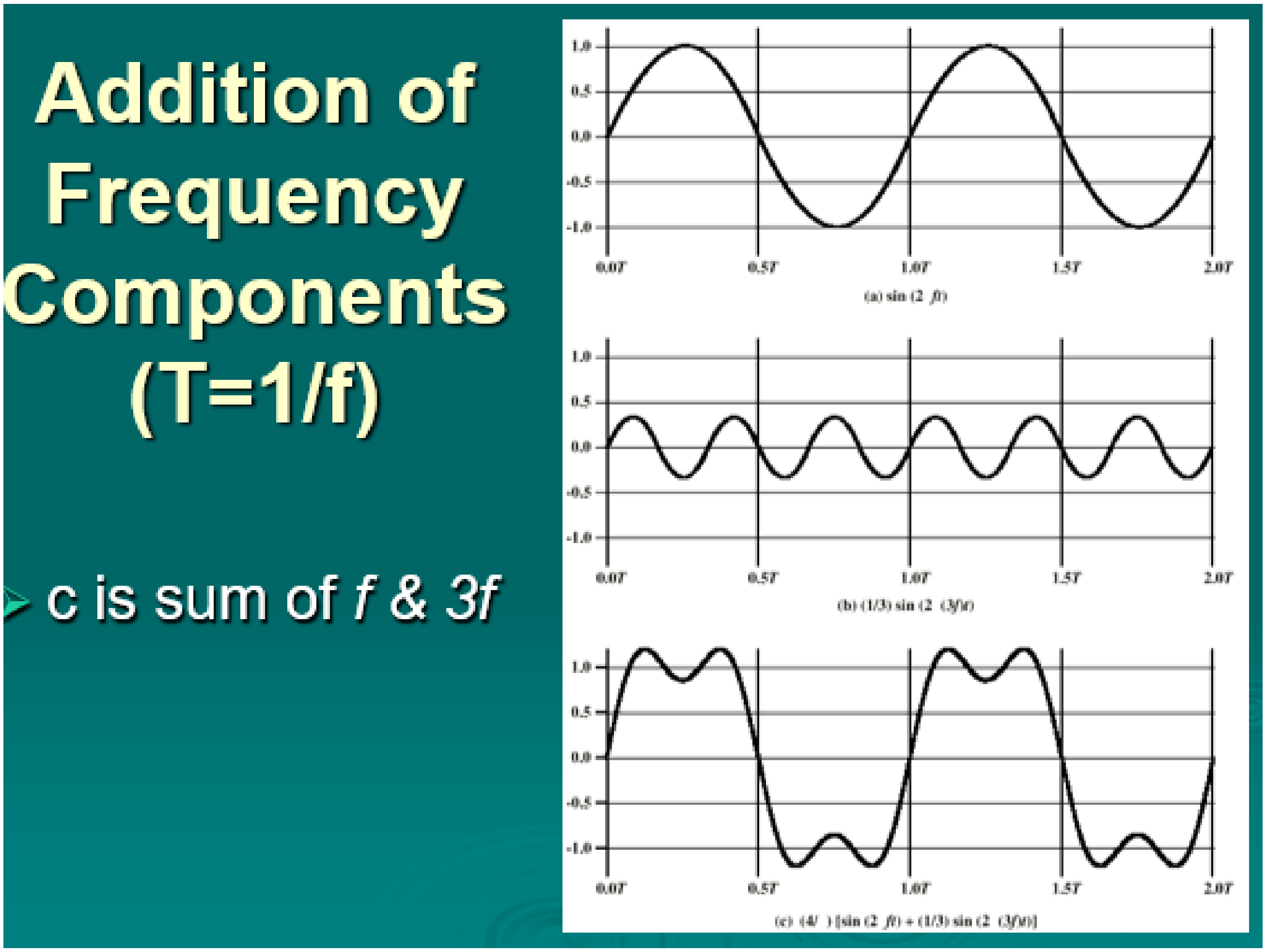
Equivalently, $\lambda f = v$.

Of particular relevance to this discussion is the case where $v = c$, the speed of light in free space, which is approximately 3×10^8 *m/s*.

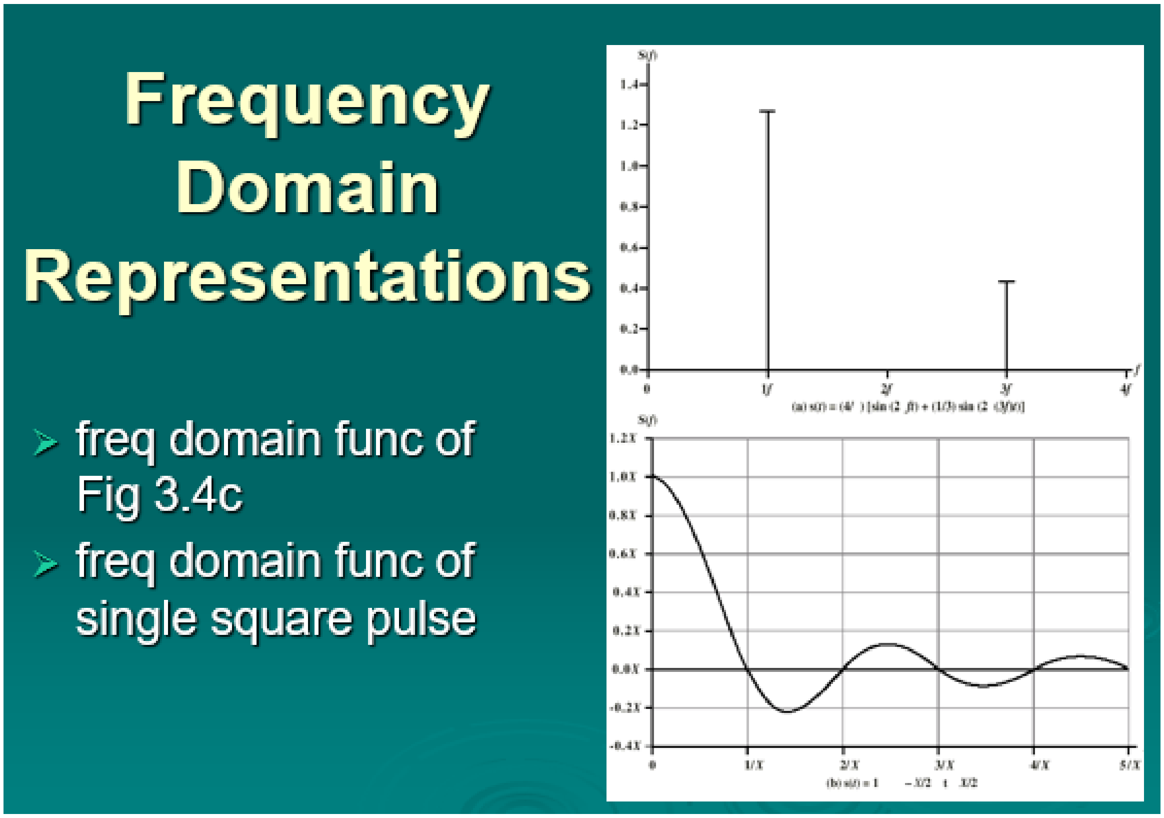
In practice, ***an electromagnetic signal will be made up of many frequencies***.

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.



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



For each signal, there is **a time domain function *s(t)*** that **specifies the amplitude of the signal at each instant in time**.

There is **a frequency domain function S(f)** that **specifies the peak amplitude of the constituent frequencies of the signal.**

The **spectrum** of a signal is the range of frequencies that it contains.

In the following diagram, it extends **from *f* to 3*f*.**

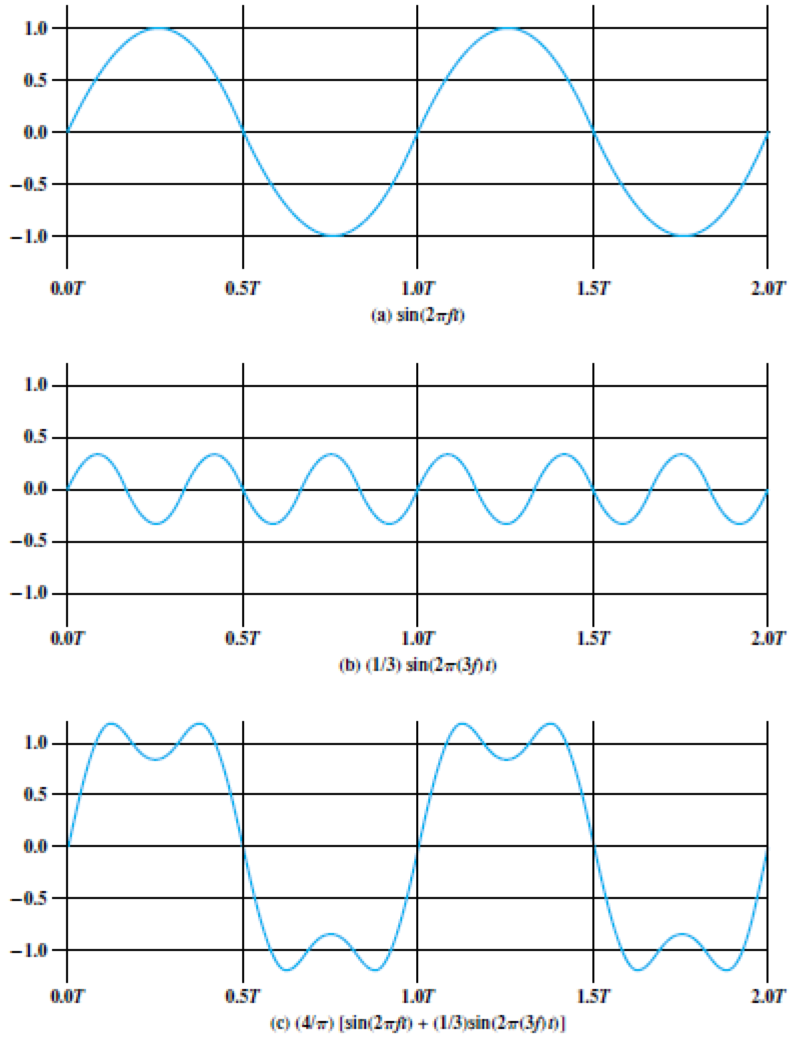


Figure 3.4 Addition of Frequency Components ($T = 1/f$)

- 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**.
- 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.

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.

The **Spectrum** of a signal is the range of frequencies that it contains.

For the signal of Figure 3.4c, the spectrum extends from f to $3f$.

The **absolute bandwidth** of a signal is the width of the spectrum.

In the case of Figure 3.4c, the bandwidth is $2f$.

Many signals, such as that of Figure 3.5b, have an infinite bandwidth.

However, most of the energy in the signal is contained in a relatively narrow band of frequencies. This band is referred to as the **effective bandwidth**, or just **bandwidth**.

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

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.

- **Data:** Entities that convey meaning, or information.
- **Signals:** Electric or electromagnetic representations of data.
- **Signaling**: The physical propagation of the signal along a suitable medium.
- **Transmission:** The communication of data by the propagation and processing of signals.

Analog data take on continuous values in some interval, the most familiar example being **audio**, which, in the form of acoustic sound waves, can be perceived directly by human beings.

Figure below shows the acoustic spectrum for human speech and for music (note log scales).

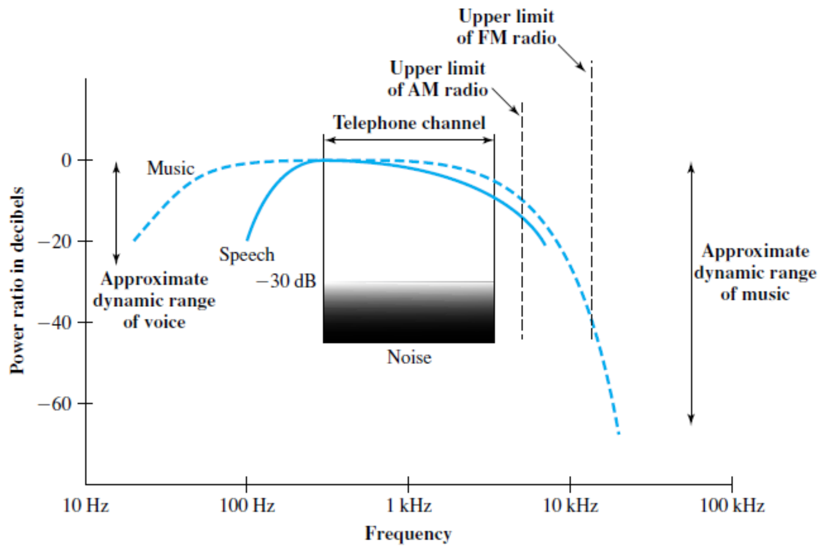


Figure 3.9 Acoustic Spectrum of Speech and Music [CARN99a]

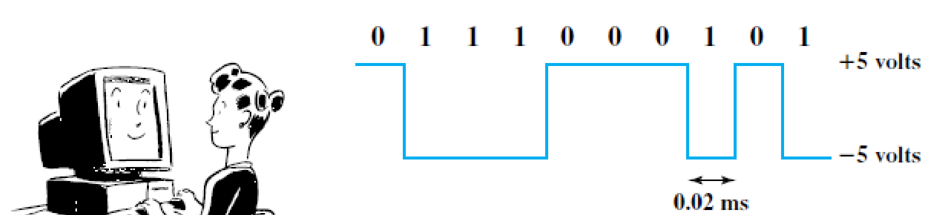
Frequency components of typical speech may be found between approximately 100 Hz and 7 kHz, and has a dynamic range of about 25 dB (a shout is approx. 300 times louder than whisper).

Another common example of analog data is **video**, as seen on a TV screen.

DIGITAL DATA

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 the figure below:.



User input at a PC is converted into a stream of binary digits (1s and 0s). In this graph of a typical digital signal, binary one is represented by −5 volts and binary zero is represented by +5 volts. The signal for each bit has a duration of 0.02 ms, giving a data rate of 50,000 bits per second (50 kbps).

Figure 3.13 Conversion of PC Input to Digital Signal

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 shown in the figure below, 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.

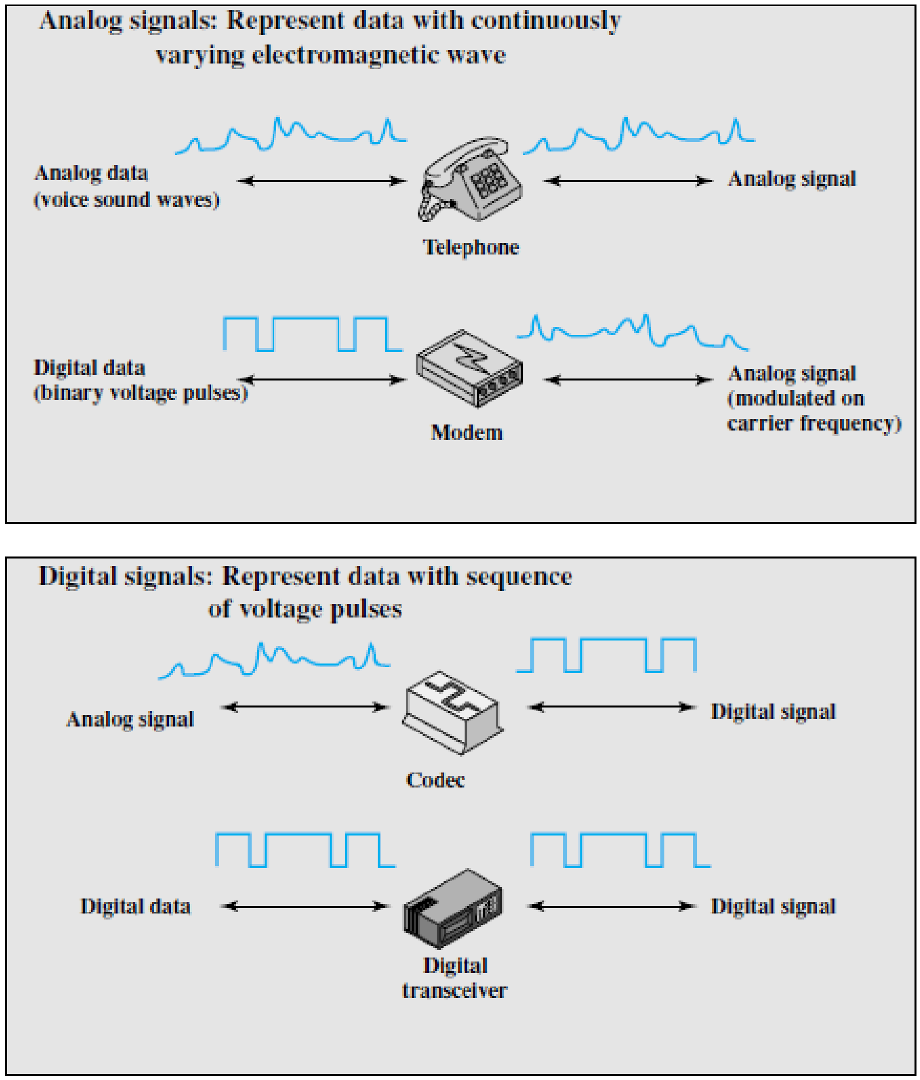


Figure 3.14 Analog and Digital Signaling of Analog and Digital Data

However, analog signals 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***.

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.

Merits of Digital Signaling:

- Generally cheaper than analog signaling, and
- Less susceptible to noise interference.

Principal Demerit of Digital Signals: Suffer more from attenuation than do analog signals

<u>Transmission Impairments</u>

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.

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.

Illustration of Attenuation.

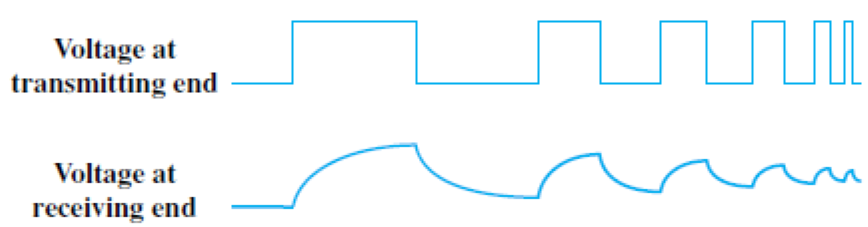


Figure 3.11 Attenuation of Digital Signals

Attenuation introduces **three considerations** for the transmission engineer:

(1) A received signal must have sufficient strength so that the electronic circuitry in the receiver can detect the signal.
(2) The signal must maintain a level sufficiently higher than noise to be received without error.
(3) 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 because the velocity of propagation of a signal through a guided medium varies with frequency.

For a bandlimited signal, the velocity tends to be highest near the center frequency and fall off toward the two edges of the band.

Thus various frequency components of a signal will arrive at the receiver at different times, resulting in phase shifts between the different frequencies.

Delay distortion is particularly critical for digital data, because ***some of the signal components of one bit position will spill over into other bit positions***, causing **intersymbol interference**.

This is **a major limitation to maximum bit rate over a transmission channel**.

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, referred to as noise, that are inserted somewhere between transmission and reception.

Noise is a major limiting factor in communications system performance.

Noise may be divided into the following **four categories**:

Thermal Noise:

- Caused due to thermal agitation of electrons.
- Present in all electronic devices and transmission media and is a function of temperature.
- Is uniformly distributed across the bandwidths typically used in communications systems and hence is often referred to as **white noise**.
- Cannot be eliminated and therefore places an upper bound on communications system performance, and. is particularly significant for satellite communication.

Intermodulation Noise:

- Caused when ***signals at different frequencies share the same transmission medium***.
- The effect is to produce signals at a frequency that is the sum or difference of the two original frequencies or multiples of those frequencies, thus possibly interfering with services at these frequencies.
- ***Produced by nonlinearities in the transmitter, receiver, and/or intervening transmission medium.***

Crosstalk:

- An unwanted coupling between signal paths.
- May occur by electrical coupling between nearby twisted pairs or, rarely, coax cable lines carrying multiple signals.
- Can also occur when microwave antennas pick up unwanted signals; although highly directional antennas are used, microwave energy does spread during propagation.
- It is of the same order of magnitude as, or less than, thermal noise.

Impulse Noise:

- Is noncontinuous, consisting of irregular pulses or noise spikes of short duration and of relatively high amplitude.
- Generated from a variety of causes, including external electromagnetic disturbances, such as lightning, and faults and flaws in the communications system.
- Generally, only a minor annoyance for analog data.
- It is the primary source of error in digital data communication.

For example, a sharp spike of energy of 0.01 s duration would not destroy any voice data but would wash out about 560 bits of data being transmitted at 56 kbps.

Like the data they represent, signals can be either **analog** or **digital**.

An **analog signal** has infinitely many levels of intensity over a period of time.

As the wave moves from value *A* to value *B*, it passes through and includes an infinite number of values along its path.

A **digital signal**, on the other hand, can have only a limited number of defined values.

Although each value can be any number, it is often as simple as 1 and 0.

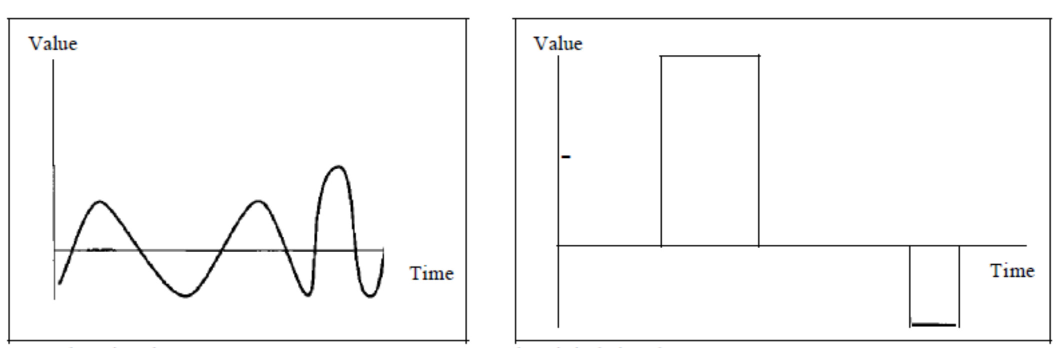
The simplest way to show signals is by plotting them on a pair of perpendicular axes.

The **vertical axis represents** the value or strength of a signal.

The **horizontal axis** represents time.

Figure 3.1 illustrates an analog signal and a digital signal.

Figure 3.1 *Comparison of analog and digital signals*



a. Analog signal b. Digital signal

The ***curve representing the analog signal passes through an infinite number of points***.

The ***vertical lines of the digital signal, however, demonstrate the sudden jump that the signal makes from value to value***.

Periodic and Non-periodic Signals

Both analog and digital signals can take one of two forms: *periodic* or *nonperiodic* (sometimes refer to as *aperiodic*, because the prefix *a* in Greek means "non").

A **periodic signal** completes a pattern within a measurable time frame, called a **period**, and ***repeats that pattern over subsequent identical periods***.

The completion of one full pattern is called a **Cycle**.

A **nonperiodic signal** changes without exhibiting a pattern or cycle that repeats over time.

Both analog and digital signals can be periodic or nonperiodic.

In data communications, we commonly use ***Periodic Analog Signals*** (because they need less bandwidth) and ***nonperiodic digital signals***.

PERIODIC ANALOG SIGNALS

Can be classified as **Simple** or **Composite**.

Simple periodic analog signal: A sine wave that cannot be decomposed into simpler signals.

Composite periodic analog signal: Composed of multiple sine waves.

Sine Wave

- The most fundamental form of a periodic analog signal.
- When we visualize it as a simple oscillating curve, its change over the course of a cycle is smooth and consistent, a continuous, rolling flow.

Figure 3.2 *A sine wave*