Lecture 3: Basics of Digital Communications

(09/09/2009)

Lecture Outline

1. Introduction
2. Transmission Media
3. Sources of attenuation and distortion
4. Signal propagation and transmission delays
5. Transmission control schemes

1. Introduction

(1) NIC (network interface card) and network access: an NIC is a device consisting of hardware with built-in firmware that performs the relevant network interface functions.

(a) Network data: regardless of the applications that send the data and the nature of the data, to an NIC all data transferred over the network is treated simply as a string of one or more blocks.

(b) An NIC transmits sequences of raw bits (0 and 1) between two DTEs across the network.

(2) **Baseline transmission** vs **modulated transmission** (Fig.1.8,p.22)

a. Baseline transmission: Fig.1.8(a), p.22: The signal output by the NIC simply varies between two voltage levels (+V and -V are two possible values) at a rate determined by the transmission bit rate.

(a) With networks that provide a digital interface (eg a LAN, an ISDN and a PSDN), baseline transmission is also used over the access lines to the network.

b. Modulated transmission: Fig.1.8(b)

(a) With analog networks such as PSTN, analog transmission is used over the access lines. The presence of a transformer in the PSTN means DC digital (for example a long string of binary 0s and 1s) will not be discernible.

(b) On the other hand, the bandwidth used over these lines is from 200Hz to 3400Hz. If we use baseline transmission, then we are limited to the bit rate supported by that limited bandwidth.

(c) Modulating: is the process of mixing or multiplying the binary signal to be transmitted by a single-frequency signal chosen from within the bandwidth used for the analog signal. This single-frequency signal is known as **carrier signal**. The mode of transmission is known as modulated transmission.

(d) The device that performs the modulation and demodulation is called **modem**.

(3) Problems of signal tranmissions

a. Problems

(a) Attenuated – signals become weaker and their amplitudes smaller due to electronic resistance

(b) Distorted – due to limited bandwidth, unequal amplification of different signals, different delays of different frequencies.

(c) Noise – due to outside noise, internal or outside interference.

b. Factors that affect signal impairments:

(a) The type of transmission medium

(b) The length of the tranmission medium

(c) The bandwidth of the medium

(d) The bit rate of the data being transmitted

(4) Transmission media

a. Copper/metal media

b. Optical media

c. Wireless/radio media

(5) Transmission control schemes

a. Asynchronous transmission

b. Synchronous transmission

2. <u>Transmission media</u>

(0) DTE (Data Terminal Equipment, eg. a computer, a terminal) and DCE (Data Circuit-Terminating Equipment, eg. a modem).

(1) **Two-wire open lines** – two lines connecting two DTEs – the simplest transmission medium (Fig.2, in Lecture 1). Features and problems

a. Capable of connecting DTEs up to 50m apart and provides moderate bit rates (up to 19.2kbps)

b. Problems:

(a) Limited distance and bit rate;

(b) **crosstalk**, caused by *capacity coupling* between two lines;

(c) Electromagnetic interference, caused by the open structure.

(2) **Twisted pair lines** – the two lines are twisted: Fig. 1.10(a), p.26

a. Twisting significantly reduces crosstalks

b. Classical twisted pairs

(a) Can provide up to 1Mbps and 100m connection;

(b) A insulating cover or shield can be added (**shielded twisted pair, STP**). Multiple twisted pairs can be placed inside an insulating outer cover to further reduce interference and noise: Fig.1.10(b), p.26

(c) Main problem: the **skin effect** – as the bit rate (hence the frequency) increases, the current flowing in the wires tends to flow only on the outer surface of the wire. This increases the electrical resistance of higher frequency signals, leading to higher attenuation.
More sophisticated driver and receiver are needed for higher bit rates.

c. Category 3, 5 (Fig.21), 6, and 7 UTP (unshielded twisted pairs). UTP is newer, better, and easier to cope with than the previous STP (invented by IBM).

(a) Catetory 3 UTP: consists of two insulated wires gently twisted together. Most offices prior 1988 have one category 3 UTP running from a central wiring closet on each floor into it. Can provide data rate up to 16 Mbps

(b) Catetory 5 UTP. Invented 1988. Similar to category 3 UTP, but with more twists per centimeter. Less crosstalk and better quality. Can provide data rate up to 100 Mbps

(c) Catetory 6 and 7 UTP. More sophisticated. Can support data rates of 250Mbps and 600Mbps respectively.

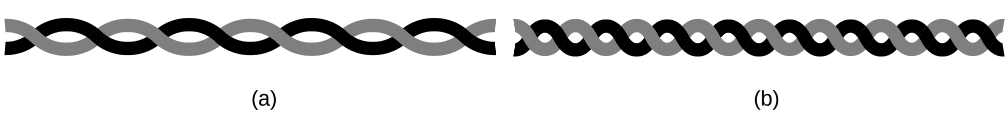

Figure 21: Category 3 and 5 UTP

(3) **Coaxial cable** (Fig.1.10(c), p.26 and Fig.22)

a. A coaxial cable consists of a copper core, a layer of insulating material, a layer of braided outer conductor, and a protective plastic covering.

b. It is more expensive, but also is more immune against **skin effect** (and its induced attenuation), and **radiation effect** (loss of signal power due to loss of energy to the environment). It hence provides high data rate and is more reliable – from 10Mbps to 1Gbps over several hundred meters.

c. Applicable to both point-to-point and multipoint topologies. Has been widely used in cable TV systems and MANs. However optical fibers are being used in place of coaxial cables.

d. Two types of coaxial cables: 50-ohm cable and 75-ohm cable. Due to historical reasons. The 50-ohm cable is intended for digital transmission only while the latter is for both analog and digital transmissions. The 75-ohm cable becmoes more important in Internet over cable technology (due to the use of the existing 4:1 impedance matching transformers).

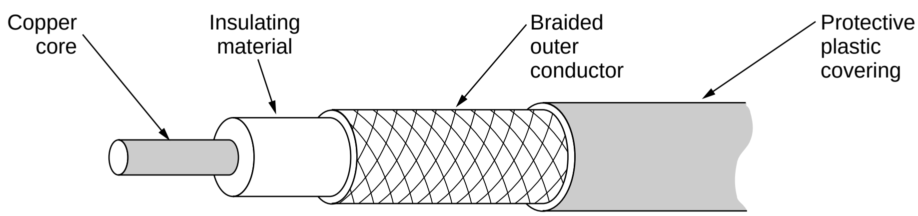

Figure 22: A coaxial cable

(4) **Optical fiber** (Fig.1.11, p.28) it carries the transmitted information in the form of a fluctuating beam of light in a glass fiber.

a. Even coaxial cable is limited in bit rate. Optical fibre eliminates the conventional limitations (electromagnetic radiation, noise and etc.).

b. Could reach hundred Mbps and very high frequency. More reliable (not easy to tap).

c. Transmitters: **LED** (*light emitting diode*), or **ILD** (*injection laser diode*).
Receivers: **photodiode** or **photo transistor**.

d. Three different principles:

(a) **Multimode stepping index** (Fig.1.11(b), top) – the cladding and the core material each has a different but uniform refractive index.

i. All light emitted by the diode at an angle less than the critical angle is reflected at the cladding surface and propagates a long the core by means of multiple (internal) reflections.

ii. Light will take a different amount of time to travel along the cable, depending upon the emitted angle – the received signal thus has a wider pulse width than the input signal, thus decreases the bit rated. This effect is called **dispersion**.

(b) **Multimode graded index fibre** (Fig.1.11(b), middle) – the core has a variable refractive index. The light is reflected by an increased amount as it moves away from the core. Thus it narrows the pulse width and hence increases the bit rate.

(c) **Singlemode fibre** (Fig.1.11(b), bottom) – the core diameter is reduced to that of a single wavelength (3-10μm) and light travel along a single straight line. Can reach several hundred Mbps.

(5) **Communication Satellites** (Fig.1.12,p.31).

a. Characteristics

(a) There are no physical lines, data is transmitted using electromagnetic (radio) waves.

(b) Communications are **geostationary**. That means the satellite orbits the earth once every 24 hours in synchronization with the earth's rotation.

(c) Can provide high bit rate communications.

b. Sharing of the high bandwidth: TDM (**time division multiplexing**) technique is used to allow multiple applications to share the high bandwidth.

(6) **Terrestrial microwaves** – similar to satellites.

a. Used in extreme circumstances or where radio signal transmission or satellite signal reception is not possible.

b. Similar to satellites, but not as reliable.

(7) **Radio** (**Cordless links**)

a. Low frequency radios are used for cordless links within a small geographical area where installing fixed-wire cables is not feasible or not economical.

b. Base stations and cells. Each base station covers a fixed area that forms a cell.

3. <u>Sources of attenuation and distortion</u>

(1) **Baud rate** and **bit rate**:

a. Baud rate: is the number of line signal transitions per second. Or in other words, it is the number of signal samplings per second.

b. Bit rate: is the number of bits sent per second.

c. If every signal sample only contains 1 bit, buad rate is equal to bit rate. However, if every signal sample contains more than 1 bit, bit rate is higher than buad rate.

   (a) If 2 signal level is used, each signal transition represents exactly 1 bit. In this case bit rate is equal to baud rate.

   (b) If 8 signal level used: 0, 1, 2, 3, 4, 5, 6, 7. Then each signal can represent 3 bits. Than the data rate is 3 times the baud rate.

d. In general, if the number of signal levels is M, then bit rate is equal to the baud rate multiplying $\log M$.

(2) Attenuations and distortions: Fig.1.9, p.24

a. The Top diagram is the original signal

b. The second diagram is the signal after attenuation

c. Limited bandwidth, delay distortion, and line noise eventually result in a bit error.

(3) The Fourier series

a. Any reasonably behaved periodic signal $v(t)$ with period T can be expressed by summing a (possibly infinite) number of sines and cosines:

$$v(t) = a_0 \ + \ \sum_{n=1}^{\infty} a_n cos(n\omega_0 t) \ + \ \sum_{n=1}^{\infty} b_n sin(n\omega_0 t)$$

where

   (a) ω_0 is the fundamental frequency

   (b) a_n and b_n are the sine and cosine amplitudes of the nth **harmonics** (terms)

   (c) The amplitudes a_n, b_n and constant c can be calculated by the formula:

   $$\begin{aligned} a_n &= \tfrac{2}{T} \textstyle\int_0^T v(t) sin(2\pi n f t) dt \\ b_n &= \tfrac{2}{T} \textstyle\int_0^T v(t) cos(2\pi n f t) dt \\ a_0 &= \tfrac{2}{T} \textstyle\int_0^T v(t) dt \end{aligned}$$

   (d) $T = 2\pi/\omega_0$ is the period of the waveform in seconds.

   (e) a_0 is the mean of the signal over the period T and is known as the **DC component**.

b. Illustration of harmonics: Fig.23.

(a) As more harmonics are received, the receiver can more faithfully reproduce the original signal.

(b) More harmonics means higher bandwidth of the channel is needed.

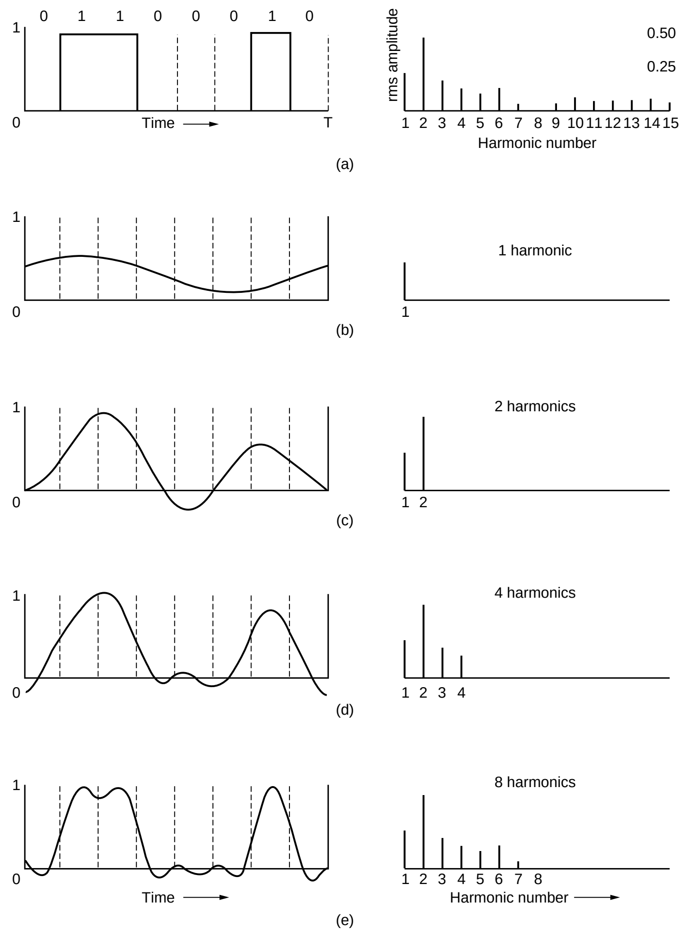

Figure 23: Illustration of Fourier series: (a) A binary signal and its root-mean-square Fourier amplitudes. (b)-(e) Successive approximations to the original signal.

b. **Worst case sequence**

(a) Binary bit sequences vary in their shapes. Sequences 10101010... repeat subsequence 10 in a period of two bit-cell intervals, while 110110110... and 111011101110 repeat subsequence 110 and 1110 in a period of three bit-cell and four bit-cell intervals respectively.

(b) The sequence 101010... has the shortest period and will yeild the highest fundamental frequency. It is hence called the worst case sequence.

c. **Unipolar** and **bipolar** binary signals

(a) **Unipolar** binary signals: the signal amplitude varies between zero and a positive voltage (say $+V$). They are also called **return-to-zero (RZ)** signals. The mean signal level for such signals is $V/2$ and the amplitude variation is V.

(b) **Bipolar** binary signals: the signal amplitude varies between an negative voltage (say $-V$) and a positive volt (say $+V$). They are also called **non-return-to-zero (NRZ)** signals and have a mean signal level zero and amplitude variation $2V$.

(c) Due to their special properties, the Fourier series for unipolar and bipolar signals are of the form:

Unipolar: $v(t) = \frac{V}{2} + \frac{2V}{\pi}\{cos\omega_0 t - \frac{1}{3}cos3\omega_0 t + \frac{1}{5}cos5\omega_0 t - \cdots\}$

Bipolar: $v(t) = \frac{4V}{\pi}\{cos\omega_0 t - \frac{1}{3}cos3\omega_0 t + \frac{1}{5}cos5\omega_0 t - \cdots\}$

where:

$v(t)$ is the voltage signal representation as a function of time
ω_0 is the fundamental frequency component in radians per second
$f_0 = \omega_0/2\pi$ is the fundamental frequency in Hz
$T = 1/f_0$ is the period of the fundamental frequency in seconds

d. For periodic binary sequences, we can conclude:

(a) They are made of and infinite series of sinusoidal signals including the fundamental frequency component f_0, a third harmonic component $3f_0$, a fifth harmonic component $5f_0$ and so on.

(b) Example (2.2, p.36, of 4th ed.). A binary singal of rate 500bps is to be transmitted over a communications channel. Derive the minimum bandwidth required assuming (a) the fundamental frequency only, (b) the fundamental and the 3rd harmonic, (c) the fundamental, plus the 3rd and the 5th harmonics of the worst-case sequence are to be received.

· The worst-case sequence is 101010... At 500bps the fundamental frequency is 250Hz. Hence the 3rd harmonic is 750Hz and the 5th harmonic is 1250Hz.

· Therefore the minimum bandwidth required to perform the required transmissions are (a) 0-250Hz; (b) 0-750Hz; (c) 0-1250 Hz.

(4) Attenuation

a. The amplitude of signals decreases as it passes communication medium (lines).

b. A thresh hold value is set. **Repeaters** are used to amplify signals.

(5) Distortion I: caused by attenuation

a. Signals attenuate as a function of frequency. Attenuation hence also causes distortion.

b. **Equalizers** are used to equalize the attenuation across defined band of frequencies.

c. Both attenuation and amplifications (done by repeaters) are measured in **dB (decibels)**. Let P_1 and P_2 be the transmitted power and received power respectively, both in the unit of **watt**. Then

   (a) Attenation $= 10 \log_{10} \frac{P_1}{P_2}$ dB

   (b) Amplification $= 10 \log_{10} \frac{P_2}{P_1}$ dB

   amplification is also called **gain**.

d. Use of **dB** allows attenuation/amplification of multiple sections in a transmission to be summed together.

e. Example (p.33, of 4th ed.). Three sections in a transmission line between two DTEs. The 1st section introduces 16dB attenuation, the 2nd has a 20dB gain, while the 3rd has a 10dB attenuation. Assume the mean transmission power is 400mW. Determine the mean output power level of the transmission line.
   Method 1:

   (a) For 1st section: $16 = 10 \log_{10} \frac{400}{P_2}$. We have $P_2 = 10.0475$mW.

   (b) For 2nd section: $20 = 10 \log_{10} \frac{P_2}{10.0475}$. We have $P_2 = 1004.75$mW.

   (c) For 3rd section: $10 = 10 \log_{10} \frac{1004.75}{P_2}$. We have $P_2 = 100.475$mW.

   That says the mean output level is 100.475mW.
   Method 2:

   (a) The overall attenuation is: $(16 - 20) + 10 = 6$dB

   (b) Hence, $6 = 10 \log_{10} \frac{400}{P_2}$, $P_2 = 100.475$mW

(6) Distortion II: due to limited bandwidth

a. Only those components of a signal that are within the bandwidth of the transmission medium can pass the medium. Others are strongly diminished. The higher the bandwidth, the more freq. components are passed, hence the more faithful reproduction of the original signal is transmitted.

b. The *Nyquist* formula: B – the bandwidth (in Hz); M – the level of signals; C – the maximum data transfer rate:

$$C = 2B \log_2 M \text{ bps}$$

Example (p.34) B = 3100, M = 8
$C = 2B \log_2 M = 2 \times 3100 \times 3 = 18600$bps

(7) Distortion III: delay distortion

The rate of propagation of a sinusoidal signal along a transmission line varies with the frequency of the signal. Consequently, the varies of frequency components of a given signal arrive at the receiving end with various delays – this is called **delay distortion**.

(8) Noise.

a. **Line noise level** – the amount of noise on the transmission line when the line is idle. Ideally should be zero.

b. S – the signal level; N – the noise level. The ratio S/N is known as **signal-to-noise ratio**. Measured in dB:

$$D = 10 \log_{10}(\frac{S}{N})\text{dB}$$

c. **Shannon-Hartley Law**: determines the theoretical maximum data rate of a transmission medium:

$$C = B \log_2(1 + \frac{S}{N})\text{bps}$$

d. **Example** (p.40-41, of 4th ed.) Known D = 20 dB; bandwidth = 3000 Hz; determine the maximum theoretical data rate.
$D = 10 \log_{10}(\frac{S}{N}) = 20 ==>$
$\frac{S}{N} = 100 ==>$
$C = B \log_2(1 + \frac{S}{N}) = 3000 \times \log_2(1 + 100) = 19963$bps

4. Signal propagation and transmission delays

(1) Propagation delay T_p is the time that it takes for a signal (electrical or optical) to travel from one end of a transmission medium to another end in a communication.

a. Formally:
$T_p = \text{propagation delay} = \frac{\text{physical separation S in meters}}{\text{velocity of propagation V in meters per second}}$

b. In theory, V is limited by the speed of light (3×10^8 m/s). In most media V is about 2×10^8 m/s.

(2) Transmission delay T_x is the time that it takes for a block of signals completely be in the transmission medium from a DTE.

a. Transmission delay is determined by the *link bit rate*, which is the rate at which a link can take (absorb) the signals from a DTE. Higher quality links normally have higher link bit rates and hence have lower transmission delays.

b. Formally:
T_x = transmission delay = $\frac{\text{number of bits to be transmitted N}}{\text{link bit rate R in bits per second}}$

(3) **Round-trip time** (RTT, also called round-trip delay):

a. In many communications, a frame sent by a sender will be acknowledged by a corresponding receiver. The total time for such a sending-acknowledging action is important for many reasons.

b. The RTT associated with a communication link is defined as the total time between the first bit of a block being transmitted by a sender and the last bit of its acknowledgment being received.

c. T_p and T_x are two most important components in RTT. There are several other components to be discussed later.

(4) The ratio between propagation delay T_p and transmission delay T_x:

$$a = \frac{T_p}{T_x}$$

is important to discuss the link utilization (More discussions about link utilization will be in next lecture).

a. If a is less than 1, then the RTT is primarily determined by the transmission delay;

b. If a is equal to 1, then both delays have same effect;

c. If a is larger than 1, then the RTT is primarily determined by the propagation delay;

d. Example (Example 1.4, p.33): A 1000-bit block of data is to be transmitted between two computers. Determine the ratio a for the following types of data link:

(a) 100 m of twisted-pair wire with a transmission rate 10kbps;

(b) 10 km of coaxial cable with a transmission rate 1 Mbps;

(c) 50,000 km of satellite (free space) link with a transmission rate 10 Mbps.

Assume that the velocity of propagation of an electrical signal within each type of cable is $2 \times 10^8 ms^{-1}$ and that of free space is $3 \times 10^8 ms^{-1}$.

(a) $T_p = \frac{S}{V} = \frac{100}{2\times10^8} = 5\times10^{-7}$s
$T_x = \frac{N}{R} = \frac{1000}{10\times10^3} = 0.1$s
$a = \frac{T_p}{T_x} = \frac{5\times10^{-7}}{0.1} = 5\times10^{-6}$

(b) $T_p = \frac{S}{V} = \frac{10\times10^3}{2\times10^8} = 5\times10^{-5}$s
$T_x = \frac{N}{R} = \frac{1000}{1\times10^6} = 1\times10^{-3}$s
$a = \frac{T_p}{T_x} = \frac{5\times10^{-5}}{1\times10^{-3}} = 5\times10^{-2}$

(c) $T_p = \frac{S}{V} = \frac{5\times10^7}{3\times10^8} = 1.67\times10^{-1}$s
$T_x = \frac{N}{R} = \frac{1000}{10\times10^6} = 1\times10^{-4}$s
$a = \frac{T_p}{T_x} = \frac{1.67\times10^{-1}}{1\times10^{-4}} = 1.67\times10^3$

e. The **bandwidth/delay product**: this is the product of the link bandwidth (i.e. link transmission rate) and the propagation delay.

(a) For the last case in above example, where $a = 1.67\times10^3$, if the bit signals for blocks are transmitted congiguously, this product will be:

$$10\times10^6\times1.67\times10^{-1} = 1.67\times10^6\text{bits}$$

where means that there will be 1.67×10^6 bits in transition between the two computers at the two ends at any time. Namely the sending computer will have to send 1.67×10^6 bits before the first bit arrives at the receiving computer.

(b) The notion of bandwidth/delay product is important when discussing link utilizations.

5. Transmission control schemes

(1) Signal **encode** and **decode**: an (alpha or numeric) character is *encoded* by the underlying electronics when entered into a computer via a keyboard, and *decoded* upon received.

a. **Codewords**: the coded bit patterns for each character.

b. Two most widely used codes: **EBCDIC** (Extended Binary Coded Decimal Interchange Code), and **ASCII** (American Standards Committee for Information Interchange) (Fig.1.14, p.36 shows ASCII character set).

(a) EBCDIC uses 8 bits, and is used in most IBM equipments.

(b) ASCII uses 7 bits, and is the same as that defined by CCITT (IA5, International Alphabet Number 5). Adopted by ISO as ISO 645.

(c) **Printable** and **control** (BS – backspace, LF – line feed, CR – carriage return, SP – space, ES – escape) charaters.

(2) **Bit-serial transmission** vs. bit-parallel transmission

a. Within end systems all data is transferred between different components in word or byte parallel (bit-parallel) mode.

b. However, signals exchanged between end systems over a computer network are usually transferred in bit-serial mode.

c. Each NIC (network interface card) must perform conversions between bit-parallel to bit-serial modes:

   (a) Parallel-to-serial conversion of each character or byte in preparation for its transmission on the line. This is done by a PISO (parallel-in, serial-out) shift register.

   (b) Serial-to-parallel conversion of each received character or byte in preparation for its storage and processing in the receiving end system. This is done by a SIPO (serial-in, parallel-out) shift register.

   Both PISO and SIPO registers are shown in Fig.1.15(a),p.37.

(3) Conditions for data transmissions: in order for a receiving end system to successfully receive and decode the signals sent by a sending end systems, three levels of synchronization are needed. To be able to receive (decode and interpret) the bit pattern correctly, the receiving DTE must be able to determine:

a. the start of each bit cell (in order to sample the incoming signal in the middle of the bit cell)

b. the start and end of each element (character or byte)

c. the start and end of each complete message block (also called as *frame*)

d. The above tree tasks are known as **bit** (or **clock**) **synchronization**; **character** (or **byte**) **synchronization**; **block** (or **frame**) **synchronization**.

e. Synchronization can be achieved in two different ways: whether the transmitter and receiver clocks are independent.

   (a) If the two clocks are not synchronized, the transmission mode is called *asynchronous.*

   (b) If the two clocks are synchronized, the transmission mode is called *synchronous.*

(4) **Asynchronous transmission**

a. Primarily used when the data to be transmitted is generated randomly. The data generation rate is *indeterminate* (primarily designed for transmission between a keyboard and a computer). Hence data rate required is relatively low.

   (a) As data is randomly generated, the transmission line can be idle for long intervals

   (b) When a new data character (or characters) is keyed in by a user, the receiver must be able to resynchronize at the start of each new character received.

   (c) For each character, one **start bit** and one or more **stop bits** are needed (Fig.3.3,p.102).

   (d) The transmission line is in idle (**marking**) state when no transmissions.

   (e) The polarities of the start and stop bits are different to ensure there is always a minimum one transition ($1 \rightarrow 0 \rightarrow 1$) between each successive character.

   Asynchronous transmission can also be used for block transmissions: the start bit of each subsequent character immediately follows the stop bits of the previous character.

b. Principles: Fig.1.15(a), (b), p.37

   (a) The receiver maintains a clock (noted as $R \times C$) whose frequency is N times of the sender's transmit clock (noted as $T \times C$).

   (b) Fig.1.15(b) shows the timing principle. The receiver relies on the clock $R \times C$ to regularly sample the incoming signal $R \times D$.

   (c) Synchronizations: Each frame and character (byte) are treated independently for frame (block), clock (bit), and character (byte) synchronization purposes and the receiver resynchronizes at the start of each character received.

c. Bit synchronization: the ratio N plays a critical role in bit synchronization

   (a) The receiver can determine the state of each transmitted bit in the character by sampling the received the signal approximately at the center of each bit cell period. The receiver achieves this by using a clock $R \times C$ whose frequency is N times higher than the transmitted bit rate freq ($N = 16$ usually).

   (b) Fig.1.16,p.39 provides three example $R \times C$ values:

      · The frequency of the receiver clock $R \times C$ in Fig.1.16(a) is the same as that of the transmit clock $T \times C$ (hence the notion $R \times C(\times 1)$).

      · The frequency of the receiver clock $R \times C$ in Fig.1.16(b) is 4 times of that of the transmit clock $T \times C$.

      · The frequency of the receiver clock $R \times C$ in Fig.1.16(c) is 16 times of that of the transmit clock $T \times C$.

d. **Example** of bit synchronization (Example 1.5, p.40). A block of data is to be transmitted across a serial data link. A clock of 19.2kHz is available to the receiver. Deduce the suitable clock rate and estimate the worst-case deviations from the nominal bit cell centers, expressed as a percentage of a bit period, for each of the following transmission rates:

(a) 1200 bps

(b) 2400 bps

(c) 9600 bps

Answer: The worst-case deviation from the nominal bit cell center is approximately plus or minus one half of one cycle of the receiver clock:

(a) At 1200bps, the maximum $R \times C$ is $\times 16$. The maximum deviation is $1/16/2 = 3.125\%$.

(b) At 2400bps, the maximum $R \times C$ is $\times 8$. The maximum deviation is $1/8/2 = 6.25\%$.

(c) At 9600bps, the maximum $R \times C$ is $\times 2$. The maximum deviation is $1/2/2 = 25\%$.

The last case is clearly unacceptable. With a low-quality line, especially one with excessive delay distortion, even the second may be unreliable. Usually a $\times 16$ clock rate ratio is used whenever possible.

e. Byte synchronization (Fig.1.15(b), p.37). Each transmitted character is encapsulated between an additional **start bit** and one or more **stop bits**.

f. Frame synchronization: Fig.1.17,p.41.

(a) Basic idea: special characters are used to encapsulate frames (or blocks) when blocks of characters are transmitted, so that the receiver can determine when a new frame arrives and when the new frame ends.

- In the simplest form, a **STX** (start-of-text) control character and **ETX** (end-of-text) control character are used to encapsulate a data frame: Fig.1.17(a).
- Problems: a string of binary bytes that is part of a frame belonging to a multimedia application may contain the ETX within it, causing the receiver to terminate the reception process prematurely.

(b) **DLE** (data link escape) plus STX and ETX.

- Another special byte DLE is used to precede each STX and ETX when transmitting binary bit streams: Fig.1.17(b).

- **Byte buffering**: After sending out the DLE-STX start-of-frame sequence, the transmitter inspects each byte before sending it out. If it sees a DLE byte, it will insert another DLE byte after the DLE byte in the data before sending the next byte.
- The receiver will perform the reverse operation. If it sees a DLE byte, it will check to see if the next byte is ETX byte. If it is, then it the end-of-frame sequence. If it is another DLE, the receiver simply drops it. If it is a byte other than ETX or DLE byte, the receiver just receives it.

g. Misc.

(a) Overhead. Due to the extra start bit and stop bits, actual data rate differs from the ransmission rate:

(i) Each character has 1 start bit, 2 stop bits. Transmission rate is 1200bps, while data rate is $1200 \times 8/11$.

(ii) Frame synchronization presents additional overhead. DLE, STX, and ETX each has 8 bits. Byte buffering induces additional overhead.

(b) As data rate increases, the requirements of $\times N$ on the receiver's clock becomes more difficult to implement reliably. Therefore asynchronous transmission mode usually can only provide data rate around 19.2kbps, which is inadequate for many applications.

(5) **Synchronous transmissions**

a. Motivations:

(a) Asynchronous transmission is relatively inefficient due to the overhead.

(b) Asynchronous transmission becomes less reliable as bit rate increases. When bit rate increases, a small shift in the sampling point can result in errors and accurately sampling the nominal center of a bit becomes increasingly difficult.

(c) Synchronous transmission is used for transmitting large blocks of data at higher bit rates.

b. Two synchronous transmission control schemes

(a) Character oriented synchronous transmission

(b) Bit oriented synchronous transmission

Both use the same bit synchronization method.

c. Bit synchronization

(a) Bit synchronization in synchronous transmission is acheived in one of two ways:

- The clock (timing) information is embedded into the transmitted signal and subsequently extracted by the receiver: Fig.1.18(a), p.43.
- The receiver has a clock that is kept in synchronism with the received signal by a device known as **DPLL** (**digital phase-lock-loop**): Fig.1.18(b), p.43. DPLL utilizes the $1 \rightarrow 0$ and $0 \rightarrow 1$ bit bit transitions in the received signal to maintain bit (clock) synchronization over an acceptably long period.
- Hybrid schemes that exploits both schemes are also used.

(b) Clock encoding and extraction: Fig.1.19, p.44. Fig.1.19(a) shows the famous Manchester encoding while Fig.1.19(b) illustrates the differential Manchester encoding.

(c) Manchester encoding: a binary 1 is always encoded as a low-to-high signal and a binary 0 is always encoded as a high-to-low signal.

- There is always a transition (high-low or low-high) at the center of each bit cell.
- This transition is the key and is used by the receiver's clock extraction circuit to produce a clock pulse that is then delayed to the center of the second half of the bit cell.
- At this point the received signal is either high (for binary 1) or low (for binary 0) and hence the correct bit is sampled and shifted into the SIPO shift register.

(d) For differential Manchester encoding, although there is still a transition at the center of each bit cell, a transition at the start of the bit cell occurs only if the next bit to be encoded is a 0.

- The above property has the effect that the encoded output signal may take on one of the two forms, depending upon the start level of the signal (high or low).
- The two forms are simply inverted version of the other, hence the term of *differential*.
- A differential driver circuit at the receiver produces a pair of differential signals and the differential receiver operates by determining the difference between these two signals.

(e) **Balanced codes**: both Manchester and Differential Manchester encoding schemes are balanced codes, as there is no mean DC value associated with them (ie the sum of DC values in the signal is zero).

- This property is guaranteed by the transition at the center of each encoded bit.
- The property is important. It allows the received signals to be **AC coupled** to the receiver electronics using a transformer. The receiver electronics can then operate using its own power supply since this is effectively isolated from the power supply of the transmitter.

d. Digital phase-lock-loop

(a) Basic idea:

- the receiver maintains a clock that is N times faster than the incoming bit rate (similar to asynchronous transmission, usually $N \leq 16$). This clock is kept in synchronism with the incoming bit rate with the help of the incoming bit stream.
- Similar to bit synchronization, DPLL uses bit transitions to synchronize the receiver's clock with the bit rate.

(b) NRZI (non-return-to-zero-inverted) encoding: Fig.1.20(a), p.46.

- The signal level (for bit 1 or 0) does not change for the transmission of a binary 1, whereas a binary 0 causes a change.
- The above property means that if there are no contiguous streams of binary 1s there will always bit transitions in the incoming signal of a NRZI waveform.
- If a bit-oriented scheme with zero insertion is used, an active line will always have a binary 0 in the transmitted bitstream at least every five bit cells. As a result, the corresponding waveform will contain a guaranteed number of transitions (remember: every bit 0 causes one transition).
- It is these transitions that will enable the receiver to keep its clock in synchronism with the incoming bit stream.

(c) The DPLL circuit: Fig.1.20(b).

- A fairly accurate crystal clock is connected to the DPLL. The clock is used by DPLL to derive the timing interval between successive samples of the received bit stream.
- The frequency of the clock is ususally 32 times the incoming bit rate.

(d) Operation when the clock is in synchronism with the incoming bit stream: Fig.1.20(c).

- In this mode, the operation is quite simple. The state of incoming signal on the line will be sampled (and hence clocked into the SIPO shift register) at the center of each bit cell with exactly 32 clock periods between each sample.

· Note: there is actually no such an operation mode, as the DPLL device has to continuously to monitor any clock drift (against the incoming bit rate).

(e) When the receiver's clock and the incoming bitstream drift out of synchronism, the DPLL will adjust its sampling instance in discrete increments, as shown in Fig.1.20(d).

· If there are no transitions on the line, synchronization is not possible. In this case the DPLL will simply generates sampling puse every 32 clock periods after the previous one.

· However, whenever a transition ($1 \to 0$ or $0 \to 1$) is detected, the DPLL will determine the time interval between the previous sampling pulse and the next sampling pulse by comparing the position of the current transition with the position where DPLL thought it should occur.

· To accurately adjust the synchronism, each bit period is divided into five segments, shown as A, B, C, D, and E.

· When a transition occurs within segment A, it indicates that the previous sampling pulse was too close to the next transition and hence late. In order to sample the nominal center of next bit cell the next sampling pulse should be shortened to 30 clock periods.

· When a transition occurs within segment E, it indicates that the previous sampling pulse was too earlier relative to next transition. In order to sample the nominal center of next bit cell the next sampling pulse should be lengthened to 34 clock periods.

· Similarly, when a transition occurs within segment B or D, the adjustment will be -1 or $+1$, respectively.

· Hence, successive adjustments keep the generated sampling pulses close to the nominal center of each bit cell.

· In practice, the widths of the five segments are not equal. The two outer segments A and E are made longer than the three inner segments. A typical arrangements will have $A = E = 10$, $B = D = 5$, and $C = 2$. Thus in the worst case the DPLL requires 10 bit transitions to converge to the nominal center of a waveform: 5 bit periods of coarse adjustments (± 2), and 5 bit periods of fine adjustments (± 1).

· It is usual before transmitting the first frame, a sender will first transmit a number of bytes to provide a minimum of 10 bit transitions. Two bytes of all 0s would provide 16 transitions. This action will ensure that the DPLL at the receiver will generate sampling pulses at the nominal center of each bit cell by the time the opening character or byte of a frame is received.

· In practice, only minor adjustments are needed during the reception of a frame.

(f) **Modulation rate**

· Fact: the maximum rate at which a NRZI encoded waveform changes polarity is one half that of the Manchester encoding. (Why?)

· Fact: if the bit period is T, then the maximum rate with NRZI encoding is 1/, while the maximum rate with Manchester encoding is 2/. The maximum rate is defined as **modulation rate**.

· In terms of Fourier series, the highest fundamental frequency component of each scheme is $1/T$ and $2/T$ respectively.

· Implication: higher maximum rates requires higher transmission bandwidths.

· Consequently, Manchester and differential Manchester encodings are both used more extensively in applications such as LANs that have relatively high bandwidth and short cable runs. These applications operate at high bit rates and the attenuation and bandwidth are generally not a problem.

· On the other hand, in networks such as ISDN where twisted-pair cable is often used with relatively high bit rates and over distances of several kilometers, encoding schemes such as NRZI are used with each bit represented by a full width of pulse.

e. Character-oriented transmission

(a) Both character-oriented and bit-oriented synchronous transmissions use the same bit-synchronization scheme just discussed. The main differences of the two are in character and frame synchronization.

(b) Character-oriented transmission is used primarily for transmission of larage blocks of characters such as files.

(c) Frame synchronization in character-oriented transmission: Fig.1.21(a) and (c), p.49

· Similar to asynchronous transmissions, a pair of STX-ETX characters are used. However the STX character is preceded with several SYN characters.

- DLE (data link escape) characters are added when transmitting binary data streams: Fig.1.21(c).

(d) Character synchronization in character-oriented transmission

- There is no start bit or stop bit for each character.
- A special control character known as **synchronization idle** or (**SYN**) is sent several times before each block of characters. These SYN characters have two purposes: they allow the receiver to obtain synchronization; and they allow the receiver to start to interpret the received bitstream on the correct character boundaries (namely to achieve character synchronization).
- The sender transmits a number of SYN characters. The SYN characters allow the receiver to achieve character synchronization before the STX character arrives.
- Once the receiver has obtained bit synchronization, it enters what is known as **hunt mode**.
- Within hunt mode, the receiver starts to interpret the received bitstream in a window of eight bits. It checks to see if the last eight bits form a known SYN character. If not, it will repeat the check. If they do, the receiver knows that if has found the correct boundary and hence the following characters are then read after each subsequent eight bits have been received.
- The receiver then searches for the STX character, and then receives each data byte, and eventually will find the ETX character.

f. Bit-oriented transmission

(a) Limitations of character-oriented transmissions

- Character-oriented transmissions requires a number of characters before and after each frame. Hence it is relatively inefficient for transmitting binary data.
- The special control characters (SYN, STX, ETX, and DLE) are character-set dependent.

Bit-oriented transmission overcomes these problems and is preferred scheme of synchronous transmission.

(b) Bit-oriented transmission scheme differs from character-oriented transmission mainly in the way the start and end of each frame is signaled: Fig.1.22, p.50.

- The start and end of a frame are both signaled by the same unique 8-bit pattern 01111110, known as **flag byte** or **flag pattern**.

- The term *bit-oriented* comes from the fact that received bitstream is searched by the receiver on a bit-by-bit basis both for the start-of-frame and end-of-frame flags.
- To maintain bit synchronization the sender sends a string of **idle bytes** (each consisting of 01111111) preceding the start-of-frame flag.
- On receipt of the opening flag, the received frame contents are read and interpreted on 8-bit (byte) boundaries until the end-of-frame flag is detected.
- To prevent the flag byte from appearing within the frame contents, the **zero bit insertion** (also called **bit stuffing**) is used. The circuit that performs this action is attached to the PISO register (Fig.1.22(b)).
  The circuit is enabled only when transmitting frame contents. If enabled, whenever it detects there is five consecutive 1s, it will insert a 0. This way, the pattern 01111110 will never appear in the frame contents.
- A corresponding circuit at the receiver located prior to the input of the SIPO shift register performs the reverse action. Whenever it detects five 1s followed by a 0, it will remove that trailing zero. Fig.1.22(c) shows an example.