DATA ENCODING - 10.02.2006 [1]

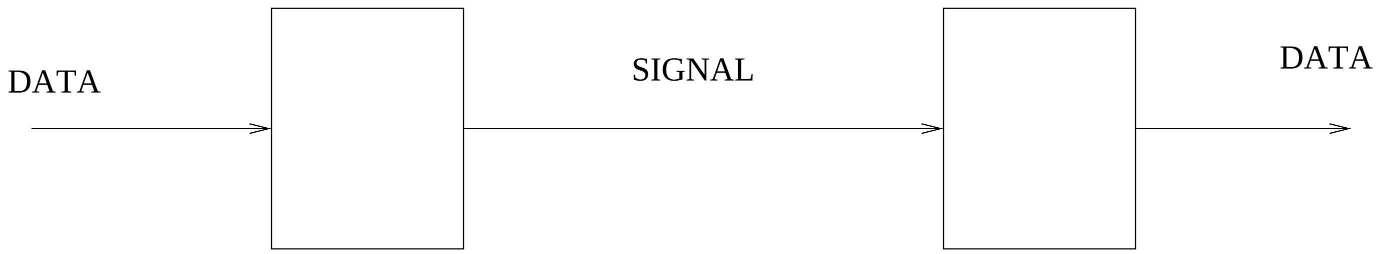
Parang Saraf[1] Shravya Reddy Konda[2]

[1]03CS1008 {parang.saraf@iitkgp.ac.in}
[2]03CS1010 { shravya_k03@yahoo.co.in }

[1]

Data and Signals

- Data may be Analog or Digital.
- We may use either Analog or Digital signals to represent either form of data.
- Various combinations are possible.



Which combination to use ?

- Digital Data, Digital Signal
  - encoding/decoding equipments are less complex/expensive.
- Analog Data, Digital Signal
  - many existing signals are analog; this permits the use of modern transmission and switching equipment.
- Digital Data, Analog Signal
  - some transmission media like optical media and wireless only propagate analog signals.
- Analog Data, Analog Signal
  - modulation is easy and cheap; FDM can be used to share bandwidth.

[1]Ref: Data and Computer Communications by William Stallings

I. Digital Data, Digital Signals

- A digital signal is a sequence of discrete, discontinuous voltage pulses.
  - Each pulse is a signal element.
  - Binary data are transmitted by encoding each data bit into signal elements.
- For proper reception, the receiver must know the timing and signal level of each bit.
- Following statements are true in general:
  - An increase in data rate increases bit error rate.
  - An increase in signal-to-noise ratio decreases bit error rate.
  - An increase in bandwidth allows an increase in data rate.

Digital Signal Encoding Formats

- **Nonreturn-to-Zero-Level (NRZ-L)**
  - 0 = High Level
  - 1 = low level
- **Nonreturn-to-Zero-Inverted (NRZI)**
  - 0 = no transition at beginning of interval (one bit time)
  - 1 = transition at the beginning of interval
- **Bipolar Alternate-Mark-Inversion (AMI)**
  - 0 = no line signal
  - 1 = positive or negative level, alternating for successive ones
- **Pseudoternary**
  - 0 = positive or negative level,alternating for successive zeros
  - 1 = no line signal
- **Manchester**
  - 0 = transition from high to low in middle of interval
  - 1 = transition from low to high in middle of interval
- **Differential Manchester**
  - Always a transition in middle of interval
  - 0 = transition at beginning of interval
  - 1 = no transition at beginning of interval

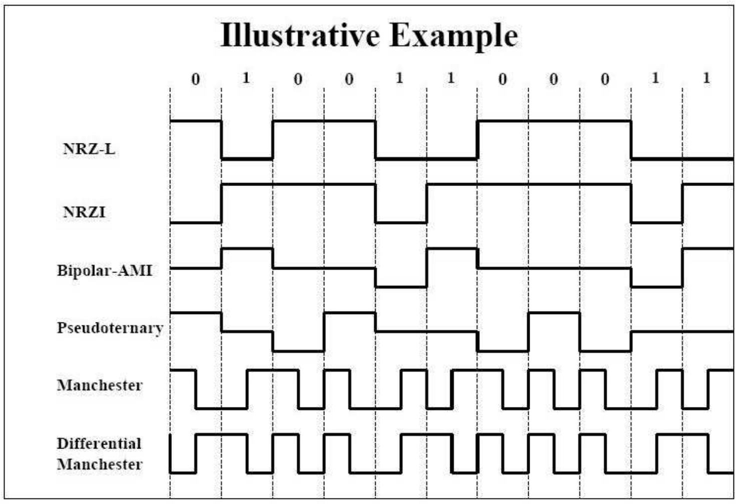


Fig. 1. Illustrative Example

Comparative Study

- NRZ-L is the simplest.
- NRZI is a type of differential encoding.
  - It may be more reliable to detect a transition in the presence of noise than to compare a value to a threshold.
  - Even if the polarities are accidentally reversed, it will work.
- The main limitation of NRZ codes are the presence of a dc component and the lack of synchronization capability.
  - Long string of 1's or 0's may lead to synchronization loss.
- Bipolar-AMI (Pseudoternary)
  - No loss of synchronization if a long string of 1's(0's) occur.
  - Long string of 0's,(1's) however, will cause problem.
  - No dc component.
- Manchester and Differential Manchester
  - In Manchester code, the mid-bit transition serves as a clocking mechanism and also as data.
  - In Differential Manchester code, the mid-bit transition is used only to provide clocking.
  - Bandwidth required is greater as compared to NRZ
    - as many as two transitions are possible in a bit time.
  - No dc component.
  - Also collectively called **biphase techniques**.

Modulation Rate

- Data rate is expressed in bits per second.
- Modulation rate is expressed in bauds.
- General relationship:
  - $D = R/L = R/log_2M$
    - D is the modulation rate in bauds.
    - R is the data rate in bps.
    - M is the number of differential signal elements($= 2^L$)
    - L is the number of bits per signal element.

Encoding	**Minimum**	**101010...**	**Maximum**
NRZ-L	0	1.0	1.0
NRZI	0	0.5	1.0
Binary-AMI	0	1.0	1.0
Pseudoternary	0	1.0	1.0
Manchester	1.0	1.0	2.0
Differential Manchester	1.0	1.5	2.0

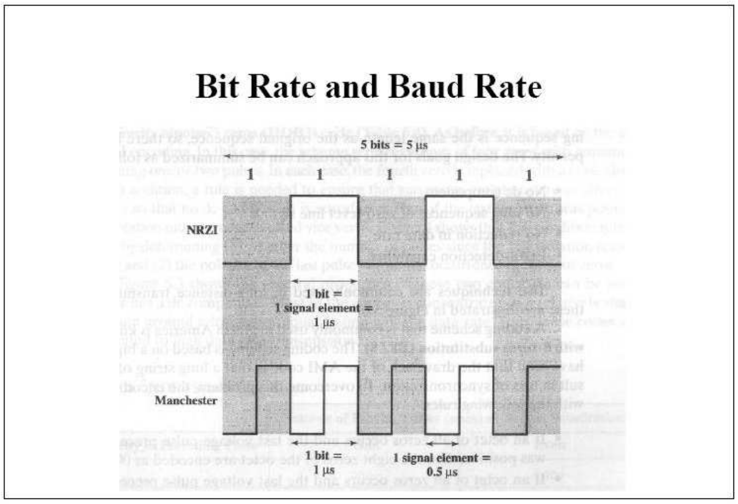


Fig. 2. Bit Rate and Baud Rate

Scrambling Techniques

- Altough the biphase techniques are widely used in LANs, they are not suitable for long-distance applications.
  - May require a high signaling rate relative to the data rate.
- Scrambling schemes may be used:
  - Sequences that would result in a constant voltage level on the line are replaced by filling sequences that will provide sufficient transitions for the receiver's clock to maintain synchronization.
- Two commonly used scrambling techniques:
  - **Bipolar with 8 zeros substitution (B8ZS)**
    - Based on bipolar-AMI.
    - Commonly used in North America.
    - If an octet of all 0's occurs and the last voltage pulse preceding this octet was positive, then the eight 0's are encoded as $000 + -0 - +$
    - If an octet of all 0's occurs and the last voltage pulse preceding this octet was negative, then the eight 0's are encoded as $000 - +0 + -$
- **High density bipolar 3 zeros (HDB3)**
  - Also based on bipolar-AMI.
  - Commonly used in Europe and Japan.
  - Replaces string of 4 zeros with sequences of one or two pulses.
- **HDB3 substitution rules**

  No. of bipolar pulses(ones) since last substitution.

Odd	**Even**	**Polarity of Preceding Pulse**
000−	+00+	−
000+	−00−	+

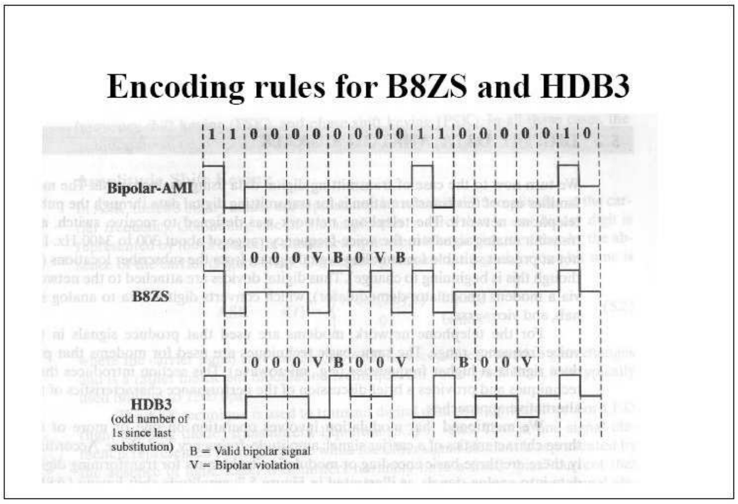


Fig. 3. Encoding Rules

Digital Signal Encoding for High Speed LANs

* 4B/5B-NRZI

- Used in both 100BASE-X and FDDI.
- Encode the binary data to guarantee the presence of transitions; then use NRZI.
- Greater efficiency than Manchester encoding.
  - 4 bits of data are encoded into a symbol with 5 code bits.
  - 80% efficiency.
  - For example, 100 Mbps can be achieved with 125 Mbauds.

* 8B/10B

- Used in Fibre Channel and Gigabit Ethernet.
- Concept similar to above.

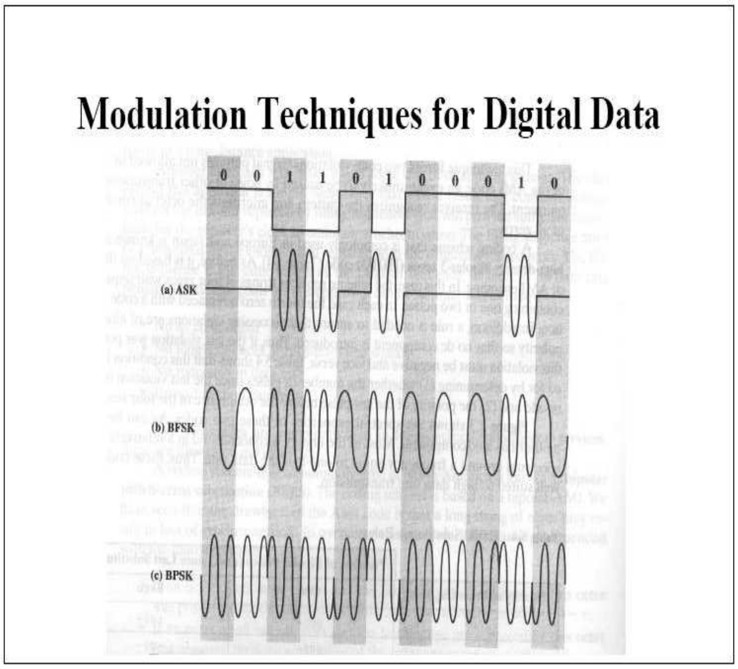
4B/5B Code Table

Data Input	Code Group
0000	11110
0001	01001
0010	10100
0011	10101
0100	01010
0101	01011
0110	01110
0111	01111
1000	10010
1001	10011
1010	10110
1011	10111
1100	11010
1101	11011
1110	11100
1111	11101

II. Digital Data, Analog Signal

- Most familiar use is the transmission of digital data through the public telephone network.
  - Can transmit analog signals in the 300 to 3400 Hz range.

- Three basic encoding or modulation techniques:
  - **Amplitude-Shift Keying (ASK)**
    - The two binary values are represented by two different amplitudes of the carrier frequency.
    - Used to transmit digital data over optical fiber.
  - **Frequency-shift keying (FSK)**
    - The two binary values are represented by two different frequencies near the carrier frequency.
    - Less susceptible to error than ASK.
  - **Phase-shift keying (PSK)**
    - The phase of the carrier signal is shifted to represent data.
    - Phase shift of 180 degree $\Longrightarrow$ each signal element represents one bit.
    - Phase shifts are multiples of 90 degrees $\Longrightarrow$ each signal element represents two bits.



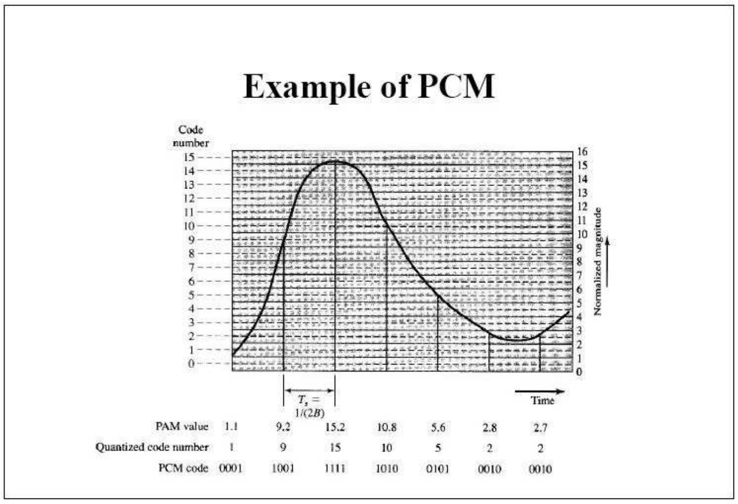
III. Analog Data, Digital Signals

- The device used for converting analog data into digital form for transmission, and subsequently recovering the original analog data from the digital, is known as a **CODEC (coder-decoder)**

- Two principal techniques used in CODEC:
  - **Pulse code modulation**
  - **Delta modulation**

Pulse Code Modulation

- **Basic principle**
  - The allowed voltage excursion is divided into n (say, 256) equally spaced amplitude levels ranging from 0 to V volts.
  - Samples are taken at regular intervals of time, and the nearest discrete amplitude level is chosen as the one transmitted.
  - Receiver reconstructs the signal using a low-pass filter.
  - This is actually called **Pulse Amplitude Modulation (PAM).**
  - Quality of reproduction depends on the fineness of the sampling process.
- The PAM samples can be quantized, each sample represents as $log_2 n$ bits.
  - **Pulse Code Modulation (PCM)**
- The PCM technique is based on the **Sampling Theorem**:
  - If a signal is limited to a bandwidth of B Hz, then the signal can be reproduced without distortion if it is sampled at a rate of 2B Hz.
- **Example**
  - A color video signal having 6 MHz bandwidth, digitized using
    - 256 quantization levels (8 bits per sample).
    - sampling rate of 12 MHz.
  - The resulting digital signal would require a 96 Mbps transmission rate.

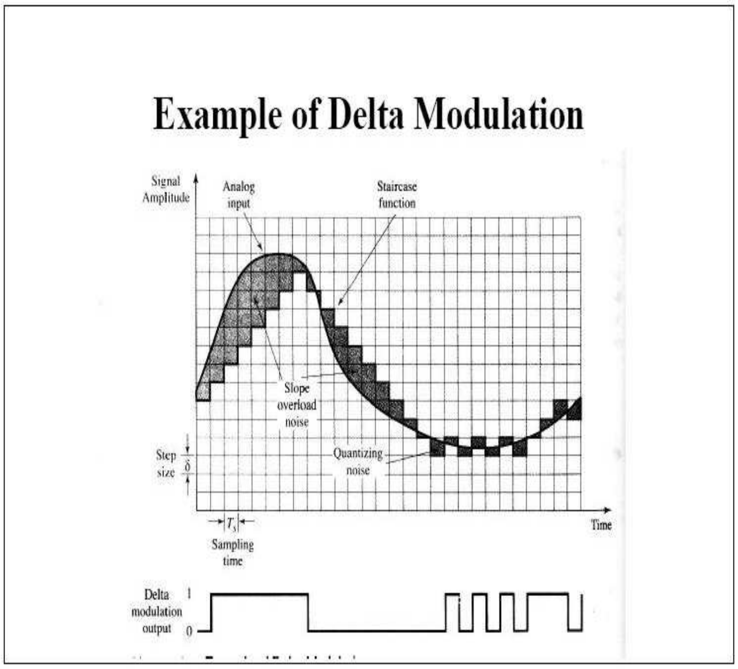


Differential PCM

- For voice signals, statistical techniques can be used to reduce the number of bits to be transmitted.
- In differential PCM, instead of sending the digitized amplitude, the difference between the current value and the previous one is sent.
  - Since jumps of more than 16 steps are uncommon on a scale of 256, 5-bits should suffice instead of 8.
  - If the signal does jumps wildly occasionally, the encoding logic may require several sampling periods to catch up.

Delta Modulation

- This is a special case of differential pulse code modulation.
- Each sampled value differs from its predecessor by either $+1$ or -1
- A single bit is transmitted, telling whether the new sample is above or below the previous one.
- The main advantage is simplicity of the implementation.



IV. Analog Data, Analog Signal

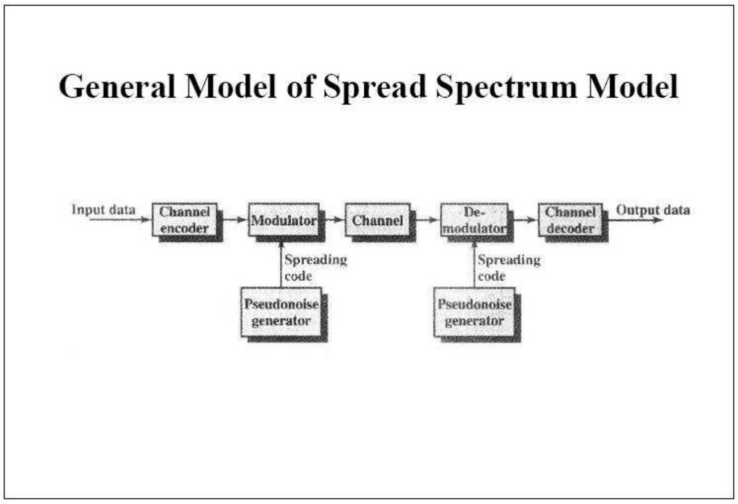
- **Three principal techniques:**
  - Amplitude modulation.
  - Frequency modulation.
  - Phase modulation.

Spread Spectrum

- This technique can be used to transmit either analog or digital data, using an analog signal.
  - Does not directly fit into any of the four categories discussed.
  - Essential idea is to spread the information signal over a wider bandwidth to make jamming and interception more difficult.
  - Various techniques have been proposed:
    * **Frequency-hopping.**
    * **Direct sequence.**

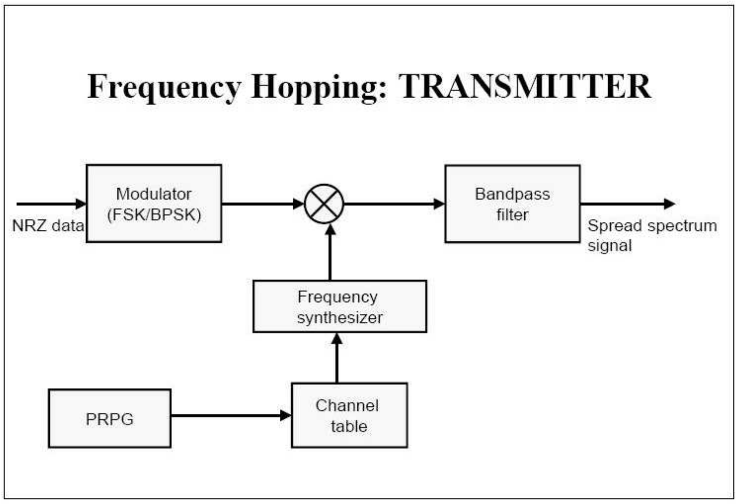
Key Characteristics of any Spread Spectrum System

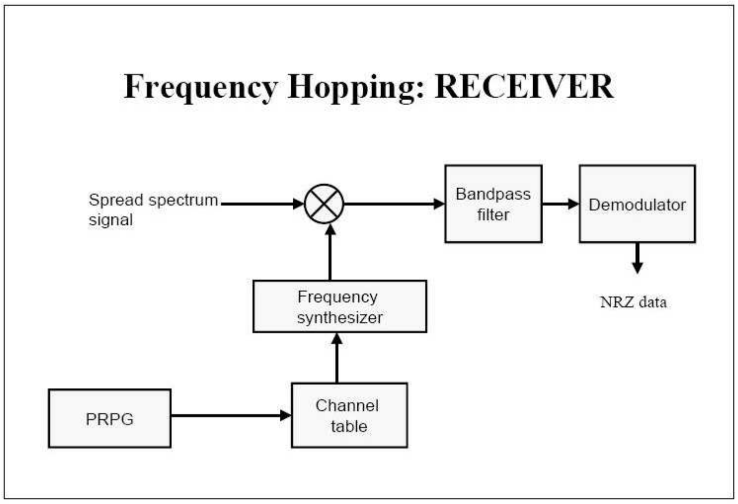
- Input is fed to a channel encoder that produces an analog signal with a relatively narrow bandwidth around some center frequency.
- This signal is further modulated using a pseudorandom sequence.
  - The effect of this modulation is to significantly increase the bandwidth(spread spectrum) of the signal to be transmitted
- On the receiving end, the same pseudorandom sequence is used to demodulate the spread spectrum signal.
- Finally, the signal is fed into a channel decoder to recover the data.



Frequency-Hopping Technique

- The signal is broadcast over a seemingly random series of radio frequencies, hopping from frequency to frequency at split-second intervals.
- The receiver, hopping between frequencies in synchronization with the transmitter, picks up the message.
- Attempts to jam the signal succeed only at knocking out a few bits of it.
- For both the transmitter and the receiver, a pseudorandom number source serves as an index into a table of frequencies.
  - At each successive interval, a new frequency is selected from the table.





Direct Sequence Technique

- Here, each bit in the original signal is represented by multiple bits in the transmitted signal, known as a *chipping code.*
- The chipping code spreads the signal across a wider frequency band in direct proportion to the number of bits used.
- The digital information stream is combined with a pseudorandom bit stream using an exclusive or.
  - The pseudorandom bit stream is clocked at a rate higher than the information rate.
  - The combination bit stream has a wider bandwidth than the information stream.