Department of Electrical Engineering
University of Arkansas

ELEG5663 Communication Theory
Ch. 2 Formatting and Baseband Signaling

Dr. Jingxian Wu
wuj@uark.edu

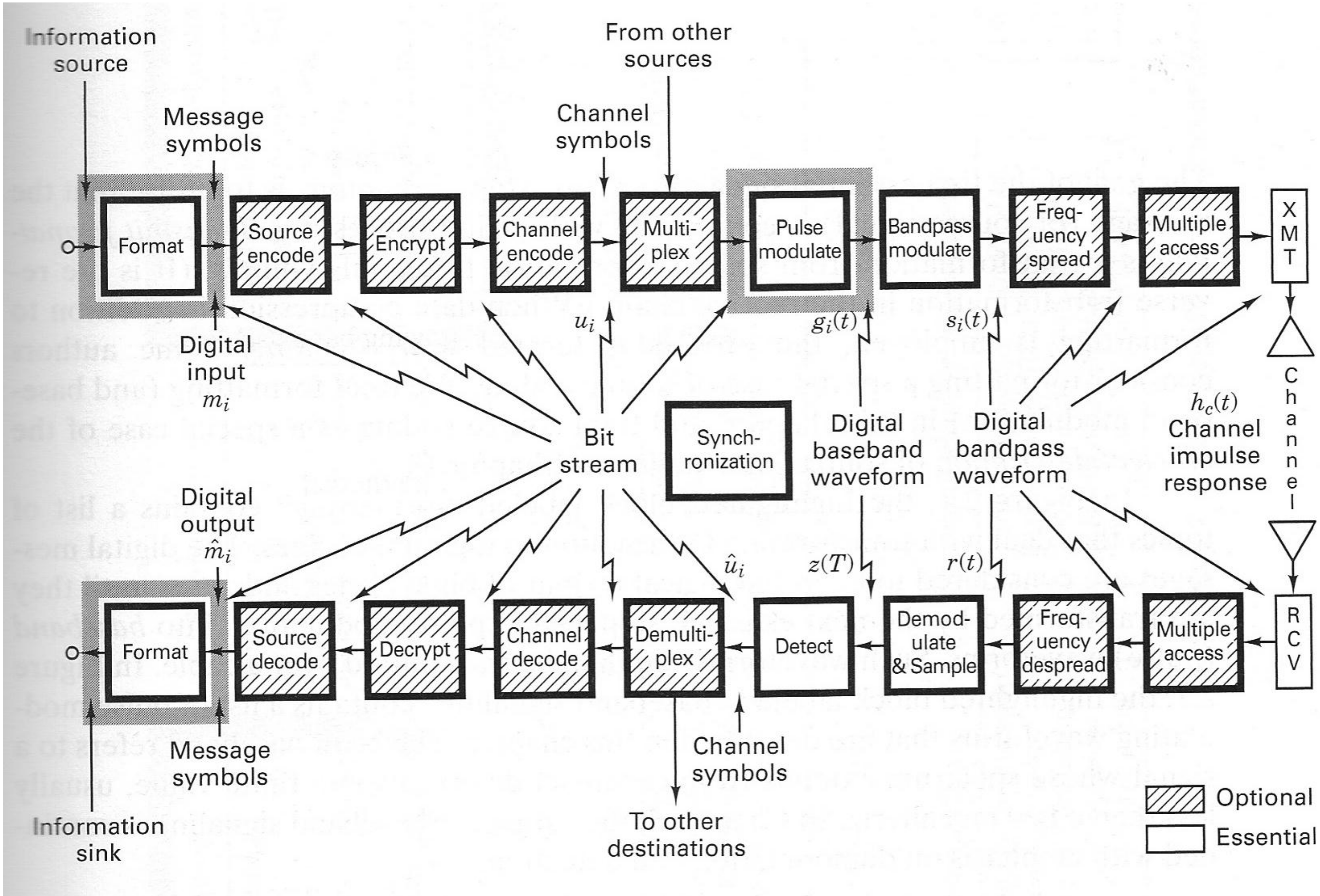
OUTLINE

- **Baseband system**
- **Formatting**
- **Baseband modulation**
- **Correlative coding**

BASEBAND SYSTEM

- **Formatting**
  - The process of transforming source information into logical digital symbols (0s and 1s).
    - formatting textual information
      - Character encoding
    - Formatting analog information
      - Sampling, quantization, pulse code modulation (PCM)
  - The output of formatting are logic 0s and 1s.
- **Baseband Signaling**
  - The process of transforming logic 0s and 1s into baseband electric waveforms that are compatible to the transmission media (channel).
  - Baseband signal: the signal whose spectrum extends from (or near) DC up to some finite value, usually less than a few MHz.
- **Baseband system**
  - The communication system with baseband signal directly transmitted in the channel (e.g. Ethernet).

BASEBAND SYSTEM

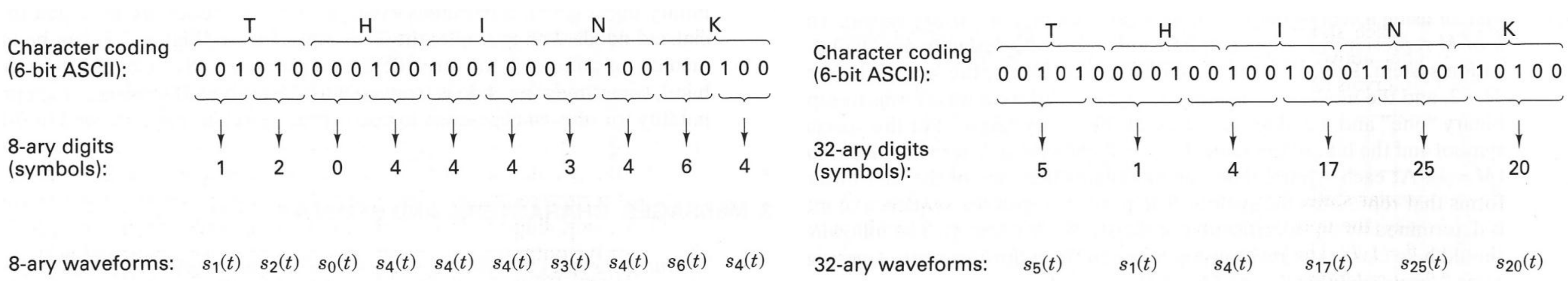


OUTLINE

- **Baseband system**
- **Formatting**
  - Formatting textual messages
  - Formatting analog information
- **Baseband modulation**
- **Correlative coding**

FORMATTING: TEXTUAL DATA

- **Character encoding**
  - The process of transforming text data into binary digits (bits).
  - Examples:
    - ASCII (7-bits, Table 2.3 of Sklar's book).
- **Message, bits, and symbols**
  - Message: the original textual message. ('T')
  - Bit: the encoded binary bits (e.g. 001010)
  - Symbols: a combination of *k* bits can form $M = 2^k$ symbols
    - M-ary system

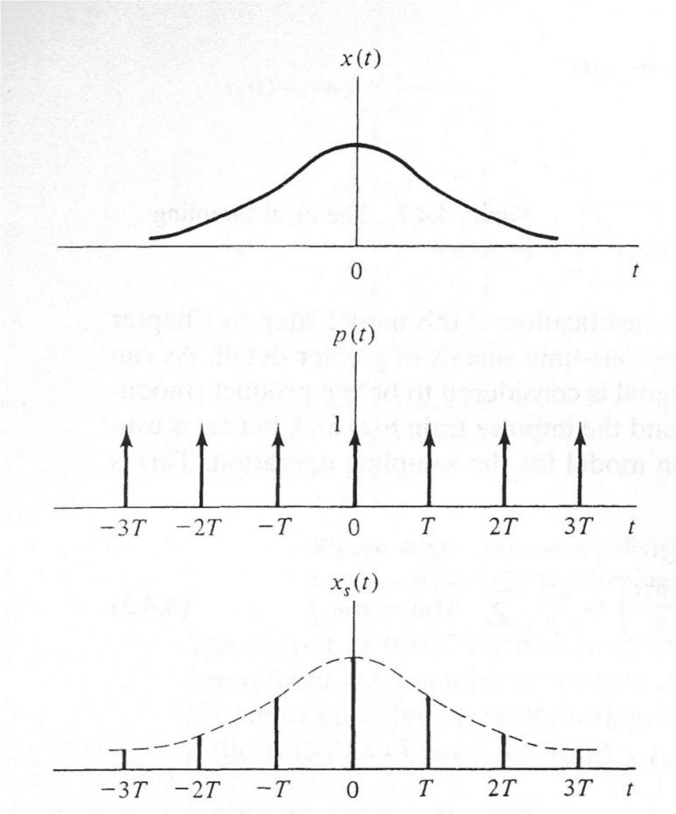


FORMATTING: ANALOG INFORMATION

- **Formatting analog information**
  - Transforming analog waveform into bits
    - analog to digital conversion (ADC)
    - Some information will be lost during this process.
  - Sampling, quantization, PCM
- **Sampling**
  - The process of taking samples of the analog waveform in the time domain.
    - Impulse sampling (ideal)
    - Natural sampling
    - Sample-and-hold

FORMATTING: SAMPLING

- **Impulse sampling**
  - Multiplication of the analog signal with a periodic train of unit impulse function.
  - Time domain:



$x(t)$

$$p(t)=\sum_{n=-\infty}^{+\infty}\delta(t-nT_s)$$

sampling period

$$x_s(t)=x(t)p(t)=\sum_{n=-\infty}^{+\infty}x(nT_s)\delta(t-nT_s)$$

FORMATTING: SAMPLING

- **Impulse sampling: frequency domain**
  - Fourier transform of the impulse train
    - impulse train is periodic

Fourier series

$$p(t)=\sum_{n=-\infty}^{+\infty}\delta(t-nT_s)=\frac{1}{T_s}\sum_{n=-\infty}^{+\infty}1\times e^{j\frac{2n\pi}{T_s}t}$$

    - Find Fourier transform on both sides

$$P(f)=\frac{1}{T_s}\sum_{n=-\infty}^{+\infty}\delta(f-nf_s)$$

    - Time domain multiplication ➔ Frequency domain convolution

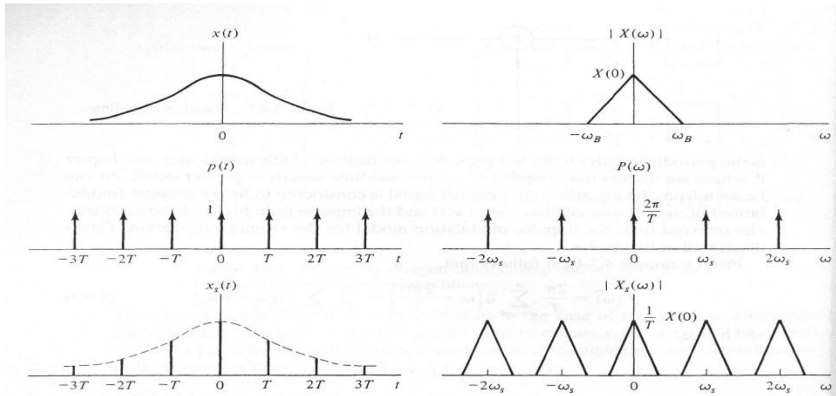
$$x(t)p(t)\Leftrightarrow\left[X(f)\otimes P(f)\right]$$

$$x(t)p(t)\Leftrightarrow\frac{1}{T}\sum_{n=-\infty}^{+\infty}X(f-nf_s)$$

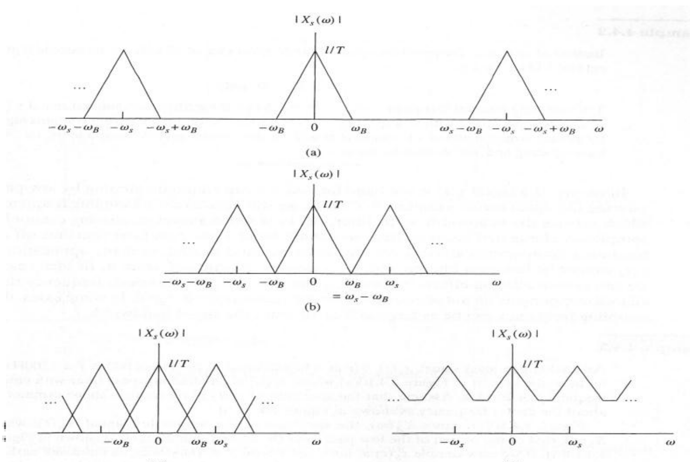
UNIVERSITY OF ARKANSAS

FORMATTING: SAMPLING

- **Impulse sampling**
  - Sampling in time domain ➔ Repetition in frequency domain



- **Sampling theorem**
  - If the sampling rate is twice of the bandwidth, then the original signal can be perfectly reconstructed from the samples. $f_s > 2f_B$



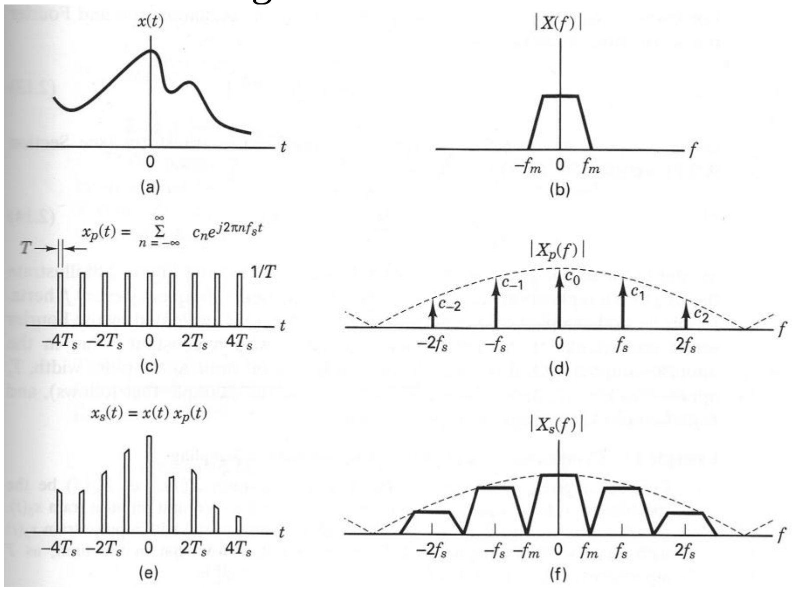
$f_s > 2f_B$ oversampling

$f_s = 2f_B$ Nyquist rate

$f_s < 2f_B$ aliasing

FORMATTING: SAMPLING

- **Natural sampling**
  - Multiplication of analog signal with a periodic train of rectangular pulses.
    - More practical than impulse sampling
    - Equivalent to using a switch.



$$x_p(t) = \sum_{n=-\infty}^{+\infty} \frac{1}{T} rect\left(\frac{t - nT_s}{T}\right) = \sum_{n=-\infty}^{+\infty} c_n e^{j2\pi n f_s t}$$

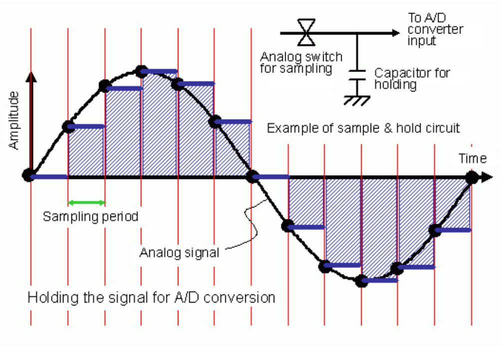
$$X_p(f) = \sum_{n=-\infty}^{+\infty} c_n \delta(f - nf_s)$$

$$X_s(f) = X(f) \otimes X_p(f) = \sum_{n=-\infty}^{+\infty} c_n X(f - nf_s)$$

FORMATTING: SAMPLING

- **Sample-and-Hold (flattop sample)**
  - The simplest and most popular sampling method
  - Can be described by the convolution of impulse sampled signal with a unity amplitude rectangular pulse of pulse width T_s

$$x_s(t) = rect\left(\frac{t}{T_s}\right) \otimes \left[\sum_{n=-\infty}^{+\infty} x(nT_s)\delta(t - nT_s)\right]$$



$$X_s(f) =$$

  - The spectrum is similar to that of natural sampling

FORMATTING: SAMPLING

- **Engineering consideration of sampling**
  - To avoid aliasing, a prefilter can be used to limit the bandwidth of signal.
  - Realizable filters usually have a transition between the passband and stop band: transitional bandwidth
    - Tansitional bandwidth is usually 10-20% of the signal bandwidth
  - Engineer's version of Nyquist sampling rate:

$$f_s \geq 2.2 f_m$$

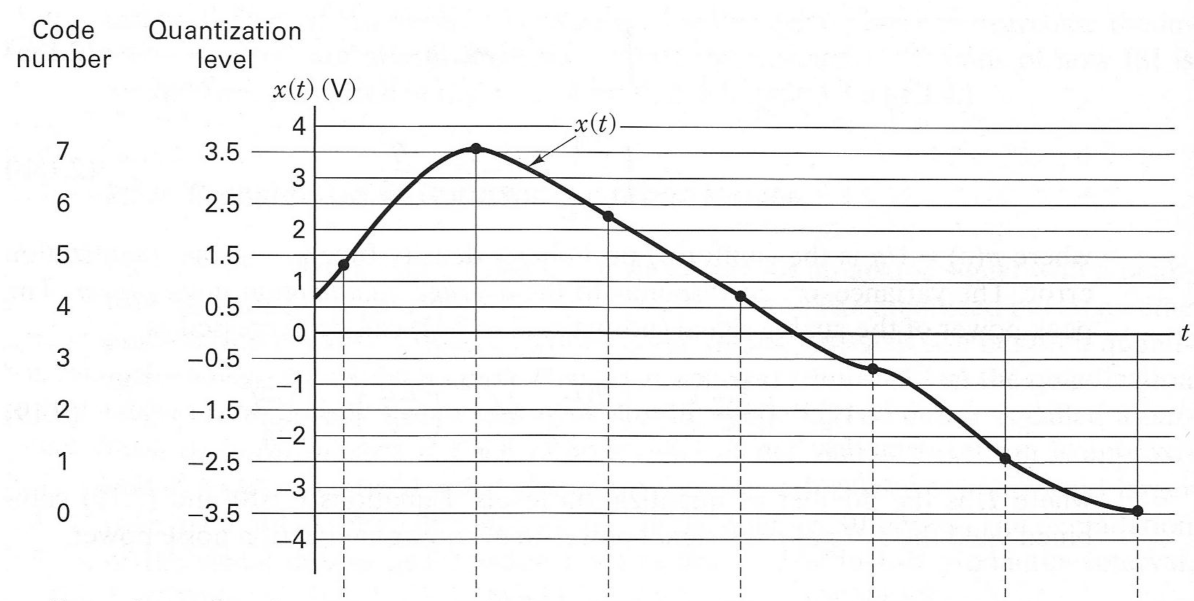
    - Example: music bandwidth: 20 kHz, the engineer's Nyquist rate is 44 ksamples/sec. The actual sampling rate used in CD: 44.1 ksamples/sec.
  - Oversampling:
    - Self-study Section 2.4.3 of Sklar's book.
  - Signal interface for a digital system
    - Self-study Section 2.4.4 of Sklar's book.

FORMATTING: QUANTIZATION

- **Quantization**
  - Why quantization?
    - The output of sampler can still have infinite number of values.
  - Quantization is the process to limit the samples to finite values
    - E.g. 3-bit quantization ➔ 8 possible values.
    - Quantized samples
    - Some information will lost during quantization ➔ quantization noise.
  - Types of quantization:
    - Uniform quantization
    - Non-uniform quantization (used to accommodate the properties of human speech).

FORMATTING: QUANTIZATION

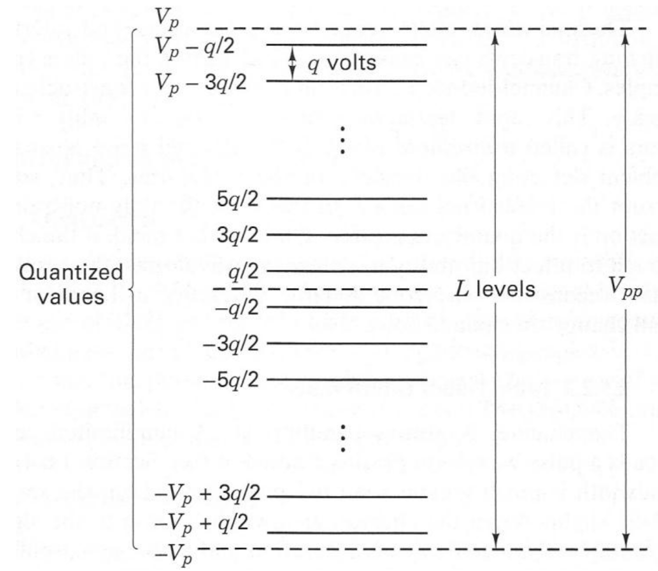
- **Uniform quantization**
  - The range of interest is divided into L levels.
    - Quantile interval: step size between two adjacent levels.
    - uniform quantization: quantile intervals are the same throughout.
    - Each level is quantized into $l = \log_2 L$ bits



Natural sample value	1.3	3.6	2.3	0.7	−0.7	−2.4	−3.4
Quantized sample value	1.5	3.5	2.5	0.5	−0.5	−2.5	−3.5
Code number	5	7	6	4	3	1	0
PCM sequence	101	111	110	100	011	001	000

FORMATTING: QUANTIZATION

- **Quantization noise: relationship with SNR**
  - An analog signal with peak-to-peak voltage range V_{pp} is uniformly quantized into L levels ➔ Quantile interval q = :
  - Quantization noise: the difference between actual value and quantized value $e = x(t) - x_Q(t)$
    - The maximum value of e:
    - The minimum value of e:
    - e is uniformly distributed in $[-q/2, q/2]$
  - pdf of quantization noise:
  - Quantization noise variance (noise power):
  - Signal power:
  - Signal-to-noise ratio (SNR): $SNR = \dfrac{L^2 q^2/4}{q^2/12} = 3L^2$



FORMATTING: QUANTIZATION

- **Quantization noise: relationship with** V_{pp}
  - The magnitude of the quantization distortion error can be specified as a fraction *p* of the peak-to-peak analog voltage V_{pp}

$$|e| \leq pV_{pp}$$

  - The relationship between maximum quantization error and V_{pp}

  - The relationship between quantization level *L* and *p*:

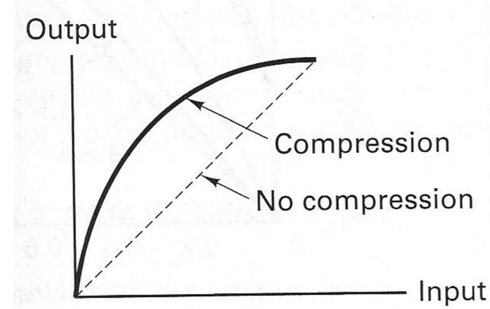
$$L \geq \frac{1}{2p}$$

FORMATTING: PCM

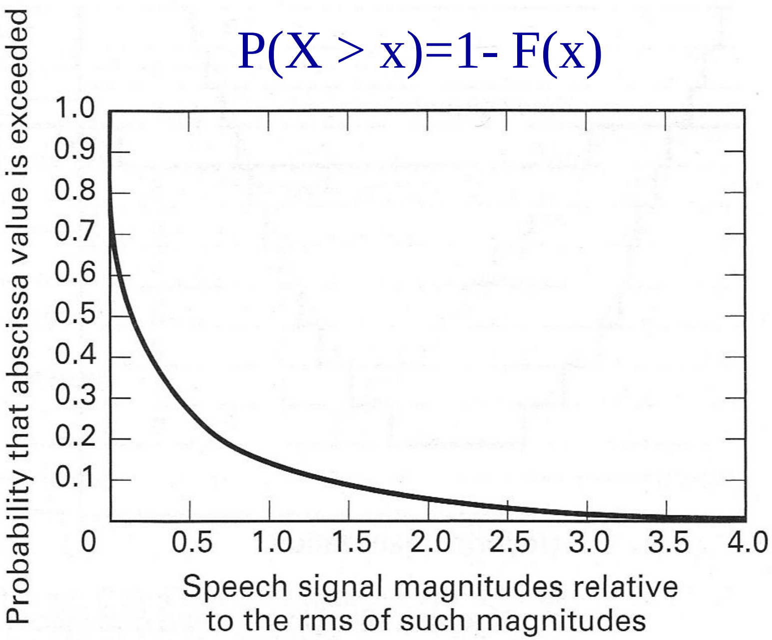
- **Pulse code modulation (PCM)**
  - 1. Sampling
    - E.g. music: 44.1ksamples/sec
  - 2. Quantization and encoding
    - E.g. 8 levels ➔ encoded into 3 bits/sample
  - Data rate:
  - More quantization levels ➔ higher data rate
  - More quantization levels ➔ higher quantization SNR ➔ better quality

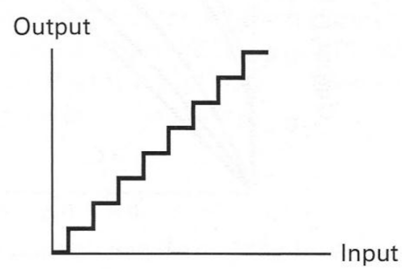
FORMATTING: QUANTIZATION

- **Non-uniform quantization**
  - The statistical property of speech signal
    - 50% of the time, the amplitude is less than ¼ of the rms (root mean square) value.
    - Only 15% of the time does the voltage exceed the rms value.
  - Small quantile levels for small value, large quantile levels for large value
    - Non-uniform quantization

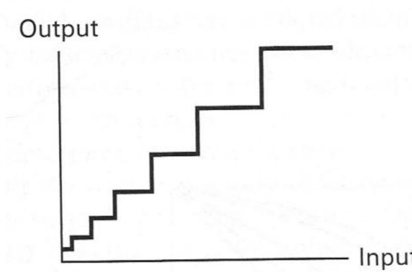


$P(X > x)=1- F(x)$





uniform quantization

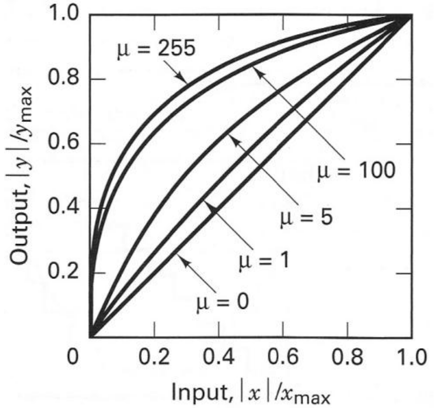


nonuniform quantization

FORMATTING: QUANTIZATION

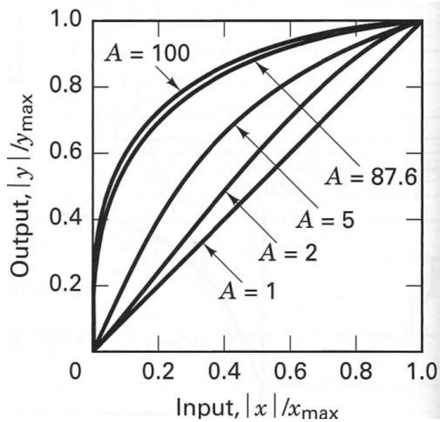
- **Non-uniform quantization**
  - 1. distort the original signal:
    - Steeper slope for small magnitude signal ➔ a smaller change in low magnitude signal results in more steps.
  - 2. the distorted signal is passed through a uniform quantizer.
  - μ - law distortion (North America):

$$y = y_{\max} \frac{\log_e\left[1+\mu(|x|/x_{\max})\right]}{\log_e(1+\mu)} \operatorname{sgn} x$$



  - A-law ditortion (Europe)

$$y = \begin{cases} y_{\max} \dfrac{A(|x|/x_{\max})}{1+\log_e A} \operatorname{sgn} x & 0 < \dfrac{|x|}{x_{\max}} \le \dfrac{1}{A} \\ y_{\max} \dfrac{1+\log_e(A|x|/x_{\max})}{1+\log_e A} \operatorname{sgn} x & \dfrac{1}{A} < \dfrac{|x|}{x_{\max}} < 1 \end{cases}$$

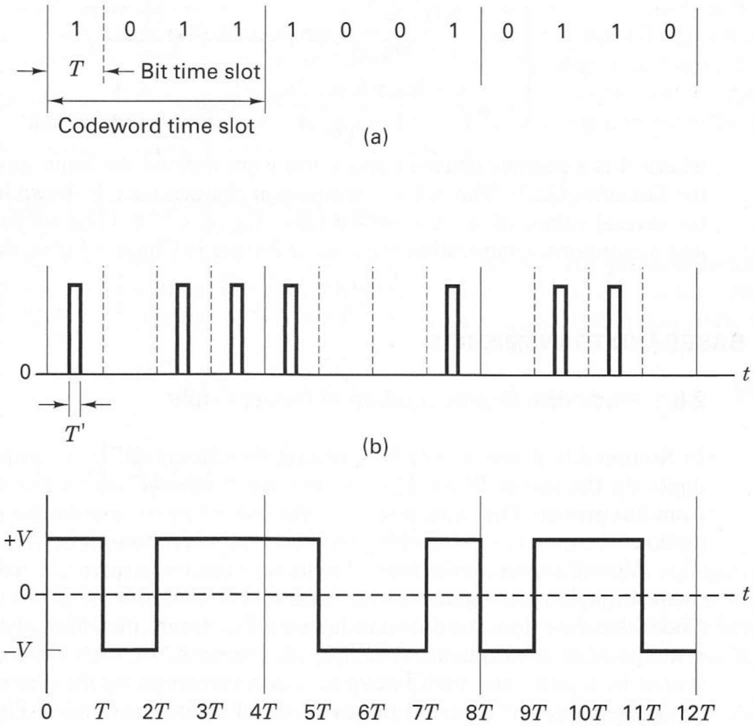


OUTLINE

- Baseband system
- Formatting
- **Baseband transmission**
- Correlative coding

BASEBAND TRANSMISSION

- **Why baseband modulation**
  - After formatting, messages are converted to logic ‘0’s and ‘1’s.
  - Baseband modulation: converting logic ‘0’s and ‘1’s to baseband electric signals that are compatible with baseband channel (pulse modulation)
  - PCM waveform: when pulse modulation is applied to a binary symbol, the resulting binary waveform is called a PCM waveform
  - M-ary pulse-modulation waveform: when pulse modulation is applied to a nonbinary symbol.

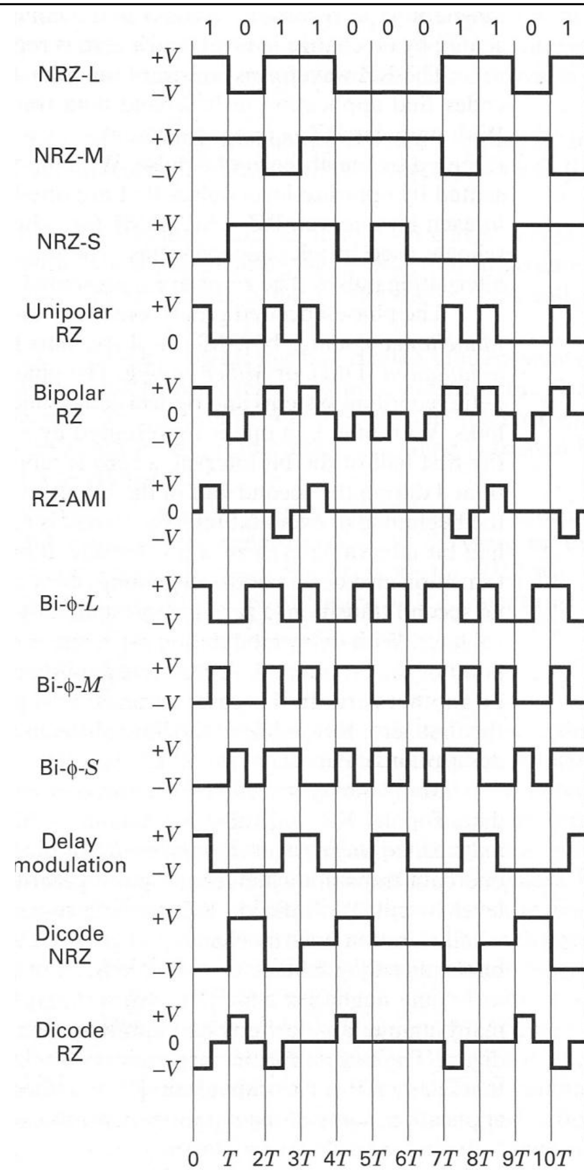


UNIVERSITY OF ARKANSAS

BASEBAND TRANSMISSION: PCM

- **PCM waveforms (line codes)**
  - Nonreturn-to-zero (NRZ)
    - NRZ-L (level) (digital logics)
    - NRZ-M (mark, '1') (magnetic tape)
    - NRZ-S (space, '0')
  - Return-to-zero (RZ) (baseband Tx)
    - Unipolar RZ
    - Bipolar RZ
    - RZ-AMI (alternate mark inversion)
  - Phase encoded
    - $bi-\phi-L$ (manchester encoding) (Ethernet)
    - $bi-\phi-M$ (mark, '1')
    - $bi-\phi-S$ (space, '0')
    - DM (delay modulation, miller coding)
  - Multilevel binary
    - Dicode NRZ
    - Dicode RZ
    - Bipolar RZ
    - RZ-AMI

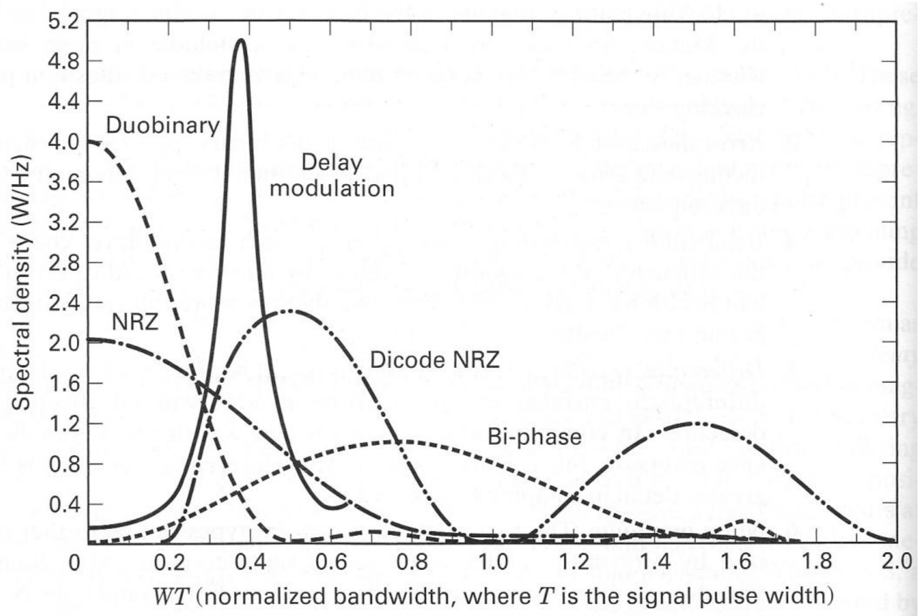
Differential encoding



UNIVERSITY OF ARKANSAS

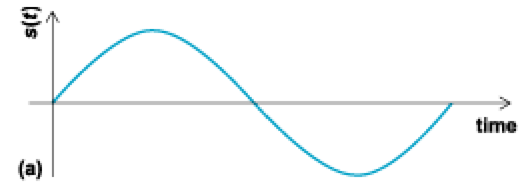
BASEBAND TRANSMISSION: PCM

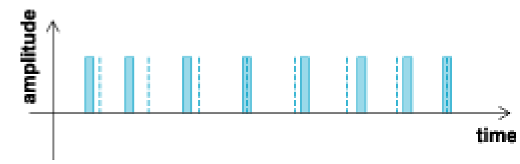
- **Why so many PCM waveforms?**
  - Self-clocking:
    - Manchester code has a transition in middle for synchronization.
  - Bandwidth
    - Less transitions ➔ less bandwidth
  - Noise immunity
    - NRZ waveforms have better error performance than unipolar RZ.
  - DC components (mean)
    - Unipolar RZ has DC components, while bipolar RZ doesn’t



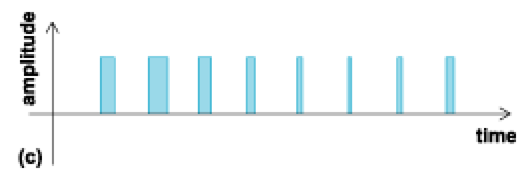
BASEBAND: M-ARY PULSE MODULATION

- **M-ary pulse modulation**
  - The analog signal are sampled, quantized to *M*-ary symbols, and then modulated on to pulses.
  - Classification
    - Pulse amplitude modulation (PAM)
      - Use the amplitude of the pulse to carry information.
      - M = 2 ➔ PCM (special case of PAM when M = 2)
    - Pulse position modulation (PPM)
      - Using the start position of a pulse to carry information
    - Pulse duration modulation (PDM), pulse width modulation (PWM)
      - Using the width of a pulse to carry information.





PPM



PWM

BASEBAND: M-ARY PULSE MODULATION

- **Why multilevel?**
  - $M = 2^k$ levels: each level (symbol) can represent k bits of information.
  - If the data rate is R bit/sec, then the symbol rate is R/k symbols/sec.
  - Pros:
    - The transitions of signals becomes slower ➔ bandwidth becomes smaller
    - Signal bandwidth is proportional to symbol rate.
    - Bandwidth efficiency: symbol rate/bandwidth
  - Cons:
    - The receiver needs to distinguish between M different symbols instead of 2 symbols in binary modulation
    - Signals are more sensitive to noise.

Tradeoff between bandwidth and performance.

BASEBAND: M-ARY PULSE MODULATION

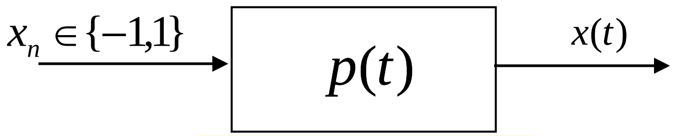
- **Example**
  - An analog information with maximum frequency $f_m = 3kHz$, is to be transmitted over an M-ary PAM system, where the number of pulse levels is M = 16. The quantization distortion is specified not to exceed 1% of the peak-to-peak analog signal.
    - How many quantization levels are required?
    - What is the minimum sampling rate?
    - What is the bit rate?
    - What is the symbol rate?
    - If the actual transmission bandwidth is 12 kHz, determine the bandwidth efficiency.

OUTLINE

- Baseband system
- Formatting
- Baseband transmission
- **Correlative coding**

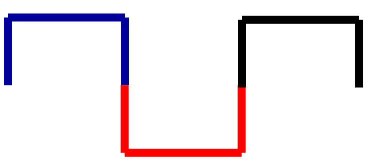
BASEBAND: CORRELATIVE CODING

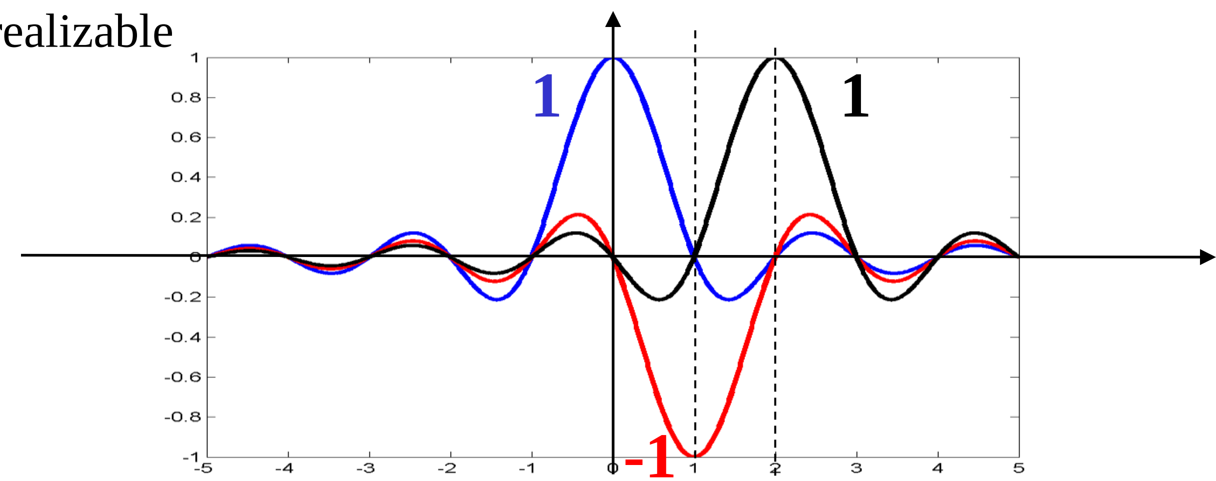
- **Baseband signaling can be represented by passing data through a filter**



$$x(t) = \sum_n x_n p(t - nT_s)$$

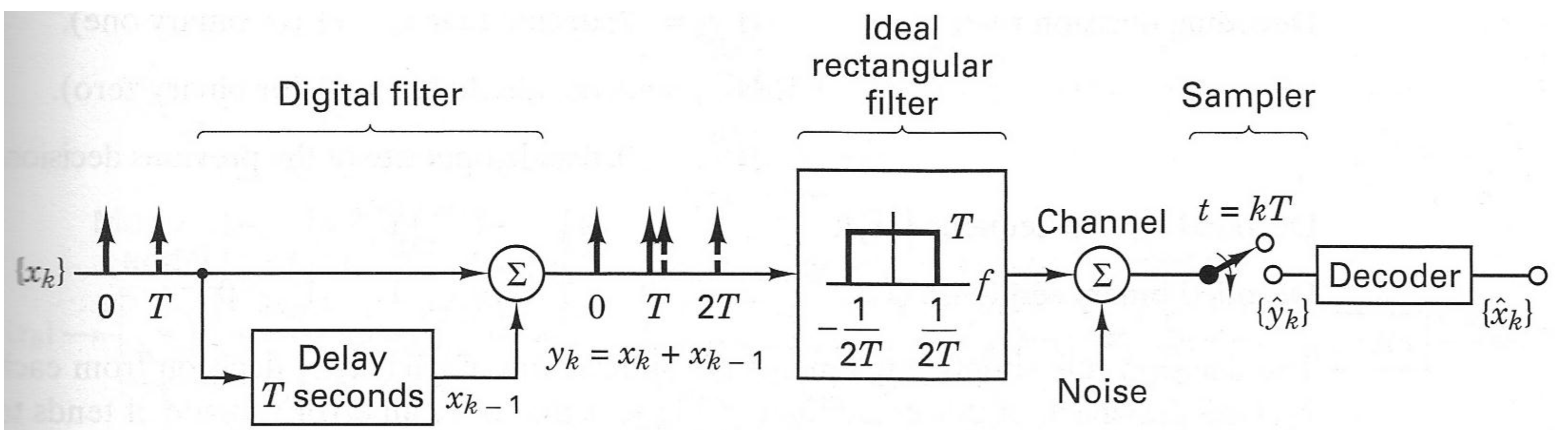
- Eg. 1. Rectangular waveform: $p(t) = u(t) - u(t - T_s) = rect\left(\frac{t - T_s/2}{T_s}\right)$
  - Bipolar PCM (unlimited bandwidth)
  - 0-to-null BW:
- Eg. 2. Ideal low pass filter: $p(t) = \sin c\left(\frac{t}{T_s}\right)$ $\quad P(f) = T_s rect\left(\frac{f}{F_s}\right)$
  - 0-to-null BW:
  - Ideal filter is unrealizable





BASEBAND: CORRELATIVE CODING

- **Correlative coding**
  - also referred to as: duobinary signaling, partial response signaling.
  - It can use a baseband bandwidth of W Hz to support a symbol rate of 2W symbols/sec without resorting to unrealizable ideal filter.



BASEBAND: CORRELATIVE CODING

- **Correlative encoding and decoding**
  - Encoder:
    - 1. Differential encoding of binary sequence $w_k = x_k \oplus w_{k-1}$
      - Modular 2 operation:

$$0 \oplus 0 = 0 \qquad 0 \oplus 1 = 1 \qquad 1 \oplus 0 = 1 \qquad 1 \oplus 1 = 0$$

Binary digit sequence $\{x_k\}$:	0	0	1	0	1	1	0
Precoded sequence $w_k = x_k \oplus w_{k-1}$:	0	0	1	1	0	1	1

      - Differential encoding is performed over 0's and 1's: it will not affect signal bandwidth.
    - 2. Mapping binary sequence to bipolar sequence

Bipolar sequence $\{w_k\}$:	−1	−1	+1	+1	−1	+1	+1

    - 3. Duobinary encoding $y_k = w_k + w_{k-1}$

Coding rule: $y_k = w_k + w_{k-1}$:	−2	0	+2	0	0	+2

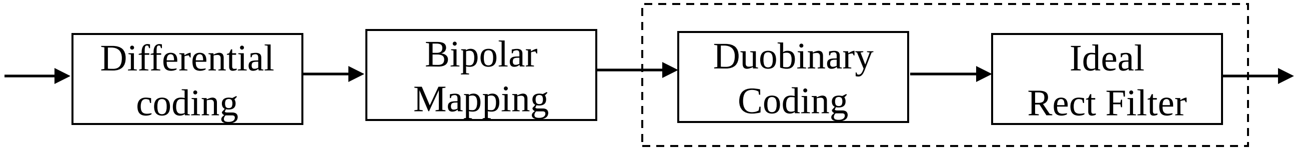
  - Decoder:
    - 1. Decision rule:
      - If $\hat{y}_k < -1$ or $\hat{y}_k > 1$, $\hat{x}_k = 0$
      - If $|\hat{y}_k| < 1$, $\hat{x}_k = 1$

UNIVERSITY OF ARKANSAS

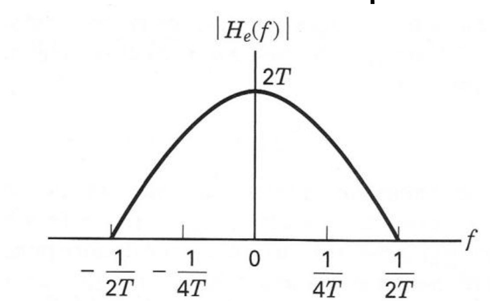
BASEBAND: CORRELATIVE CODING

- **Frequency domain analysis**

$$H_e(f) = H_1(f)H_2(f)$$



- Time domain of duobinary coding: $y(t) = w(t) + w(t-T)$
- Frequency domain: $Y(f) = W(f) + W(f)e^{-j2\pi fT}$
- Transfer function of duobinary coding: $H_1(f) = \dfrac{Y(f)}{W(f)} = 1 + e^{-j2\pi fT}$
- Transfer function of ideal rect filter: $H_2(f) = T, \ |f| < \dfrac{1}{2T}$
- Overall response of the encoder: $H_e(f) = T(1 + e^{-j2\pi fT}), \ |f| < \dfrac{1}{2T}$



Baseband bandwidth: 1/(2T)

Even though $H_2(f)$ is an ideal filter and cannot be implemented, $H_e(f)$ can be easily approximated. In reality, duobinary coding and ideal rect filter are approximated by ONE filter.

UNIVERSITY OF ARKANSAS

BASEBAND: CORRELATIVE CODING

- **Duobinary Coding v.s. PCM**
  - Bandwidth for transmission data rate of 1/T bps:
    - binary PAM: 0-to-null is 1/T Hz, absolute bandwidth is infinity
    - Duobinary: absolute bandwidth: 1/(2T)
  - Performance:
    - Binary PAM: two levels
    - Duobinary: three levels (worse performance compared to binary PAM)
  - Tradeoff between bandwidth and performance.