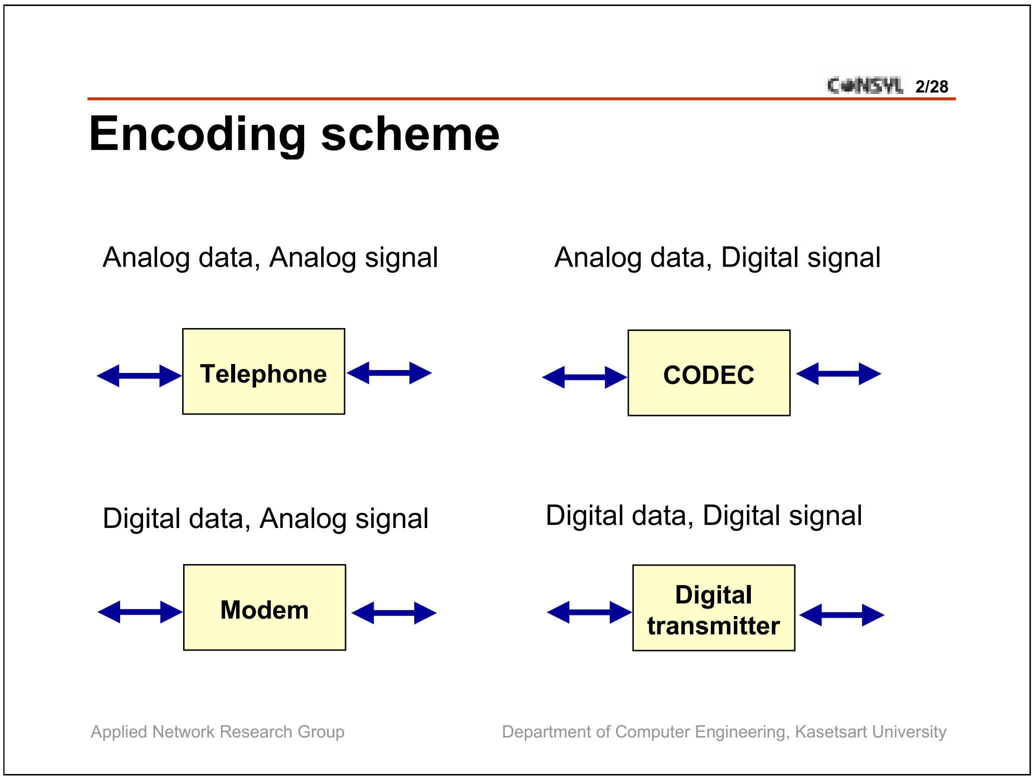
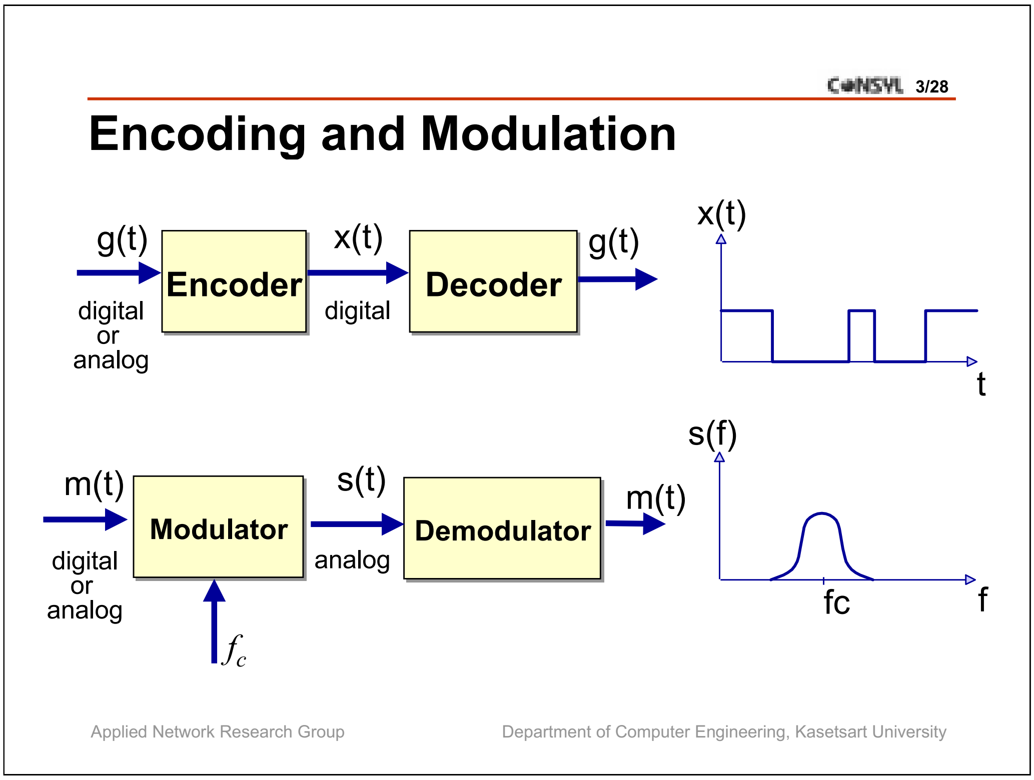
# Data Encoding

*Surasak Sanguanpong*
*nguan@ku.ac.th*
*http://www.cpe.ku.ac.th/~nguan*
Last updated: May 19, 1999




Analog transmission means a transmitting analog signals without regard to their content; the signal may represent analog data (e.g. voice) or digital data (e.g. binary data). Digital transmission is concerned with a transmitting binary signal




For digital signal, a data source *g(t)*, which may be either digital or analog, is encoded into a digital signal *x(t)*.

Analog transmission uses a continuous constant-frequency signal known as the *carrier signal*. The frequency of the carrier signal is chosen to be compatible with the transmission medium. Data is transmitted using a carrier signal by *modulation*, which is the process of encoding source data onto a carrier signal with frequency $f_c$

# Why encoding?

- **Three factors determine successfulness of receiving signal**
  - S/N
  - data rate
  - bandwidth
- **More factor can be used to improve**
  - encoding scheme



With other factors held constant, the following statements are true.

- An increase in data rate increases bit error rate.
- An increase S/N decreases bit error rate.
- An increase in bandwidth allows an increase in data rate [Stalling, p98].

CONSYL 5/28

# Encoding evaluation factors

- Signal spectrum
- Clocking
- Error detection
- Signal interference& noise immunity
- Cost and complexity

Applied Network Research Group    Department of Computer Engineering, Kasetsart University

Five factors are used to evaluate the various encoding scheme:

*Signal spectrum* : a lack of high-frequency components means that less bandwidth is required for transmission. No dc component is desirable.

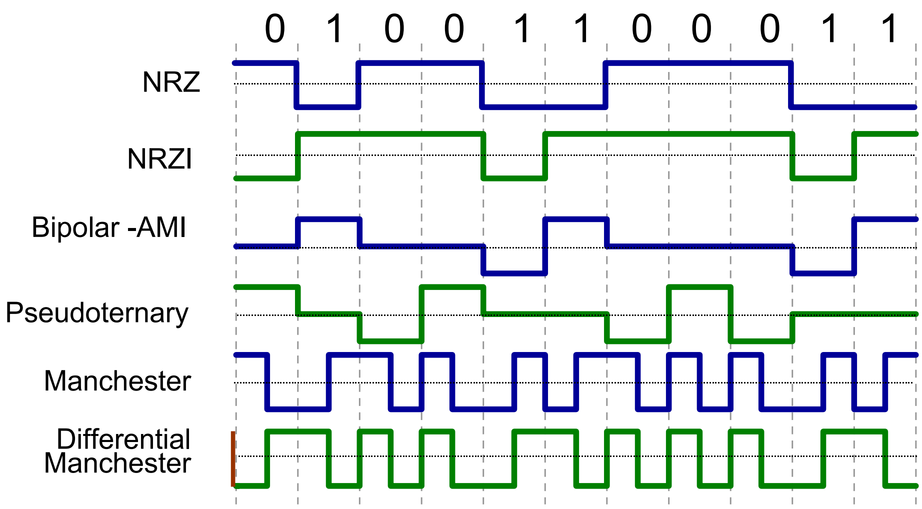
*Clocking* : suitable encoding provide some synchronization mechanism to determine the beginning and end of each bit position.

*Error detection* : some error detection can be built into the encoding scheme.

*Signal interference & noise immunity* : some encoding scheme has superior performance in the presence of noise.

*Cost and complexity* : higher signaling rate to achieve a greater data rate results expensive devices.

# Digital data, Digital signal






**Definition of Digital Signal Encoding Formats**

**Nonreturn-to-Zero-Level (NRZ-L)**
0 = high level
1 = low level l

**Nonreturn to Zero Inverted (NRZI)**
0 = no transition at beginning of interval (one bit time)
1 = transition at beginning of interval

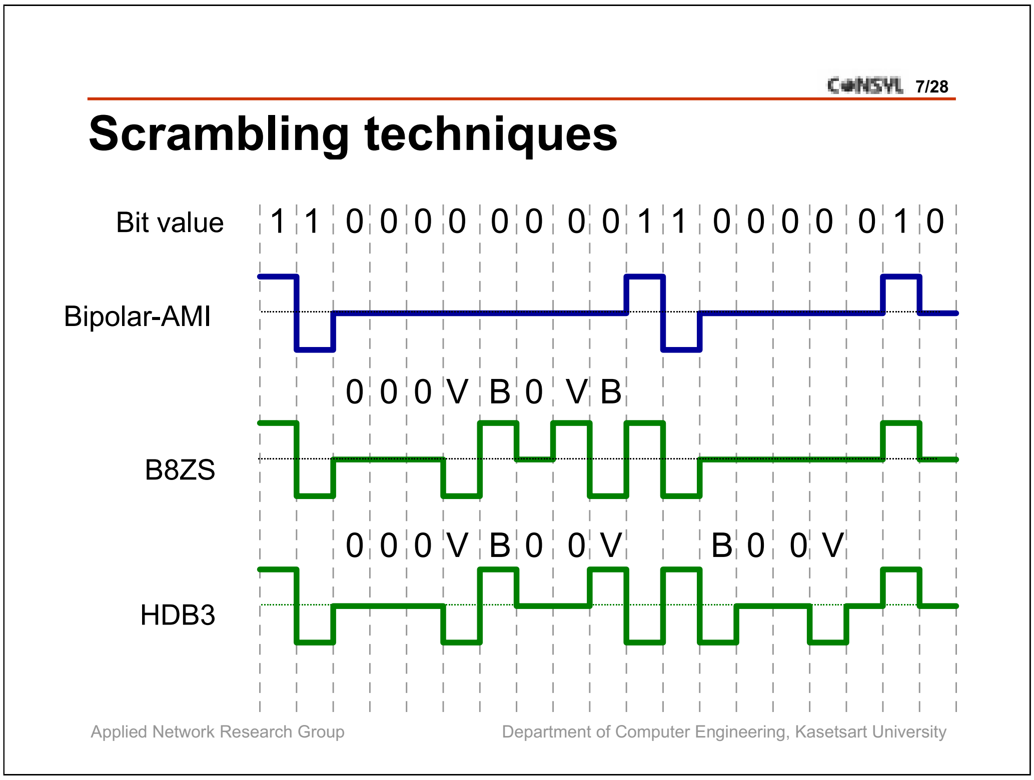
**Bipolar-AMI**
0 = no line signal
1 = positive or negative level, alternating for successive ones

**Pseudoternary**
0 = positive or negative level, alternating for successive zeroes
1 = no line signal

**Manchester**
0 = transition from high to low in middle of interval
1 = transition from high to low in middle of interval

**Differential Manchester**
Always a transition in middle of interval
0 = no transition at beginning of interval
1 = transition at beginning of interval

[Stallings, p99,100]




To maintain synchronization for the receiver's clock using bipolar.

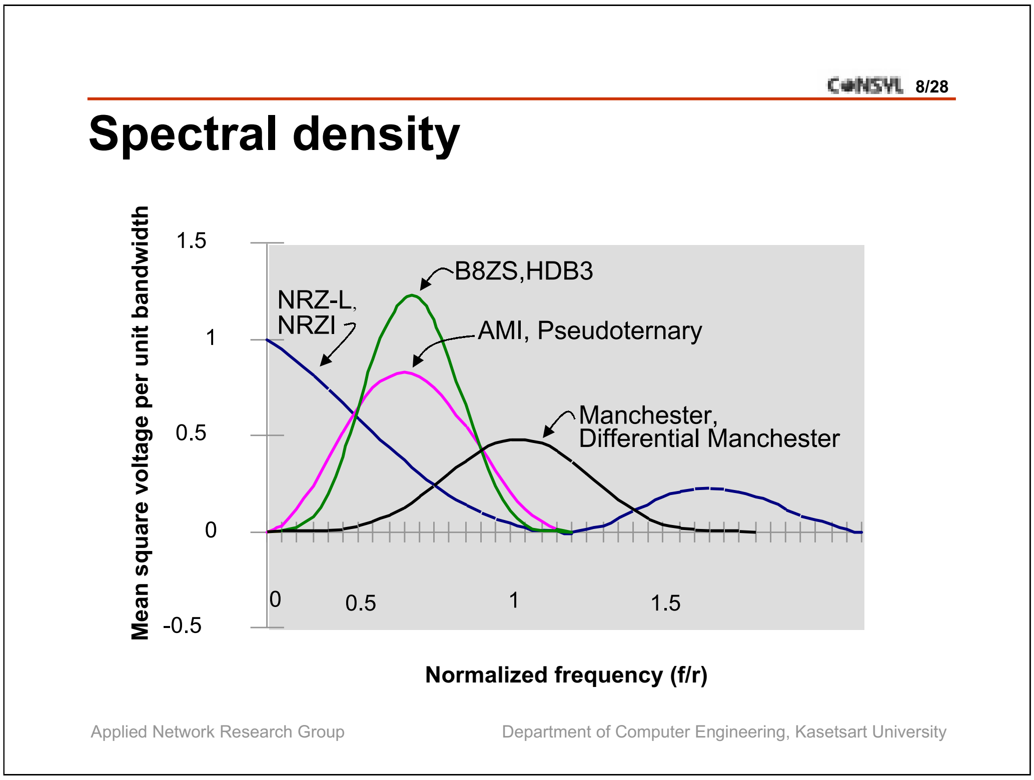
**B8ZS**

Same as bipolar AMI, except that any string of eight zeros is replaced by a string with two code violations.

- If an octet of all zeros occur and the last voltage pulse preceding this octet was positive, the eight zeros of the octet are encode as 00+-0-+
- If an octet of all zeros occur and the last voltage pulse preceding this Oct was positive, the eight zeros of the octet are encode as 00-+0+-

**HD3B**

Same as bipolar AMI, except that any string of four zeros is replaced by a string with one code violation. The scheme replace strings of four zeros with the sequence B00V. [Stallings, p106]




NRZ make efficient use of bandwidth. most of the frequency in NRZ and NRZI signals are between dc and half the bit rate.

Manchester& Different Manchester has the bulk of the energy between one-half and one times the bit rate. Thus the bandwidth is reasonably narrow and contain no dc component.

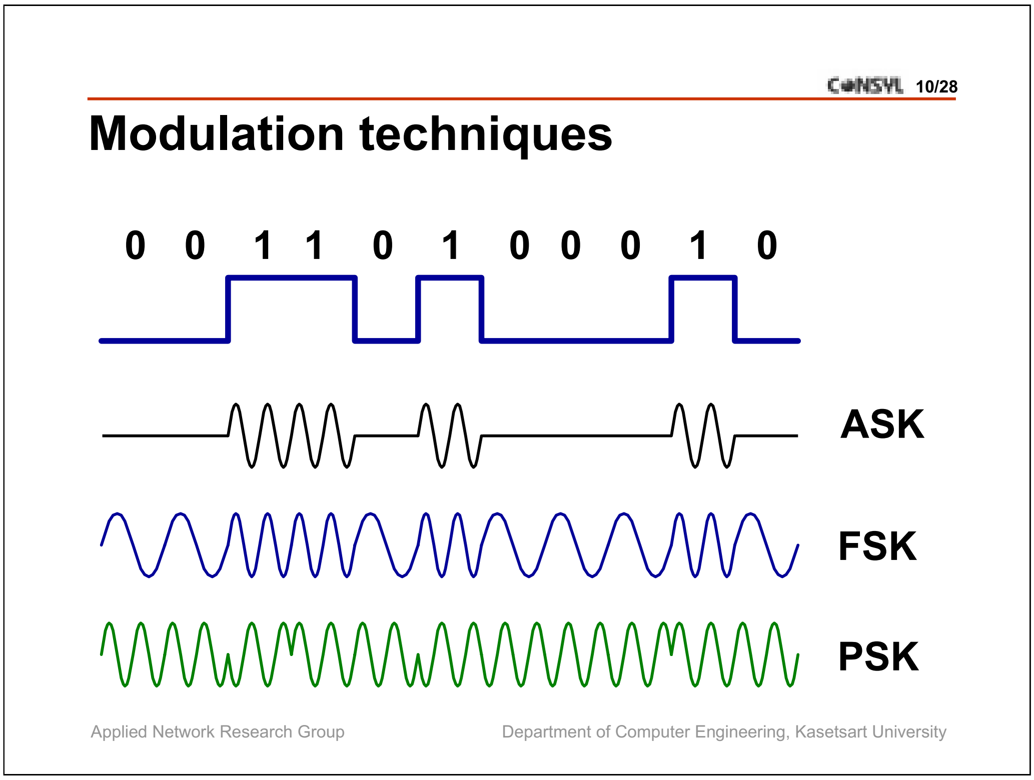
AMI make use of bandwidth less than the bandwidth of NRZ [Stallings, p102].

## Digital data, Analog signal

- **Transmitting digital data through PSTN**
- **Modem is used to convert digital data to analog signal and vice versa**



The most familiar use is for transmitting digital data through the public telephone network. The telephone network was designed to support analog signals in the voice-frequency range 300-3400 Hz. It is not suitable for handling digital signal. Modem is used to convert digital to analog signal and vice versa.

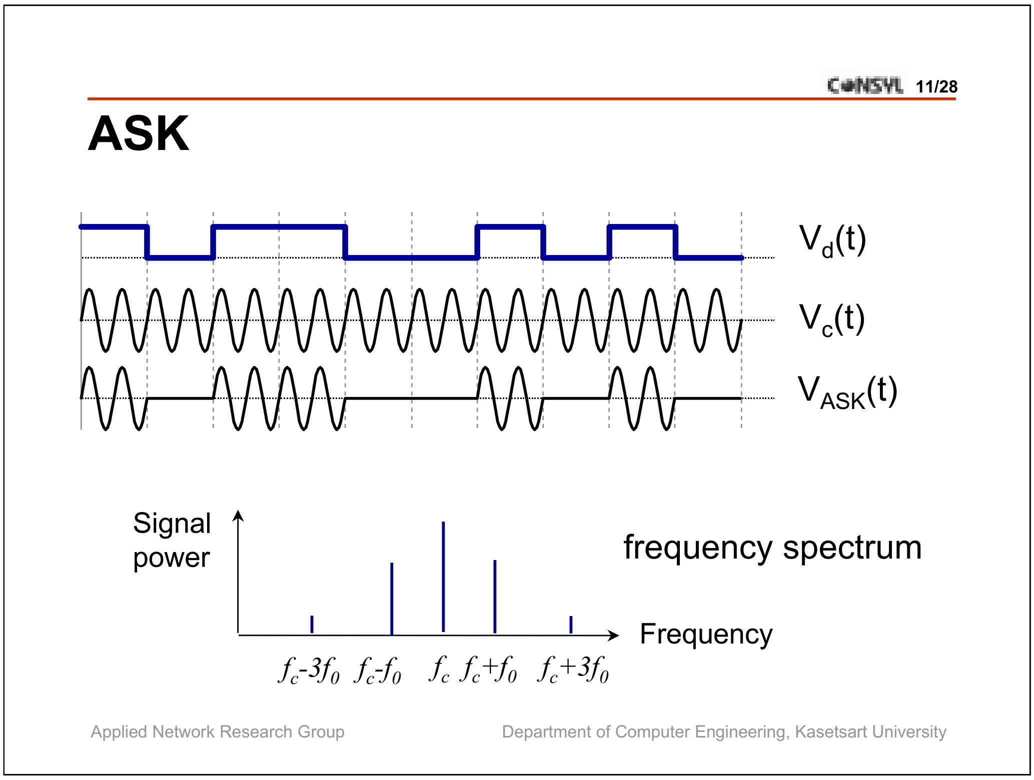



There are three basic modulation techniques for transforming digital data into analog signals:

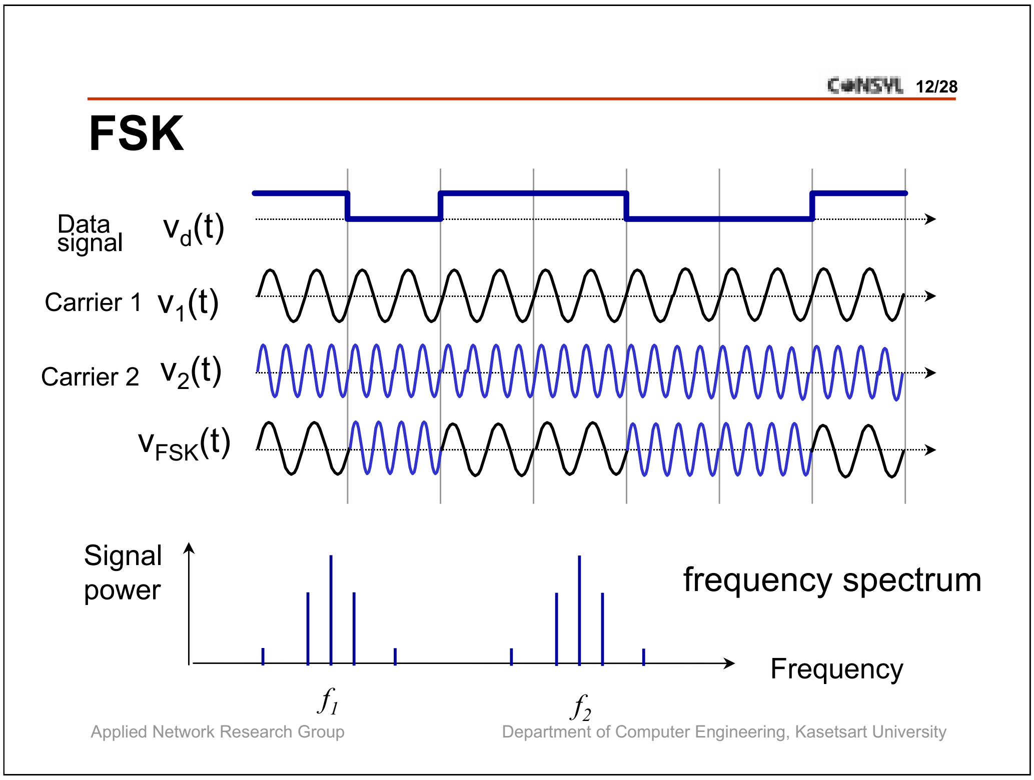
Amplitude-shift keying (ASK)

Frequency-shift keying (FSK)

Phase-shift keying (PSK)

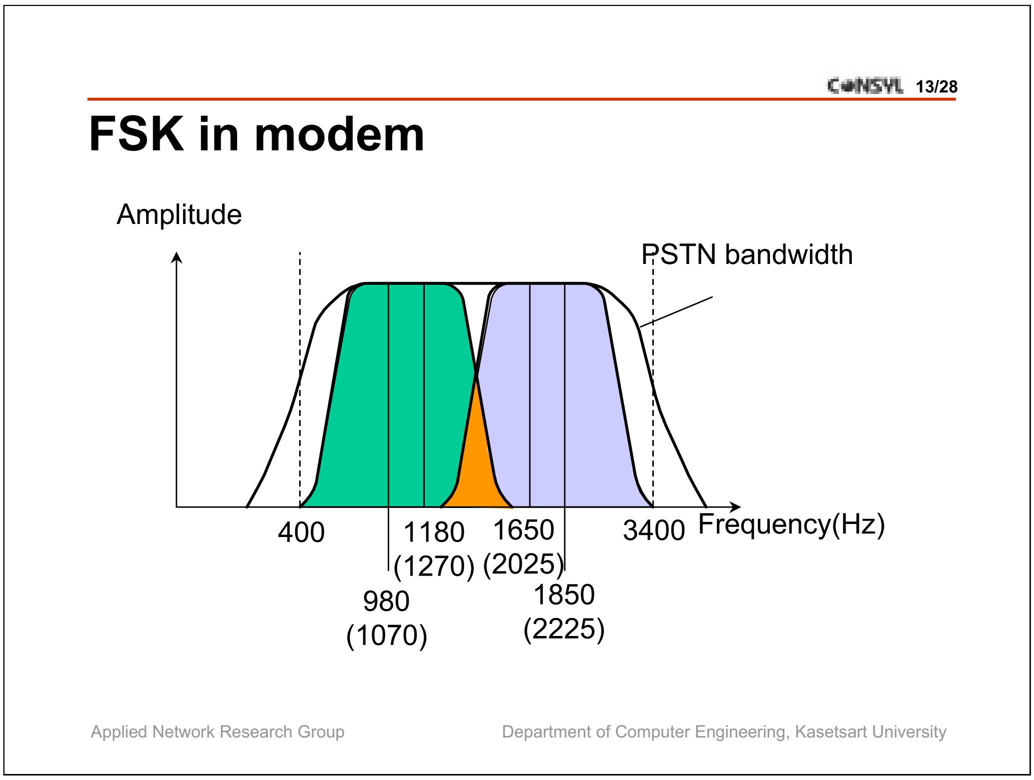



In ASK, the amplitude of a single-frequency known as the carrier frequency is switched between two levels at a rate determined by the bit rate of the transmitted binary data signal. Bandpass filter is used to limit the band of frequencies based on Nyquist's theorem. [Halsall, p.59]

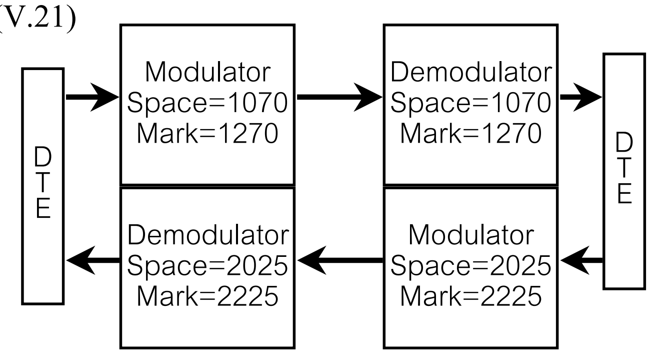



In FSK, two fixed amplitude carrier signals are used, one for a binary 0 and the other for a binary 1. The different between the two carriers is known as the frequency shift. The modulation operation is equivalent to summing together the outputs of two separate ASK modulators.

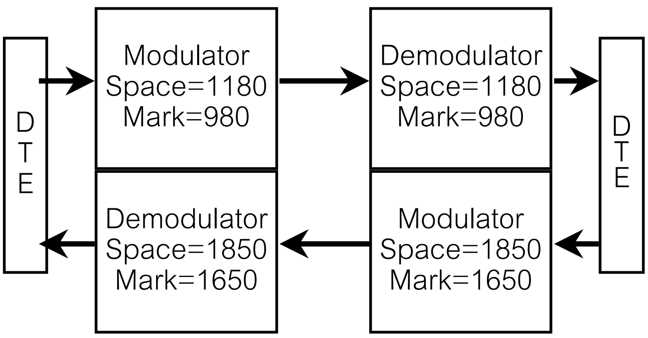
FSK is the modulation method that was used in all early low bit rate modems. [Halsall p.62]



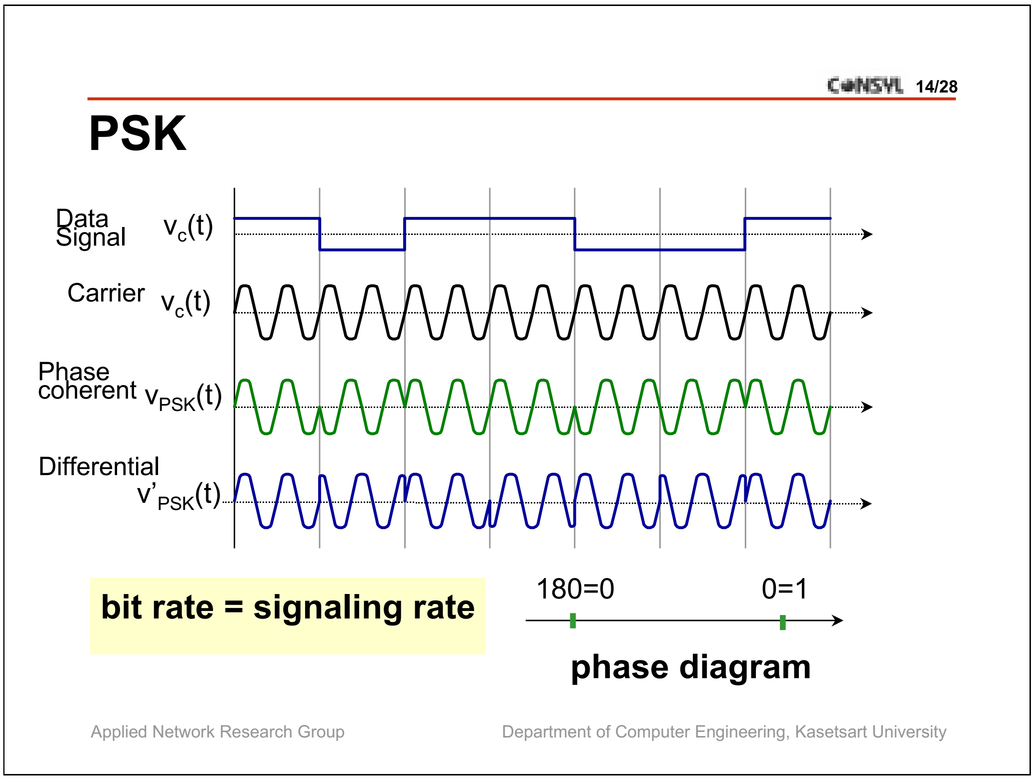

The above figure illustrate the frequency assignments that are used for two types of FSK modem to provide a full-duplex 300 bps link between two DTEs. One set of frequency assignments is defined by EIA and the other by ITU-T (V.21)




**EIA frequency assignment**




**ITU-T frequency assignment (V.21)**

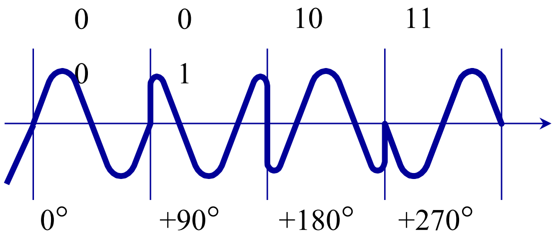



In PSK, the phase of the carrier signal is shift to represent data. Two type of PSK are used.

**Phase coherent PSK** : Used two fixed carrier signals to represent a binary 0 and 1 with a 180° phase different. The disadvantage of this scheme us that a reference carrier signal is required at the receiver against with the phase of the received signal is compared, this required more complex demodulation circuit.

**Differential PSK** : A phase shift of 90° relative to the current signal indicates a binary 0 is the next bit while a phase shift 270° indicates a binary 1. The demodulation circuitry need determine only the magnitude of each phase shift rather than its absolute value. [Halsall, p.64]

# Multilevel modulation method

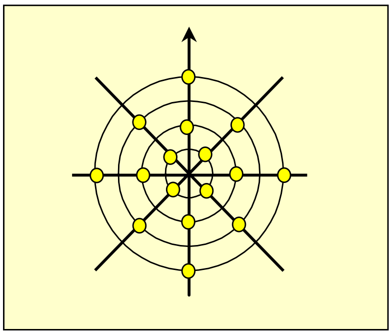



bit rate = n x signaling rate

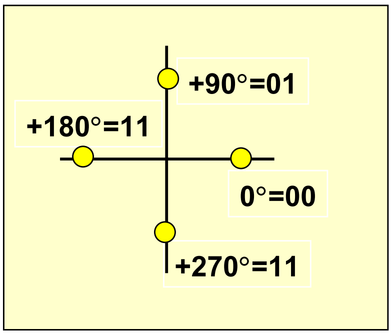


More sophisticated modulation methods are used which involve either multiple signal levels or a mix of the basic scheme, particularly amplitude and phase. More bit rate can be achieved if signaling element represent more than one bit.

# Multilevel modulation method



16-QAM phase diagram




4-PSK phase diagram



**QPSK ( Quadrature PSK or 4-PSK) :** Four different phase changes (0°, 90°, 180°, 270°) to enables each phase change to convey 2 bits (bit rate=2*signaling rate).

**QAM (Quadrature Amplitude Modulation or 16-QAM)** : Phase and amplitude changes, 16 levels per signal element and hence 4 bit symbols (bit rate=4*signaling rate).

# Bit rate and Baud rate

- **Bit rate** : **A number of bits that are transmitted in a second**
- **Baud rate** : **A number of line signal changed variation per second**

*If a modem transmits 1 bit for every signal change*

**bit rate = baud rate**

*If a signal change represents 2 or more or n bits*

**bit rate = baud rate*n**



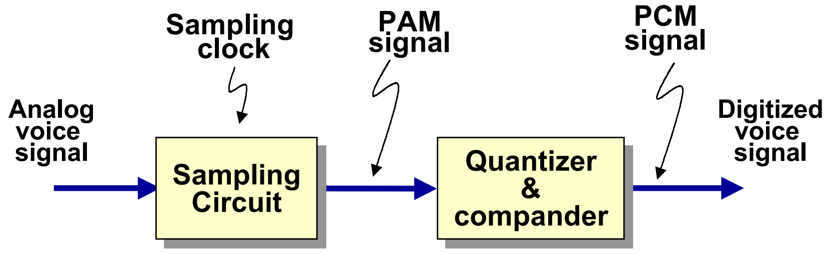
The relationship between bit transfer rate and baud rate depends on the number of bit values that are encoded in a signal. When each signal represent one bit, the bit and baud rate are the same. When a signal encodes multiple bits, the bit rate is a multiple of the baud rate.

In modem, encoding techniques are employed to make a signal change represent 2 or more bits.

The term baud comes from *Baudot*, who developed an encoding scheme for the French Telegraph system in 1877.

# Analog data, Digital signal

- **Two principle techniques used**
  - PCM
  - DM






Digitization is a process of converting analog data into digital data. The digital signal is converted back into analog signal at the receiver.

The device used for converting analog data into digital form, and recovering the original analog data is known as CODEC (Coder-Decoder).

## Nyquist theorem

*" In a perfectly noiseless channel, if f is the maximum frequency the medium can transmit, the receiver can completely reconstruct a signal by sampling it 2*f times per second"*

***Nyquist, 1920***

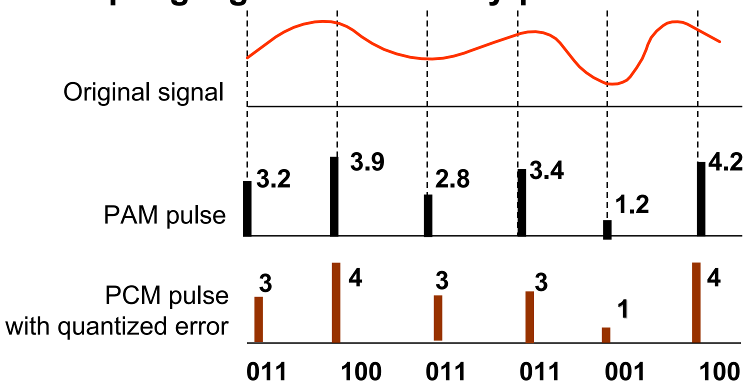


In 1920, Harry Nyquist developed his classic theory. Nyquist showed that original signal must be sampled at a maximum rate of greater than twice the highest frequency component to send to receiver to completely reconstruction. For example, to convert a voice signal which 4 kHz highest frequency into digital form, it must be sampled at a rate of 8000 times per second.

C●NSYL 20/28

# PCM

- **Sampling signal based on nyquist theorem**



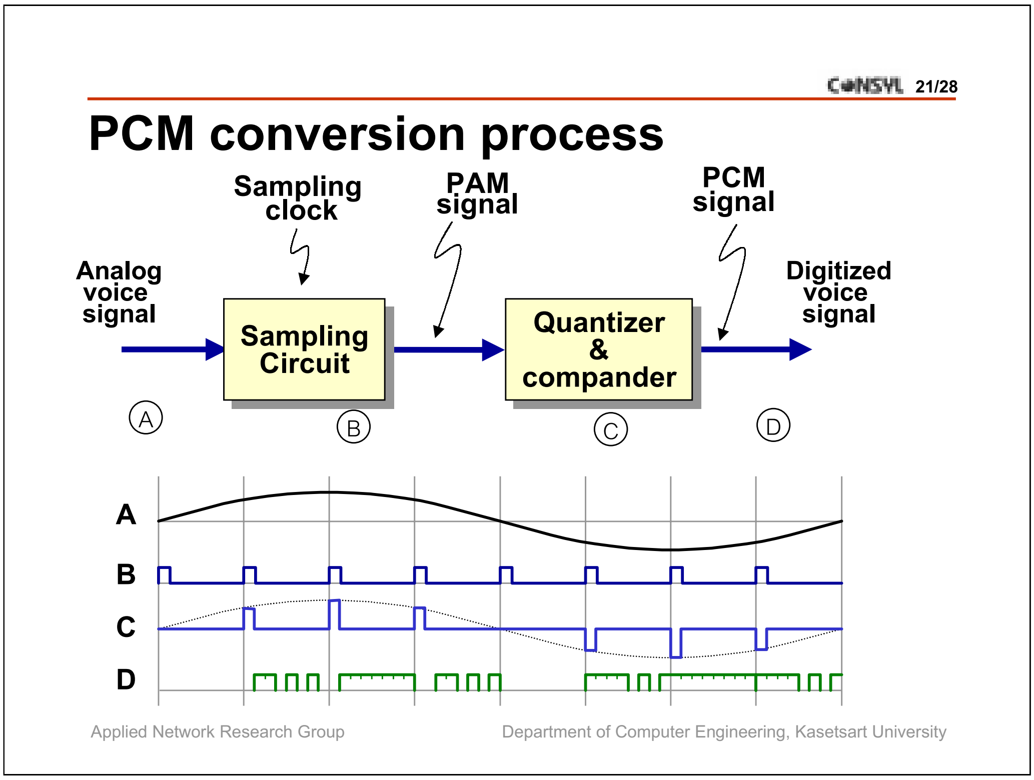

**PCM output** **011100011011001100**

Applied Network Research Group

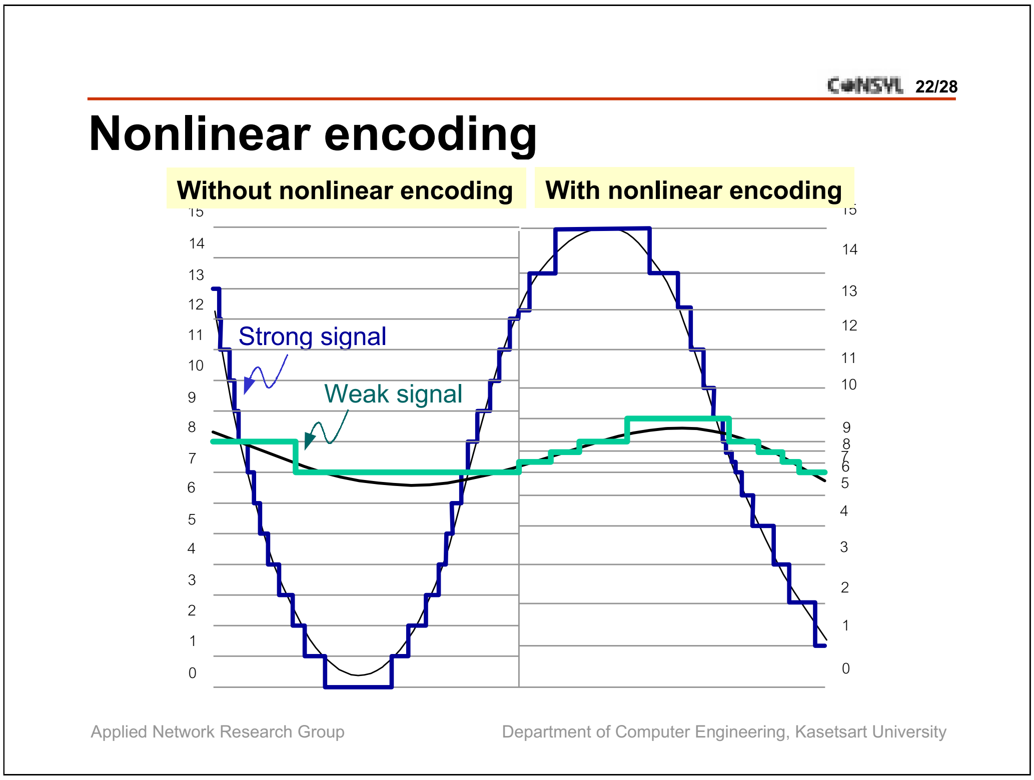
Department of Computer Engineering, Kasetsart University

PCM is based on the sampling theorem. The original signal is assumed to be band limited with a bandwidth of B. Signal is sampled at a rate 2B. Samples signal are represents as narrow pulse whose amplitude is proportional to the value of the original signal and is known as PAM (Pulse Amplitude Modulation).

The amplitude of each PAM pulse is approximated by an n-bit integer,. In the sample above, n=3. Thus $8=2^3$ levels are used for approximating the PAM pulses.



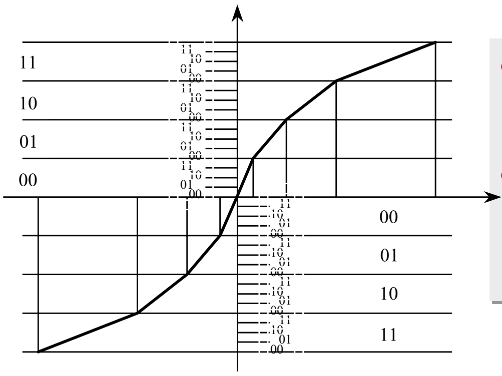

The process starts with an analog signal, which is sampled by PAM sample. the resulting pulse are quantized to produced PCM pulses and then encoded to produce bit stream. At the receiver end, the process is reversed to reproduce the analog signal. [Halsall, p.69]




Quantization levels are not necessary equally spaced. The problem with equal spacing is that the mean absolute error for each sample is the same, regardless the signal level. Lower amplitude values are relatively more distorted. [Stallings, p.118]

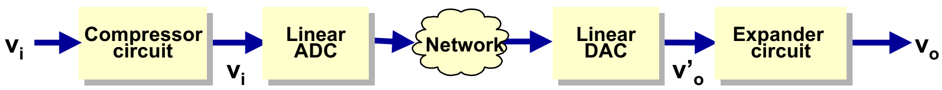
CoNSYL 23/28

# Companding process




**linear quantization interval**

- Implement nonlinear encoding via companding process
- Companding = Compressing Expanding




Applied Network Research Group

Department of Computer Engineering, Kasetsart University

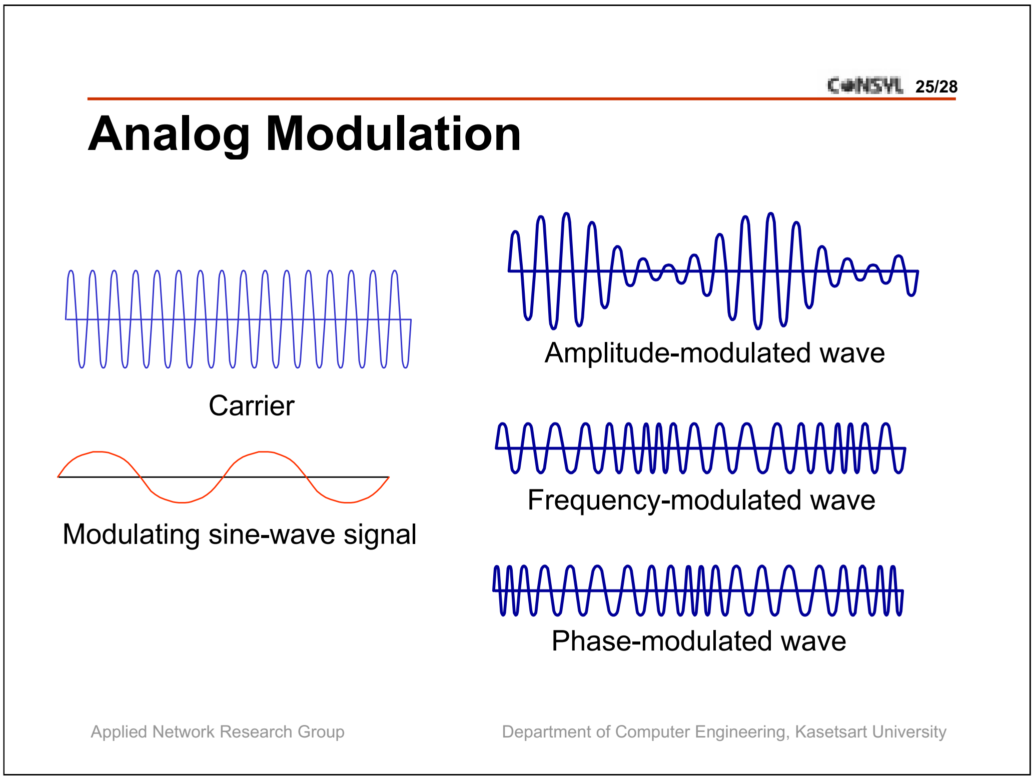
Prior to the input signal being sampled and converted by ADC into a digital form, it is passed through a circuit known as a *compressor*. Similarly, at the destination, the reverse operation is perform on the output of the DAC by a circuit known as *expander*. [Halsall, p. 71]

## Analog Data, Analog signal

- Use Modulation techniques
- Need a high frequency for effective transmission
- Modulation permits frequency division multiplex



Modulation has been defined as the process of combining an input signal and a carrier signal.




Amplitude modulation is the simplest form of modulation. The modulated signal has constant frequency but its amplitude is vary with the input signal.

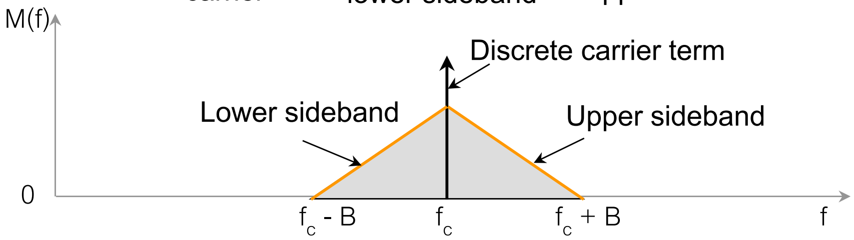
The envelope of the resulting signal is $1+n_a x(t)$ and as long as $n_a<1$, the envelope is an exact reproduction of the original signal. If $n_a>1$ , the envelope will cross the time axis and information is lost. [Stallings, p.123]

# AM Spectrum

case $x(t) = \cos 2\pi f_m t$

$$s(t) = [1 + n_a \cos 2\pi f_m t]\cos 2\pi f_c t$$

$$s(t) = \underbrace{\cos 2\pi f_c t}_{\text{carrier}} + \underbrace{\frac{n_a}{2}\cos 2\pi (f_c - f_m)t}_{\text{lower sideband}} + \underbrace{\frac{n_a}{2}\cos 2\pi (f_c + f_m)t}_{\text{upper sideband}}$$




- each sideband contains the complete spectrum of s(t) !



The resulting signal has a component at the original carrier frequency plus a pair of components each spaced $f_m$ Hz from the carrier.

From the equation above, it can be seen that AM involves the multiplication of the input signal by the carrier. [Stallings, p.123]

# AM power saving

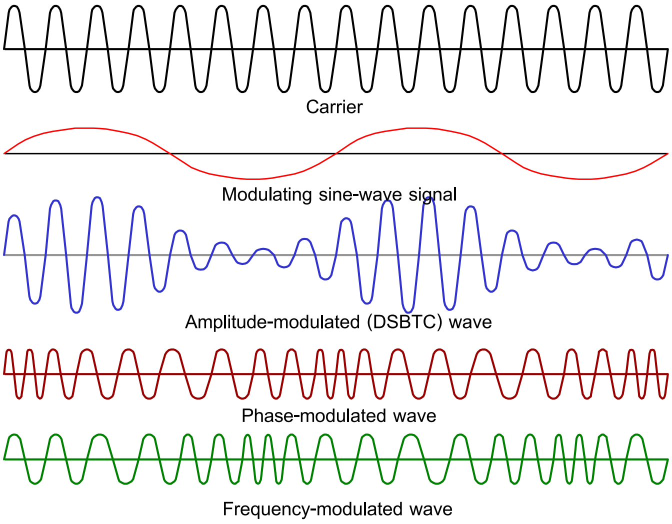
- SSB (Single sideband)
- DSBSC (Double sideband suppresses carrier)
- VSB (Vestigial sideband)



It should be clear that s(t) contains unnecessary components, since each of the sideband contains the complete spectrum. SSB, DSBSC and VSB are methods to save power and bandwidth.

The disadvantage of suppressing the carrier is that the carrier can be used for synchronization purposed. A constant carrier provides a clocking mechanism.A compromise approach is VSB, which uses one sideband and a reduced-power carrier. [Stallings, p.124]

# FM and PM


Carrier
Modulating sine-wave signal
Amplitude-modulated (DSBTC) wave
Phase-modulated wave
Frequency-modulated wave