MODULATION IN TELEMETRY AND OTHER DATA PROCESSING SYSTEMS

Dr. Bruce C. Gabrielson
Security Engineering Services
PO Box 550
Chesapeake Beach, MD 20732

INTRODUCTION

Relative to low power signal transmission, not only must an engineer be able to understand standard modulation concepts, but he/she must also be familiar with telemetry systems and the way information appears in both its intended and unintended forms. This paper describes generic modulation concepts as they are normally understood in the standard transmission of communication signals.

DEFINITIONS

Several terms are regularly used in modulation theory. Since some of the definitions appear to be circular in that terms used in one definition are themselves defined using the other definition -- e.g., carrier and modulation -- the standard Institute of Electrical and Electronic Engineers (IEEE) definitions have been augmented with some clarifications.

CARRIER *A wave having at least one characteristic that may be varied from a known reference value by modulation*.

The carrier frequency, f_c, is normally the frequency to which a receiver is tuned in order to extract modulation. In broadcast communications, f_c is much higher than the modulation frequency, f_m. This is, however, not always true as in the case of powerline communications where $f_m > f_c$.

COMPLEX MODULATION *Any combination of modulation techniques applied to a single carrier.*

Complex modulation may be achieved by imposing multiple modulation techniques on a single carrier, as in quadriture phase shift key - amplitude modulation (QPSKAM) and the use of the FM carrier for both the FM modulation and the subsidiary communication authorization (SCA) modulation, or by using a succession of modulated carriers to modulate a higher level carrier, as in Bell System frequency domain multiplexing (FDM). Unintentional complex modulation is also frequently the result of poor filtering and unintended feedback loops.

MODULATION *The process by which some characteristic of a carrier is varied in accordance with a modulating wave.*

As discussed below, there are only three ways the carrier can be modulated: amplitude (AM), frequency (FM), and phase (PM).

MULTIPLEX *To interleave or simultaneously transmit two or more messages on a single channel.*

The two most common forms of multiplexing are time domain multiplex (TDM) and frequency domain multiplex (FDM). Multiplexing only occurs intentionally.

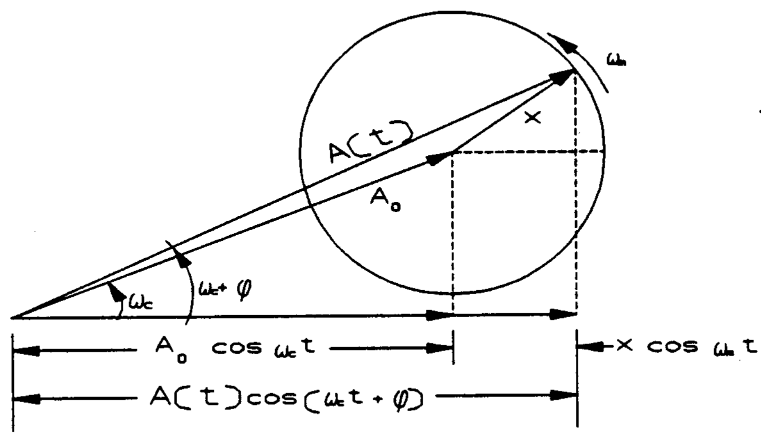

Figure 1 - Fixed Reference Vector

STANDARD MODULATION CONCEPTS

Figure 1 is a vectorial representation of an instantaneous modulated signal. Mathematically, this signal is described by:

$$s(t) = A(t) \cos[ù_c t + ö(t)] \quad (1)$$

where $A(t) \equiv$ the envelope of the modulated carrier
$ù_c \equiv$ the angular frequency of the carrier
$ö(t) \equiv$ the modulated phase.

Modulation can generally be considered to take one of three forms:

1. *Amplitude Modulation (AM)* in which the signal amplitude is varied by the modulating signal while the phase remains unchanged:

$$s_{AM}(t) = A_c[1 + x(t)] \cos ù_c t \quad (2)$$

where
$A_c \equiv$ the envelope of the unmodulated carrier

$x(t) \equiv$ the time varying modulation.

The simplest way to achieve AM is to add the carrier and the modulation in an amplifier. Typical AM receivers operate on the principle of envelope detection, in which the receiver responds to variations in the peak amplitude of the RF signal without regard to minor changes in the carrier frequency.

2. *Angle modulation* in which the carrier amplitude remains unchanged while the phase is varied by the modulation:

$$s_a(t) = A_c \cos [\omega_c t + \phi(t)]. \quad (3)$$

Pure angle modulation receivers amplify and clip the received signal before demodulation. They then track the phase or frequency variations of the signal with respect to the frequency of a reference oscillator.

Angle modulation is accomplished in two forms:

a. *Frequency Modulation (FM)* in which the frequency deviation of the carrier is proportional to the message signal:

$$s_{FM}(t) = A_c \cos [\omega_c t + k_f \int_{-\infty}^{t} x(\tau)\, d\tau] \quad (4)$$

where

$k_f \equiv$ the frequency deviation constant.

FM is easily achieved by varying the voltage across a Voltage Controlled Oscillator (VCO).

b. *Phase Modulation (PM)* in which the phase deviation of the carrier is proportional to the message signal:

$$s_{PM}(t) = A_c \cos [\omega_c t + k_p x(t)] \quad (5)$$

where

$k_p \equiv$ the phase deviation constant.

Analog PM is essentially a phase shifted FM. Digital PM usually requires discrete phase shifters and switches to select the desired output phases.

3. Some combination of 1 and 2 above:

$$s(t) = A_c [1 + x(t)] \cos [\omega_c t + \phi(t)]. \quad (6)$$

Figure 2 compares AM, FM, and PM for simple analog and digital modulating signals. Note that for analog modulating signals, the difference between FM and PM is essentially a phase shift of the modulator. For digital modulators, however, the resulting FM and PM signals are extremely different.

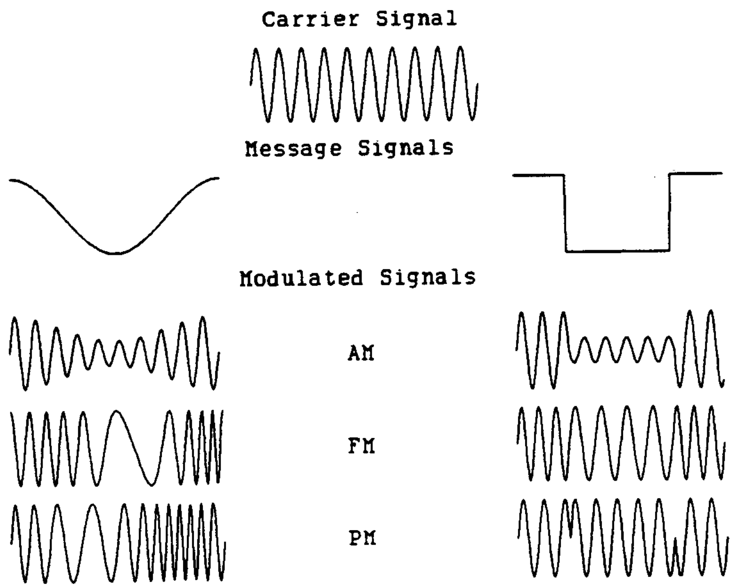

Figure 2 - AM, FM, and PM Modulation

Figure 3 illustrates simple techniques for accomplishing each type of modulation. Note that both AM, FM, and analog PM can occur easily in normal circuits using amplifiers and oscillators. Digital PM is much more difficult to generate and almost never occurs unintentionally.

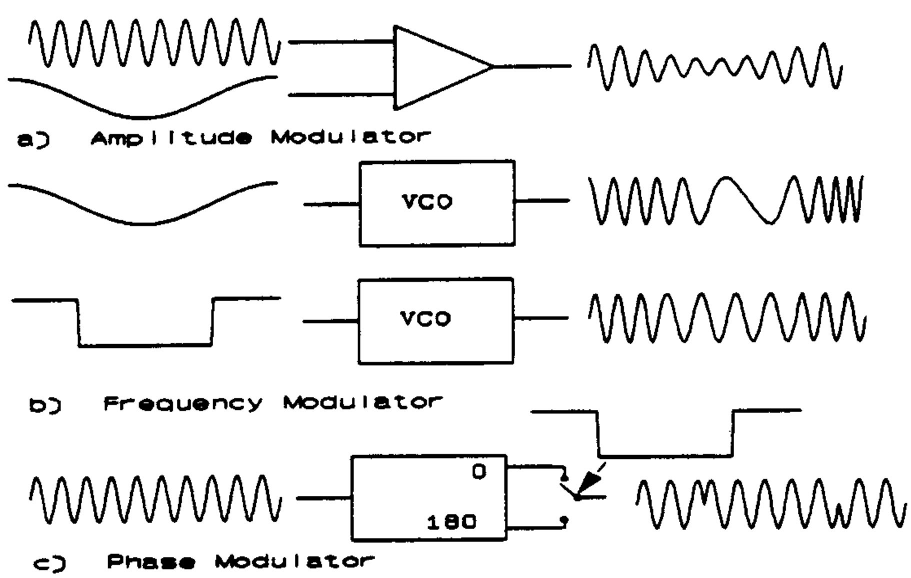

Figure 3 - Simple Modulation Techniques for AM, FM, and PM.

MEASUREMENT TOOLS

Figures 2 and 3 represent the results in the time domain, as they would be seen on an oscilloscope. The oscilloscope gives the observer a good presentation of the amplitude variations of a signal over some small finite period of time. Unfortunately, it is usually incapable of providing a reliable depiction of a signal's instantaneous frequency (f), or its time analog: wave period (p). When directly observing AM, one can often freeze the carrier sine wave on the display, measure the period, and calculate the frequency ($f = 1/p$). When directly observing FM or PM, it is usually extremely difficult to freeze the signal so that the time variation of frequency and instantaneous frequency can be measured. The frequency bandwidth of the oscilloscope is usually fixed and the user cannot accurately measure signals with components whose frequency is greater than that bandwidth.

A spectrum analyzer can be used to view the same signals in the frequency domain. It displays signal amplitude versus frequency. Normally the spectrum analyzer is allowed to scan slowly across a band of frequencies. In this way, it will display nearly the peak signal amplitude at each of the signal's component frequencies. (The ideal spectrum analyzer would have nearly zero bandwidth so that it would display each component frequency with no variation. Real receivers, however, have some finite bandwidth and display the vector sum of the frequencies within that passband at any given time.) The typical commercially available spectrum analyzer cannot simultaneously display, or measure, the time variation of a signal at each of its component frequencies. In its fixed frequency mode, the spectrum analyzer displays signal amplitude versus time in a narrow band of frequencies as determined by the analyzer's bandwidth setting. The maximum receiver bandwidth of a spectrum analyzer is usually less than the minimum bandwidth of an inexpensive oscilloscope.

A third instrument frequently used in signal analysis is a frequency selective voltmeter (FSVM). While it is not identical to a spectrum analyzer, the differences are beyond the scope of this paper.

SPECTRUM ANALYSIS

When two frequencies, f_1 and f_2, are mixed in a nonlinear junction, such as a diode or transistor, the resulting output signals consist of all the following:

a. the two fundamental frequencies: f_1 and f_2
b. all the harmonics of both fundamentals: $2f_1$, $2f_2$, $3f_1$, $3f_2$, $4f_1$, $4f_2$,....
c. all the intermodulation frequencies: $f_1 \pm f_2$, $f_1 \pm 2f_2$, $2f_1 \pm f_2$,...., $mf_1 \pm nf_2$, ..., where m and n are integers and $-\infty < m,n < \infty$.

The amplitude of each of these mixing products is a function of the amplitude of the two input signals and certain properties of the mixer. Figure 4 illustrates the effect.

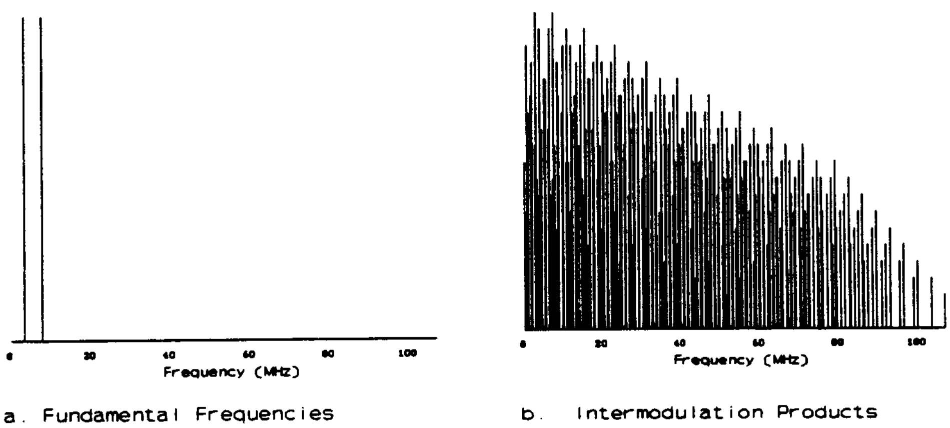

Figure 4 - Frequency Domain Representation of Intermodulation

By carefully filtering the modulator output, the radio frequency (RF) engineer can restrict the emission spectrum to only those frequencies necessary to convey the message. Most RF systems, even accidental mixers, contain some inherent filtering which naturally limits the output spectrum.

If $f_1 >> f_2$, the usual situation in intentional RF communications, the RF engineer can easily band limit the system so that it transmits only a narrow band of frequencies of the form: $f_1 + nf_2$, where n is an integer and $-\infty < n < \infty$. This is normally the situation for RF communications. Normal AM, in which the modulation is a single tone, f_m, contains only three frequencies: f_c, $f_c + f_m$, and $f_c - f_m$. If the modulation is complex, such as voice or music, the side tones, $f_c \pm f_m$, will become sidebands. This spectrum is shown in Figure 5. The spectrum resulting from analog FM or PM is much more complex than AM and the number of mixing components in the sidebands is a function of the modulation index as discussed below.

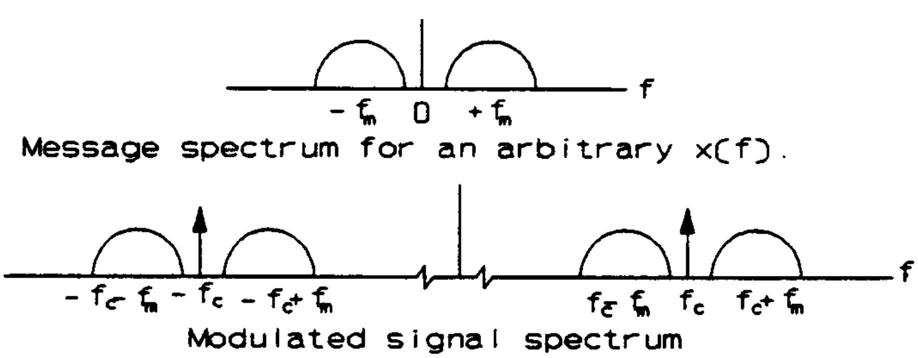

Figure 5 - Spectrum for an AM Signal

MODULATION INDEX

AMPLITUDE MODULATION[a]

As shown above the instantaneous AM signal can be described as

$$s_{AM}(t) = A_c[1 + x(t)] \cos ù_c t = A(t) \cos ù_c t \qquad (2)$$

where $x(t) = m_a \cos ù_m t$
and $A_c \cos ù_c t \equiv$ unmodulated carrier signal.

The AM *modulation index*, m_a, is defined as

$$m_a = \frac{[A(t)]_{max} - [A(t)]_{min}}{[A(t)]_{max} + [A(t)]_{min}} . \qquad (7)$$

If $s_{AM}(t)$ is the transmitted signal, it can be shown that[b]

$$S_T = S_C + S_C S_M \qquad (8)$$

where

$$S_C = A_c^2/2 \equiv \text{Carrier Power} \qquad (9)$$

$$S_M \equiv [x_{avg}(t)]^2 \qquad (10)$$

$$S_C S_M \equiv \text{Modulation Power}$$

If x(t) is a sinewave, then

$$S_M \equiv [x_{avg}(t)]^2 = \frac{m_a^2}{2} \qquad (11)$$

The reader will note in figure 5 that the modulation power is equally distributed into two sidebands; therefore, the power in each sideband is:

$$S_S = \frac{A_c^2[x_{avg}(t)]^2}{4} \qquad (12)$$

Equation (12) shows that the maximum power in either sideband is half of the carrier power.

AM is the simplest form of linear (amplitude) modulation and is

---

[f] *For purposes of this discussion, $|x(t)| \leq 1$. If $|x(t)| > 1$, the AM signal becomes overmodulated and A_c contains higher order modulation products. This is similar to the condition that would occur if x(t) describes a complex modulation. AM modulation index is usually measured using simple sine wave modulation to eliminate such confusion. As will be seen later, this constraint does not apply to the modulation index for angle modulation.*

[b] *Shanmugam, K. Sam, Digital and Analog Communication Systems, John Wiley and Sons, New York, 1979, pp 265-6.*

the form normally found in unintentional modulation and propagation situations. Other forms are:

Double Sideband (DSB) which appears similar to AM but without the carrier,

Single Sideband (SSB) which has only one sideband and no carrier, and

Vestigal Sideband (VSB) which has one sideband greatly attenuated.

Digital AM is a special case of linear AM in which the instantaneous amplitude, x(t), is constrained to a finite number of discrete values, such as pulse modulation, On-Off Keying (OOK), or amplitude shift keying (ASK).

Figure 6 shows both the time domain and frequency domain representations of a rectangular pulse train. The spectrum and the pulse train have several characteristic features:

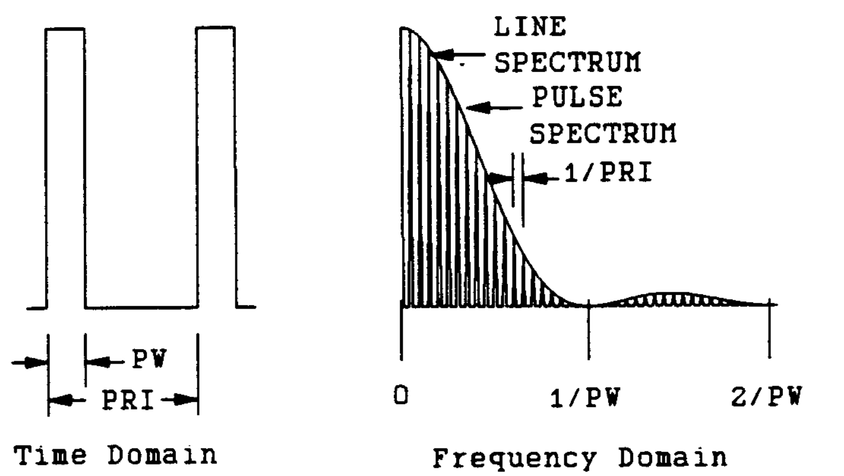

Figure 6 - Time and Frequency Domain Representations of a Rectangular Pulse

<u>Pulse Width</u> (PW, T) -- the length of time the pulse is ON. The frequency described by the inverse pulse width (1/PW) is the frequency distance between spectral nulls. The main lobe of the spectrum is 2/PW wide with its peak at f_c (0 Hz if the pulse is not modulating a carrier). This is the minimum receiver bandwidth for accurate demodulation of the pulse.

<u>Pulse Repetition Rate (PRR)</u>, (also pulse repetition freqency [PRF]) -- the number of pulses transmitted in one second. This frequency represents the separation between the discrete spectral

lines in the pulse spectrum. The time from the start of one pulse to the start of the next pulse is the pulse repetition interval (PRI), PRI = 1/PRF = τ.

ANGLE MODULATION

As previously shown the angle modulated signal can be represented by

$$s_a(t) = A_c \cos [\omega_c t + \phi(t)] \quad (3)$$

$$\text{where } \phi(t) = \begin{cases} k_p x(t) & \text{for PM} \quad (13) \\ k_f \int_{-\infty}^{t} x(\tau)\, d\tau & \text{for FM} \quad (14) \end{cases}$$

If $x(t) = A_M \cos \omega_m t$ where A_M is the magnitude of the modulating signal,

$$\text{then} \quad \phi(t) = \begin{cases} k_p A_M \cos \omega_m t & \text{for PM} \quad (15) \\ \dfrac{k_f A_M}{\omega_m} \sin \omega_m t & \text{for FM} \quad (16) \end{cases}$$

Equations (15) and (16) show that for tone modulation, PM and FM have essentially the same form but with a 90^o phase shift. Therefore, for either we could write

$$s_a(t) = A_c \cos [\omega_c t + \beta \sin (\omega_m t + \delta)] \quad (17)$$

$$\text{where} \quad \beta = \begin{cases} k_p A_M & \text{for PM} \quad (18) \\ \dfrac{k_f A_M}{\omega_m} & \text{for FM} \quad (19) \end{cases}$$

$$\text{and} \quad \delta = \begin{cases} -\pi/2 \ (-90^o) & \text{for PM} \quad (20) \\ 0 & \text{for FM} \quad (21) \end{cases}$$

β is the *modulation index* for angle modulation and has meaning

ONLY for tone modulation.

To find the spectrum of s(t), we can represent it as a Fourier series

$$s_a(t) = A_c \overset{\infty}{\underset{n=-\infty}{Ó}} J_n(â) \cos [(ù_c + nù_n)t] \qquad (22)$$

where $J_n(â)$ are Bessel functions of the first kind. Figure 7 shows a plot of the values of $J_n(â)$ for $J_0(â)$ and the first 12 sidebands. Once â is known, enter the graph in Figure 7 with a vertical line at that value and extract the values of $J_n(â)$ intersecting the vertical line. [$J_{-n}(â) = (-1)^n J_n(â)$.] These values are also provided in tabular form in many mathematics and engineering handbooks.

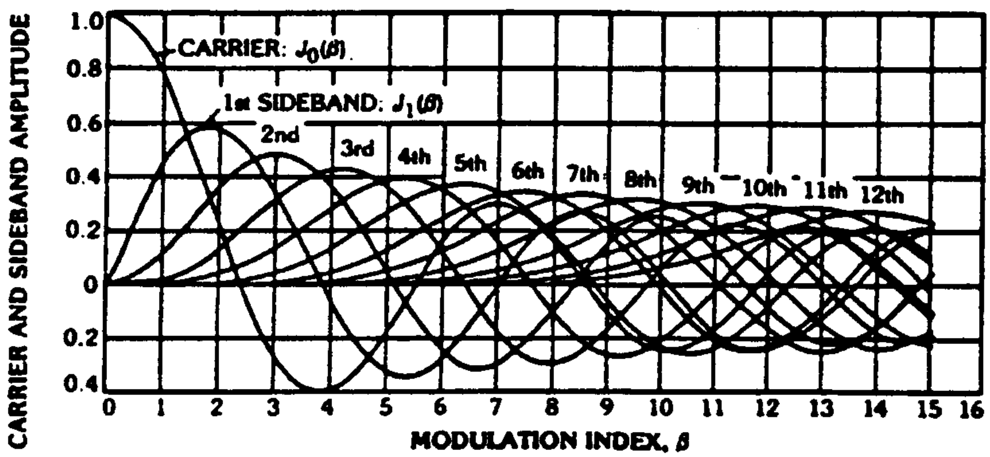

Figure 7 - Plot of Bessel Functions of the First Kind as a Function of â. (From P. F. Panter, *Modulation, Noise, and Spectral Analysis*, Fig. 7-3, McGraw-Hill Book Co., 1965.)

The spectrum has the following components:

1. The carrier, f_c, ($f = 2ðù$)
2. An infinite set of sidebands at $f_c \pm nf_m$.

Figure 8 shows the composition of an FM signal. In order to band limit the transmitted signal, the RF engineer usually limits the number of sidebands which will give ≥ 98% of the available power as determined by the inequality:

$$\tfrac{1}{2}[J_0(â)]^2 + [J_1(â)]^2 + [J_2(â)]^2 + \ldots \geq 0.49. \qquad (23)$$

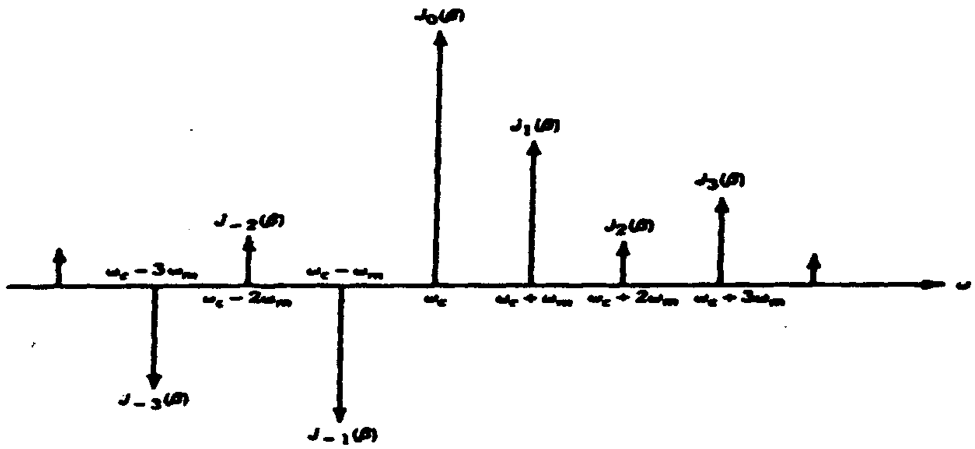

Figure 8 - Composition of an FM signal into sidebands. (From P. F. Panter, *Modulation, Noise, and Spectral Analysis*, Fig. 7-3, McGraw-Hill Book Co., 1965.)

Digital angle modulation appears distinctly different from analog angle modu

DIGITAL ANGLE MODULATION

In the frequency domain, digital angle modulation appears distinctly different from analog angle modulation. The basic unit for both digital FM and PM is the ASK rectangular pulse discussed above. Figure 9 shows the spectrum of a pulse whose data rate, frequency, is r_b, the single bit rate. A receiver bandwidth of $2r_b$ will obviously capture 90% of the energy in the pulse. The only way to increase the data rate is to increase the bit rate, which increase the signal bandwidth proportionately.

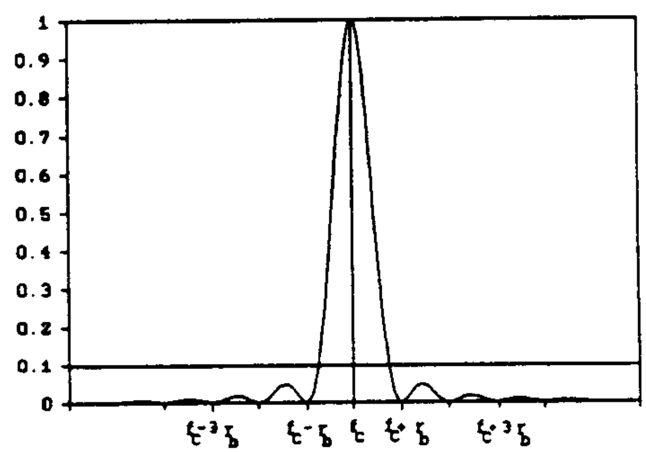

Figure 9 - Spectrum of a Single Frequency Pulse

Phase Shift Keying (PSK)

When digital PM, or phase shift keying (PSK), is monitored with a spectrum analyzer, its pulse spectrum is nearly identical to the rectangular pulse shown in Figure 9 and the necessary receiver bandwidth remains at $2r_b$. Using phase modulation, one can increase the data rate without increasing the bandwidth by sending a parallel bit stream, or MPSK, where M denotes the number of possible states at any instance. (QPSK, quadrature PSK, has four orthogonal states; 8PSK has eight states each separated by 450; etc.) The number of bits in each transmitted word, byte, is half the possible number of states. One further enhancement is achieved by adding ASK to the MPSK, normally referred to as MPSKAM. If the modulating data is essentially a random bit stream, the resulting pulse spectrum will have no line spectrum but will appear as a pulse of random noise. MPSK schemes have found great popularity in the last decade because of their ability to pass very large quantities of data over the extremely limited telephone bandwidths. Modulator-demodulators (modems) designed for 64 PSK and higher are commercially available. Such a device can handle data at 32 times that of binary ASK using the same bandwidth.

Digital FM

Digital FM, or frequency shift keying (FSK), appears as a pair of pulse modulation spectra centered about the carrier frequency. The frequency separation between the individual frequency maxima away from the carrier, f_d is arbitrarily chosen by the designer. Normally this decision is based on the available bandwidth and the-ability of the demodulator to distinguish between frequencies. Figure 10 shows the spectra for the component bit pulses and the resulting FSK pulse when using $f_d = r_b/2$. The reader will note that at this frequency dispersion, the effective BW is still approximately 2r b and the probability of distinguishing between the two bits is good. The amplitude of the composite pulse is proportional to the sum of the amplitudes of the component pulses.

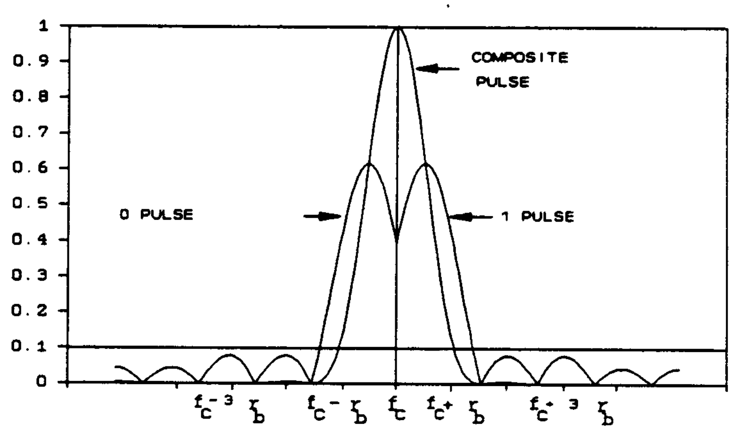

Figure 10 - Composite FSK Pulse and Component Bit Pulses for $f_d = r_b/2$

Figure 11 shows the power spectral densities, psd, for several different FSK frequency dispersions and a single pulse. A little inspection shows that only a small increase in bandwidth is necessary for $f \leq 0.7r_b$ and a bandwidth of $4r_b$ will contain $f_d = r_b$. All of these FSK schemes suffer the same problem: they only transmit data at the bit rate, r_b, and they require more sophisticated modulators and demodulators than binary ASK.

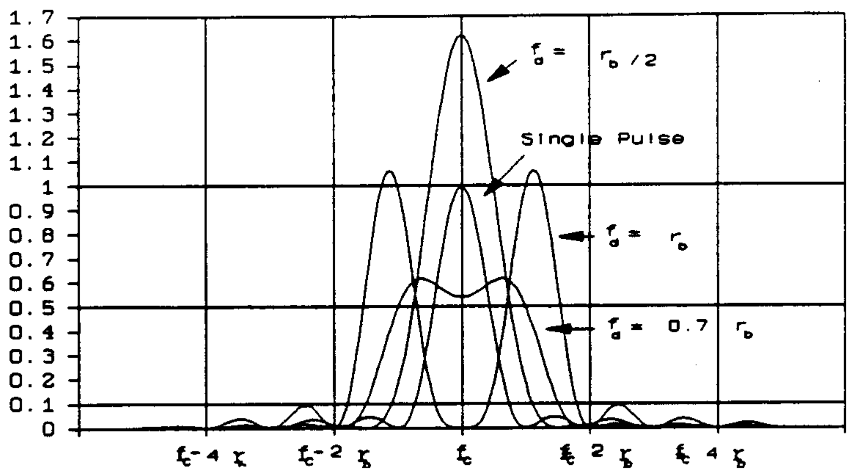

Figure 11 - Power Spectral Densities for Several FSK Signals Relative to a Single Pulse

A solution to the FSK problem is frequency domain multiplexing, FDM, using binary ASK. For the two-tone examples shown in Figure 11, the result is a doubling of the data rate using essentially the same equipment and bandwidth. Each extra tone, channel, adds another simultaneous bit and a proportional increase in bandwidth. Even greater efficiencies can be achieved if MPSK modulates the tone rather than ASK.

SPREAD SPECTRUM

Spread spectrum commonly is used to refer to any signaling technique that uses more spectral bandwidth than that actually required to transfer the information. The most common spread spectrum techniques are FDM and frequency agile radar. FDM uses multiple narrow bandwith channels each with its own preassigned carrier frequency (subcarrier) to modulate a single carrier for simultaneous transmission of a wide bandwidth signal. In FDM, the carrier frequency remains constant and the multiplexing allows for efficient spectrum utilization. Frequency agile radar changes its carrier frequency periodically in order to enhance targets or to avoid detection by the enemy. At any instance the radar is only using a narrow bandwidth, but over a short period, it will appear to be occupying a very wide bandwidth.

Much public attention has been focused on "spread spectrum communications." This term has been used to refer to a new class of military and civilian communications systems that do not remain on a fixed channel, but jump-about in some seemingly random fashion. Some, such as cellular telephone, remain on a given channel for the entire period of a communication, but change frequency when they change cells or at the start of a new communication. Other systems, mostly still experimental, change frequency frequently within a communication, similar to frequency agile radar. The theory for these systems is that they only use a particular frequency for a very short period; therefore, they cause very little interference to the regularly assigned user. By operating in short bursts on many frequencies, the system is able to use an entire operating band without significantly degrading the service to any of the current users. Additionally, if the individual frequency bursts are very short, it is virtually impossible to "lock-on" to the spread spectrum user to determine its location or listen to its transmissions.