Modulation Techniques

Prof. Tzong-Lin Wu
EMC Laboratory
Department of Electrical Engineering
National Taiwan University
• Analog modulation: AM, FM and PM.
• Modern wireless systems use digital modulation.
• More spectrum efficient, less power required and error correction
• Better immunity to fading.
• Study the effect of noise of digital modulation, the modulation and de-modulation.
9.1 Analog modulation

Because the IF signal in a receiver contains noise from the antenna and receiver circuitry, as well as the desired signal, the characteristics of the demodulator play a critical role in the overall performance of the wireless system.

**FIGURE 9.1** Illustrating the function of modulation and demodulation in transmitter and receiver systems: (a) modulator, (b) demodulator.
Single-sideband modulation

The upper sideband is selected from the mixer output with a bandpass filter, although the lower sideband could be selected, if preferred.

the spectrum occupied by the SSB signal extends from $f_{\text{IF}}$ to $f_{\text{IF}} + f_M$, where $f_M$ is the maximum frequency of $m(t)$.

The IF bandpass filter has a passband extending from $f_{\text{IF}}$ to $f_{\text{IF}} + f_M$,

Note that the demodulator LO is identical in frequency and phase with the modulator LO; this is called a synchronous, or coherent, demodulator.

FIGURE 9.2  Block diagram of a SSB modulator and demodulator.
The SSB input voltage to the demodulator can therefore be expressed as

\[ v_i(t) = A \cos(\omega_{IF} + \omega_m)t + n(t), \quad (9.1) \]

\( n(t) \) is limited to a narrow frequency band

\[ v_i(t) = A \cos(\omega_{IF} + \omega_m)t + x(t) \cos \omega_{IF}t - y(t) \sin \omega_{IF}t, \quad (9.2) \]

The output voltage of the mixer

\[ v_1(t) = v_i(t) \cos \omega_{IF}t \]

\[ = \frac{A}{2} \cos \omega_m t + \frac{A}{2} \cos(2\omega_{IF} + \omega_m)t + \frac{1}{2} x(t)(1 + \cos 2\omega_{IF}t) - \frac{1}{2} y(t) \sin 2\omega_{IF}t \quad (9.3) \]

After low-pass filtering, the final output voltage

\[ v_o(t) = \frac{A}{2} \cos \omega_m t + \frac{1}{2} x(t), \quad (9.4) \]
The average input signal power of the SSB signal of (9.1) is found by time-averaging:

\[ S_i = \frac{A^2}{2}. \]  \hspace{1cm} (9.5)

The average power of the output signal

\[ S_o = \frac{1}{2} \left( \frac{A}{2} \right)^2 = \frac{A^2}{8} = \frac{S_i}{4}. \]  \hspace{1cm} (9.6)

The power of the narrowband input noise of (9.1) and (9.2) is given by

\[ N_i = E\{n^2(t)\} = E\{x^2(t)\} = E\{y^2(t)\}, \]  \hspace{1cm} (9.7)

according to (3.47). Then the noise power of the output voltage of (9.4) is

\[ N_o = E\left\{ \left[ \frac{1}{2} x(t) \right]^2 \right\} = \frac{1}{4} E\{x^2(t)\} = \frac{N_i}{4}. \]  \hspace{1cm} (9.8)

So the output SNR is, from (9.6) and (9.8),

\[ \frac{S_o}{N_o} = \frac{S_i}{4} \frac{4}{N_i} = \frac{S_i}{N_i}, \]  \hspace{1cm} (9.9)
mixing noise has the effect of reducing its power by half.

in agreement with our earlier result of (9.8).

\[ N_i = 2f_M \left( \frac{n_0}{2} \right) = n_0 f_M. \quad (9.10) \]

\[ N_1 = 4f_M \left( \frac{n_0}{8} \right) = \frac{n_0 f_M}{2}. \quad (9.11) \]

\[ N_o = 2f_M \left( \frac{n_0}{8} \right) = \frac{n_0 f_M}{4} = \frac{N_i}{4}. \quad (9.12) \]

\[ \frac{S_o}{N_o} = \frac{S_i}{n_0 f_M}. \quad (9.13) \]

FIGURE 9.3  Power density spectrums of noise in a SSB demodulator. (a) Power spectral density of the noise at the demodulator input. (b) Power spectral density of the noise at the output of the mixer. (c) Power spectral density of the noise after low-pass filtering.
Double-sideband suppressed-carrier modulation

If the mixer were ideal, the carrier frequency \((f_w)\) would not be present in the output, so this modulation is referred to as double-sideband suppressed carrier (DSB-SC).

Input voltage to the demodulator can then be expressed as

\[
v_i(t) = \frac{A}{\sqrt{2}} \cos(\omega_{IF} - \omega_m)t + \frac{A}{\sqrt{2}} \cos(\omega_{IF} + \omega_m)t + n(t).
\] (9.14)

We have normalized the input voltage so that the total DSB-SC input power is identical to the input power of the SSB case.
\[ S_i = \frac{1}{2} \left( \frac{A}{\sqrt{2}} \right)^2 + \frac{1}{2} \left( \frac{A}{\sqrt{2}} \right)^2 = \frac{A^2}{2}. \]  \hfill (9.15)

narrowband representation for \( n(t) \):

\[ v_i(t) = \frac{A}{\sqrt{2}} \cos(\omega_{IF} - \omega_m)t + \frac{A}{\sqrt{2}} \cos(\omega_{IF} + \omega_m)t + x(t) \cos \omega_{IF}t - y(t) \sin \omega_{IF}t. \]  \hfill (9.16)

The output of the mixer is

\[ v_1(t) = v_i(t) \cos \omega_{IF}t \]

\[ = \frac{A}{\sqrt{2}} \cos \omega_m t + \frac{A}{\sqrt{2}} \cos(2\omega_{IF} - \omega_m)t + \frac{1}{2} x(t)(1 + \cos 2\omega_{IF}t) - \frac{1}{2} y(t) \sin 2\omega_{IF}t \]  \hfill (9.17)

After the low-pass baseband filter,

\[ v_o(t) = \frac{A}{\sqrt{2}} \cos \omega_m t + \frac{1}{2} x(t). \]  \hfill (9.18)

The output signal power is

\[ S_o = \frac{1}{2} \left( \frac{A}{\sqrt{2}} \right)^2 = \frac{A^2}{4} = \frac{S_i}{2}. \]  \hfill (9.19)

\( S_o = S_i/2 \), whereas in the SSB case we had \( S_o = S_i/4 \).
This improvement is due to the effective doubling of signal power.
A more realistic result is found by expressing the input noise power in terms of the power spectral density of the white noise at the input to the IF stage.

\[ N_i = 4f_M \left( \frac{n_0}{2} \right) = 2n_0f_M, \quad (9.22) \]

where the factor of four accounts for positive and negative frequency portions of the filter response.

\[ \frac{S_o}{N_0} = \frac{S_i}{n_0f_M}, \quad (9.23) \]

which is seen to be identical to the result for the SSB case given in (9.13). This implies that the coherent SSB and DSB-SC demodulators have the same SNR performance, when expressed in terms of a uniform input white noise level; this is a more meaningful comparison than that of (9.21)
EXAMPLE 9.1 SNR FOR A DSB-SC DEMODULATOR

Derive the output SNR for a DSB-SC demodulator with an arbitrary modulating signal, \( m(t) \), where the input signal and noise is represented by

\[ v_i(t) = m(t) \cos \omega_{IF}t + n(t). \]

Consider \( m(t) \) to be a random process with \( E\{m(t)\} = 0 \) and \( E\{m^2(t)\} = m^2/2 \).

Solution

Assuming an LO voltage of \( \cos \omega_{IF}t \), the output of the mixer is

\[ v_1(t) = v_i(t) \cos \omega_{IF}t \]

\[ = m(t) \cos^2 \omega_{IF}t + x(t) \cos^2 \omega_{IF}t - y(t) \cos \omega_{IF}t \sin \omega_{IF}t \]

\[ = \frac{1}{2} m(t)(1 + \cos 2\omega_{IF}t) + \frac{1}{2} x(t)(1 + \cos 2\omega_{IF}t) - \frac{1}{2} y(t) \sin 2\omega_{IF}t \]

After low-pass filtering the output voltage is

\[ v_o(t) = \frac{1}{2} m(t) + \frac{1}{2} x(t). \]

The average power of the input signal is found by time-averaging the expected value of the square of the input signal voltage:

\[ S_i = \frac{1}{2} E\{m^2(t)\} = \frac{m^2}{4}. \]
The output signal power is

\[ S_o = E \left\{ \left( \frac{1}{2} m(t) \right)^2 \right\} = \frac{m^2}{8} = \frac{S_i}{2}. \]

The noise power of the demodulator output is

\[ N_o = E \left\{ \left( \frac{1}{2} x(t) \right)^2 \right\} = \frac{1}{4} E\{x^2(t)\} = \frac{N_i}{4}. \]

Then the output signal-to-noise ratio is

\[ \frac{S_o}{N_o} = \frac{S_i}{2} \frac{4}{N_i} = 2 \frac{S_i}{N_i}, \]

which is the same result as was obtained for a sinusoidal modulating signal. \( \square \)
Double-sideband large-carrier modulation

If the double-sideband signal of the previous case is transmitted without suppression of the carrier wave, it is referred to as double-sideband large-carrier (DSB-LC) modulation.

This is an advantage in that the carrier signal, even if much lower in amplitude than the sidebands, can be used as a reference signal to phase-lock the local oscillator to synchronization with the incoming signal.

\[
v_i(t) = A[1 + m(t)] \cos \omega_{IF} t + n(t)
\]

\[
= A \cos \omega_{IF} t + \frac{mA}{2} \left[ \cos(\omega_{IF} - \omega_m)t + \cos(\omega_{IF} + \omega_m)t \right] + n(t)
\]  

(9.24)

The amplitude, \( m \), of the modulating signal relative to the carrier is called the modulation index.

\[v_i(t) = A[1 + m(t)] \cos \omega_{IF} t + n(t)\]

\[= A \cos \omega_{IF} t + \frac{mA}{2} \left[ \cos(\omega_{IF} - \omega_m)t + \cos(\omega_{IF} + \omega_m)t \right] + n(t)\]
Double-sideband large-carrier modulation

The signal power associated with this input voltage

\[ S_i = \frac{A^2}{2} + \frac{m^2 A^2}{4}, \quad (9.25) \]

The carrier power thus increases the total input power, but does not directly contribute to the output power after demodulation since it does not contain any modulation information.

Mixing the input voltage of (9.24) with the LO and low-pass filtering gives the demodulator output voltage as

\[ v_o(t) = \frac{mA}{2} \cos \omega_m t + \frac{1}{2} x(t). \quad (9.26) \]

the output signal power

\[ S_o = \frac{m^2 A^2}{8} = \frac{m^2 A^2}{8} \left( \frac{A^2}{2} + \frac{m^2 A^2}{4} \right) = \frac{m^2 S_i}{2(2 + m^2)}, \quad (9.27) \]
Using (9.20) and (9.22) with (9.27) gives the output SNR

\[
\frac{S_o}{N_o} = \frac{S_i}{N_i} \frac{2m^2}{2 + m^2} = \frac{m^2 S_i}{n_0 f_M (2 + m^2)}.
\] (9.28)

Note that if \( m \gg 1 \), so that the carrier power is negligible relative to the sideband power, (9.28) reduces to (9.23). When the carrier power is not negligible, however, the output SNR is less than that for either SSB or DSB-SC. For example, when \( m = 1 \) (a modulation index of 100\%), we have

\[
\frac{m^2}{2 + m^2} = \frac{1}{3},
\]

which implies a reduction in SNR of 4.8 dB.
Envelope detection of double-sideband modulation

SSB and DSB-SC demodulators require a synchronous local oscillator for proper operation.

An advantage of DSB with a carrier component (DSB-LC) is that detection can be done without a local oscillator and mixer, by using an envelope detector (noncoherent demodulator).

It is the preferred method for broadcast AM radio, where it is desired to make the receiver as inexpensive as possible.

**The DSB-LC signal waveform**

\[ v_i(t) = A[1 + m(t)] \cos \omega_{IF}t = r(t) \cos \omega_{IF}t, \quad (9.29) \]

**FIGURE 9.5** Modulation envelope of a double-sideband amplitude modulated carrier.
The basic envelope detector circuit

The RC time constant should **be large enough** so that the capacitor voltage does not decay too quickly before the next carrier peak arrives, but **small enough** so the output can track the envelope when it is decreasing.

The output of the envelope detector shown in Figure 9.6a has a DC level that must be removed with a series capacitor. This may limit the low frequency response of the detector.

Thus broadcast radio systems must ensure that the modulation index is always less than 100% in order to avoid signal distortion.

![Diagram of basic envelope detector circuit](image)

**FIGURE 9.6** Envelope detection of an AM signal. (a) Basic envelope detector circuit. (b) Output waveform of the envelope detector.
Derive the Output SNR of the Envelope Detector

![Block diagram of a DSB-LC modulator and demodulator using envelope detection.](image)

**FIGURE 9.7** Block diagram of a DSB-LC modulator and demodulator using envelope detection.

\[ v_i(t) = A[1 + m(t)] \cos \omega_{IF}t + n(t) \]
\[ = [A + Am \cos \omega_m t + x(t)] \cos \omega_{IF}t - y(t) \sin \omega_{IF}t \]
\[ = r(t) \cos[\omega_{IF}t + \theta(t)] \quad (9.30) \]
the envelope of the carrier

\[ r(t) = \sqrt{[A + Am \cos \omega_m t + x(t)]^2 + y^2(t)}. \] \hspace{1cm} (9.31)

The phase function

\[ \theta(t) = \tan^{-1} \frac{y(t)}{A + Am \cos \omega_m t + x(t)}. \] \hspace{1cm} (9.32)

The signal input power

\[ S_i = \frac{A^2}{2} + \frac{m^2 A^2}{4}, \] \hspace{1cm} (9.33)

The output signal power is given by the time-average of the square of the envelope voltage with the noise terms set to zero:

\[ S_o = A^2 + \frac{m^2 A^2}{2}. \] \hspace{1cm} (9.34)

The first term in (9.34) represents power of a DC component, which would generally be filtered with a DC block.
Evaluation of the noise power contained in the envelope voltage is complicated by the presence of the square root in (9.31).

by assuming \( x(t) \ll A \) and \( y(t) \ll A \) (9.31) reduces to

\[
r(t) = \left[ A + Am \cos \omega_m t + x(t) \right] \sqrt{1 + \frac{y^2(t)}{A + Am \cos \omega_m t + y(t)}}
\]

\[
\approx A + Am \cos \omega_m t + x(t)
\]

(9.35)

\[
N_o = E\{x^2(t)\} = N_i,
\]

(9.36)

\[
\frac{S_o}{N_o} = \frac{S_i}{N_i} \frac{2m^2}{2 + m^2} \quad \text{(for large } S_i/N_i)\]

(9.37)

This is identical to the result of (9.28) for the synchronous demodulation of a DSB-LC signal, showing that no degradation in SNR occurs with the much simpler method of envelope detection for high input SNR.
When the input SNR is not large, however, the output SNR for envelope detection will be reduced. In fact, if we expand (9.31) as

\[ r(t) = \sqrt{A^2 + 2A^2 m(t) + 2Ax(t) + 2Am(t)x(t) + x^2(t) + A^2m^2(t) + y^2(t)}, \quad (9.38) \]

we see that the envelope voltage does not contain a term proportional to the modulating signal \( m(t) \), but involves products of the signal and noise, the square of the signal, and the square root function. This means that the output of the envelope detector is seriously distorted and unusable for small input SNR. Figure 9.8 shows a graph of the output SNR.

**FIGURE 9.8** Output versus input SNR for synchronous and envelope detection of a DSB-LC AM signal.
Frequency modulation (FM)

Better SNR then AM at expense of bandwidth

An FM waveform has the general form

\[ v(t) = A \cos \theta(t), \]

where

\[ \theta(t) = \omega_{IF} t + k \int_{\tau=0}^{t} m(\tau) \, d\tau. \quad \text{rad} \]

\[ \omega = \frac{d\theta(t)}{dt} = \omega_{IF} + km(t) \quad \text{rad/sec.} \]
For the special case of a sinusoidal modulating signal \( m(t) = m \cos \omega_m t \), (9.40) reduces to

\[
\theta(t) = \omega_{IF}t + \frac{km}{\omega_m} \sin \omega_m t. \tag{9.42}
\]

If we define \( \Delta \omega = km \) as the maximum frequency deviation, and define \( \beta = \Delta \omega / \omega_m \) as the modulation index, then (9.39) can be written simply as

\[
v(t) = A \cos(\omega_{IF}t + \beta \sin \omega_m t). \tag{9.43}
\]
The FM signal of (9.43) has an **infinite number of sidebands**, in contrast to the finite-bandwidth AM signals discussed above. We can analyze the frequency spectrum of (9.43) by expanding the waveform into a **Fourier series** that is periodic in multiples of $\omega_m$. The easiest way to do this is to write (9.43) as the real part of a **complex exponential**,

$$v(t) = A \text{Re}\{e^{j\theta(t)}\} = A \text{Re}\{e^{j\omega_m t} e^{j\beta \sin \omega_m t}\}, \quad (9.44)$$

$$e^{j\beta \sin \omega_m t} = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega_m t}.$$  

$$C_n = \frac{1}{T} \int_{t=-T/2}^{T/2} e^{j\beta \sin \omega_m t} e^{-jn\omega_m t} dt, \quad (9.45)$$

$$C_n = \frac{1}{2\pi} \int_{x=-\pi}^{\pi} e^{j(\beta \sin x - nx)} \, dx = J_n(\beta), \quad (9.46)$$

where $J_n(\beta)$ is the Bessel function of the first kind of order $n$. 
Thus there are sidebands spaced at multiples of $\omega_m$ on either side of the carrier at $\omega_{IF}$, with amplitudes given by $AJ_n(\beta)$. Typical spectra for three values of $\beta$ are shown in Figure 9.10.

$$v(t) = A \Re \left\{ e^{j\omega_{IF}t} \sum_{n=-\infty}^{\infty} J_n(\beta)e^{jn\omega_m t} \right\} = A \sum_{n=-\infty}^{\infty} J_n(\beta) \cos(\omega_{IF} + n\omega_m)t.$$  (9.47)
Observe that the sideband magnitude decreases for large $n$, so the practical bandwidth is not infinite.

In fact, good transmission fidelity generally only requires those sidebands whose amplitude is 1% or larger relative to the unmodulated carrier (the $n = 0$ term)

for all $n$ such that $|J_n(\beta)| \geq 0.01$.

For large $\beta$, $n \approx \beta$ will satisfy this condition.

$$B = 2n\omega_m \approx 2\beta\omega_m = 2\Delta\omega, \text{ (large } \beta)$$

In the case of small $\beta$, $n \approx 1$ for the last required sideband.

$$B = 2n\omega_m \approx 2\omega_m. \text{ (small } \beta)$$

![Magnitude of the spectrum of a frequency modulated signal, for various values of the modulation index $\beta$. (a) $\beta = 0.1$, (b) $\beta = 1.0$, and (c) $\beta = 5.0$. The center frequency is $f_0 = \omega_0/2\pi$, and the spacing of all sidebands is $f_s = \omega_m/2\pi$.](image)
A convenient way to combine these two results is with the approximation known as Carson's rule:

\[ B \approx 2(\Delta \omega + \omega_m) = 2\omega_m(1 + \beta). \quad (9.48) \]

In telecommunication, Carson's bandwidth rule defines the approximate bandwidth requirements of communications system components for a carrier signal that is frequency modulated by a continuous or broad spectrum of frequencies rather than a single frequency.
S/N for FM modulation

\[ v_i(t) = A \cos \theta(t) = A \cos(\omega_{IF} t + \beta \sin \omega_m t). \]  \hfill (9.49)

\[ S_i = A^2 / 2. \]  \hfill (9.50)

\[ v_1(t) = \frac{d}{dt} v_i(t) = -(\omega_{IF} + \beta \omega_m \cos \omega_m t) \sin(\omega_{IF} t + \beta \sin \omega_m t). \]  \hfill (9.51)

After envelope detection (and DC blocking of the \( \omega_{IF} \) term) the output voltage is

\[ v_o(t) = \beta \omega_m \cos \omega_m t, \]  \hfill (9.52)

\[ S_o = \beta^2 \omega_m^2 / 2. \]  \hfill (9.53)

**FIGURE 9.9** Block diagram of an FM modulator and demodulator.
Noises

Next we consider the presence of noise, and set $\beta = 0$ to compute the output noise

$$s_i(t) + n_i(t) = A \cos \omega_{\text{IF}} t + x(t) \cos \omega_{\text{IF}} t - y(t) \sin \omega_{\text{IF}} t$$

$$= r(t) \cos[\omega_{\text{IF}} t - \phi(t)]$$

(9.54)

$$r(t) = \sqrt{[A + x(t)]^2 + y^2(t)},$$

$$\phi(t) = \tan^{-1} \frac{y(t)}{A + x(t)}.$$ 

If we assume the noise voltage is small relative to the carrier amplitude,

$A \gg x(t)$ and $y(t)$

$$\phi(t) \approx \frac{y(t)}{A}.$$ 

(9.55)
The limiter will remove the effect of the time variation of \( r(t) \), so the noise output of the discriminator will be

\[
\begin{align*}
n_1(t) &= \frac{d}{dt} \cos[\omega_{IF} t - \phi(t)] \\
&= -\left(\omega_{IF} - \frac{d\phi(t)}{dt}\right) \sin[\omega_{IF} t - \phi(t)] \quad (9.56)
\end{align*}
\]

After envelope detection and DC blocking, the output noise voltage is

\[
n_o(t) = \frac{d\phi(t)}{dt} = \frac{1}{A} \frac{dy(t)}{dt}. \quad (9.57)
\]

Since \( H(\omega) = j\omega \) for a differentiator (see Appendix B), the power spectral density of the output noise can be found as

\[
S_{n_o}(\omega) = |H(\omega)|^2 S_y(\omega) = \frac{\omega^2}{A^2} n_0 \quad \text{for } |\omega| < \omega_m. \quad (9.58)
\]
Then the output noise power over the IF passband from $-\omega_m$ to $\omega_m$ is

$$N_o = \frac{1}{2\pi} \int_{-\omega_m}^{\omega_m} S_{n_0}(\omega) d\omega = \frac{n_0 \omega_m^3}{3\pi A^2}.$$  \hspace{1cm} (9.59)

The input noise power over the same passband is

$$N_i = \frac{n_0}{2} (2\omega_m) \frac{1}{2\pi} = \frac{n_0 \omega_m}{2\pi}.$$  \hspace{1cm} (9.60)

Finally, the output SNR of the FM demodulator is

$$\frac{S_o}{N_o} = \frac{3\pi A^2 \beta^2}{2n_0 \omega_m} = \frac{3\pi \beta^2}{n_0 \omega_m} S_i = \frac{3}{2} \beta^2 \frac{S_i}{N_i},$$  \hspace{1cm} (9.61)

Note that (9.61) indicates that the SNR of a demodulated FM signal can be improved by a factor of $3\beta^2/2$. 
Comparing this with the corresponding result for AM demodulation by using (9.28) with a 100% modulation index \( m = 1 \)

\[
\frac{S_o}{N_o} \bigg|_{\text{AM}} = \frac{2\pi S_i}{3n_0\omega_m},
\]

\[
\frac{S_o}{N_o} \bigg|_{\text{FM}} = \frac{9}{2} \beta^2.
\] (9.62)

Using

\[ B_{\text{FM}} = 2\beta f_m \quad B_{\text{AM}} = 2f_m. \]

\[
\frac{S_o}{N_o} \bigg|_{\text{FM}} = \frac{9}{2} \left( \frac{B_{\text{FM}}}{B_{\text{AM}}} \right)^2.
\] (9.63)

The results of (9.62) and (9.63) clearly illustrate the main advantage of FM relative to AM, which is that FM allows an improvement in SNR at the expense of increased bandwidth, while AM does not. For example, with a modulation index of \( \beta = 4 \) the SNR for FM is 72 times better than that for AM, while the bandwidth is four times larger.
9.2 Binary digital modulation

Digital Mod. Vs Analog mod.

Improved performance in the presence of noise and fading, lower transmit power requirements, and better suitability for transmission of digital data with error correction and encryption.

Amplitude shift keying (ASK)
Frequency shift keying (FSK)
Phase shift keying (PSK)

FIGURE 9.11 Binary baseband data and modulated waveforms for ASK, FSK, and PSK.
Binary Signals: RZ, NRZ

**FIGURE 9.12** Binary signaling methods. (a) On-off, or return-to-zero (RZ), coding. (b) Non-return-to-zero (NRZ) coding. (c) Polar NRZ coding.
**Amplitude shift keying**

This requires the LO to **have precisely the same phase** and frequency as the incoming signal.

When the local oscillator does not have perfect phase coherence or frequency synchronization, distortion may be introduced.

\[
\begin{align*}
v(t) &= m(t) \cos \omega_0 t, \\
v_1(t) &= v(t) \cos \omega_0 t \\
&= \frac{1}{2} m(t)(1 + \cos 2\omega_0 t) \\
v_o(t) &= \frac{1}{2} m(t),
\end{align*}
\]  \hspace{1cm} (9.64)
ASK can be demodulated coherently using a synchronous local oscillator and mixer, but it is also possible to demodulate ASK with an envelope detector.

Although envelope detection cannot be used to demodulate a DSB-SC signal, it can be used with an ASK DSB-SC waveform because the modulating signal m(t) is never negative.
Another noncoherent demodulator is the rectifier, or square-law, detector shown in Figure 9.14.

**FIGURE 9.14** Circuit for a rectifier detector.
EXAMPLE 9.2  SQUARE-LAW DETECTOR

Another noncoherent demodulator is the rectifier, or square-law, detector shown in Figure 9.14. It is similar to the envelope detector, with the important difference that a capacitor is not used. If the diode is biased to operate in its square-law region, we can assume that the voltage across the resistor is proportional to the square of the input voltage. If the ASK waveform of (9.64) is applied to the input of the square-law detector, derive the output voltage from the low-pass filter.

Solution

Assuming a square-law dependence, the resistor voltage can be written as

\[ v_R(t) = C v^2(t) = C m^2(t) \cos^2 \omega_0 t \]

\[ = \frac{C}{2} m^2(t)(1 + \cos 2\omega_0 t) \]

After low-pass filtering, we then have the output

\[ v_o(t) = \frac{C}{2} m^2(t) = \frac{C}{2} m(t), \]

where the last step follows because \( m^2(t) = m(t) \) when \( m = 0 \) or \( 1 \).
Frequency shift keying involves switching a sinusoidal carrier wave between two frequencies.

\[ \omega_1 = \omega_0 - \Delta \omega \]
\[ \omega_2 = \omega_0 + \Delta \omega \]

\( \Delta \omega \) is known as the frequency deviation.

**FSK modulated waveform**

\[ v(t) = \cos \omega t, \quad \omega = \begin{cases} \omega_1 & \text{for } m(t) = 1 \\ \omega_2 & \text{for } m(t) = 0 \end{cases} \quad (9.66) \]

In general the spectrum of an FSK signal is complicated because of the essentially random switching between two frequency states.

It can be shown

the effective bandwidth of an FSK signal is \( B = 2(\Delta f + 2/T) \).
Modulation and demodulation of FSK signals.

Assume we have the incoming FSK waveform of (9.66) with $\omega = \omega_1$.

\begin{align}
v_1(t) &= v(t) \cos \omega_1 t = \cos^2 \omega_1 t = \frac{1}{2} (1 + \cos 2\omega_1 t) \\
v_2(t) &= v(t) \cos \omega_2 t = \cos \omega_1 t \cos \omega_2 t \\
&= \frac{1}{2} \left[ \cos(\omega_1 - \omega_2) t + \cos(\omega_1 + \omega_2) t \right]
\end{align}

(9.67a) (9.67b)

After low-pass filtering only the DC term from (9.67a) remains, resulting in a positive pulse at the output of the summer, indicating that a "1" has been received.

**FIGURE 9.15** Modulation and demodulation of FSK signals. (a) FSK modulator. (b) Synchronous demodulation of FSK.
FSK can also be demodulated using an envelope detector, thereby avoiding the requirement for two coherent local oscillators.

The outputs of the envelope detectors are combined with a summer to form a polar **NRZ** output.

---

**FIGURE 9.16** Envelope detection of FSK. (a) FSK envelope detector. (b) Decomposition of FSK signal into two ASK signals.
Phase shift keying

In PSK modulation the phase of the carrier wave is switched between two states, usually 0° and 180°.

\[ v(t) = m(t) \cos \omega_0 t, \]  

where \( m(t) = 1 \) or \(-1\).

The spectrum of the PSK waveform is relatively wide in bandwidth due to the sharp transitions caused by phase reversal. These are usually smoothed by filtering, but the resulting bandwidth is usually still wide enough that PSK is impractical for multichannel wireless systems.

**FIGURE 9.17** Modulation and demodulation of PSK signals. (a) PSK modulator, (b) Synchronous demodulation of PSK.
Synchronous demodulation of PSK

Since the PSK waveform has a constant envelope, it cannot be demodulated with an envelope detector.

\[
v_1(t) = v(t) \cos \omega_0 t = \frac{1}{2} m(t)[1 + \cos 2\omega_0 t],
\]

\[
v_o(t) = \frac{1}{2} m(t),
\]  \hspace{1cm} (9.69)
Carrier synchronization

Synchronism implies that both frequency and phase are identical, a result that is generally difficult to achieve in practice.

As an example, if we set a criteria of requiring less than 45" phase error at a carrier frequency of 1 GHz, synchronization of the LO to the carrier must be better than $T/8 = 0.125$ nS. A free-running local oscillator will virtually never exhibit such synchronism because of frequency drift, Doppler effects, and the arbitrary (and sometimes variable) distance between the transmitter and receiver.

Two ways in which a local oscillator can be synchronized with an incoming carrier wave: transmit a pilot carrier, or use a carrier-recovery circuit. Transmitting a low-level carrier is probably the easiest way, as this signal can be used to phase-lock the local oscillator.

ASK and FSK can use envelop detector but bad BER.
Carrier recovery circuits use phase or frequency information from the received signal to synchronize the local oscillator. (a phase-locked loop)

A clever alternative carrier-recovery circuit for **PSK demodulation**

![Diagram of carrier recovery circuit]

**FIGURE 9.18** A carrier-recovery circuit suitable for synchronous PSK demodulation.
9.3 Error probability for binary modulation

**PCM signals and detections**

In general, PCM codes a binary "1" as a signal voltage $s_1(t)$, and a binary "0" as a signal voltage $s_2(t)$, each of bit duration $T$.

In the transmitter these PCM signals are used to modulate a carrier by amplitude, frequency, or phase modulation.

In the receiver, as we have shown, synchronous demodulation or envelope detection (for ASK or FSK) can be used to recover $s(t) = s_1(t)$ or $s_2(t)$.

The demodulated signal is then used to make the decision as to whether a binary "1" or "0" has been received. **In the absence of noise** this can simply be done by setting a detection threshold, where "1" or "0" is decided using a comparator circuit.

### TABLE 9.1 PCM Signals for ASK, FSK, and PSK

<table>
<thead>
<tr>
<th>Modulation Type</th>
<th>$s_1(t)$</th>
<th>$s_2(t)$</th>
<th>Detection Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASK</td>
<td>V</td>
<td>0</td>
<td>V/2</td>
</tr>
<tr>
<td>FSK</td>
<td>V</td>
<td>$-V$</td>
<td>0</td>
</tr>
<tr>
<td>PSK</td>
<td>V</td>
<td>$-V$</td>
<td>0</td>
</tr>
</tbody>
</table>
PCM signals and detections

When noise is present, the problem of detection is more difficult. It can be shown in this case that optimum detection of PCM signals can be made with a correlation receiver, consisting of an integrator (or low-pass filter) followed by a sampler and a comparator.

Although the effect of the (zero mean) noise voltage is minimized by averaging over the bit period, the random nature of the noise results in occasional errors, and the likelihood of an error increases with the power of the input noise.

The likelihood that a single bit is received incorrectly is called the probability of error, or the bit error rate (BER).

![Diagram of correlation receiver](image)

**FIGURE 9.19** Optimum correlation receiver for PCM detection. (a) Correlation detector circuit. (b) Input and output signal voltage waveforms.
• Probability of error or bit error rate (BER)

\[ v_o(T) = \int_{t=0}^{T} [s(t) + n(t)] \, dt = s_o(T) + n_o(T), \quad (9.70) \]

\[ s_o(T) = \int_{t=0}^{T} s(t) \, dt = \begin{cases} VT & \text{if } s(t) = V \\ -VT & \text{if } s(t) = -V \\ 0 & \text{if } s(t) = 0 \end{cases} \quad (9.71) \]

\[ n_o(T) = \int_{t=0}^{T} n(t) \, dt. \quad (9.72) \]

**FIGURE 9.19** Optimum correlation receiver for PCM detection. (a) Correlation detector circuit. (b) Input and output signal voltage waveforms.
Bit energy as the energy of the (non-zero) signal voltage over one bit period:

\[ E_b = \int_{t=0}^{T} s^2(t) \, dt = V^2 T. \quad (9.73) \]

The noise power output from an integrator with a bandlimited white noise input was derived in (3.38) as

\[ N_o = \frac{n_0 T}{2} = \sigma^2, \quad (9.74) \]

where \( n_0/2 \) is the two-sided power spectral density of the white noise, and \( \sigma^2 \) is the variance of the Gaussian probability density function (don't confuse \( n_o(T) \) with \( n_0 \)). We can now
Synchronous ASK

Assume that a binary "0" has been transmitted in an ASK system, in the presence of bandlimited Gaussian white noise. An error will occur if the output signal and noise voltage is greater than the threshold $VT/2$.

\[ P_e^{(0)} = P \left\{ s_o(T) + n_o(T) > \frac{VT}{2} \right\} = P \left\{ n_o(T) > \frac{VT}{2} \right\} \]

\[ = \int_{n_0 = \frac{VT}{\sqrt{2\pi} \sigma}}^{\infty} e^{-n_0^2/2\sigma^2} \frac{e^{-n_0^2/2\sigma^2}}{\sqrt{2\pi} \sigma} dn_0 \]

\[ P_e^{(0)} = \frac{1}{\sqrt{\pi}} \int_{x = \frac{VT}{2\sqrt{2\sigma}}}^{\infty} e^{-x^2} dx = \frac{1}{2} \text{erfc} \left( \frac{VT}{2\sqrt{2\sigma}} \right). \quad (9.75) \]

using (9.73) and (9.74):

\[ \frac{VT}{2\sqrt{2\sigma}} = \sqrt{\frac{E_b}{T}} \sqrt{\frac{T}{2\sqrt{2}}} \sqrt{\frac{2}{n_0 T}} = \sqrt{\frac{E_b}{4n_0}}. \]

\[ P_e^{(0)} = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{E_b}{4n_0}} \right). \quad (9.76) \]

\[ P_e = \frac{1}{2} P_e^{(0)} + \frac{1}{2} P_e^{(1)} = P_e^{(0)} = P_e^{(1)} = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{E_b}{4n_0}} \right). \quad (9.77) \]
Synchronous PSK

Now assume that a binary "0" is transmitted in a PSK system, in the presence of bandlimited Gaussian white noise.

Then we have \( s(t) = s_2(t) = -V \) and \( s_o(T) = -VT \), with a detection threshold of zero.

An error will occur if the sampled output voltage from the integrator is greater than the threshold.

\[
P_e^{(0)} = P\{s_o(T) + n_o(T) > 0\} = P\{n_o(T) > VT\}
\]

\[
= \int_{n_0=VT}^{\infty} \frac{e^{-n_0^2/2\sigma^2}}{\sqrt{2\pi \sigma^2}} \, dn_0 = \frac{1}{2} \text{erfc} \left( \frac{VT}{\sqrt{2\sigma}} \right) = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{E_b}{n_0}} \right)
\]

(9.78)

Due to the symmetry of the PSK signal and the demodulator, \( P_e^{(1)} = P_e^{(0)} \).
ASK v.s. PSK

Observe that the bit energy-to-noise ratio in the argument of the complementary error function for the PSK case differs by a factor of four (6 dB) compared to the result for ASK.

This implies that, for the same probability of error, PSK requires only one-fourth the power of an ASK system.

Since an ASK signal is off half the time, however, the average transmit power of an ASK system is half that of a PSK system, for the same peak power (same signal voltage, V).

Thus, in terms of average transmit power, the PSK result is better by a factor of two (3 dB), compared with ASK.
Synchronous FSK

Next we consider a synchronous FSK receiver, and assume that a binary "0" is transmitted in the presence of bandlimited Gaussian white noise.

In the synchronous FSK demodulator output noise consists of the difference between noise that has passed through both the \( \omega_1 \) and \( \omega_2 \) channels.

\[
n(t) = n_1(t) - n_2(t). \tag{9.79}
\]

The variance of \( n(t) \) can be computed as (\( n_1 \) and \( n_2 \) are uncorrelated)

\[
E\{n^2(t)\} = E\{n_1^2(t) - 2n_1(t)n_2(t) + n_2^2(t)\} \\
= E\{n_1^2(t)\} + E\{n_2^2(t)\} = 2\sigma^2 \tag{9.80}
\]

The result of (9.80) shows that the total noise power of the FSK demodulator is doubled relative to the synchronous ASK or PSK demodulator.
Synchronous FSK

The probability of error for synchronous FSK is calculated in the same way as for PSK, but with $\sigma^2$ replaced with $2\sigma^2$:

$$P_e^{(0)} = P\{s_o(T) + n(T) > 0\} = P\{n_1(T) - n_2(T) > VT\} = \frac{1}{2} \text{erfc} \left( \frac{VT}{2\sigma} \right) = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{E_b}{2n_0}} \right)$$

(9.81)

Observe that synchronous FSK requires 3 dB more signal power than an equivalent PSK system for the same probability of error, and 3 dB less power than an ASK system on a peak power basis.

FSK and ASK, however, have equal error rates when compared in terms of average transmit power, since an ASK system transmits power only half the time.
Envelope detection of ASK

The derivation of the probability of error is more difficult because of the effect of the nonlinearity of the envelope detector on the noise.

\[ v(t) = m(t) \cos \omega_0 t + n(t), \quad (9.82) \]

Narrowband representation of noise

\[ v(t) = [m(t) + x(t)] \cos \omega_0 t - y(t) \sin \omega_0 t \]
\[ = r(t) \cos[\omega_0 t + \theta(t)] \quad (9.83) \]

\[ r^2(t) = [m(t) + x(t)]^2 + y^2(t), \quad (9.84) \]

\[ \tan \theta(t) = \frac{y(t)}{m(t) + x(t)}. \quad (9.85) \]

Statistics of \( r(t) \) depends on \( m(t) \), We will discuss \( m(t)=0 \) and \( m(t)=V \), respectively.
\[ m(t) = 0 \]

\( x(t) \) and \( y(t) \) are independent, joint probability distribution function

\[
    f_{xy}(x, y) = f_x(x)f_y(y) = \frac{e^{-(x^2+y^2)/2\sigma^2}}{2\pi \sigma^2},
\]

\[ r^2(t) = x^2(t) + y^2(t), \]

\[ x(t) = r(t) \cos \theta(t) \]
\[ y(t) = r(t) \sin \theta(t) \]

\[
    f_{r\theta}(r, \theta) = f_{xy}(x, y) = \frac{e^{-r^2/2\sigma^2}}{2\pi \sigma^2}.
\]

Rayleigh probability distribution

\[
    f_r(r) = \int_{\theta=0}^{2\pi} f_{r\theta}(r, \theta) r \, d\theta = \frac{r e^{-r^2/2\sigma^2}}{\sigma^2},
\]
detection threshold is set at \( r_0 \), the probability of error when sending \( m(t) = 0 \) is

\[
P_e^{(0)} = P\{r(t) > r_0\} = \int_{r=r_0}^{\infty} f_r(r) \, dr = -e^{-r^2/2\sigma^2} \bigg|_{r=r_0}^{\infty} = e^{-r_0^2/2\sigma^2}.
\]

(9.89)
m(t) = V

Let \( x' = x + V \) the pdf for \( x' \) is

\[
f_{x'}(x') = \frac{e^{-(x'-V)^2/2\sigma^2}}{\sqrt{2\pi \sigma^2}}. \tag{9.90}\]

\[
r^2(t) = [V + x(t)]^2 + y^2(t) = x'^2(t) + y^2(t), \tag{9.91}\]

\[
\tan \theta(t) = \frac{y(t)}{V + x(t)} = \frac{y(t)}{x'(t)}. \tag{9.92}\]

\[
x'(t) = r(t) \cos \theta(t)
\]
\[
y(t) = r(t) \sin \theta(t),
\]

\[
f_{r\theta}(r, \theta) = f_{x'y'}(x', y) = f_{x'}(x') f_y(y) = \frac{e^{-[(x'-V)^2+y^2]/2\sigma^2}}{2\pi \sigma^2}
= \frac{e^{-A^2/2\sigma^2} e^{-(r^2-2r \cos \theta)/2\sigma^2}}{2\pi \sigma^2}
\tag{9.93}\]
Rewrite in terms of the modified Bessel function of zero order:

\[
f_r(r) = \frac{r}{\sigma^2} e^{-(V^2+r^2)/2\sigma^2} I_0 \left( \frac{Vr}{\sigma^2} \right), \tag{9.95}
\]

The peak of (9.95) occurs when \( r = V \).

Note that the rician distribution reduces to the Rayleigh pdf when \( V = 0 \), as expected; it can be shown that the rician pdf approaches a Gaussian distribution when the argument is large.
Rician probability distribution

The probability of error when sending $m(t) = V$

$$P_{e}^{(1)} = P\{r(t < r_0)\} = \int_{r=0}^{r_0} f_r(r) \, dr = \int_{r=0}^{r_0} \frac{r}{\sigma^2} e^{-\frac{V^2+r^2}{2\sigma^2}} I_0\left(\frac{Vr}{\sigma^2}\right) \, dr, \quad (9.96)$$

The integral in (9.96) cannot be evaluated in closed form, but must be calculated numerically.

The overall probability of error for envelope detected ASK is

$$P_e = P_0 P_{e}^{(0)} + P_1 P_{e}^{(1)}. \quad (9.97)$$

The bit error probabilities $P_{e}^{(0)}$ and $P_{e}^{(1)}$ depend on the detection threshold $r_0$, as well as the ratio $V/\sigma$; these expressions can be written in terms of the (peak) bit energy-to-noise ratio using the relation that $V^2/\sigma^2 = 2E_b/n_0$. 
the usual case where $P_0 = P_1 = 0.5$, the probability of error will be minimized when the detection threshold is chosen at the intersection point of the probability distribution functions for $P_e^{(0)}$ and $P_e^{(1)}$ (see Problem 9.20). In general, the optimum threshold is a function of $V$ and $\sigma$, and must be found numerically, but for the case where the input SNR is very large it can be shown that an approximate value is $r_0 = V/2$. The fact that the threshold depends on the received signal level is a disadvantage of ASK, particularly when fading is present.

**FIGURE 9.20** Envelope detection of ASK. The optimum detection threshold occurs at the intersection of the Rayleigh and rician pdf.
Derivation of the probability of error for noncoherent FSK is similar to the procedure used for noncoherent ASK, since the two channels of the FSK demodulator of Figure 9.16a essentially decompose the FSK signal into two ASK signals.

If frequency $\omega_1$ is transmitted for a "1",

$$r_1(t) = V + n_1(t), \quad (9.98a)$$

$$r_2(t) = n_2(t). \quad (9.98b)$$

The probability of error when a "1" is sent can be written as

$$P_e^{(1)} = P\{r_1(t) - r_2(t) < 0\} = P\{r_2(t) > r_1(t)\}. \quad (9.99)$$
In this expression both \( r_1(t) \) and \( r_2(t) \) are random variables. Since \( r_1(t) \) consists of signal plus noise, it has a rician pdf; \( r_2(t) \) has a Rayleigh pdf since it consists of noise only.

If we temporarily assume a fixed value of \( r_1(t) \), the probability of error can be computed as

\[
\int_{r_2=r_1}^{\infty} f_{r_2}(r_2) \, dr_2.
\]

The overall probability of error:

\[
P_e^{(1)} = \int_{r_1=0}^{\infty} f_{r_1}(r_1) \int_{r_2=r_1}^{\infty} f_{r_2}(r_2) \, dr_2 \, dr_1
\]

\[
= \int_{r_1=0}^{\infty} \frac{r_1}{\sigma^2} e^{-\left(V^2 + r_1^2\right)/2\sigma^2} I_0 \left( \frac{V r_1}{\sigma^2} \right) \int_{r_2=r_1}^{\infty} \frac{r_2}{\sigma^2} e^{-r_2^2/2\sigma^2} \, dr_2 \, dr_1
\]

(9.100)
The inner integral can be evaluated directly as

$$\int_{r_2=r_1}^{\infty} \frac{r_2}{\sigma^2} e^{-r_2^2/2\sigma^2} \, dr_2 = e^{-r_1^2/2\sigma^2},$$

$$P_e^{(1)} = \int_{r_1=0}^{\infty} \frac{r_1}{\sigma^2} e^{-V^2/2\sigma^2} e^{-r_1^2/2\sigma^2} I_0 \left( \frac{V r_1}{\sigma^2} \right) \, dr_1. \quad (9.101)$$

Now use the change of variable $x = \sqrt{2} r_1$, which gives

$$P_e^{(1)} = \frac{1}{2} e^{-V^2/4\sigma^2} \int_{x=0}^{\infty} \frac{x}{\sigma^2} e^{-(V^2/2+x^2)/2\sigma^2} I_0 \left( \frac{Vx}{\sqrt{2}\sigma^2} \right) \, dx. \quad (9.102)$$

The integrand of (9.102) is identical to the rician pdf of (9.95), if $V^2$ is replaced with $V^2/2$. Thus the integral in (9.102) must be unity. This gives the final result that

$$P_e^{(1)} = P_e^{(0)} = \frac{1}{2} e^{-V^2/4\sigma^2} = \frac{1}{2} e^{-E_b/2n_0}. \quad (9.103)$$
Bit rate and bandwidth efficiency

$E_b/n_0$, the ratio of bit energy to noise power spectral density. The dimensions of $E_b$ are W·sec, while the dimensions of $n_0$ are W/Hz, so the ratio is dimensionless. In practice it is more convenient to write the bit energy in terms of signal power and the data rate.

**Signal power:** $S = R_b \times E_b$

Let $R_b$ be the bit rate of the binary message signal, with dimensions of bits per second (bps).

$$\frac{E_b}{n_0} = \frac{S}{n_0 R_b}.$$  \hspace{1cm} (9.104)

Thus the probability of error is determined solely by the carrier power, the bit rate, and the PSD of the input noise.
Bit rate and bandwidth efficiency

We can also express the bit energy-to-noise density ratio in terms of the SNR of the receiver.

If the receiver has an IF bandwidth $\Delta f$, then the noise power is $N = (2\Delta f)(n_0/2) = n_0\Delta f$, for noise with a two-sided PSD of $n_0/2$. Then we have

$$\frac{E_b}{n_0} = \frac{S}{N} \frac{\Delta f}{R_b}.$$  \hspace{1cm} (9.105)

Depending on the type of modulation, the required receiver bandwidth may range from one to several times the bit rate.
Bit rate and bandwidth efficiency

Each of the above binary modulation methods transmit one bit during each bit period, and are therefore said to have a bandwidth efficiency of 1 bps/Hz.

Thus, for example, an analog telephone circuit having a bandwidth of 2400 Hz (600 Hz-3 kHz) is limited to a maximum data rate of (2400 Hz)(1 bps/Hz) = 2400 bps when used with binary modulation.

We will see later in this chapter that substantially greater data rates can be achieved with multilevel modulation methods that transmit more than one bit per period. Such methods have bandwidth efficiencies greater than 1 bps/Hz.

Thus the terms data rate and bandwidth are not synonymous, in spite of the often-heard (but incorrect) use of the word "bandwidth" when referring to high data rates.
Comparison of ASK, FSK, and PSK systems

**FIGURE 9.21** Comparison of bit error rates for coherent (ASK, FSK, and PSK) and noncoherent (ASK and FSK) demodulation. The results for ASK are based on peak bit energy.
Comparison of ASK, FSK, and PSK systems

Note that the lowest error rate occurs for the coherent detection of PSK, but the price paid for this performance is the need for a synchronized LO and a relatively wide signal bandwidth.

Next in order of performance is coherent FSK, which requires 3-4 dB more power than PSK for the same error rate. The error rate for noncoherent (envelope detected) FSK is only slightly worse (about 1 dB) than for coherent FSK, and is better than coherent ASK.

This conclusion only applies, however, when using the peak power for ASK, when the average power of the ASK signal is used the error rate for coherent ASK and coherent FSK are identical.

The worst performance is obtained with noncoherent ASK.

<table>
<thead>
<tr>
<th>Modulation Method</th>
<th>$E_b/n_0$ for $P_e = 10^{-5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>coherent PSK</td>
<td>9.6 dB</td>
</tr>
<tr>
<td>coherent FSK</td>
<td>12.6 dB</td>
</tr>
<tr>
<td>noncoherent FSK</td>
<td>13.4 dB</td>
</tr>
<tr>
<td>coherent ASK</td>
<td>15.6 dB</td>
</tr>
<tr>
<td>noncoherent ASK</td>
<td>16.4 dB</td>
</tr>
</tbody>
</table>
Comparison of ASK, FSK, and PSK systems

The error rate that is required in practice depends on the application, as well as tradeoffs between bandwidth, transmitter power, range, and receiver complexity and cost. It may range from $10^{-2}$ (image data from early space probes) to $10^{-8}$ for high-data rate computer networks.

ASK transmitters are very simple, and efficient since no power is radiated when no data is being sent. ASK receivers are also simple, if envelope detection is used. Because the bit error rate is poor in comparison to other modulation methods, ASK is limited to low data rates. In addition, the fact that the detection threshold depends on the received signal level makes ASK performance very poor in a fading environment. For these reasons, ASK applications are usually limited to short-range, low-cost telemetry and RFID.

An FSK transmitter is only slightly more complicated than for ASK, and an FSK receiver using envelope detection can be made simply and inexpensively.

Noncoherent FSK has found widespread historical application in a wide variety of both baseband and modulated data transmission systems, such as data modems, teletype, and fax.
PSK also requires a relatively wide signal spectrum, typically ranging from $2R_b$ to $4R_b$. For these reasons, applications of PSK are generally limited to high performance systems such as space and satellite communications.
EXAMPLE 9.3  COMPARISON OF FSK AND ASK MODULATION

A wireless local area network operating at 2.44 GHz transmits data at the rate of 1.15 Mbps, with a desired error rate of $10^{-5}$. The transmitter and receiver each use a monopole antenna with a gain of 4.5 dB. The transmit power is 0.5 W, and the receiver and receive antenna have a combined system noise temperature of 600 K. Compare the maximum possible operating range for coherent FSK and noncoherent ASK modulation.

Solution

From Table 9.2, $E_b/n_0 = 12.6$ dB = 18.2 for coherent FSK with $P_e = 10^{-5}$. The wavelength is

$$\lambda = \frac{300}{2440} = 0.123 \text{ m.}$$
Using (9.104) to find the required receive carrier power gives

\[ P_r = \frac{E_b}{n_0} (n_0 R_b) = \frac{E_b}{n_0} k T_{sys} R_b = (18.2)(1.38 \times 10^{-23})(600)(1.15 \times 10^6) \]
\[ = 1.73 \times 10^{-13} \text{ W} \]

Then the Friis formula gives the range as

\[ R = \sqrt{\frac{P_t G_t G_r \lambda^2}{(4\pi)^2 P_r}} = \sqrt{\frac{(0.5)(2.82)^2(0.123)^2}{(4\pi)^2(1.73 \times 10^{-13})}} = 47 \text{ km. (coherent FSK)} \]

For the case of noncoherent ASK, \( E_b/n_0 = 16.4 \text{ dB} = 43.6 \) for \( P_e = 10^{-5} \). The required receive carrier power is now

\[ P_r = \frac{E_b}{n_0} (n_0 R_b) = \frac{E_b}{n_0} k T_{sys} R_b = (43.6)(1.38 \times 10^{-23})(600)(1.15 \times 10^6) \]
\[ = 4.15 \times 10^{-13} \text{ W} \]

Then the maximum range is

\[ R = \sqrt{\frac{P_t G_t G_r \lambda^2}{(4\pi)^2 P_r}} = \sqrt{\frac{(0.5)(2.82)^2(0.123)^2}{(4\pi)^2(4.15 \times 10^{-13})}} = 30 \text{ km. (noncoherent ASK)} \]

We see that the use of noncoherent ASK modulation leads to a reduction of about 25% in maximum range, compared with coherent FSK. Because of less than ideal propagation conditions, such as blockage, attenuation, and fading effects, these ranges would not likely be achieved in practice.
9.4 Effect of Rayleigh Fading of Bit Error Rates

The presence of multiple signal paths between transmitter and receiver lead to Rayleigh fading.

This causes sharp reductions in received signal power over short time intervals, or with movement over small distances, due to phase cancellation.

Although the average receive power is not greatly reduced, fading has a major impact on error rates because the presence of brief but deep drops in received power results in very large error rates for short periods.

\[ s_o(t) = r s_i(t) + n(t) \]

**FIGURE 9.22** Model for a propagation channel having additive white noise and fading.

tive noise. The input signal is \( s_i(t) \), and the multiplicative factor \( r(t) \), with \( 0 \leq r(t) < \infty \), represents the change in amplitude due to Rayleigh fading. White gaussian noise is added
r(t) is a random variable with a Rayleigh PDF given by

$$f_r(r) = \frac{r}{\alpha^2} e^{-r^2/2\alpha^2}, \quad (9.106)$$

The rms value of the distribution of $r(t)$ is $\sqrt{2\alpha}$.

**FIGURE 9.22** Model for a propagation channel having additive white noise and fading.
Account for Rayleigh fading by considering the conditional probability of error for a fixed signal amplitude, then integrating over the Rayleigh probability distribution function for this amplitude.

If we consider \( r(t) \) to be a fixed multiplier, then the probability of error derived in (9.78) can be modified to give

\[
P_e^{(0)}(E_b | r) = P\{n_o(T) > rVT\} = \frac{1}{2} \text{erfc}\left(\sqrt{\frac{r^2E_b}{n_0}}\right). \tag{9.107}
\]

The overall probability of error

\[
P_e^{(0)} = \int_{r=0}^{\infty} P_e^{(0)}(E_b | r) f_r(r) \, dr = \frac{1}{2} \int_{r=0}^{\infty} \text{erfc}\left(\sqrt{\frac{r^2E_b}{n_0}}\right) \frac{r}{\alpha^2} e^{-r^2/2\alpha^2} \, dr. \tag{9.108}
\]

using

\[
\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-x^2} \, dx,
\]
Define average received bit energy-to-noise power spectral density ratio of the faded received signal as:

\[ P_e^{(0)} = \frac{n_0}{4E_b\alpha^2} \int_{u=0}^{\infty} \text{erfc}(\sqrt{u})e^{-un_0/2\alpha^2E_b} du \]

\[ = \frac{n_0}{2E_b\alpha^2\sqrt{\pi}} \int_{u=0}^{\infty} \int_{x=\sqrt{u}}^{\infty} e^{-x^2} e^{-un_0/2\alpha^2E_b} dx du \quad (9.109) \]

\[ \Gamma = \frac{2\alpha^2E_b}{n_0}, \quad (9.110) \]

\[ P_e^{(0)} = \frac{1}{\sqrt{\pi\Gamma}} \int_{u=0}^{\infty} \int_{x=\sqrt{u}}^{\infty} e^{-x^2} e^{-u/\Gamma} dx du. \quad (9.111) \]
Using integration by parts:

\[ U = \int_{x=\sqrt{u}}^{\infty} e^{-x^2} \, dx \quad \text{and} \quad dV = e^{-u/\Gamma} \, du \]

\[ \int U \, dV = UV - \int V \, dU \]

\[ P_e^{(0)} = \frac{1}{\sqrt{\pi} \Gamma} \left\{ - \Gamma e^{-u/\Gamma} \int_{x=\sqrt{u}}^{\infty} e^{-x^2} \, dx \bigg|_{u=0}^{\infty} - \Gamma \int_{u=0}^{\infty} \frac{e^{-u}}{2\sqrt{u}} e^{-u/\Gamma} \, du \right\} \]

\[ = \frac{1}{\sqrt{\pi} \Gamma} \left\{ \frac{\sqrt{\pi} \Gamma}{2} - \Gamma \int_{u=0}^{\infty} \frac{e^{-(1+1/\Gamma)u}}{2\sqrt{u}} \, du \right\} \]

\[ = \frac{1}{2} \left[ 1 - \sqrt{\frac{\Gamma}{1+\Gamma}} \right] \quad (9.112) \]

energy-to-noise ratio. While the probability of error for large \( E_b/n_0 \) for the nonfaded case decreases exponentially (see Problem 9.17), (9.113) decreases much more slowly. Also
The derivation is easier because the non-faded expression for probability of error in (9.113) is simpler.

\[
P_e(E_b | r) = \frac{1}{2} e^{-r^2 E_b / n_0}.
\]

(9.113)

\[
P_e = \int_{r=0}^{\infty} P_e(E_b | r) f_r(r) \, dr = \frac{1}{2} \int_{r=0}^{\infty} \frac{r}{\alpha^2} e^{-r^2 E_b / 2n_0} e^{-r^2 / 2\alpha^2} \, dr
\]

\[
= \frac{1}{2 \left( 1 + \frac{\alpha^2 E_b}{n_0} \right)} = \frac{1}{2 + \Gamma}
\]

(9.114)

The required integral in (9.114) is listed in Appendix B.
Comparison of faded and nonfaded error rates

**FIGURE 9.23** The effect of Rayleigh fading on the bit error rates of coherent PSK and noncoherent FSK.

For the nonfaded cases we let $2\alpha^2 = 1$. 
Note the dramatic increase in the probability of error when fading is present. For example, for $P_e = 10^{-5}$, fading has the effect of increasing the required bit energy-to-noise ratio by approximately 30 dB for both PSK and FSK. Since transmit power is usually limited by regulation, fading has the ultimate effect of reducing the usable range of a wireless system, for a given data rate, error rate, and noise level. Fortunately, because most of the errors on a fading channel occur in short bursts, error-correcting codes can be used very effectively to improve the error rate.
EXAMPLE 9.4  EFFECT OF RAYLEIGH FADING

A cellular phone base station using BPSK transmits at 882 MHz with an EIRP of 50 W, and an antenna height of 50 m. The mobile receive antenna is at a height of 2 m, with a gain of $-1.0$ dBi and a noise temperature of 300 K. Find the maximum operating range if the required error rate is $10^{-5}$, for a nonfaded and a flat Rayleigh faded channel with $2\alpha^2 = -10$ dB. Assume a data rate of 24.3 kbps, and a receiver noise figure of 8 dB.

Solution
The equivalent noise temperature of the receive antenna and the receiver is

$$T_{sys} = T_A + (F - 1)T_0 = 300 + (6.3 - 1)(290) = 1837 \text{ K}$$

since the receiver noise figure is $F = 8$ dB = 6.3.

Without fading, $E_b/n_0 = 9.6 \text{ dB} = 9.12$ for PSK with $P_e = 10^{-5}$. So from (9.104) the required receiver input power is

$$P_r = \frac{E_b}{n_0} (n_0 R_b) = \frac{E_b}{n_0} (kT_{sys} R_b) = (9.12)(1.38 \times 10^{-23})(1837)(24.3 \times 10^3)$$

$$= 5.6 \times 10^{-15} \text{ W}$$
Since the propagation channel involves multipath scattering, we will use the link formula for ground reflections to include a realistic path loss factor. So from (3.58) we calculate the maximum range as

\[ R = \left[ \frac{P_t G_t G_r h_1^2 h_2^2}{P_r} \right]^{1/4} = \left[ \frac{(50)(0.8)(50)^2(2)^2}{5.6 \times 10^{-15}} \right]^{1/4} = 9.2 \times 10^4 \text{ m}. \]

For the faded case, (9.112) gives \( \Gamma = 44 \text{ dB} \) for \( P_e = 10^{-5} \). Since \( 2\alpha^2 = -10 \text{ dB} \), we have \( E_b/n_0 = 44 + 10 \text{ dB} = 54 \text{ dB} = 2.51 \times 10^5 \). Then the receiver input power is

\[ P_r = (2.51 \times 10^5)(1.38 \times 10^{-23})(1837)(24.3 \times 10^3) \]
\[ = 1.55 \times 10^{-10} \text{ W} \]

and the maximum range is now

\[ R = \left[ \frac{(50)(0.8)(50)^2(2)^2}{1.55 \times 10^{-10}} \right]^{1/4} = 7.1 \times 10^3 \text{ m}, \]

which is reduced by a factor of more than 10 from the nonfaded result.
9.5 M-ary digital modulation

Here we consider the more general case of M-ary modulation methods, where more than one bit may be transmitted per signaling interval. This allows greater bit rates for the same bandwidth, at the expense of a more complex system.

If we transmit $M = 2^n$ symbols for each signaling interval, a bandwidth efficiency of $n$ bps/Hz can be achieved. This can be done by using multiple discrete amplitude levels, or with multiple phase states, or with a combination of these. If the symbol rate is $R_s$, then the effective bit rate is $R_b = nR_s$. In the binary case we have $n = 1$, with $M = 2$ symbols.

**FIGURE 9.24** Example of an M-ary signal with $n = 5$. There are $M = 2^5 = 32$ symbols, each consisting of 5 bits. The symbol period is $T = 5/R_b$, and the bit period is $T/5 = 1/R_b$. 
Quadrature phase shift keying

If we use four states, where \( n = 2 \) and \( M = 4 \), then we can transmit two bits, or four symbols, for each signaling interval. This is called quadrature phase shift keying (QPSK).

\[
\begin{align*}
    s_0(t) &= A \cos(\omega_0 t + 45^\circ) \\
    s_1(t) &= A \cos(\omega_0 t + 135^\circ) \\
    s_2(t) &= A \cos(\omega_0 t - 135^\circ) \\
    s_3(t) &= A \cos(\omega_0 t - 45^\circ)
\end{align*}
\] (9.115a, b, c, d)

\[
s_i(t) = A \cos(\omega_0 t + \phi_i), \quad \phi_i = (2i + 1) \frac{\pi}{4}, \quad \text{for } i = 0, 1, 2, 3.
\] (9.116)

The horizontal axis represents the in-phase (I) component, while the vertical axis represents the quadrature (Q), component that is shifted 90° in phase.

![Phase states for QPSK modulation. (a) Phasor representation. (b) Constellation diagram.](image)
The In-phase and Quadrature components of the QPSK signal

\[ s_i(t) = A_I \cos \omega_0 t + A_Q \sin \omega_0 t, \]

\[ (9.117) \]

<table>
<thead>
<tr>
<th>( i )</th>
<th>( \phi_i )</th>
<th>( A_I )</th>
<th>( A_Q )</th>
<th>Binary Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>45°</td>
<td>1</td>
<td>1</td>
<td>1,1</td>
</tr>
<tr>
<td>1</td>
<td>135°</td>
<td>-1</td>
<td>1</td>
<td>0,1</td>
</tr>
<tr>
<td>2</td>
<td>-135°</td>
<td>-1</td>
<td>-1</td>
<td>0,0</td>
</tr>
<tr>
<td>3</td>
<td>-45°</td>
<td>1</td>
<td>-1</td>
<td>1,0</td>
</tr>
</tbody>
</table>
Modulator

Each QPSK phase state can then be used to represent two bits of data, as shown in the above table. The incoming serial binary data stream must be multiplexed into groups of two bits, which separately modulate the I and Q components of the carrier.

The QPSK modulator thus takes the form shown in Figure 9.26. A 90° hybrid divider can be used to provide the I and Q components of the LO from a single oscillator.

Because the average transition between phase states is 90°, the bandwidth of the QPSK spectrum is narrower than the spectrum of a BPSK signal.

**FIGURE 9.26** Block diagram of a QPSK modulator.
FIGURE 9.27  Data and timing diagrams for the QPSK modulator.
Demodulation of QPSK requires **coherent detection**. The demodulator uses two mixers with quadrature LO components to recover the I and Q signals. PCM detectors, each consisting of an integrator and a sampler, are used to detect the I and Q NRZ data, which is then decoded with a parallel-to-serial converter to provide serial binary output.

Since QPSK is a constant envelope modulation, the detectors can use a zero threshold.

**FIGURE 9.28** Coherent QPSK demodulator.
Demodulator

\[ V_I = \int_{t=0}^{T} A^2 \cos(\omega_0 t + \phi_i) \cos \omega_0 t \, dt \]
\[ = \frac{A^2}{2} \int_{t=0}^{T} [\cos(2\omega_0 t + \phi_i) + \cos \phi_i] \, dt = \frac{A^2 T}{2} \cos \phi_i \quad (9.118a) \]

\[ V_Q = -\int_{t=0}^{T} A^2 \cos(\omega_0 t + \phi_i) \sin \omega_0 t \, dt \]
\[ = -\frac{A^2}{2} \int_{t=0}^{T} [\sin(2\omega_0 t + \phi_i) - \sin \phi_i] \, dt = \frac{A^2 T}{2} \sin \phi_i \quad (9.118b) \]

The constant \( C \) is defined as \( A^2 T/2\sqrt{2} \).

**TABLE 9.4**  I and Q Outputs of a QPSK Demodulator

<table>
<thead>
<tr>
<th>( s_i(t) )</th>
<th>( \phi_i )</th>
<th>( V_I )</th>
<th>( V_Q )</th>
<th>Binary Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_0 )</td>
<td>45°</td>
<td>C</td>
<td>C</td>
<td>1,1</td>
</tr>
<tr>
<td>( s_1 )</td>
<td>135°</td>
<td>-C</td>
<td>C</td>
<td>0,1</td>
</tr>
<tr>
<td>( s_2 )</td>
<td>-135°</td>
<td>-C</td>
<td>-C</td>
<td>0,0</td>
</tr>
<tr>
<td>( s_3 )</td>
<td>-45°</td>
<td>C</td>
<td>-C</td>
<td>1,0</td>
</tr>
</tbody>
</table>
The assignment of bits pairs to symbols given in Table 9.3 is not unique.

But the above choice has the very useful property that when an error occurs in the detection of a symbol, it is most likely that only one of the bits will be in error, rather than both bits.

This is because an error is most likely to result in a shift from the correct phase state to the immediately adjacent phase, rather than the diametrically opposite phase.

**FIGURE 9.25** Phase states for QPSK modulation. (a) Phasor representation. (b) Constellation diagram.
By symmetry the error rates for all $s_i$ are equal, so let us assume that $s_1$ is received

$$s(t) + n(t) = s_1(t) + n(t) = A \cos(\omega_0 t + 135^\circ) + n(t). \quad (9.119)$$

$$V_I(T) = \frac{-A^2 T}{2\sqrt{2}} + n_o(T) = \frac{-E_s}{\sqrt{2}} + n_o(T) \quad (9.120a)$$

$$V_Q(T) = \frac{A^2 T}{2\sqrt{2}} + n_o(T) = \frac{E_s}{\sqrt{2}} + n_o(T) \quad (9.120b)$$

where $E_s = \frac{A^2 T}{2}$ is the symbol energy.

$$n_o(T) = \int_{t=0}^{\infty} A n(t) \cos \omega_0 t \, dt, \quad (9.121)$$

which is a gaussian random variable.
We can find the variance of $n_o(T)$ as follows:

\[
\sigma^2 = E\left\{n_o^2(T)\right\} = E\left\{ \int_{t=0}^{T} \int_{s=0}^{T} A^2 n(t)n(s) \cos \omega_0 t \cos \omega_0 s \, dt \, ds \right\}
\]

\[
= A^2 \int_{t=0}^{T} \int_{s=0}^{T} E\{n(t)n(s)\} \cos \omega_0 t \cos \omega_0 s \, dt \, ds
\]

\[
= \frac{A^2 n_0}{2} \int_{t=0}^{T} \int_{s=0}^{T} \delta(t-s) \cos \omega_0 t \cos \omega_0 s \, dt \, ds
\]

\[
= \frac{A^2 n_0}{2} \int_{s=0}^{T} \cos^2 \omega_0 s \, ds = \frac{A^2 T n_0}{4} = \frac{E_s n_0}{2}
\]

(9.122)

The detection threshold for both detectors is zero, so the probability of a symbol error at the I correlator is

\[
P_{\text{e}(I)} = P\{V_I > 0\} = p\left\{n_o(T) > \frac{E_s}{\sqrt{2}}\right\}
\]

\[
= \int_{E_s/\sqrt{2}}^{\infty} \frac{e^{-n_0^2/2\sigma^2}}{\sqrt{2\pi \sigma^2}} \, dn_0 = \frac{1}{\sqrt{\pi}} \int_{E_s/2\sigma}^{\infty} e^{-x^2} \, dx
\]

\[
= \frac{1}{2} \text{erfc}\left(\frac{E_s}{2\sigma}\right) = \frac{1}{2} \text{erfc}\left(\sqrt{\frac{E_s}{2n_0}}\right)
\]

(9.123a)
By symmetry we will also obtain the same expression for the probability of a symbol error at the Q correlator:

$$P_e^{(Q)} = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{E_s}{2n_0}} \right).$$  \hspace{1cm} \text{(9.123b)}

Since the overall probability that a symbol is received correctly is the product of the probabilities that each correlator operates correctly, the overall probability of error for a symbol is

$$P_e^{(s)} = 1 - (1 - P_e^{(I)})(1 - P_e^{(Q)}) = 1 - \left[-2P_e^{(I)} + (P_e^{(I)})^2\right]$$

$$\approx 2P_e^{(I)} = \text{erfc} \left( \sqrt{\frac{E_s}{2n_0}} \right)$$  \hspace{1cm} \text{(9.124)}

where we made the approximation that $P_e \ll 1$.

Twice data rate at same bandwidth and error rate as BPSK
If, as is the usual case in practice, we use QPSK with Gray coding, then we can assume that a symbol error is most likely to cause only a single bit error.

Then since each symbol contains two bits, the bit error rate for QPSK will be one-half the symbol error rate:

$$P_e^{(e)} = \frac{1}{2} \text{erfc} \left( \frac{\sqrt{E_s}}{\sqrt{2}n_0} \right).$$  \hspace{1cm} (9.125)

Note that because the symbol period $T$ is twice the bit period, $E_s = 2E_b$, where $E_b$ is the bit energy.

This shows that the expression of (9.125) is equivalent to the probability of error for BPSK as given in (9.78).
EXAMPLE 9.5 COMPARISON OF BPSK AND QPSK MODULATION

A QPSK system transmits data at 20 Mbps over a radio link with an average transmit power of 2 W. The total link loss is 110 dB, and the two-sided power spectral density of the noise is $n_0/2 = 4 \times 10^{-20}$ W/Hz. If Gray coding is used, find the bit error rate, and compare with a similar system using BPSK.

Solution

The receive power is $P_r = 2 \times 10^{-110/10} = 2 \times 10^{-11}$ W. For QPSK the symbol energy-to-noise ratio is

$$\frac{E_s}{n_0} = \frac{P_r}{n_0 R_s} = \frac{2P_r}{n_0 R_b} = \frac{2(2 \times 10^{-11})}{(8 \times 10^{-20})(20 \times 10^6)} = 25.$$

The probability of a bit error is given by (9.125)

$$P_e = \frac{1}{2} \text{erfc} \left( \frac{\sqrt{E_s}}{2n_0} \right) = \frac{1}{2} \text{erfc} \left( \frac{\sqrt{25}}{2} \right) = \frac{1}{2} \text{erfc}(3.535) = 2.87 \times 10^{-7}.$$

For BPSK the bit energy is

$$\frac{E_b}{n_0} = \frac{P_r}{n_0 R_b} = \frac{2 \times 10^{-11}}{(8 \times 10^{-20})(20 \times 10^6)} = 12.5,$$

and the probability of a bit error is given by (9.78):

$$P_e = \frac{1}{2} \text{erfc} \left( \frac{\sqrt{E_b}}{n_0} \right) = \frac{1}{2} \text{erfc}(\sqrt{12.5}) = \frac{1}{2} \text{erfc}(3.535) = 2.87 \times 10^{-7}.$$

As expected, the bit error rates for QPSK and BPSK are identical.
M-ary phase shift keying

\[ s_i(t) = A \cos(\omega_0 t + \phi_i), \quad (9.126) \]

where

\[ \phi_i = \frac{2\pi i}{M}, \text{ for } i = 0, 1, 2, \ldots M - 1. \]

It is straightforward to show that the probability of a symbol error for M-PSK is approximately given by

\[ P_e^{(s)} = \text{erfc} \left( \sqrt{\frac{E_s}{n_0} \sin^2 \frac{\pi}{M}} \right). \]

Valid for \( M > 2 \) and \( P_e < 10^{-3} \)

If Gray coding is used:

\[ P_e = \frac{1}{n} P_e^{(s)} = \frac{1}{n} \text{erfc} \left( \sqrt{\frac{E_s}{n_0} \sin^2 \frac{\pi}{M}} \right), \quad (9.128) \]
Quadrature amplitude modulation (QAM)

We can further generalize M-ary modulation by allowing the amplitudes of the I and Q components of the carrier as given in (9.117) to vary arbitrarily.

\[ s_i(t) = a_i \cos \omega_0 t + b_i \sin \omega_0 t. \]  
\[ (a_0, b_0) = \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \]
\[ (a_1, b_1) = \left( \frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \]
\[ (a_2, b_2) = \left( \frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} \right) \]
\[ (a_3, b_3) = \left( \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}} \right) \]

FIGURE 9.30  Constellation diagram for 16-QAM.

We can therefore say that 4-QAM is identical to QPSK, if the signal amplitudes are constant.
It can be shown that an approximate expression for the symbol error rate of 16-QAM is given by

\[
P_e = \frac{3}{2} \text{erfc} \left( \sqrt{\frac{2E_b}{5n_0}} \right).
\]  

(9.130)

Because of the high bandwidth efficiency that can be obtained with QAM, it is increasingly being used in modem wireless systems, including point-to-point microwave radios, LMDS systems, and the DVB-C digital video cable broadcasting system.

Table 9.5 summarizes the ideal performance of several types of coherent digital modulation methods.
### TABLE 9.5  Summary of Performance of Various Digital Modulation Methods

<table>
<thead>
<tr>
<th>Modulation Type</th>
<th>$E_b/n_0$ (dB) for $P_e = 10^{-5}$</th>
<th>Bandwidth Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>binary ASK</td>
<td>15.6</td>
<td>1</td>
</tr>
<tr>
<td>binary FSK</td>
<td>12.6</td>
<td>1</td>
</tr>
<tr>
<td>binary PSK</td>
<td>9.6</td>
<td>1</td>
</tr>
<tr>
<td>QPSK (4-QAM)</td>
<td>9.6</td>
<td>2</td>
</tr>
<tr>
<td>8-PSK</td>
<td>13.0</td>
<td>3</td>
</tr>
<tr>
<td>16-PSK</td>
<td>18.7</td>
<td>4</td>
</tr>
<tr>
<td>16-QAM</td>
<td>13.4</td>
<td>4</td>
</tr>
<tr>
<td>64-QAM</td>
<td>17.8</td>
<td>6</td>
</tr>
</tbody>
</table>
Channel capacity

It is possible to achieve as low an error rate as is desired once $E_b/n_0$ is above a critical value.

Since $E_b = S/R_b$, this also implies that, for a fixed signal power $S$, there is a critical value of the data rate $R_b$ for which the error rate can be made as small as desired.

This particular value is called the channel capacity, and is given by a formula derived by Claude Shannon:

Shannon limit:

$$C = B \log_2 \left(1 + \frac{S}{n_0B}\right),$$  \hspace{1cm} (9.131)

$C$: maximum data rate in bps
$B$: bandwidth in Hz
$S$: signal power in W
$n_0/2$: two-sided noise power
The Shannon channel capacity formula gives the upper bound on the maximum data rate that can be achieved for a given channel in the presence of additive Gaussian noise.

Practical modulation methods usually perform at only a fraction of this value, but the use of error correcting codes can provide performance close to the Shannon limit.