Thus the minimum tone spacing for *coherent* FSK signaling occurs for $n = 1$ as follows:

$$f_1 - f_2 = \frac{1}{2T} \tag{4.56}$$

Therefore, for the same symbol rate, coherently detected FSK can occupy less bandwidth than noncoherently detected FSK and still retain orthogonal signaling. We can say that coherent FSK is more *bandwidth efficient.* (The subject of bandwidth efficiency is addressed in greater detail in Chapter 9.)

The required tone spacings are now closer than in part (a) because when we align two periodic waveforms so that their starting phases are the same, we achieve orthogonality by virtue of an even-versus-odd symmetry in the respective waveforms over one symbol time. This is unlike the way orthogonality was achieved in part (a), where we paid no attention to phase. In the coherent case here, the phase alignment in the correlator stages means that we can bring the tones closer together in frequency and still maintain orthogonality among the set of FSK tones. Prove it to yourself by plotting two sine waves (or cosine waves or square waves). Start them off at the same phase (0 radians is convenient). Using quadrille paper, choose a simple time scale to represent one symbol time T. Then, plot a tone at one cycle per T. Below that, with the same starting phase, plot another tone at one-and-a-half cycles per T. Perform product-summation over period T, and verify that these waveforms (spaced $1/2T$ hertz apart) are indeed orthogonal.

4.6 COMPLEX ENVELOPE

The description of real-world modulators and demodulators is facilitated by the use of complex notation which began in Section 4.2.1 and continues here. Any real bandpass waveform $s(t)$ can be represented using complex notation as

$$s(t) = \mathrm{Re}\{g(t)e^{j\omega_0 t}\} \tag{4.57}$$

where $g(t)$ is known as the *complex envelope,* expressed as

$$g(t) = x(t) + jy(t) = |g(t)|e^{j\theta(t)} = R(t)e^{j\theta(t)} \tag{4.58}$$

The magnitude of the complex envelope is then

$$R(t) = |g(t)| = \sqrt{x^2(t) + y^2(t)} \tag{4.59}$$

and its phase is

$$\theta(t) = \tan^{-1}\frac{y(t)}{x(t)} \tag{4.60}$$

With respect to Equation (4.57), we can call $g(t)$ the baseband message or data in complex form, and $e^{j\omega_0 t}$ the carrier wave in complex form. The product of these two represents modulation, and $s(t)$, the real part of this product, is the transmitted waveform. Therefore, using Equations (4.4), (4.57), and (4.58), we can express $s(t)$ as follows:

$$
\begin{aligned}
s(t) &= \mathrm{Re}\{[x(t) + jy(t)][\cos \omega_0 t + j \sin \omega_0 t]\} \\
&= x(t) \cos \omega_0 t - y(t) \sin \omega_0 t
\end{aligned} \tag{4.61}
$$

Note that the modulation of signals expressed in the general form of $(a + jb)$ times $(c + jd)$ yields a waveform with a sign inversion (at the quadrature term of the carrier wave) of the form *ac-bd*.

4.6.1 Quadrature Implementation of a Modulator

Consider an example of a baseband waveform $g(t)$, described by a sequence of ideal pulses $x(t)$ and $y(t)$ appearing at discrete times $k = 1, 2, \ldots$. Thus, $g(t)$, $x(t)$, and $y(t)$ in Equation (4.58) can be written as g_k, x_k, and y_k, respectively. Let the pulse-amplitude values be $x_k = y_k = 0.707A$. The complex envelope can then be expressed in discrete form as

$$
g_k = x_k + jy_k = 0.707A + j0.707A \tag{4.62}
$$

We know from complex algebra that $j = \sqrt{-1}$, but from a practical point of view, the j term can be regarded as a "flag," reminding us that we may not use ordinary addition in combing the terms in Equation (4.62). In preparation for the next step, inphase and quadrature modulation, x_k and y_k are treated as an ordered pair. A quadrature-type modulator is shown in Figure 4.21, where we see that the x_k pulse is multiplied by $\cos \omega_0 t$ (inphase component of the carrier wave), and the y_k pulse is multiplied by $\sin \omega_0 t$ (quadrature component of the carrier wave). The modulation process can be described succinctly as multiplying the complex envelope by $e^{j\omega_0 t}$, and then transmitting the real part of the product. Hence, we write

$$
\begin{aligned}
s(t) &= \mathrm{Re}\{g_k e^{j\omega_0 t}\} \\
&= \mathrm{Re}\{(x_k + jy_k)(\cos \omega_0 t + j \sin \omega_0 t)\} \\
&= x_k \cos \omega_0 t - y_k \sin \omega_0 t \\
&= 0.707A \cos \omega_0 t - 0.707A \sin \omega_0 t \\
&= A \cos\left(\omega_0 t + \frac{\pi}{4}\right)
\end{aligned} \tag{4.63}
$$

Again, note that the quadrature term of the carrier wave has undergone a sign inversion in the modulation process. If we use $0.707A \cos \omega_0 t$ as a reference, then the transmitted waveform $s(t)$ in Equation (4.63) leads the reference waveform by $\pi/4$.

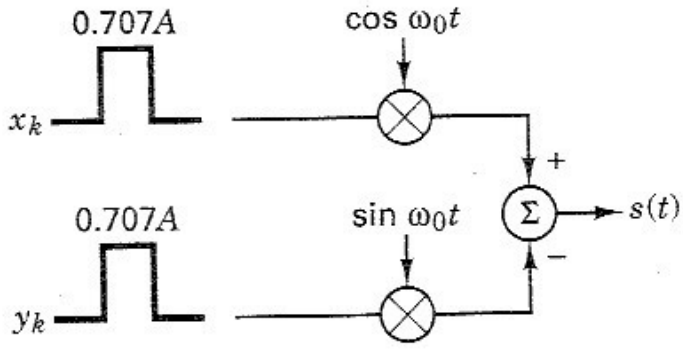


Figure 4.21 Quadrature type modulator.

Similarly, if we consider $-0.707A \sin \omega_0 t$ as the reference, then $s(t)$ in Equation (4.63) lags the reference waveform by $\pi/4$. These relationships are illustrated in Figure 4.22.

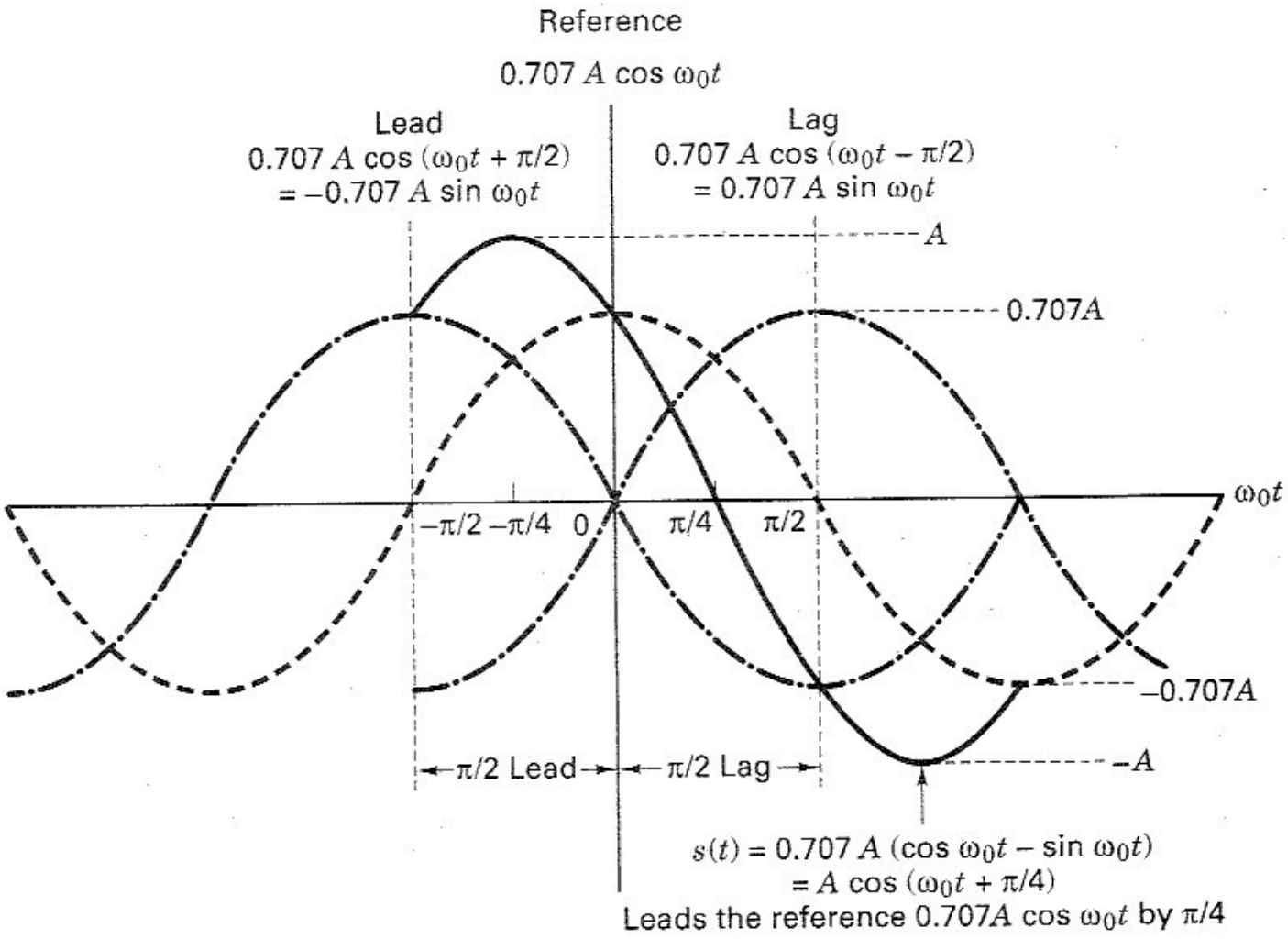


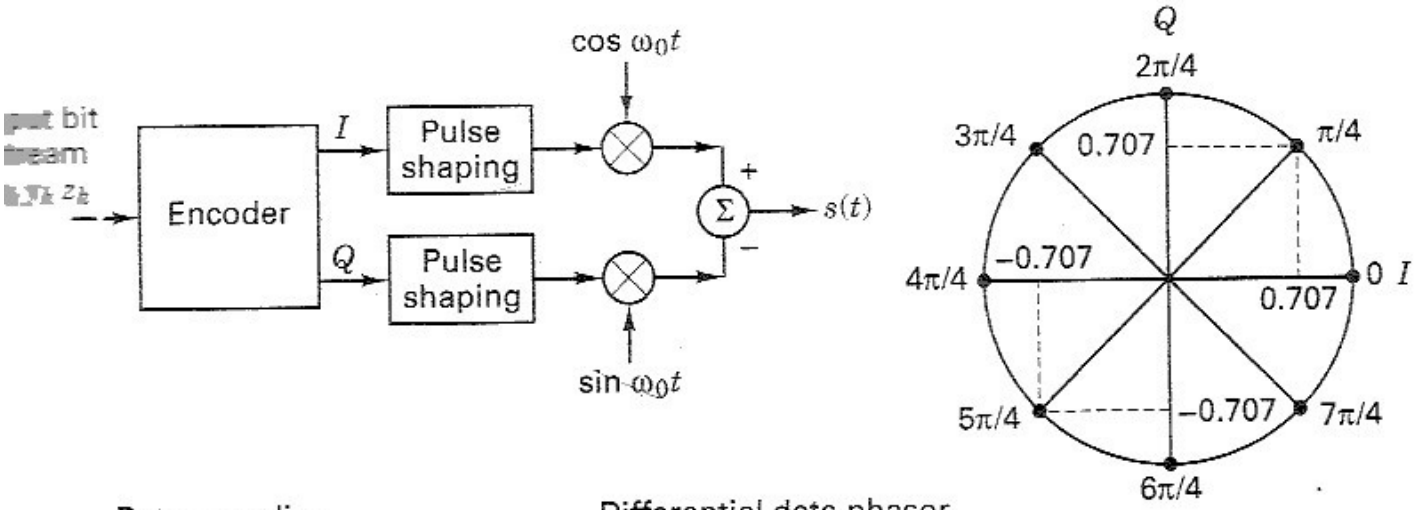
Figure 4.22 Lead/Lag relationships of sinusoids.

4.6.2 D8PSK Modulator Example

Figure 4.23 depicts a quadrature implementation of a differential 8-PSK (D8PSK) modulator. Because the modulation is 8-ary, we assign a 3-bit message (x_k, y_k, z_k) to each phase $\Delta\phi_k$. Because the modulation is differential, at each kth transmission time we send a data phasor ϕ_k, which can be expressed as

$$\phi_k = \Delta\phi_k + \phi_{k-1} \tag{4.64}$$

The process of adding the current message-to-phase assignment $\Delta\phi_k$ to the prior data phase ϕ_{k-1} provides for the differential encoding of the message. A sequence of phasors created by following Equation (4.64) yields a similar differential encoding as the procedures described in Section 4.5.2. You might notice in Figure 4.23 that the assignment of 3-bit message sequences to $\Delta\phi_k$ does not proceed along the natural binary progression from 000 to 111. There is a special code being used here called a *Gray code.* (The benefits that such a binary to M-ary assignment provides is explained in Section 4.9.4.)



Data encoding

x_k	y_k	z_k	$\Delta\phi_k$
0	0	0	0
0	0	1	$\pi/4$
0	1	1	$2\pi/4$
0	1	0	$3\pi/4$
1	1	0	$4\pi/4$
1	1	1	$5\pi/4$
1	0	1	$6\pi/4$
1	0	0	$7\pi/4$

Differential data phasor

$\phi_k = \phi_{k-1} + \Delta\phi_k$

Let $\phi_0 = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$x_k\, y_k\, z_k$	110	001	110	010
$\Delta\phi_k$:	π	$\pi/4$	π	$3\pi/4$
ϕ_k:	π	$5\pi/4$	$\pi/4$	π
I:	−1	−0.707	0.707	−1
Q:	0	−0.707	0.707	0

Figure 4.23 Quadrature implementation of a D8PSK modulator.

For the modulator in Figure 4.23, let the input data sequence at times $k = 1, 2, 3, 4$, be equal to 110, 001, 110, 010, respectively. Next, we use the date-encoding table shown in Figure 4.23 and Equation (4.64), with the starting phase at time $k = 0$ to be $\phi_0 = 0$. At time $k = 1$, the differential data phase corresponding to $x_1\, y_1\, z_1 = 110$ is $\phi_1 = 4\pi/4 = \pi$. Taking the magnitude of the rotating phasor to be unity, the inphase (I) and the quadrature (Q) baseband pulses are −1 and 0, respectively. As indicated in Figure 4.23, these pulses are generally shaped with a filter (such as a root-raised-cosine).

For time $k = 2$, the table in Figure 4.23 shows that the message 001 is assigned $\Delta\phi_2 = \pi/4$. Therefore, following Equation (4.64), the second differential data phase is $\phi_2 = \pi + \pi/4 = 5\pi/4$, and for time $k = 2$, the I and Q baseband pulses are $x_k = -0.707$ and $y_k = -0.707$, respectively. The transmitted waveform follows the form of Equation (4.61), rewritten as

$$\begin{aligned} s(t) &= \mathrm{Re}\{(x_k + jy_k)(\cos\omega_0 t + j\sin\omega_0 t)\} \\ &= x_k \cos\omega_0 t - y_k \sin\omega_0 t \end{aligned} \tag{4.65}$$

For a signaling set that can be represented on a phase-amplitude plane, such as MPSK or MQAM, Equation (4.65) provides an interesting observation. That is, quadrature implementation of the transmitter transforms all such signaling types

to simple amplitude modulation. Any phasor on the plane is transmitted by amplitude-modulating its inphase and quadrature projections onto the cosine and sine wave components of the carrier, respectively. For ease of notation, we neglect the pulse shaping; that is, we assume that the data pulses have ideal rectangular shapes. Then, using Equation (4.65), for time $k = 2$, where $x_k = -0.707$ and $y_k = -0.707$, the transmitted $s(t)$ can be written as follows:

$$\begin{aligned} s(t) &= -0.707 \cos \omega_0 t + 0.707 \sin \omega_0 t \\ &= \sin\left(\omega_0 t - \frac{\pi}{4}\right) \end{aligned} \tag{4.66}$$

4.6.3 D8PSK Demodulator Example

In the previous section, the quadrature implementation of a modulator began with multiplying the complex envelope (baseband message) by $e^{j\omega_0 t}$, and transmitting the real part of the product $s(t)$, as described in Equation (4.63). Using a similar quadrature implementation, demodulation consists of reversing the process—that is, multiplying the received bandpass waveform by $e^{-j\omega_0 t}$ in order to recover the baseband waveform. The left side of Figure 4.24 shows the modulator of Figure 4.23 in simplified form, and the waveform $s(t) = \sin(\omega_0 t - \pi/4)$ that was launched at time $k = 2$ for that example. We continue this same example in this section, and also show a quadrature implementation of the demodulator on the right side of Figure 4.24.

Notice the subtle difference between the $-\sin \omega_0 t$ term at the modulator and the $-\sin \omega_0 t$ term at the demodulator. At the modulator, the minus sign stems from taking the real part of the complex waveform (product of the complex envelope and complex carrier wave). At the demodulator, $-\sin \omega_0 t$ stems from multiplying the bandpass waveform by the conjugate $e^{-j\omega_0 t}$ of the modulator carrier wave; demodulation is coherent if phase is recovered. To simplify writing the basic relationships of the process, the noise is neglected. After the inphase multiplication by $\cos \omega_0 t$ in the demodulator, we get the signal at point A:

$$\begin{aligned} A &= (-0.707 \cos \omega_0 t + 0.707 \sin \omega_0 t) \cos \omega_0 t \\ &= -0.707 \cos^2 \omega_0 t + 0.707 \sin \omega_0 t \cos \omega_0 t \end{aligned} \tag{4.67}$$

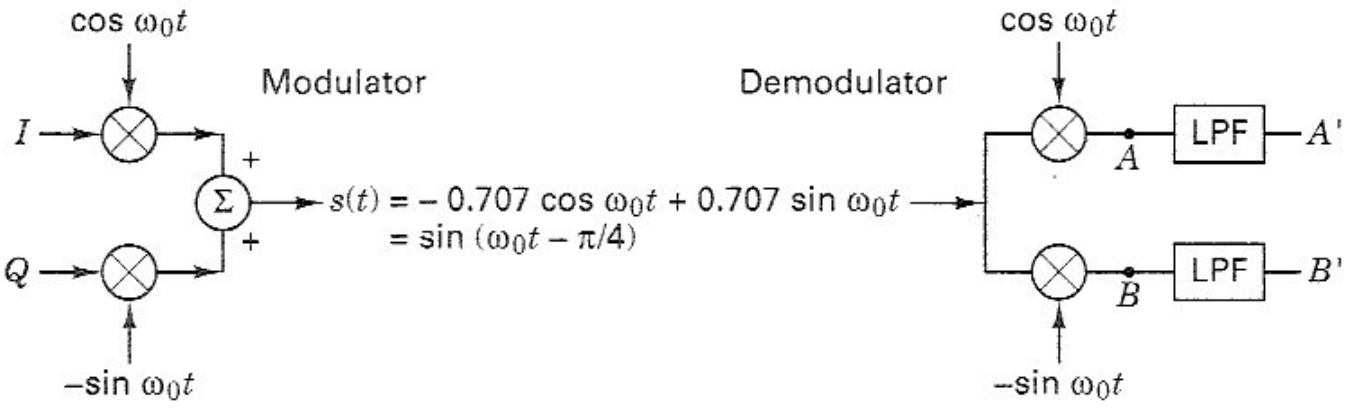


Figure 4.24 Modulator/demodulator example.

Using the trigonometric identities shown in Equation (D.7) and (D.9), we get

$$A = \frac{-0.707}{2}(1 + \cos 2\omega_0 t) + \frac{0.707}{2}\sin 2\omega_0 t \tag{4.68}$$

After filtering with a low-pass filter (LPF) we recover at point A' an ideal negative pulse, as follows:

$$A' = -0.707 \text{ (times a scale factor)} \tag{4.69}$$

Similarly, after the quadrature multiplication by $-\sin \omega_0 t$ in the demodulator, we get the signal at point B, as follows:

$$\begin{aligned} B &= (-0.707 \cos \omega_0 t + 0.707 \sin \omega_0 t)(-\sin \omega_0 t) \\ &= \frac{0.707}{2}\sin 2\omega_0 t - \frac{0.707}{2}(1 - \cos 2\omega_0 t) \end{aligned} \tag{4.70}$$

After the LPF we recover at point B' an ideal negative pulse, as follows:

$$B' = -0.707 \text{ (times a scale factor)} \tag{4.71}$$

Thus, we see from the demodulated and filtered points A' and B' that the differential (ideal) data pulses for the I and Q channels are each equal to -0.707. Since the modulator/demodulator is differential, then for this $k = 2$ example,

$$\Delta\phi_{k=2} = \phi_{k=2} - \phi_{k=1} \tag{4.72}$$

We presume that the demodulator at the earlier time $k = 1$ had properly recovered the signal phase to be π. Then, from Equation (4.72), we can write

$$\Delta\phi_{k=2} = \frac{5\pi}{4} - \pi = \frac{\pi}{4} \tag{4.73}$$

Referring back to the data-encoding table in Figure 4.23, we see that the detected data sequence is $x_2\, y_2\, z_2 = 001$, which corresponds to the data that was sent at time $k = 2$.

4.7 ERROR PERFORMANCE FOR BINARY SYSTEMS

4.7.1 Probability of Bit Error for Coherently Detected BPSK

An important measure of performance used for comparing digital modulation schemes is the probability of error, P_E. For the correlator or matched filter, the calculations for obtaining P_E can be viewed geometrically. (See Figure 4.6.) They involve finding the probability that given a particular transmitted signal vector, say $\mathbf{s}_1$, the noise vector $\mathbf{n}$ will give rise to a received signal falling outside region 1. The probability of the detector making an incorrect decision is termed the *probability of symbol error* (P_E). It is often convenient to specify system performance by the probability of bit error (P_B), even when decisions are made on the basis of symbols

for which $M > 2$. The relationship between P_B and P_E is treated in Section 4.9.3 for orthogonal signaling and in Section 4.9.4 for multiple phase signaling.

For convenience, this section is restricted to the coherent detection of BPSK modulation. For this case the symbol error probability is the bit error probability. Assume that the signals are equally likely. Also assume that when signal $s_i(t)$ $(i = 1, 2)$ is transmitted, the received signal $r(t)$ is equal to $s_i(t) + n(t)$, where $n(t)$ is an AWGN process, and any degradation effects due to channel-induced ISI or circuit-induced ISI have been neglected. The antipodal signals $s_1(t)$ and $s_2(t)$ can be characterized in a one-dimensional signal space as described in Section 4.4.1, where

$$\left.\begin{aligned} s_1(t) &= \sqrt{E}\,\psi_1(t) \\ s_2(t) &= -\sqrt{E}\,\psi_1(t) \end{aligned}\right\} 0 \le t \le T \tag{4.74}$$

and

The detector will choose the $s_i(t)$ with the largest correlator output $z_i(T)$; or, in this case of equal-energy antipodal signals, the detector, using the decision rule in Equation (4.20), decides on the basis of

$$\begin{aligned} &s_1(t) \quad \text{if } z(T) > \gamma_0 = 0 \\ &s_2(t) \quad \text{otherwise} \end{aligned} \tag{4.75}$$

and

Two types of errors can be made, as shown in Figure 4.9: The first type of error takes place if signal $s_1(t)$ is transmitted but the noise is such that the detector measures a negative value for $z(T)$ and chooses hypothesis H_2, the hypothesis that signal $s_2(t)$ was sent. The second type of error takes place if signal $s_2(t)$ is transmitted but the detector measures a positive value for $z(T)$ and chooses hypothesis H_1, the hypothesis that signal $s_1(t)$ was sent.

In Section 3.2.1.1, an expression was developed in Equation (3.42) for the probability of a bit error P_B, for this binary *minimum error* detector. We rewrite this relationship as

$$P_B = \int_{(a_1 - a_2)/2\sigma_0}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du = Q\left(\frac{a_1 - a_2}{2\sigma_0}\right) \tag{4.76}$$

where σ_0 is the standard deviation of the noise out of the correlator. The function $Q(x)$, called the *complementary error function* or *co-error function.* is defined as

$$Q(X) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{u^2}{2}\right) du \tag{4.77}$$

and is described in greater detail in Sections 3.2 and B.3.2.

For equal-energy antipodal signaling, such as the BPSK format in Equation (4.74), the receiver output signal components are $a_1 = \sqrt{E_b}$ when $s_1(t)$ is sent and $a_2 = -\sqrt{E_b}$ when $s_2(t)$ is sent, where E_b is the signal energy per binary symbol. For AWGN we can replace the noise variance σ_0^2 out of the correlator with $N_0/2$ (see Appendix C), so that we can rewrite Equation (4.76) as follows:

$$P_B = \int_{\sqrt{2E_b/N_0}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du \tag{4.78}$$

$$= Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{4.79}$$

This result for bandpass antipodal BPSK signaling is the same as the results that were developed earlier in Equation (3.70) for the matched-filter detection of antipodal signaling in general, and in Equation (3.76) for the matched-filter detection of baseband antipodal signaling in particular. This is an example of an *equivalence theorem,* described earlier. For linear systems the equivalence theorem establishes that the mathematics of detection is unaffected by a shift in frequency. Hence in this chapter, the use of matched filters or correlators in the detection of bandpass signals yields the same relationships as those developed for comparable signals at baseband.

Example 4.4 Bit Error Probability for BPSK Signaling

Find the bit error probability for a BPSK system with a bit rate of 1 Mbit/s. The received waveforms $s_1(t) = A \cos \omega_0 t$ and $s_2(t) = -A \cos \omega_0 t$, are coherently detected with a matched filter. The value of A is 10 mV. Assume that the single-sided noise power spectral density is $N_0 = 10^{-11}$ W/Hz and that signal power and energy per bit are normalized relative to a 1-Ω load.

Solution

$$A = \sqrt{\frac{2E_b}{T}} = 10^{-2}\ \text{V} \qquad T = \frac{1}{R} = 10^{-6}\ \text{s}$$

Thus,

$$E_b = \frac{A^2}{2} T = 5 \times 10^{-11}\ \text{J} \qquad \text{and} \qquad \sqrt{\frac{2E_b}{N_0}} = 3.16$$

Also,

$$P_B = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) = Q(3.16)$$

Using Table B.1 or Equation (3.44), we obtain

$$P_B = 8 \times 10^{-4}$$

4.7.2 Probability of Bit Error for Coherently Detected, Differentially Encoded Binary PSK

Channel waveforms sometimes experience inversion; for example, when using a coherent reference generated by a phase-locked loop, one may have phase ambiguity. If the carrier phase were reversed in a DPSK modulation application, what would be the effect on the message? The only effect would be an error in the bit during

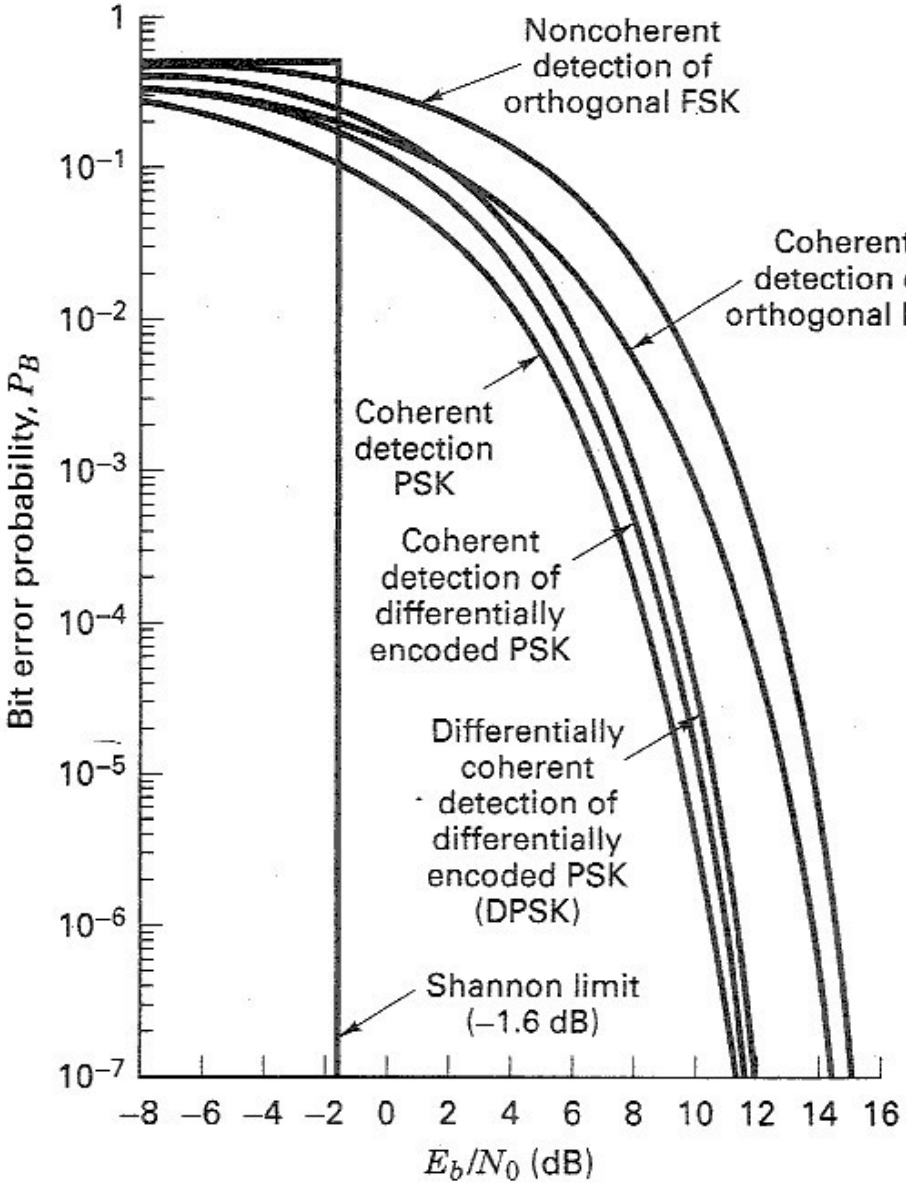


Figure 4.25 Bit error probability for several types of binary systems.

which inversion occurred or the bit just after inversion, since the message information is encoded in the similarity or difference between adjacent symbols. The similarity or difference quality remains unchanged if the carrier is inverted. Sometimes messages (and their assigned waveforms) are *differentially encoded* and *coherently detected* simply to avoid these phase ambiguities.

The probability of bit error for coherently detected, differentially encoded PSK is given by [5]

$$P_B = 2Q\left(\sqrt{\frac{2E_b}{N_0}}\right)\left[1 - Q\left(\sqrt{\frac{2E_b}{N_0}}\right)\right] \tag{4.80}$$

This relationship is plotted in Figure 4.25. Notice that there is a small degradation of error performance compared with the coherent detection of PSK. This is due to the differential encoding, since any single detection error will usually result in two decision errors. Error performance for the more popular differentially coherent detection (DPSK) is covered in Section 4.7.5.

4.7.3 Probability of Bit Error for Coherently Detected Binary Orthogonal FSK

Equations (4.78) and (4.79) describe the probability of bit error for coherent antipodal signals. A more general treatment for binary coherent signals (not limited to antipodal signals) yields the following equation for P_B [6]:

$$P_B = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{(1-\rho)E_b/N_0}}^{\infty} \exp\left(-\frac{u^2}{2}\right) du \tag{4.81}$$

From Equation (3.64b), $\rho = \cos\theta$ is the time cross-correlation coefficient between signal $s_1(t)$ and $s_2(t)$, where θ is the angle between signal vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ (see Figure 4.6). For antipodal signals such as BPSK, $\theta = \pi$, thus $\rho = -1$.

For orthogonal signals such as binary FSK (BFSK), $\theta = \pi/2$, since the $\mathbf{s}_1$ and $\mathbf{s}_2$ vectors are perpendicular to each other; thus $\rho = 0$, as can be verified with Equation (3.64a), and Equation (4.81) can then be written as

$$P_B = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{E_b/N_0}}^{\infty} \exp\left(-\frac{u^2}{2}\right) du = Q\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{4.82}$$

where the *co-error function* $Q(x)$ is described in greater detail in Sections 3.2 and B.3.2. The result in Equation (4.82) for the coherent detection of orthogonal BFSK plotted in Figure 4.25 is the same as the results that were developed earlier in Equation (3.71) for the matched-filter detection of orthogonal signaling, in general, and in Equation (3.73) for the matched-filter detection of baseband orthogonal signaling (unipolar pulses), in particular. The details of *on-off keying* (OOK) are not treated in this book. However, it is worth noting that on-off keying is an orthogonal signaling set (unipolar pulse signaling is the baseband equivalent of OOK). Thus the relationship in Equation (4.82) applies to the matched-filter detection of OOK, as it does to any coherent detection of orthogonal signaling.

The relationship in Equation (4.82) can also be confirmed by noting that the energy difference between the orthogonal signal vectors $\mathbf{s}_1$ and $\mathbf{s}_2$, with amplitudes of $\sqrt{E_b}$, as shown in Figure 3.10b, can be computed as the square of the distance between the heads of the orthogonal vectors, which will be $E_d = 2E_b$. Using this result in Equation (3.63) also yields Equation (4.82). If we compare Equation (4.82) with Equation (4.79), we can see that 3-dB more E_b/N_0 is required for BFSK to provide the same performance as BPSK. It should not be surprising that the performance of BFSK signaling is 3-dB worse than BPSK signaling, since for a given signal power, the distance-squared between orthogonal vectors is a factor of two less than the distance squared between antipodal vectors.

4.7.4 Probability of Bit Error for Noncoherently Detected Binary Orthogonal FSK

Consider the equally likely binary orthogonal FSK signal set $\{s_1(t)\}$, defined in Equation (4.8) as follows:

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos(\omega_i t + \phi) \qquad 0 \le t \le T, \qquad i = 1, 2$$

The phase term ϕ is unknown and assumed constant. The detector is characterized by $M = 2$ channels of bandpass filters and envelope detectors, as shown in Figure 4.19. The input to the detector consists of the received signal $r(t) = s_i(t) + n(t)$, where $n(t)$ is a white Gaussian noise process with two-sided power spectral density $N_0/2$. Assume that $s_1(t)$ and $s_2(t)$ are separated in frequency sufficiently that they have negligible overlap. For $s_1(t)$ and $s_2(t)$ being equally likely, we start the bit-error probability P_B computation with Equation (3.38) as we did for baseband signaling:

$$\begin{aligned} P_B &= \tfrac{1}{2} P(H_2|s_1) + \tfrac{1}{2} P(H_1|s_2) \\ &= \frac{1}{2} \int_{-\infty}^{0} p(z|s_1)\, dz + \frac{1}{2} \int_{0}^{\infty} p(z|s_2)\, dz \end{aligned} \tag{4.83}$$

For the binary case, the *test statistic* $z(T)$ is defined by $z_1(T) - z_2(T)$. Assume that the bandwidth of the filter W_f is $1/T$, so that the envelope of the FSK signal is (approximately) preserved at the filter output. If there was no noise at the receiver, the value of $z(T) = \sqrt{2E/T}$ when $s_1(t)$ is sent, and $z(T) = -\sqrt{2E/T}$ when $s_2(t)$ is sent. Because of this symmetry, the optimum threshold is $\gamma_0 = 0$. The pdf $p(z|s_1)$ is similar to $p(z|s_2)$; that is,

$$p(z|s_1) = p(-z|s_2) \tag{4.84}$$

Therefore, we can write

$$P_B = \int_{0}^{\infty} p(z|s_2)\, dz \tag{4.85}$$

or

$$P_B = P(z_1 > z_2|s_2) \tag{4.86}$$

where z_1 and z_2 denote the outputs $z_1(T)$ and $z_2(T)$ from the envelope detectors shown in Figure 4.19. For the case in which the tone $s_2(t) = \cos \omega_2 t$ is sent, such that $r(t) = s_2(t) + n(t)$, the output $z_1(T)$ is a *Gaussian noise random variable only;* it has no signal component. A Gaussian distribution into the *nonlinear envelope detector* yields a Rayleigh distribution at the output [6], so that

$$p(z_1|s_2) = \begin{cases} \dfrac{z_1}{\sigma_0^2} \exp\left(-\dfrac{z_1^2}{2\sigma_0^2}\right) & z_1 \ge 0 \\ 0 & z_1 < 0 \end{cases} \tag{4.87}$$

where σ_0^2 is the noise at the filter output. On the other hand, $z_2(T)$ has a Rician distribution, since the input to the lower envelope detector is a sinusoid plus noise [6]. The pdf $p(z_2|s_2)$ is written as

$$p(z_2|s_2) = \begin{cases} \dfrac{z_2}{\sigma_0^2} \exp\left[-\dfrac{(z_2^2 + A^2)}{2\sigma_0^2}\right] I_0\left(\dfrac{z_2 A}{\sigma_0^2}\right) & z_2 \geq 0 \\ 0 & z_2 < 0 \end{cases} \tag{4.88}$$

where $A = \sqrt{2E/T}$, and as before, σ_0^2 is the noise at the filter output. The function $I_0(x)$, known as the modified zero-order Bessel function of the first kind [7], is defined as

$$I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp(x \cos\theta)\, d\theta \tag{4.89}$$

When $s_2(t)$ is transmitted, the receiver makes an error whenever the envelope sample $z_1(T)$ obtained from the upper channel (due to noise alone) exceeds the envelope sample $z_2(T)$ obtained from the lower channel (due to signal plus noise). Thus the probability of this error can be obtained by integrating $p(z_1|s_2)$ with respect to z_1 from z_2 to infinity, and then averaging over all possible values of z_2. That is,

$$\begin{aligned} P_B &= P(z_1 > z_2|s_2) \\ &= \int_0^\infty p(z_2|s_2) \left[\int_{z_2}^\infty p(z_1|s_2)\, dz_1\right] dz_2 \qquad (4.90) \\ &= \int_0^\infty \frac{z_2}{\sigma_0^2} \exp\left[-\frac{(z_2^2 + A^2)}{2\sigma_0^2}\right] I_0\left(\frac{z_2 A}{\sigma_0^2}\right) \left[\int_{z_2}^\infty \frac{z_1}{\sigma_0^2} \exp\left(-\frac{z_1^2}{2\sigma_0^2}\right) dz_1\right] dz_2 \qquad (4.91) \end{aligned}$$

where $A = \sqrt{2E/T}$ and where the inner integral is the conditioned probability of an error for a fixed value of z_2, given that $s_2(t)$ was sent, and the outer integral averages this conditional probability over all possible values of z_2. This integral can be evaluated [8], yielding

$$P_B = \frac{1}{2} \exp\left(-\frac{A^2}{4\sigma_0^2}\right) \tag{4.92}$$

Using Equation (1.19), we can express the filter output noise as

$$\sigma_0^2 = 2\left(\frac{N_0}{2}\right) W_f \tag{4.93}$$

where $G_n(f) = N_0/2$ and W_f is the filter bandwidth. Thus Equation (4.92) becomes

$$P_B = \frac{1}{2} \exp\left(-\frac{A^2}{4N_0 W_f}\right) \tag{4.94}$$

Equation (4.94) indicates that the error performance depends on the bandpass filter bandwidth, and that P_B becomes smaller as W_f is decreased. The result is valid only when the intersymbol interference (ISI) is negligible. The minimum W_f allowed (i.e., for no ISI) is obtained from Equation (3.81) with the filter roll-off factor $r = 0$. Thus $W_f = R$ bits/s $= 1/T$, and we can write Equation (4.94) as

$$P_B = \frac{1}{2}\exp\left(-\frac{A^2T}{4N_0}\right) \tag{4.95}$$

$$= \frac{1}{2}\exp\left(-\frac{E_b}{2N_0}\right) \tag{4.96}$$

where $E_b = (1/2)A^2T$ is the energy per bit. When comparing the error performance of noncoherent FSK with coherent FSK (see Figure 4.25), it is seen that for the same P_B, noncoherent FSK requires approximately 1 dB more E_b/N_0 than that for coherent FSK (for $P_B \leq 10^{-4}$). The noncoherent receiver is easier to implement, because coherent reference signals need not be generated. Therefore, almost all FSK receivers use noncoherent detection. It can be seen in the following section that when comparing noncoherent orthogonal FSK to noncoherent DPSK, the same 3-dB difference occurs as for the comparison between coherent orthogonal FSK and coherent PSK.

As mentioned earlier, the details of on-off keying (OOK) are not treated in this book. However, it is worth noting that the bit error probability P_B described in Equation (4.96) is identical to the P_B for the noncoherent detection of OOK signaling.

4.7.5 Probability of Bit Error for Binary DPSK

Let us define a BPSK signal set as follows:

$$\begin{aligned} x_1(t) &= \sqrt{\frac{2E}{T}}\cos(\omega_0 t + \phi) \qquad 0 \leq t \leq T \\ x_2(t) &= \sqrt{\frac{2E}{T}}\cos(\omega_0 t + \phi \pm \pi) \qquad 0 \leq t \leq T \end{aligned} \tag{4.97}$$

A characteristic of DPSK is that there are no fixed decision regions in the signal space. Instead, the decision is based on the phase difference between successively received signals. Then for DPSK signaling we are really transmitting each bit with the binary signal pair

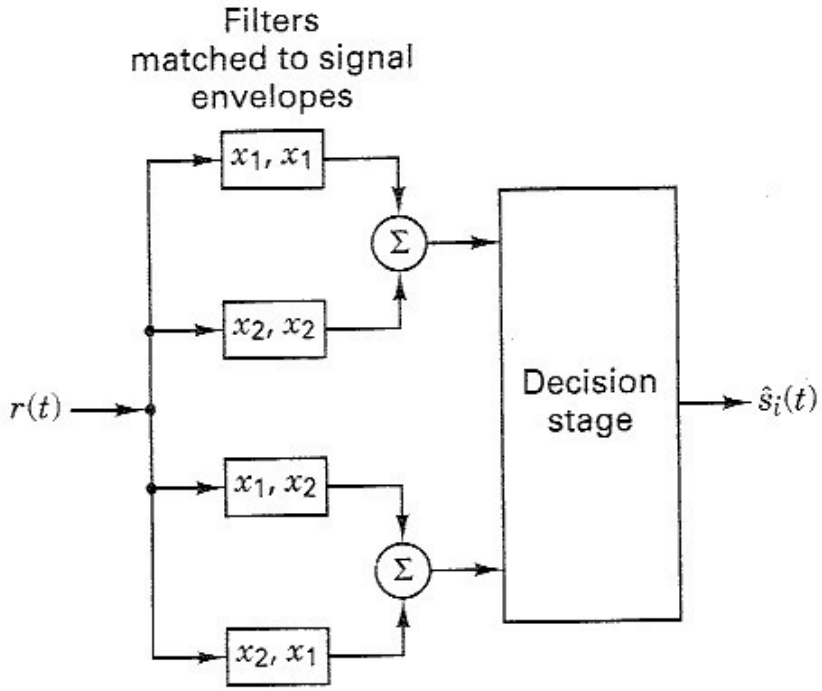
and

$$\begin{aligned} s_1(t) &= (x_1, x_1) \quad \text{or} \quad (x_2, x_2) \qquad 0 \leq t \leq 2T \\ s_2(t) &= (x_1, x_2) \quad \text{or} \quad (x_2, x_1) \qquad 0 \leq t \leq 2T \end{aligned} \tag{4.98}$$

where (x_i, x_j) $(i, j = 1, 2)$ denotes $x_i(t)$ followed by $x_j(t)$ defined in Equation (4.97). The first T seconds of each waveform are actually the last T seconds of the previous waveform. Note that $s_1(t)$ and $s_2(t)$ can each have either of two possible forms and that $x_1(t)$ and $x_2(t)$ are antipodal signals. Thus the correlation between $s_1(t)$ and $s_2(t)$ for *any combination* of forms can be written as

$$z(2T) = \int_0^{2T} s_1(t)s_2(t)\,dt = \int_0^T [x_1(t)]^2\,dt - \int_0^T [x_1(t)]^2\,dt = 0 \tag{4.99}$$

Therefore, pairs of DPSK signals can be represented as orthogonal signals $2T$ seconds long. Detection could correspond to noncoherent envelope detection with four channels matched to each of the possible envelope outputs, as shown in Figure 4.26a. Since the two envelope detectors representing each symbol are negatives of each other, the envelope sample of each will be the same. Hence we can implement the detector as a single channel for $s_1(t)$ matched to either (x_1, x_1) or (x_2, x_2), and a single channel for $s_2(t)$ matched to either (x_1, x_2) or (x_2, x_1), as shown in Figure 4.26b. The DPSK detector is therefore reduced to a standard two-channel noncoherent detector. In reality, the filter can be matched to the difference signal so that only one channel is necessary. In Figure 4.26, the filters are matched to the signal envelopes (over two symbol times). What does this mean in light of the fact that



(a)

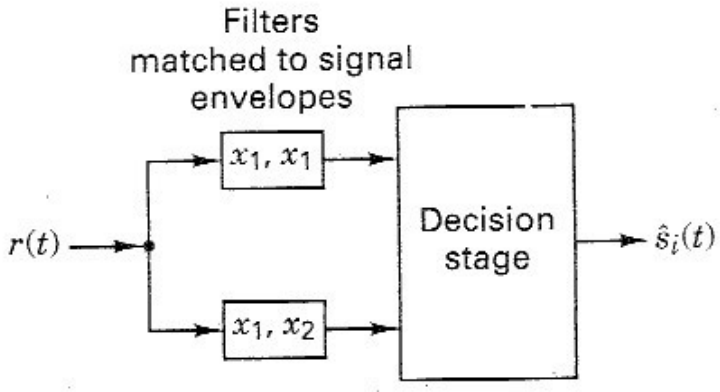


Figure 4.26 DPSK detection. (a) Four-channel differentially coherent detection of binary DPSK. (b) Equivalent two-channel detector for binary DPSK.

DPSK is a constant envelope signaling scheme? It means that we need to implement an energy detector, similar to the quadrature receiver in Figure 4.18, with each of the I and Q reference signals occuring in pairs during the time interval $(0 \le t \le 2T)$, as follows:

$$\begin{aligned} s_1(t)\ I \text{ reference:} &\quad \sqrt{2/T}\cos\omega_0 t,\ \sqrt{2/T}\cos\omega_0 t \\ s_1(t)\ Q \text{ reference:} &\quad \sqrt{2/T}\sin\omega_0 t,\ \sqrt{2/T}\sin\omega_0 t \end{aligned}$$

$$\begin{aligned} s_2(t)\ I \text{ reference:} &\quad \sqrt{2/T}\cos\omega_0 t,\ -\sqrt{2/T}\cos\omega_0 t \\ s_2(t)\ Q \text{ reference:} &\quad \sqrt{2/T}\sin\omega_0 t,\ -\sqrt{2/T}\sin\omega_0 t \end{aligned}$$

Since pairs of DPSK signals are orthogonal, such noncoherent detection operates with the bit-error probability given by Equation (4.96). However, since the DPSK signals have a bit interval of $2T$, the $s_i(t)$ signals defined in Equation (4.98) have twice the energy of a signal defined over a single-symbol duration. Thus, we may write P_B as

$$P_B = \frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right) \tag{4.100}$$

Equation (4.100) is plotted in Figure 4.25, designated as differentially coherent detection of differentially encoded PSK, or simply DPSK. This expression is valid for the optimum DPSK detector shown in Figure 4.17c. For the detector shown in Figure (4.17b), the error probability will be slightly inferior to that given in Equation (4.100) [3]. When comparing the error performance of Equation (4.100) with that of coherent PSK (see Figure 4.25), it is seen that for the same P_B, DPSK requires approximately 1 dB more E_b/N_0 than does BPSK (for $P_B \le 10^{-4}$). It is easier to implement a DPSK system than a PSK system, since the DPSK receiver does not need phase synchronization. For this reason, DPSK, although less efficient than PSK, is sometimes the preferred choice between the two.

4.7.6 Comparison of Bit Error Performance for Various Modulation Types

The P_B expressions for the best known of the binary modulation schemes discussed above are listed in Table 4.1 and are illustrated in Figure 4.25. For $P_B = 10^{-4}$, it can be seen that there is a difference of approximately 4-dB between the best (coherent PSK) and the worst (noncoherent orthogonal FSK) that were discussed here. In some cases, 4 dB is a small price to pay for the implementation simplicity gained in going from coherent PSK to noncoherent FSK; however, for other cases, even a 1-dB saving is worthwhile. There are other considerations besides P_B and system complexity; for example, in some cases (such as a randomly fading channel), a noncoherent system is more desirable because there may be difficulty in establishing and maintaining a coherent reference. Signals that can withstand significant degradation before their ability to be detected is affected are clearly desirable in military and space applications.

TABLE 4.1 Probability of Error for Selected Binary Modulation Schemes

Modulation	P_B
PSK (coherent)	$Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$
DPSK (differentially coherent)	$\frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$
Orthogonal FSK (coherent)	$Q\left(\sqrt{\frac{E_b}{N_0}}\right)$
Orthogonal FSK (noncoherent)	$\frac{1}{2}\exp\left(-\frac{1}{2}\frac{E_b}{N_0}\right)$

4.8 *M*-ARY SIGNALING AND PERFORMANCE

4.8.1 Ideal Probability of Bit Error Performance

The typical probability of error versus E_b/N_0 curve was shown to have a waterfall-like shape in Figure 3.6. The probability of bit error (P_B) characteristics of various binary modulation schemes in AWGN also display this shape, as shown in Figure 4.25. What should an *ideal* P_B versus E_b/N_0 curve look like? Figure 4.27 displays the ideal characteristic as the *Shannon limit.* The limit represents the threshold E_b/N_0 below which reliable communication cannot be maintained. Shannon's work is described in greater detail in Chapter 9.

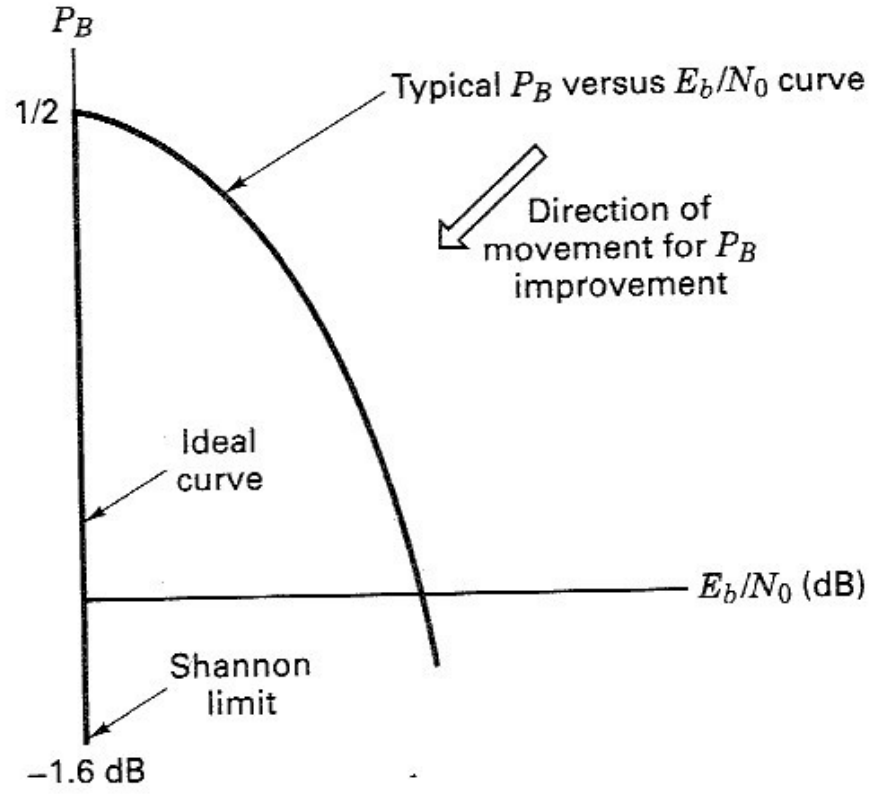


Figure 4.27 Ideal P_B versus E_b/N_0 curve.

We can describe the ideal curve in Figure 4.27 as follows. For all values of E_b/N_o above the Shannon limit of −1.6 dB, P_B is zero. Once E_b/N_0 is reduced below the Shannon limit, P_B degrades to the worst-case value of $\frac{1}{2}$. (Note that $P_B = 1$ is not the worst case for binary signaling, since that value is just as good as $P_B = 0$; if the probability of making a bit error is 100%, the bit stream could simply be inverted to retrieve the correct data.) It should be clear, by comparing the typical P_B curve with the ideal one in Figure 4.27 that the large arrow in the figure describes the desired direction of movement to achieve improved P_B performance.

4.8.2 *M*-ary Signaling

Let us review M-ary signaling. The processor considers k bits at a time. It instructs the modulator to produce one of $M = 2^k$ waveforms; binary signaling is the special case where $k = 1$. Does M-ary signaling improve or degrade error performance? (Be careful with your answer—the question is a loaded one.) Figure 4.28 illustrates the probability of bit error $P_B(M)$ versus E_b/N_0 for coherently detected *orthogonal* M-ary signaling over a Gaussian channel. Figure 4.29 similarly illustrates $P_B(M)$ versus E_b/N_0 for coherently detected *multiple phase* M-ary signaling over a Gaussian channel. In which direction do the curves move as the value of k (or M) increases? From Figure 4.27 we know the directions of curve movement for improved and degraded error performance. In Figure 4.28, as k increases, the curves move in the direction of improved error performance. In Figure 4.29, as k increases, the curve move in the direction of degraded error performance. Such movement tells us that M-ary signaling produces improved error performance with orthogonal signaling and degraded error performance with multiple phase signaling. Can that be true? Why would anyone ever use multiple phase PSK signaling if it provides degraded error performance compared to binary PSK signaling? It *is* true, and many systems do use multiple phase signaling. The question, as stated, is loaded because it implies that error probability versus E_b/N_0 is the *only* performance criterion; there are many others (e.g., bandwidth, throughput, complexity, cost), but in Figures 4.28 and 4.29 error performance is the characteristic that stands out explicitly.

A performance characteristic that is not explicitly seen in Figures 4.28 and 4.29 is the required system bandwidth. For the curves characterizing M-ary orthogonal signals in Figure 4.28, as k increases, the required bandwidth also increases. For the M-ary multiple phase curves in Figure 4.29, as k increases, a larger bit rate can be transmitted within the same bandwidth. In other words, for a fixed data rate, the required bandwidth is decreased. Therefore, *both* the orthogonal and multiple phase error performance curves tell us that M-ary signaling represents a vehicle for performing a system trade-off. In the case of orthogonal signaling, error performance improvement can be achieved at the expense of bandwidth. In the case of multiple phase signaling, bandwidth performance can be achieved at the expense of error performance. Error performance versus bandwidth performance, a fundamental communications trade-off, is treated in greater detail in Chapter 9.

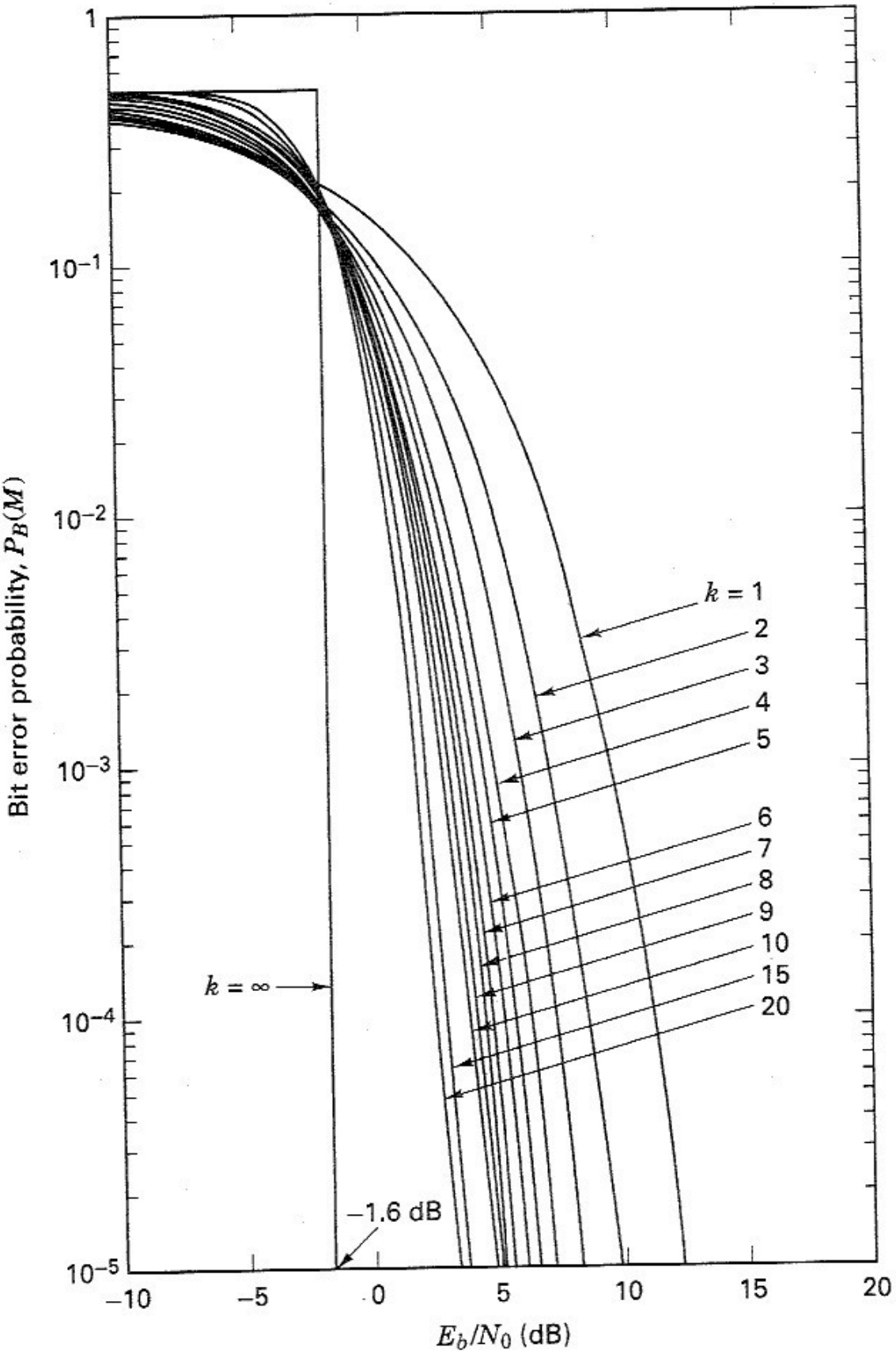


Figure 4.28 Bit error probability for coherently detected *M*-ary orthogonal signaling. (Reprinted from W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*, Prentice Hall, Inc., Englewood Cliffs, N.J., 1973, courtesy of W. C. Lindsey and Marvin K. Simon).

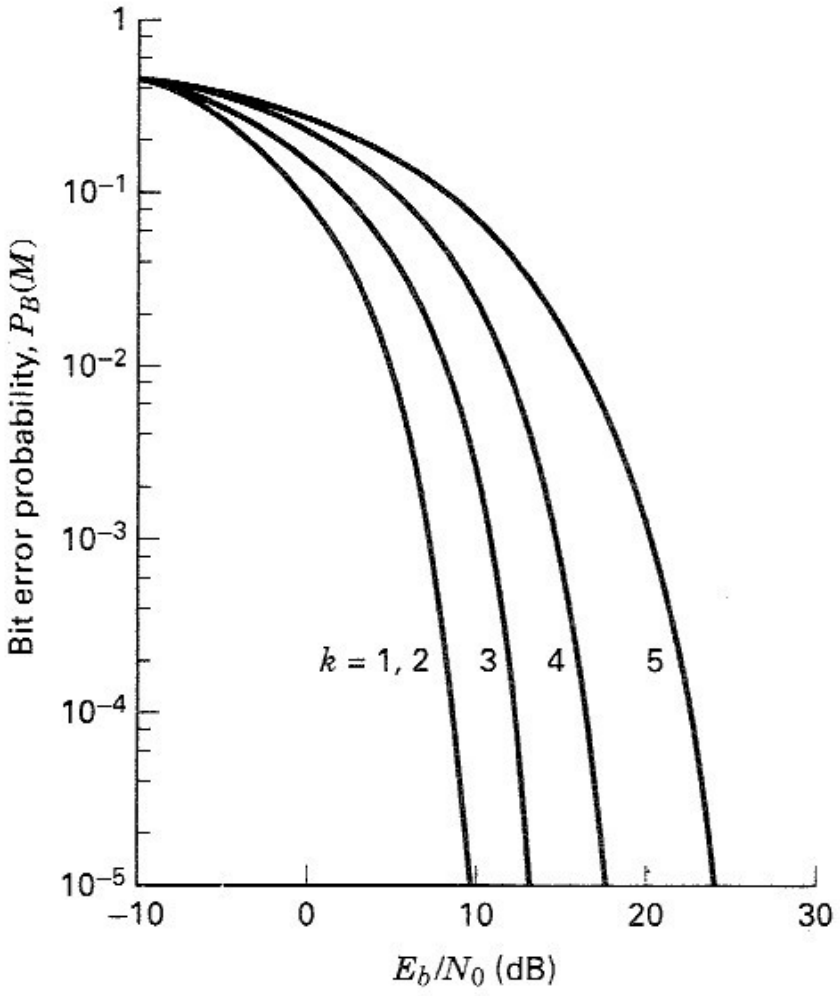


Figure 4.29 Bit error probability for coherently detected multiple phase signaling.

4.8.3 Vectorial View of MPSK Signaling

Figure 4.30 illustrates MPSK signal sets for $M = 2, 4, 8$, and 16. In Figure 4.30a we see the binary ($k = 2$, $M = 2$) antipodal vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ positioned 180° apart. The decision boundary is drawn so as to partition the signal space into two regions. On the figure is also shown a noise vector $\mathbf{n}$ equal in magnitude to $\mathbf{s}_1$. The figure establishes the magnitude and orientation of the minimum energy noise vector that would cause the detector to make a symbol error.

In Figure 4.30b we see the 4-ary ($k = 2$, $M = 4$) vectors positioned 90° apart. The decision boundaries (only one line is drawn) divide the signal space into four regions. Again a noise vector $\mathbf{n}$ is drawn (from the head of a signal vector, normal to the closest decision boundary) to illustrate the minimum energy noise vector that would cause the detector to make a symbol error. Notice that the minimum energy noise vector of Figure 4.30b is smaller than that of Figure 4.30a, illustrating that the 4-ary system is more vulnerable to noise than the 2-ary system (signal energy being equal for each case). As we move on to Figure 4.30c for the 8-ary case and Figure 4.30d for the 16-ary case, it should be clear that for multiple phase signaling, as M increases, we are crowding more signal vectors into the signal plane. As the vectors are moved closer together, a smaller amount of noise energy is required to cause an error.

Figure 4.30 adds some insight as to why the curves of Figure 4.29 behave as they do as k is increased. Figure 4.30 also provides some insight into a basic trade-