PHASE SHIFT KEYING

Digital transmission has many advantages over conventional analog systems: superior noise immunity, suitability for signal processing and multiplexing, source coding, security, error control coding, merging of different sources data, voice, and video, and relatively inexpensive and more reliable digital circuitry. Digital communication systems are rapidly replacing analog communication systems.

Many communication channels, such as radio, are unsuitable for the direct transmission of data streams. A data stream is used to modulate a high-frequency sinusoidal carrier to produce a digital passband signal, whose energy content is centered about the carrier frequency. Compared with using a baseband digital signal directly the potential advantages of digital passband signaling include:

- More suitable matching of the transmitted signal to the propagation properties of the communication channel
- Efficient radio transmission with a reasonably sized antenna
- Ability to exploit the bandwidth capabilities of the channel by the frequency multiplexing of many different signals for simultaneous transmission over a common channel
- Exploitation of the available radio spectrum

The three main categories of digital passband signals involve using the data stream to change the amplitude, frequency, or phase of the carrier, corresponding to amplitude, frequency, and phase modulation techniques, respectively. This article concentrates entirely on digital phase modulation. There are a number of concepts that influence the choice of phase modulation technique. These include

1. *Bandwidth Efficiency.* The frequency spectrum is one of the primary communication resources and, with ever-increasing demands for more channels, should be used efficiently. It is important to use spectral modulation techniques that make effective use of the available bandwidth by transmitting a maximum information rate in a minimum possible bandwidth. For this reason it is important to determine the spectral content of the digitally modulated signals from the power spectral density. Unfortunately the bandwidth is somewhat difficult to quantify and a variety of definitions of the modulation signal are possible. These include the 3 dB (or half-power) bandwidth, the noise equivalent bandwidth, the null-to-null bandwidth, and the fractional power containment bandwidth (which is the band containing 99% of the power). The reader is referred to the textbooks listed in Refs. 1–7 and to Ref. 8 for detailed treatment of the bandwidth issue. All of the above definitions are meaningful, but at least a *consistent* definition should be used when comparing modulation techniques on the basis of bandwidth efficiency. The bandwidth efficiency is defined as the ratio of the information bit rate to the occupied bandwidth and high value is desirable. This is particularly a very important issue in radio communication systems where the channel is a limited and valuable resource.
2. *Out-of-Band Radiation.* Closely related to bandwidth efficiency is the amount of transmitted signal energy lying outside the main lobe of the signal spectrum (1,4,5,9). This is a very important consideration in the choice of modulation technique for many multiuser radio communication systems, such as in mobile radio, as it represents a potential cause of adjacent channel interference (interference to the channels above and below). A modulation method that gives low out-of-band radiation is very desirable in such applications. In general, this is achieved by choosing modulation signals that avoid abrupt phase changes.
3. *Resilience to Nonlinear Distortion.* Most digital radio systems operate their main high-power amplifiers at or near saturation in order to achieve high efficiency (1,4,5,9). Such amplifiers introduce nonlinear amplitude and phase distortion, and a potential consequence of this is a spectral spreading of the transmitted signal which, in turn, increases undesirable adjacent channel interference. Filtering of signals for band-limiting and spectral shaping results in envelope variation and, when passed through nonlinear high-power amplifiers, the spectral sidelobes are reintroduced, causing spectral spreading. In order to minimize this effect it is important to choose a constant envelope modulation technique without abrupt phase transitions.
4. *Power Efficiency.* Another very important parameter, particularly in power limited systems, is the power efficiency. For example, in battery-operated mobile radio systems power efficiency is a very important consideration, as it translates into battery size, weight, and recharging intervals. The power efficiency can be characterized by the received power that is needed in the presence of noise in order to achieve an acceptably small number of bit errors. The standard performance criterion for comparing the power efficiencies of different digital communication systems in additive white Gaussian noise is the relationship between probability of bit error, P_e, and the ratio of signal energy per bit to noise spectral density, E_b/N_0 (dB) (1–7). A modulation technique that provides an acceptable P_e with a low value of E_b/N_0 is important.
5. *Resilience to Multipath.* In most radio channels the transmitted signal arrives at the receiver via multiple propagation paths and the resilience of the modulation technique to mutipath is an important factor. This subject is outside the scope of this article. The reader is referred to (4,5).
6. *Implementation Complexity.* Low complexity implementation of the modulator and demodulator and hence, cost, are also very important considerations in the choice of modulation method.

In this article we describe a number of commonly used digital phase modulation techniques for radio transmission, in terms of their modulator and demodulator, their waveforms, the geometrical representation of their waveforms, their

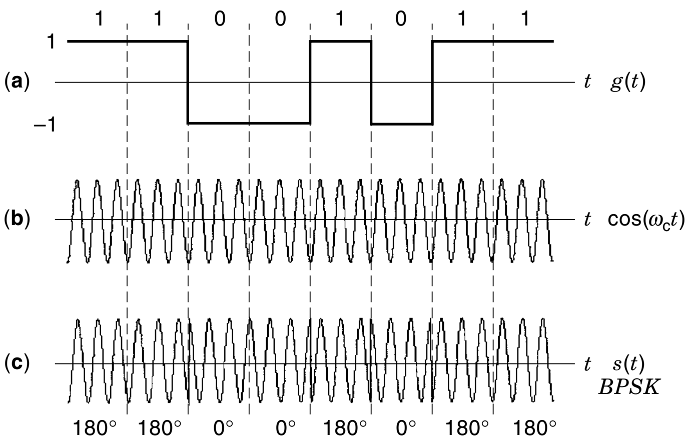

Figure 1. BPSK (a) Time-domain representation of a baseband bipolar data signal; (b) sinusoidal carrier signal; (c) the resulting BPSK signal; T_b is bit period.

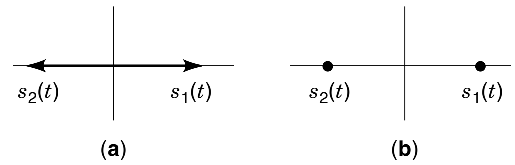

Figure 3. Geometrical representation of BPSK. (a) phasor diagram; (b) constellation points.

bandwidth efficiency, the effect on their spectral properties of transmission over a nonlinear channel, and their performance in noise. The techniques covered are binary phase shift keying, differential phase shift keying, *M*-ary phase shift keying, quadriphase shift keying, offset quadriphase shift keying, and minimum shift keying (BPSK, DPSK, *M*-ary PSK, QPSK, OQPSK, and MSK, respectively).

BINARY PHASE SHIFT KEYING (BPSK)

In binary phase modulation a constant-amplitude and fixed frequency carrier signal is permitted to have two phases and these are used to represent the two digital bits 0 and 1. Usually the two phases are chosen to differ by 180°, in which case the BPSK signal during the bit period T_b is given mathematically by:

$$\begin{aligned} s_1(t) &= A\cos(2\pi f_c t) && \text{for a data 0} \\ s_2(t) &= A\cos(2\pi f_c t + \pi) = -A\cos(2\pi f_c t) && \text{for a data 1} \end{aligned} \tag{1}$$

where A is the baseband modulating signal and f_c is the carrier frequency. The phase separation of 180° causes the maximum possible difference in phase between the two signals and leads to the lowest probability of an error occurring when additive noise is present (the highest noise immunity). Signals which are 180° apart are sometimes termed antipodal signals. An alternative way of writing Eq. (1) is:

$$s(t) = g(t)\cos(2\pi f_c t) \qquad \text{where } g(t) = \pm A \tag{2}$$

One common form of baseband data transmission is nonreturn-to-zero signaling, in which the data bit 1 is represented by a positive pulse and the data bit 0 is represented by a negative pulse, as shown in Fig. 1(a). If such a bipolar data signal is multiplied by a carrier one performs the operations of Eq. (2), and thus implements a BPSK modulator. A carrier signal and the resulting BPSK output waveform are shown in Fig. 1(b) and Fig. 1(c). A symbolic representation of the modulator is shown in Fig. 2.

It is useful to realize that the signal waveforms given by Eq. (1) can be represented by vectors on a polar plot; the vector length corresponds to the signal amplitude (constant in our case) and the vector direction corresponds to the signal phase, as shown in Fig. 3(a). If the plot is simplified still further by replacing the arrows with dots at the arrowheads, we obtain a *signal space* or *signal-state constellation* diagram, shown in Fig. 3(b). This is particularly useful in the visualization and detection of the signaling scheme.

The optimum form of detection for BPSK is to correlate the received signal in a bit period with all possible transmit signals (two in this case), and then to select that which gives the highest correlation. In the case of BPSK this correlation detection can be simplified to multiplying the incoming signal by a single signal $A\cos(2\pi f_c t)$, integrating, and then determining whether this integrated signal is greater or less than zero. This arises because a received signal $B\cos(2\pi f_c t)$ will produce $-\int_0^{T_b} A\cos(2\pi f_c t)B\cos(2\pi f_c t)dt = -AB/2$ (assuming a normalized T_b value of 1 and f_c is an integer multiple of the bit period T_b), whereas a received signal $-B\cos(2\pi f_c t)$ will produce $\int_0^{T_b} A\cos(2\pi f_c t)B\cos(2\pi f_c t)dt = AB/2$. Thus the optimum detector for BPSK is the conceptually simple coherent detection system shown in Fig. 4.

A major practical requirement in the receiver of Fig. 4 is that of obtaining a reference signal $A\cos(2\pi f_c t)$ that has the correct phase relative to that in the incoming BPSK signal. Some form of *carrier recovery* is necessary (1–4) and it is necessary that this recovered carrier should track any instabilities in the phase of the incoming signal due to changing propagation conditions. Another practical requirement is that of sampling the integrator at the correct timing instant $t = (n + 1)T_b$ where $n = 0, 1, 2, \ldots$. This is provided by what is known as symbol timing recovery circuit (2–7).

Performance of BPSK

Equation (2) shows that a BPSK signal may be considered as that special form of amplitude modulated signal where the

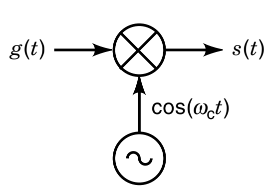

Figure 2. A symbolic representation of BPSK modulator.

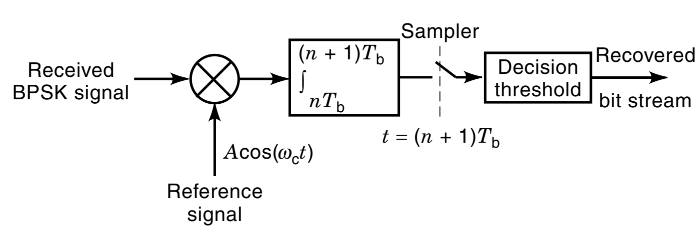

Figure 4. Block diagram of BPSK correlator detector.

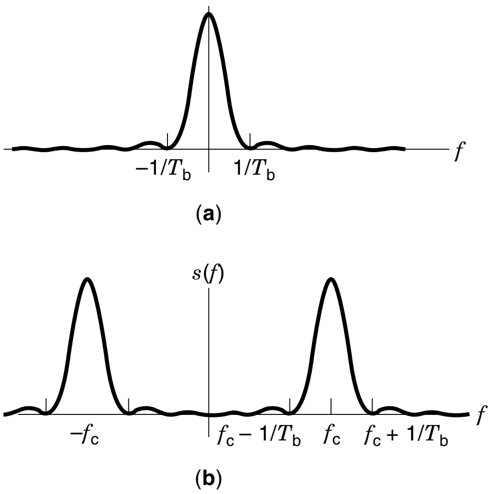

Figure 5. Power spectral density of (a) baseband signal; (b) BPSK signal.

carrier is suppressed; it is equivalent to double sideband suppressed carrier modulation. For the special case of alternate 1's and 0's one has the multiplication of a carrier by a bipolar square wave, and the spectrum is that of a square wave except for a frequency translation up and down by $\pm f_c$. However, a random sequence of 1's and 0's represented by the values $+A$ and $-A$ is a more relevant data stream. This has an autocorrelation function, which is a triangle of peak height A^2 centered at zero lag and extending from $-T_b$ to $+T_b$, where T_b is the bit period. The Fourier transform of this autocorrelation function gives the power spectral density of the data stream (1–4) as:

$$G(f) = A^2 T_b \left(\frac{\sin(\pi f T_b)}{\pi f T_b} \right)^2 \tag{3}$$

The corresponding BPSK signal has the same spectrum as this except centered at $\pm f_c$. The power spectral density of a random binary wave given by Eq. (3) and of the consequent BPSK signal are shown in Fig. 5(a) and Fig. 5(b). It can be seen that the null-to-null bandwidth of BPSK is $2/T_b$, which is basically twice that of the baseband signal.

It is generally most convenient to express the spectral properties of digital passband signals in terms of the power spectral density of its complex envelope, a result known as the *baseband power spectral density* $S(f)$ (4–7). If the power of the BPSK signal is P_c this baseband power spectral density is given by

$$S(f) = 2P_c T_b \left(\frac{\sin(\pi f T_b)}{\pi f T_b} \right)^2 \tag{4}$$

This is plotted on a dB scale in Fig. 6 where, by arbitrarily considering $2P_c T_b = 1$, the baseband power spectral density at zero frequency becomes normalized to 0 dB.

As already stated, the RF power spectral density is of the same form as the baseband power spectral density except frequency translated. Mathematically the RF power spectral density is related to the baseband power spectral density (7) by the equation

$$S_S(f) = \frac{1}{4}[S(f - f_c) + S(f + f_c)] \tag{5}$$

Using Eq. (4) or Fig. 6 it is deduced that the BPSK signal has a bandwidth to the first spectral nulls of $2/T_b$. The bit rate is $R_b = 1/T_b$ bits per second (bps) and, on the basis of this null-to-null bandwidth, the bandwidth efficiency, η, is 0.5 bps/Hz.

An important factor in judging a modulation technique is the fraction of transmitted power lying outside of an RF bandwidth B centered at f_c. From Eq. (4) one obtains

$$\text{Fractional out-of-band power} = \frac{\int_{B/2}^{\infty} S(f)\,df}{\int_0^{\infty} S(f)\,df} \tag{6}$$

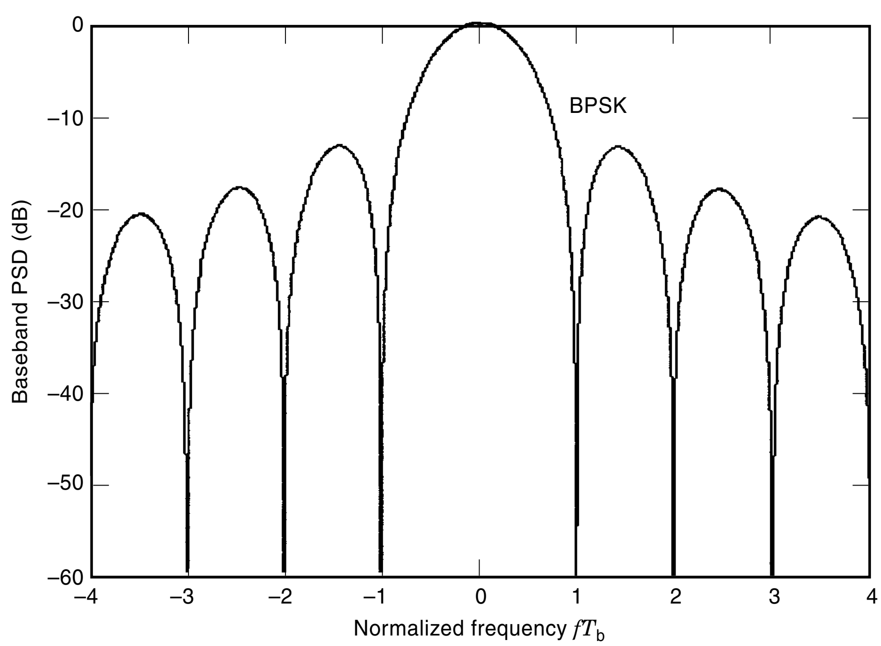

Figure 6. Power spectral density of BPSK.

Table 1. Fractional Out-of-Band Power for BPSK

RF Bandwidth, B (Hz)	Fractional Out-of-Band Power
$1/T_b$	23% (−6.4 dB)
$1.5/T_b$	11% (−9.6 dB)
$2/T_b$	9.6% (−10.2 dB)
$3/T_b$	6.7% (−11.7 dB)
$4/T_b$	4.8% (−13.2 dB)
$8/T_b$	2.3% (−16.4 dB)
$16/T_b$	1% (−20 dB)

This equation is important when evaluating the level of adjacent channel interference. For example, applying it to Eq. (6), one obtains Table 1. The null-to-null bandwidth is $2/T_b$ but a bandwidth of $16/T_b$ is required before the fractional out-of-band power is down to 1%, and this is considered particularly important since it corresponds to the Federal Communication Commission (FCC) criterion for the bandwidth containing 99% of the signal energy. It is noted that this fractional power containment bandwidth is very much greater than the null-to-null bandwidth.

One possible way of minimizing adjacent channel interference is to use a band-limiting filter in the transmitter prior to the final power amplifier. An unfortunate side-effect of this is that the BPSK signal loses its constant envelope such that if it is then passed through an efficient nonlinear high-power amplifier, envelope distortion will be introduced. This signifies the reintroduction of spectral sidelobes giving undesirable out-of-band radiation. It is apparent that a constrained bandwidth and efficient nonlinear amplification are mutually incompatible. A major motivation behind some of the other phase shift keying schemes to be considered is that they behave better in this respect by having smaller or smoother phase changes between bits. This causes them to have less out-of-band spectral content. This means that there is less envelope distortion caused by band-limiting and hence the reintroduction of spectral sidelobes due to nonlinear amplification is less.

As mentioned in the introduction, another important parameter of performance in modulation techniques is the power efficiency, which is characterized by the probability of the detector making an incorrect decision, P_e, versus the ratio of energy per bit to noise spectral density, E_b/N_0. The probability of bit error of BPSK using coherent detection in additive white Gaussian noise is given (2) as

$$P_e = \mathrm{Erfc}\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{7}$$

where Erfc(x) is the complementary error function given by $\mathrm{Erfc}(x) = 1/\sqrt{2\pi}\int_x^\infty e^{-y^2/2}\,dy$. This result is plotted in Fig. 7. It can be seen that a P_e of 10^{-6} is achieved when E_b/N_0 is about 10.6 dB. The noise performance of BPSK is discussed further in a later section, in conjunction with the other phase shift keying systems, but it is relevant to say at this stage that BPSK has the best noise immunity of all of them.

DIFFERENTIAL PHASE SHIFT KEYING (DPSK)

In DPSK the phase does not indicate directly whether a 0 or a 1 is transmitted. Instead the data stream is conveyed by

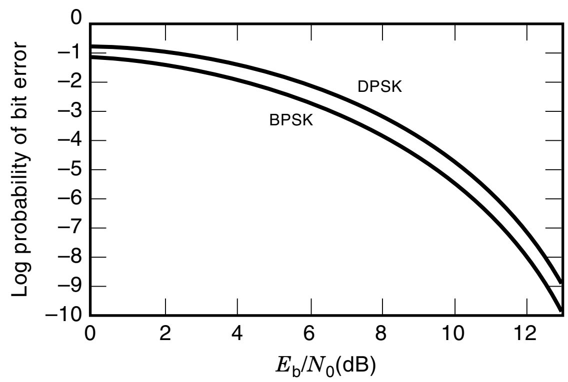

Figure 7. Probability of bit error versus E_b/N_0 for BPSK and DPSK.

the change or lack of change of phase between bits. If a 0 is to be transmitted, the phase in that bit period is changed by 180° relative to that in the previous bit period. If a 1 is to be transmitted, the phase in that bit period stays unchanged. In practice, this is done by differentially encoding the binary data stream and then applying the resultant sequence to the same PSK modulator described in Fig. 2. As an example of differentially encoded data and the consequent DPSK signal, consider an original data stream 1101011001, shown in Table 2. The differentially encoded sequence needs to commence with an extra starting bit. Assume this is 1 and that the PSK modulator produces $\phi = 0$ for a 0, and $\phi = \pi$ for a 1. The resulting phase shift is then as shown in Table 2. Mathematically the relation between the differentially encoded data c_n and the original binary data stream d_n can be expressed as

$$c_n = \overline{c_{n-1} \oplus d_n} \tag{8}$$

where $\oplus$ represents modulo-2 addition and the overbar denotes complementing.

Compared with BPSK, DPSK has several advantages and disadvantages. The first great advantage of DPSK over BPSK concerns the avoidance in the receiver of the need to recover the transmitted carrier signal. Instead the received signal is delayed by one bit period and, as shown in Fig. 8, this is used as the reference signal for detecting the present bit. The receiver, equivalently, compares the phase of the received symbol with the phase of the previous symbol and observes whether they are identical or of the opposite phase. No change is taken as indicating a 1, and a 180° change is taken as indicating a 0. This is illustrated in Table 3, assuming the transmitted phase shifts, ϕ_n, are received without any channel error.

A second advantage is that BPSK carrier recovery circuits can suffer from a 180° phase ambiguity, in which the recovered reference carrier may be either in-phase or in anti-phase with the transmitted carrier. This can lead to bit inversion of

Table 2. DPSK Modulation Mapping for the Data Bits 1101011001

Bit index n	0	1	2	3	4	5	6	7	8	9	10
Data stream d_n		1	1	0	1	0	1	1	0	0	1
Differentially encoded data c_n	1	1	1	0	0	1	1	1	0	1	1
Transmitted phase shift ϕ_n	π	π	π	0	0	π	π	π	0	π	π

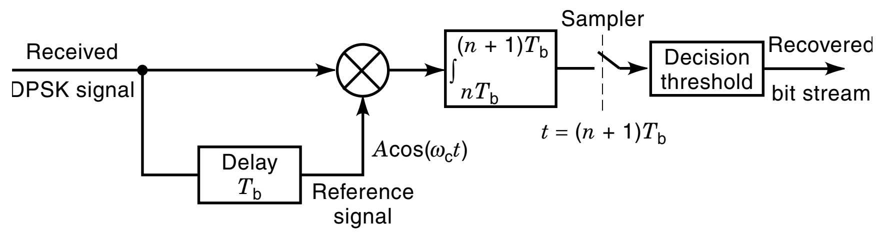

Figure 8. Block diagram of DPSK correlator detector.

the demodulated data. By using differential encoding at the transmitter and the phase of the previous bit of the received signal as the reference signal, this phase ambiguity is avoided (2–4).

A disadvantage of DPSK, however, is that the error rate due to noise is greater than for BPSK. This is because BPSK compares the received signal with a clean reference generated at the receiver, whereas DPSK compares one noisy signal with another. This is shown clearly in Table 4 by introducing one error in the fifth bit. In BPSK this results in one bit error, whereas in DPSK this results in the two errors shown underlined. As might be expected, this means that, for the same error rate, DPSK requires more signal power than BPSK. The probability of bit for DPSK is given (2–4) by

$$P_e = \frac{1}{2}\exp(-E_b/N_0) \tag{9}$$

This result is plotted in Fig. 7. It can be seen that DPSK requires approximately 1 dB more E_b/N_0 than BPSK for a probability of error of 10^{-4}.

M-ARY PHASE SHIFT KEYING (*M*-ARY PSK)

If T_b is the bit period in the original data stream a BPSK signal has two possible phases in time T_b. However, an alternative approach to transmitting the data stream can arise from partitioning the data stream into blocks that are k bits long ($k = 1$ in the binary case) and regarding the contents of each block as a symbol. Each symbol conveys one of 2^k possible states and can then be communicated by transmitting one of $M = 2^k$ different phases in what is known either as M-ary PSK or as multiple phase shift keying. A generalized expression for an M-ary PSK signal during each signaling interval of duration $T_s = kT_b$ is

$$s_i(t) = A\cos\left[2\pi f_c t + \frac{2\pi(i-1)}{M}\right] \qquad i = 1, 2, \ldots, M \tag{10}$$

Considering as an example $k = 3$, such that $M = 8$, we would divide up the data stream into blocks of 3 bits and then, for each block, transmit one of 8 phases 45° apart, depending on whether that group is binary 000, 001, 010, 011, 100, 101, 110, or 111, as shown in Fig. 9(a). The principle of partitioning a data stream into blocks is summarized by Fig. 9(b) where, if the information rate of the input is R_b bits/s, the output symbol rate of R_s symbols/s, also referred to as the baud rate, is given by

$$R_s = R_b/k \tag{11}$$

Whereas the bits in the original data stream are of duration T_b, the symbols are of duration $T_s = kT_b$. The advantages of using M-ary PSK are described next.

Bandwidth Efficiency of *M*-ary PSK

To be able to examine the spectral properties and bandwidth efficiency of M-ary PSK it is important first to consider the power spectral density. The baseband power spectral density of unfiltered M-ary PSK signals is given (6) by

$$S(f) = 2kP_c T_b\left[\frac{\sin(k\pi f T_b)}{k\pi f T_b}\right]^2 \tag{12}$$

where again P_c is the average carrier power. These results are plotted in Fig. 10 for $k = 2, 3, 4$ (i.e., 4-ary PSK, 8-ary PSK, 16-ary PSK). It can be seen from these that the null-to-null RF bandwidths for $M = 4, 8, 16$ are $1/T_b$, $2/3T_b$, and $1/2T_b$, respectively. Generalizing this,

$$\text{Null-to-null RF bandwidth} = \frac{2}{T_b\log_2 M} = \frac{2R_b}{\log_2 M} \tag{13}$$

The bandwidth efficiency of unfiltered M-ary PSK on the basis of null-to-null bandwidths, is therefore given by

$$\eta = \frac{R_b}{B} = 0.5\log_2 M \quad \text{bps/Hz} \tag{14}$$

It is apparent from Eq. (14) that the bandwidth efficiency of M-ary PSK is maximized by increasing the number of symbols M for a given bit rate R_b. For example, systems with $M = 4, 8, 16, 32$ give $\eta = 1, 1.5, 2, 2.5$ bps/Hz, respectively. In bandwidth-limited systems one needs to choose a modulation technique that gives the highest bandwidth efficiency. However, there is a penalty to be paid, in terms of the error performance, as explained next.

Table 3. DPSK Detection for the Transmitted Phase Shift Sequence in Table 1

Received phase shift ϕ_n	π	π	π	0	0	π	π	π	0	π	π
Reference phase		π	π	π	0	0	π	π	π	0	π
Recovered bit stream		1	1	0	1	0	1	1	0	0	1

Table 4. DPSK Detection for the Transmitted Phase Shift Sequence in Table 1 with Error in the Fifth Bit

Received phase shift ϕ_n	π	π	π	0	$\underline{\pi}$	π	π	π	0	π	π
Reference phase		π	π	π	0	π	π	π	π	0	π
Recovered bit stream		1	1	0	$\underline{0}$	$\underline{1}$	1	1	0	0	1

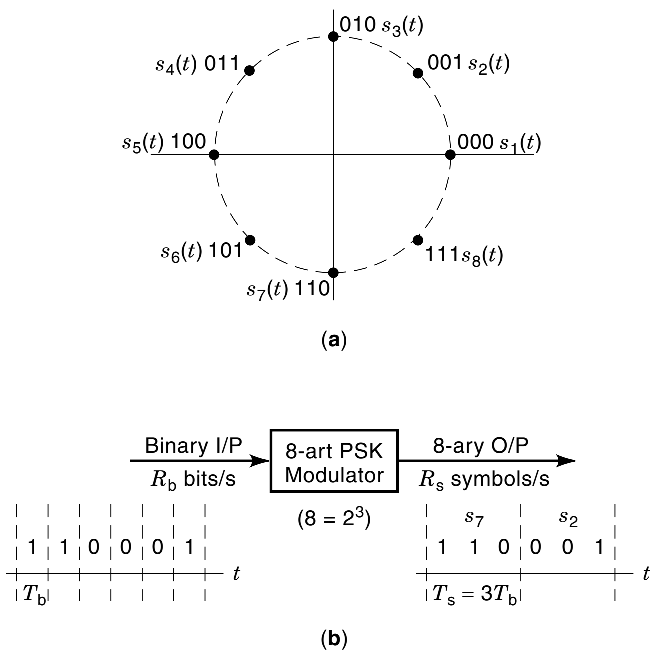

Figure 9. *M*-ary PSK. (a) constellation of 8-ary PSK; (b) input and output data streams for 8-ary PSK signaling.

Error Performance of *M*-ary PSK

Assume the received signal of M-ary PSK in the presence of additive white Gaussian noise is given by

$$r(t) = s_i(t) + n(t) \tag{15}$$

where $n(t)$ is a zero mean and two-sided power spectral density of $N_0/2$ w/Hz. According to the maximum likelihood decision rule, the observation signal space is simply partitioned into M regions, as shown in Fig. 11 for the $M = 8$ case, where

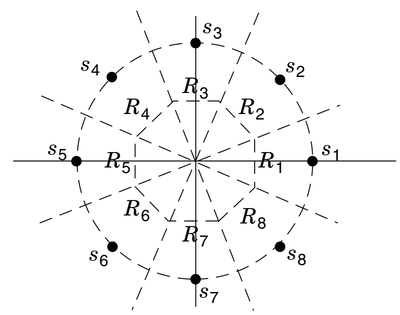

Figure 11. Signal-space constellation diagram for the detection of 8-ary PSK.

R_1 is the decision region corresponding to signal $s_1(t)$, R_2 is the decision region corresponding to $s_2(t)$, and so on. The decision boundaries are shown as lines to divide the signal space into eight equal regions, so as to minimize the probability of symbol error for equally probable symbols. The decision rule used at the receiver is simply to decide $s_i(t)$ was transmitted if the received signal point falls in region R_i. The probability of symbol error, therefore, can be found by integrating the two-dimensional probability density function of noise for each signal in turn over the corresponding error region and averaging the results. However, the probability of bit error is a more meaningful measure of performance. The bit error probability not only depends on the symbol error, but also on the bit-mapping used, because a symbol detection error translates into several bit errors. The probability of bit error expressions for M-ary PSK with any bit-mapping have been derived in (10). However, useful approximation expression for the probability of bit error in M-ary PSK using Gray coding is given (2) as

$$P_e \approx \frac{2}{\log_2 M} \mathrm{Erfc}\left[\sqrt{\frac{2E_b \log_2 M}{N_0}} \sin\left(\frac{\pi}{M}\right)\right] \qquad \text{for} \quad M \geq 4 \tag{16}$$

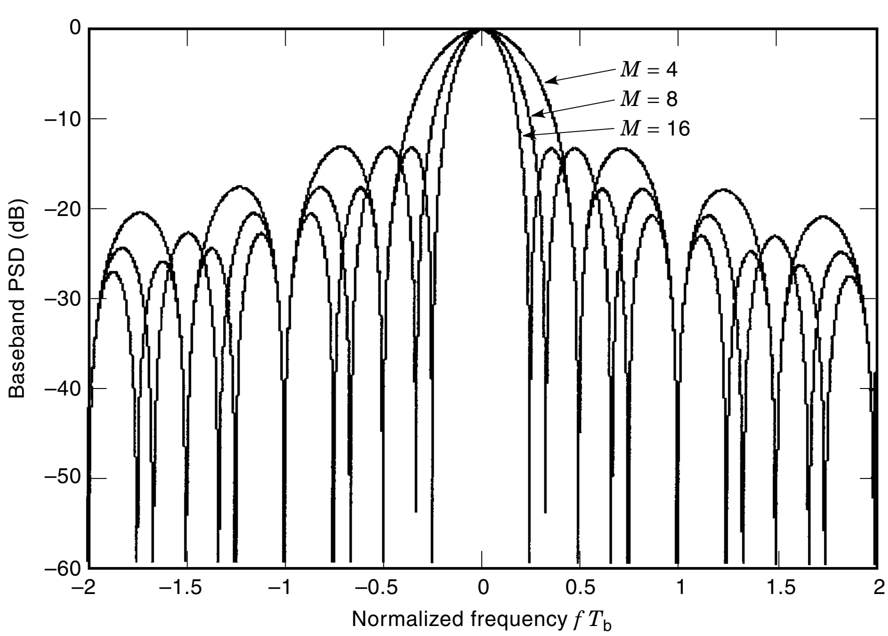

Figure 10. Power spectral density of *M*-ary PSK.

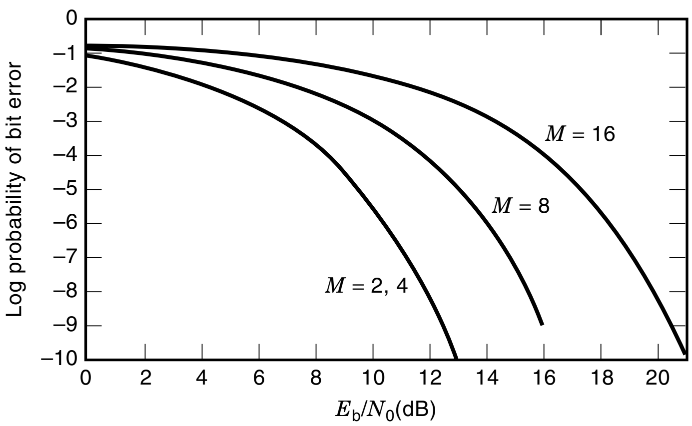

Figure 12. Probability of bit error versus E_b/N_0 for coherent detection of M-ary PSK.

It is noted that this is valid only for $M \geq 4$. For example, substitution of $M = 2$ into Eq. (16) gives twice the P_e predicted by Eq. (7) for BPSK. Figure 12 shows the probability of bit error for various values of M. It can be seen from this figure that the probability of error is the same for $M = 2$ and $M = 4$, but otherwise degrades as M increases. This is due to the fact that more phases in the same bandwidth means crowding more signals into the signal space without increasing the system bandwidth. In other words, as M increases the signals are closer together, making them become more vulnerable to noise; thus we need to increase signal power so that the probability of error is not degraded. Error performance versus bandwidth performance is a fundamental communications trade-off. Therefore, in M-ary PSK, improved bandwidth performance is achieved at the expense of error probability at a given E_b/N_0, or at the expense of increasing E_b/N_0 to keep the same probability of error.

QUADRIPHASE SHIFT KEYING (QPSK)

Quadriphase PSK is also known as quaternary PSK and quadrature PSK. Whereas there are two possible phases 0 or π in BPSK, there are four possible phases in QPSK, one of which is transmitted during every symbol interval $T_s = 2T_b$. Therefore, the input information bit stream is partitioned into blocks of length two bits, to represent each waveform. Now sine and cosine are orthogonal waveforms, which can be combined at the transmitter and completely separated at the receiver. Therefore, two PSK signals can be sent independently in the same bandwidth to represent the two information bits. The modulation process is summarized by the block diagram shown in Fig. 13, in which a QPSK signal is generated by combining an in-phase and a quadrature modulated signal. A generalized expression for the generated QPSK signal can be expressed in terms of this quadrature representation as

$$s_i(t) = d_I \cos(\omega_c t) + d_Q \sin(\omega_c t) \qquad 2nT_b \leq t \leq 2(n+1)T_b \tag{17}$$

where d_I and d_Q are the in-phase and quadrature data pulses $(= \pm 1)$ and $d_I \cos(\omega_c t)$ and $d_Q \sin(\omega_c t)$ are the in-phase and quadrature signal components, respectively. Equivalently, the QPSK signal can be expressed by a trigonometric representation as

$$s_i(t) = A \cos(\omega_c t + \phi_i) \qquad 2nT_b \leq t \leq 2(n+1)T_b \tag{18}$$

where ϕ_i are the four possible phases.

The modulation process is best illustrated by an example. Consider the bit stream 11000111 shown in Fig. 14(a). After the serial to parallel conversion the input bit stream is split into one channel of even-numbered bits, namely, 1001, and one channel of odd-numbered bits, namely, 1011. As shown in Fig. 14(b) the bit period in each of these channels is extended from the original value of T_b to become $2T_b$. The two bit streams then undergo a unipolar to bipolar conversion to produce the waveforms of levels +1 and −1. Thus, as shown in Fig. 14(c), the channel of even-numbered bits gives the sequence (+1 −1 −1 +1) and the channel of odd-numbered bits gives the sequence (+1 −1 1 1). The channel of even-numbered bits is then multiplied with the in-phase component of a carrier $\cos(\omega_c t)$ to create what is termed an in-phase or I-channel, and the channel of odd-numbered bits is multiplied with the quadrature component of a carrier, which is $\sin(\omega_c t)$, to create what is termed a quadrature or Q-channel. Finally, the two are added. The resulting four signals are thus

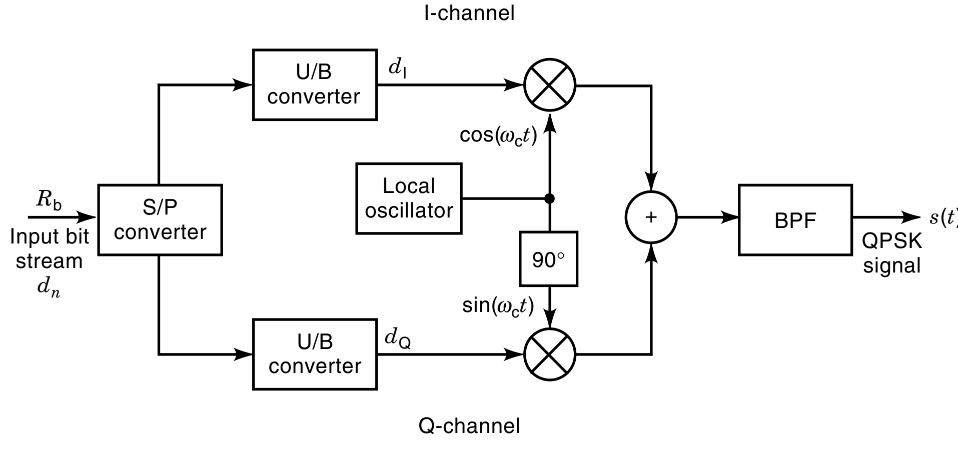

Figure 13. Block diagram of QPSK modulator.

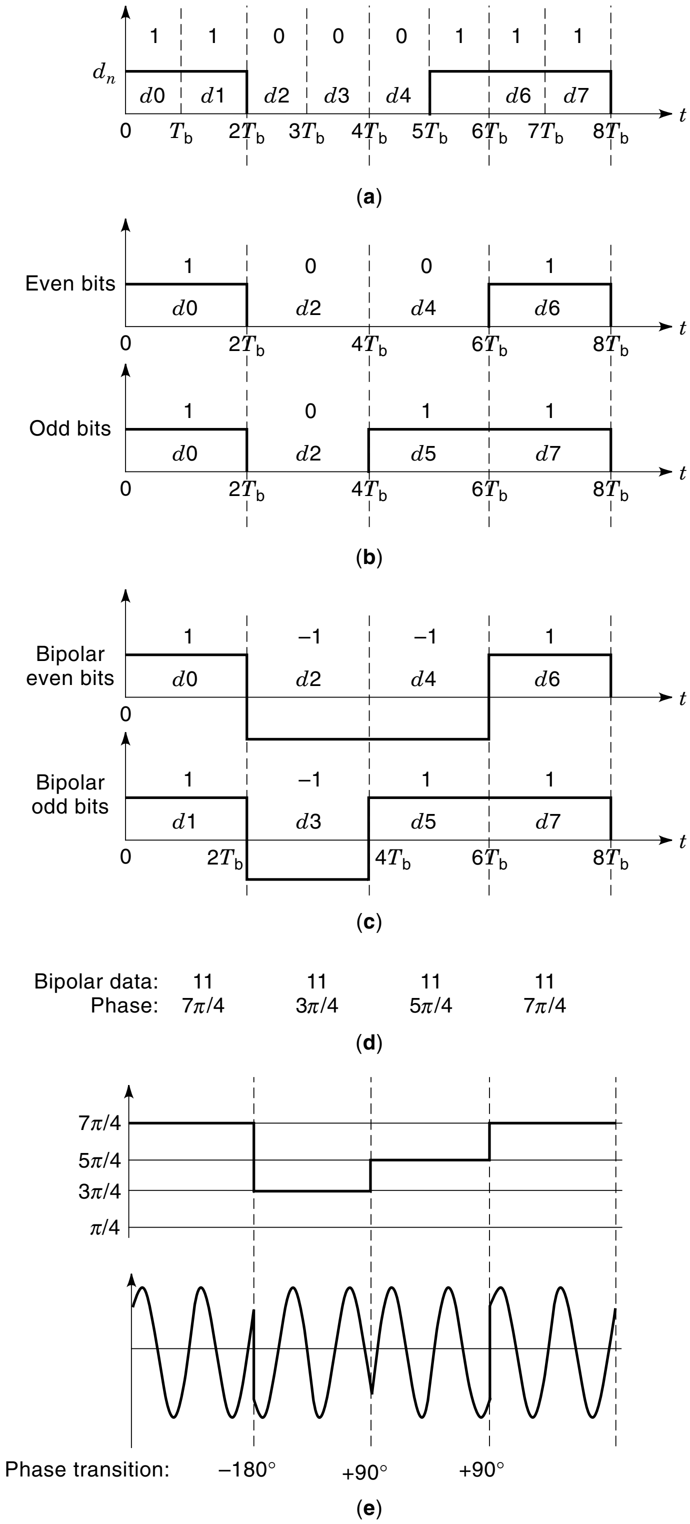

Figure 14. QPSK waveforms for the bit stream 11000111. (a) input data stream d_n; (b) serial-to-parallel conversion of data stream into I-data stream d_I (even bits) and Q-data stream d_Q (odd bits). The symbol period $T_s = 2T_b$; (c) unipolar to bipolar conversion of the I and Q streams; (d) bipolar data stream from the I- and Q-channels during T_s period and the assigned phases; (e) phase transitions and the corresponding time-domain waveform.

Table 5. Bits-to-Phase Mapping for QPSK

Bits	Phase
1 0	$\pi/4$
0 0	$3\pi/4$
0 1	$5\pi/4$
1 1	$7\pi/4$

transmitted as

$$
\begin{aligned}
\cos(\omega_c t) + \sin(\omega_c t) &= \frac{1}{\sqrt{2}} \cos\left(\omega_c t + \frac{7\pi}{4}\right) \\
-\cos(\omega_c t) - \sin(\omega_c t) &= \frac{1}{\sqrt{2}} \cos\left(\omega_c t + \frac{3\pi}{4}\right) \\
-\cos(\omega_c t) + \sin(\omega_c t) &= \frac{1}{\sqrt{2}} \cos\left(\omega_c t + \frac{5\pi}{4}\right) \\
\cos(\omega_c t) + \sin(\omega_c t) &= \frac{1}{\sqrt{2}} \cos\left(\omega_c t + \frac{7\pi}{4}\right)
\end{aligned}
\tag{19}
$$

Thus the phase of the QPSK waveform varies with time, as listed in Fig. 14(d) and illustrated in Fig. 14(e). In this particular example, only three phases occur, but, in general, the fourth phase of $\pi/4$ would also be involved. It will be noted that the binary data symbol 11 converts to $7\pi/4$, the symbol 00 to $3\pi/4$, and the symbol 01 to $5\pi/4$. The missing symbol 10 would convert to $\pi/4$. Figure 14(e) also shows the waveform itself.

The bits-to-symbol mapping used here is shown in Table 5 and the corresponding signal-state constellation for the QPSK signal is shown in Fig. 15(a). It is possible for any symbol to

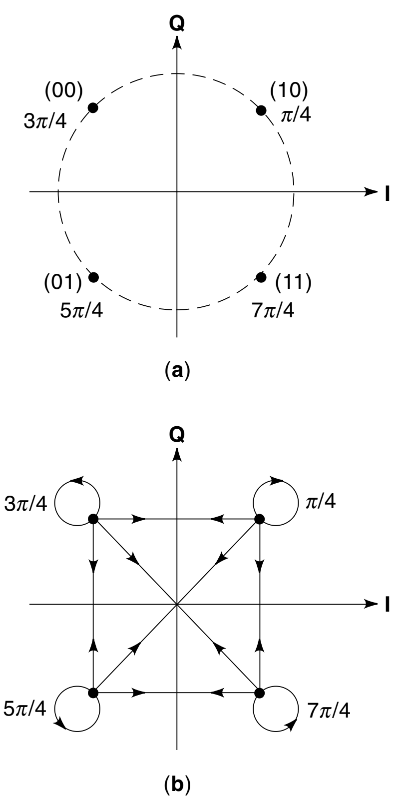

Figure 15. QPSK. (a) signal-space constellation diagram of Gray-coded QPSK signal; (b) allowed phase transitions: 0°, ±90°, ±180°.

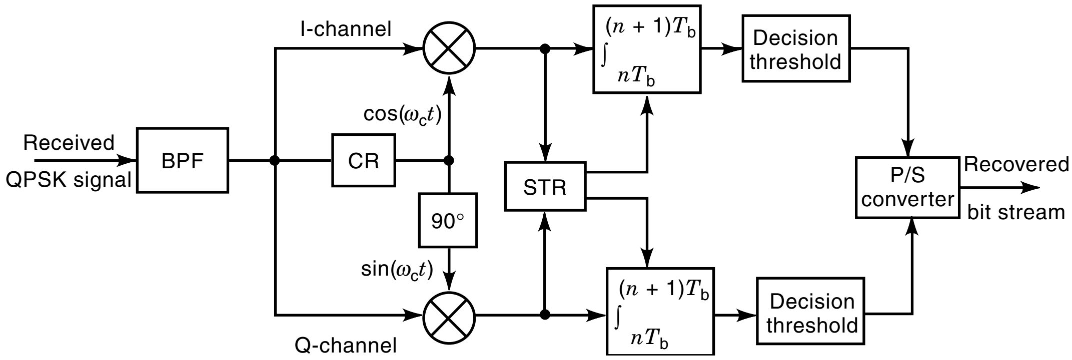

Figure 16. Block diagram of QPSK coherent detector.

be followed by any other symbol and, hence, for any phase to be followed by any other phase. This is indicated by the directional arrows in Fig. 15(b)), which demonstrate possible phase transitions of 0°, ±90°, ±180° at the end of each symbol interval. It will be noted that adjacent points around the circle only differ by 1 bit. This means that the assignments of phases to information bits is achieved here according to Gray coding (rather than binary), such that a single symbol error occurrence would correspond to a single bit error.

The detection process follows a similar procedure but in reverse, and is illustrated by the receiver block diagram shown in Fig. 16. The receiver resolves the received signal into I and Q components using two correlators. The input band-pass filter reduces out-of-band noise and adjacent channel interference. Following a carrier recovery (CR) circuit to provide $\cos(\omega_c t)$ and, hence, $\sin(\omega_c t)$ following a 90° phase shifter, the signals $\cos(\omega_c t)$ and $\sin(\omega_c t)$ are multiplied with the incoming QPSK signal. The multiplier outputs are then integrated over the symbol period using a symbol timing recovery (STR) circuit to provide the correct timing for this integration. Because the transmit signal is composed of two orthogonal channels, the output of the upper I-channel integrator in the receiver is unaffected by the Q-channel component. Similarly, the output of the lower Q-channel integrator in the receiver is unaffected by the I-channel component. At the end of each symbol period, at a time $2(n + 1)T_b$ provided by the STR circuit, the integrator outputs are sampled and held. These signals are then passed into a zero threshold circuit, which decides what was transmitted in the I- and Q-channels. By this means the two waveforms of Fig. 14(b) are regained; finally, these are recombined using a parallel-to-serial converter to produce the original bit stream.

Performance of QPSK

In contrast with possible phase transitions every T_b seconds, in the case of BPSK the possible phase transitions in a QPSK system occur every $2T_b$ seconds. It follows that the bandwidth of QPSK for the same bit rate is halved, and that the bandwidth efficiency is therefore double that of QPSK, to become 1 bps/Hz. The baseband power spectral density for QPSK has already been plotted in Fig. 10 and described with respect to BPSK and other M-ary PSK techniques. If we consider just the I-channel of the QPSK system and neglect any interference from the transmitted Q-channel component, we have exactly the same probability of error for a given ratio of E_b/N_0 as that with a BPSK system. Since the I- and Q-channels contribute independently to the final reconstituted bit stream, it follows that the complete QPSK system has the same probability of error for a given ratio of E_b/N_0 as does a BPSK system, as shown in Fig. 12. Thus it maintains error performance while achieving a greater bandwidth efficiency. The only penalty is an increase in the complexity of the modulation and demodulation process.

In BPSK the waveform contains 180° phase shifts that cause the spectrum to spread, thus introducing the risk of

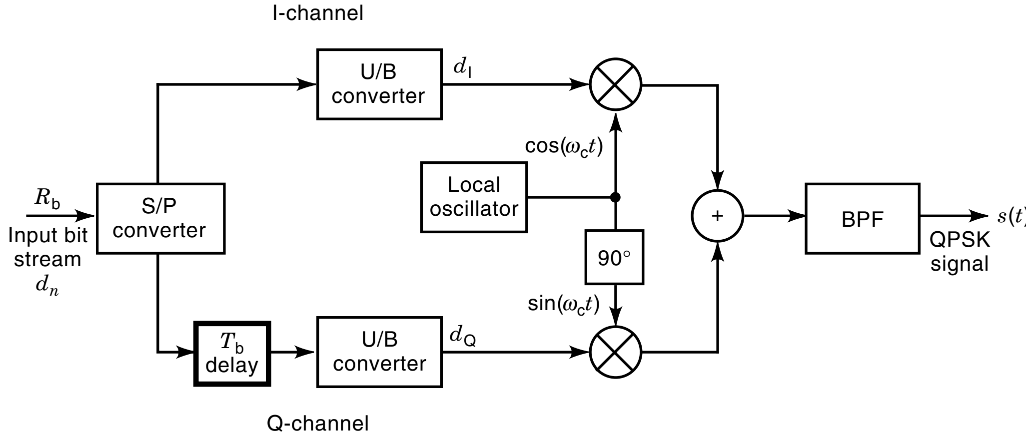

Figure 17. Block diagram of OQPSK modulator.

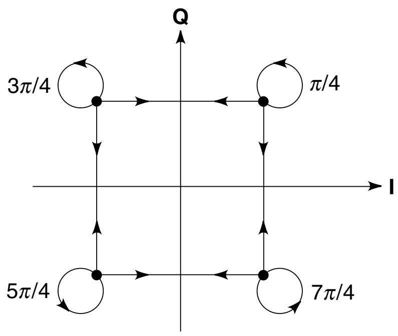

Figure 18. Signal-space constellation diagram and phase states transitions of OQPSK. Allowed phase state transitions are limited to 0°, ±90°.

adjacent channel interference. Band-limiting can be used prior to transmission, but the effect of this is to introduce envelope nulls at the phase reversals. The effect of this is that subsequent amplification by nonlinear amplifiers causes spectral spreading, thus *re*introducing the problem of adjacent channel interference. QPSK is somewhat better, in that, although many of the phase transitions are ±90°, some of the phase transitions are still 180°. A technique known as offset QPSK involves a relatively minor modification and can greatly reduce the problem of spectral spreading.

OFFSET QUADRIPHASE SHIFT KEYING (OQPSK)

OQPSK has some of the features of BPSK and some of the features of QPSK. It is like BPSK, in that the phase transmitted can change every bit period, rather than every two bit periods. It is like QPSK, in that the signal transmitted can have any one of four possible phases. The main difference of OQPSK with QPSK is that the phase is precluded from changing by more than 90° at the transition points between bits. The important advantage of this is that band-limiting in the transmitter does not cause any large change in the envelope. Subsequent amplification in an efficient nonlinear power amplifier causes some spectral spreading into adjacent channels, but very much less than with QPSK.

A block diagram of OQPSK modulator is shown in Fig. 17. This is the same as for QPSK, shown in Fig. 12, except that one bit delay is introduced in the Q channel to offset the data by one bit period relative to the I channel. As with QPSK, the channel of even-numbered bits is multiplied with a carrier $\cos(\omega_c t)$ to create an I-channel, and the channel of odd-numbered bits is multiplied with a carrier $\sin(\omega_c t)$ to create a Q channel. Again these two channels are added. Because of the offset there can be a phase transition every bit period, caused by a 180° transition in one of the two channels and no transition in the other. Phase transitions can occur every bit period but they are limited to 0°, ±90°. In terms of a signal-state constellation diagram, possible phase transitions are shown in Fig. 18. The modulation process is best illustrated by the example in Fig. 19 for the same data sequence used for QPSK. The I and Q data alignment are shown and the corresponding waveform very clearly shows the phase transitions. As shown in Fig. 20, the demodulator is similar to the QPSK demodulator of Fig. 16, except that the integration of the Q-channel is T_b later than that of the I-channel, to compensate for the offset delay of T_b that was introduced into the Q-channel of the modulator.

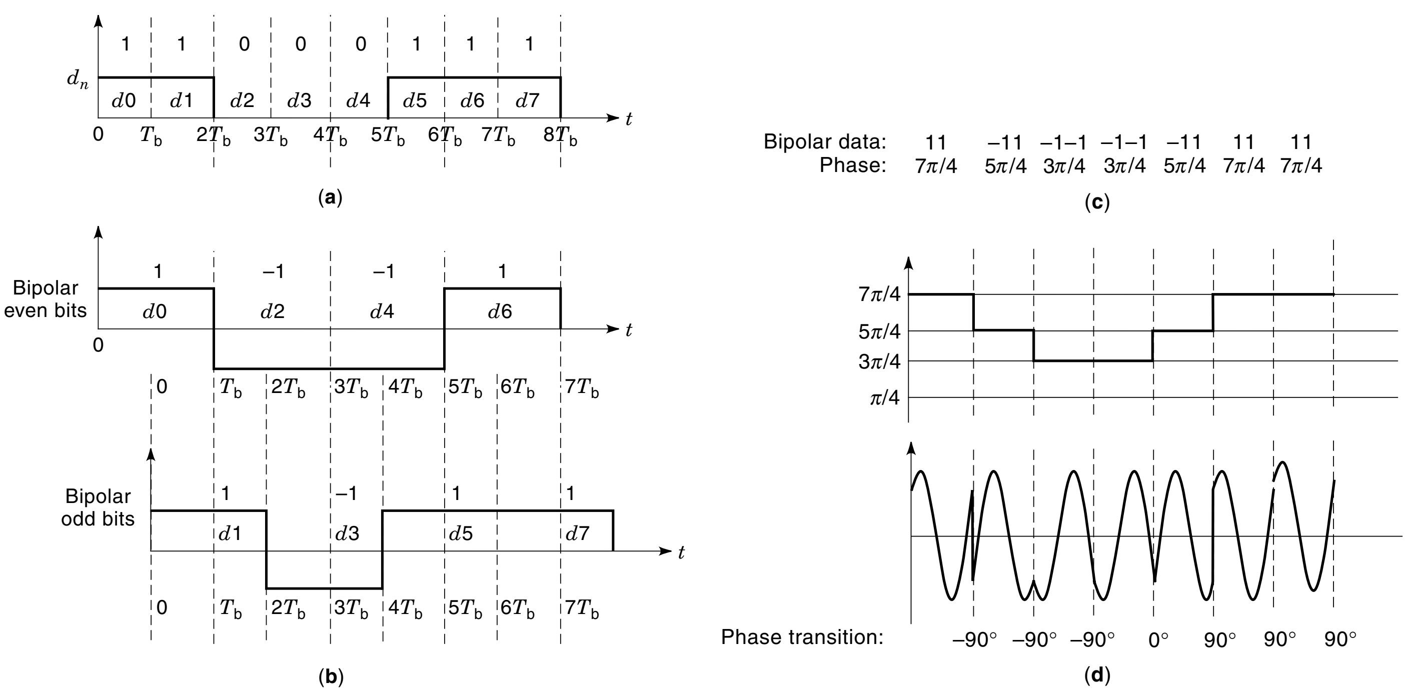

Figure 19. OQPSK waveforms for the bit stream 11000111. (a) input data stream; (b) bipolar I-data stream (even bits) and Q-data stream (odd bits) with T_b delay in the Q stream. The symbol period $T_s = 2T_b$; (c) bipolar data stream from the I- and Q-channels during every T_b period and the assigned phases; (d) phase transitions and the corresponding time-domain waveform.

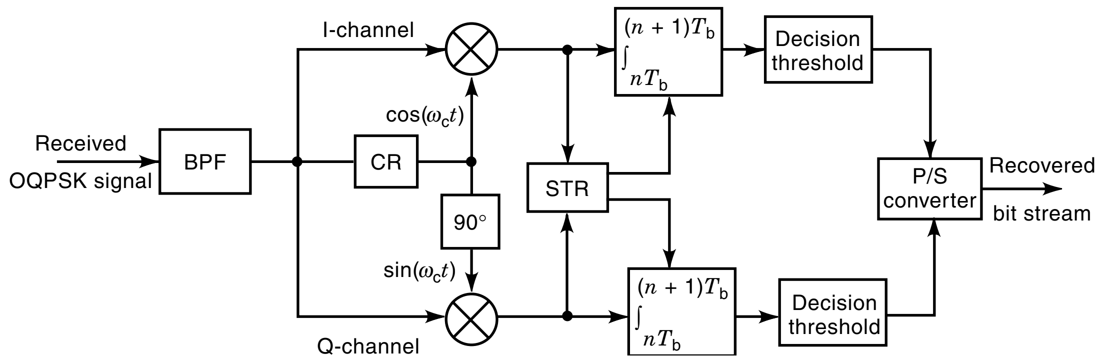

Figure 20. Block diagram of coherent detection of OQPSK.

Performance of OQPSK

The baseband power spectral density of unfiltered OQPSK is the same as of QPSK (11,12) and, using Eq. (12) with $k = 2$, is given by

$$S(f) = 4P_c T_b \left(\frac{\sin(2\pi f T_b)}{2\pi f T_b}\right)^2 \tag{20}$$

Although the spectral characteristics of QPSK and OQPSK are identical for the unfiltered case, the important difference is that the avoidance of any 180° phase changes means that band-limiting of OQPSK will not introduce any major dips in the waveform envelope, and subsequent nonlinear amplification will not cause any substantial spectral spreading to cause severe adjacent channel interference.

Integration of the power spectral density given in Eq. (20), to determine the fraction of the total transmitted power lying outside a given bandwidth, produces the results of Table 6. The final entry is considered particularly important since it corresponds to the FCC criterion for the bandwidth containing 99% of the signal energy. It should be mentioned here that, since offsetting (staggering) the bit streams does not change the orthogonality of the I and Q carriers, OQPSK has the same bit error performance as BPSK and QPSK (11,12).

MINIMUM SHIFT KEYING (MSK)

Although OQPSK provides a major improvement over QPSK in terms of out-of-band interference, there is still some caused by the abrupt phase transitions. A small modification to OQPSK can lead to the total avoidance of abrupt transitions and leads to a still further reduction of spectral spreading. Minimum Shift Keying (MSK) (13) is one method of obtaining a digital passband signal in which the phase of the waveform is continuous at all times.

Table 6. Fractional Out-of-Band Power for OQPSK

RF Bandwidth, B (Hz)	Fractional Out-of-Band Power
$1/T_b$	10% (−10 dB)
$1.5/T_b$	6.3% (−12 dB)
$2/T_b$	5.0% (−13 dB)
$3/T_b$	3.2% (−15 dB)
$4/T_b$	2.2% (−16.5 dB)
$8/T_b$	1% (−20 dB)

The modulator can be implemented using the quadrature structure shown in Fig. 21, similar to that used in OQPSK, except that before multiplying the I-channel data stream by the carrier signal $\cos(\omega_c t)$, it is first multiplied by another low-frequency cosinusoid $\cos(\pi t/2T_b)$, whose period corresponds to that of 4 bits. In a similar manner, the Q channel is multiplied by $\sin(\pi t/2T_b)$ before multiplication by the carrier signal $\sin(\omega_c t)$. The resulting MSK signal after summing the I and Q channels is

$$s(t) = d_I \cos\left(\frac{\pi t}{2T_b}\right)\cos(\omega_c t) + d_Q \sin\left(\frac{\pi t}{2T_b}\right)\sin(\omega_c t) \tag{21}$$

where d_I and d_Q represent the bipolar even and odd data bits, and $\cos(\pi t/2T_b)$ and $\sin(\pi t/2T_b)$ are the sinusoidal weighting functions. It should be noted here that this MSK signal can also be thought of as a continuous phase binary frequency shift keying (CPFSK) signal (1) with signaling frequencies $f_1 = f_c + 1/4T_b$ and $f_2 = f_c - 1/4T_b$:

$$s_i(t) = A\cos\left[2\pi\left(f_c + \frac{d_n}{4T_b}\right)t + x_n\right] \qquad nT_b \le t \le (n+1)T_b \tag{22}$$

where d_n is the bipolar data to be transmitted (±1), and x_n is a phase constant over the nth bit interval ($x_n = 0$ or π), chosen such that the phase of the waveform is continuous at $t = kT_b$. Note that when $d_n = 1$ the transmitted signaling frequency is f_1, and when $d_n = -1$ the transmitted frequency is f_2.

Figure 22 shows the MSK waveforms for the same sequence of data used in the OQPSK. After multiplication of the bipolar even bits by the carrier signal $\cos(\omega_c t)$, and of the bipolar odd bits by $\sin(\omega_c t)$, the waveforms shown in Fig. 22(c) are obtained. Finally, these modified I- and Q-channel signals are added to give the MSK waveform of Fig. 22(d). Because the envelopes of the I- and Q-channels have been attenuated to zero at the end of their symbol periods, this final waveform has the required constant envelope and a continuous variation of phase shown in Fig. 22(d). At the receiver the demodulator can be implemented using the same coherent detection method as in the QPSK and OQPSK, except that the I and Q signals are multiplied by the sinusoidal weighting functions after multiplication by the I and Q carriers, as shown in Fig. 23.

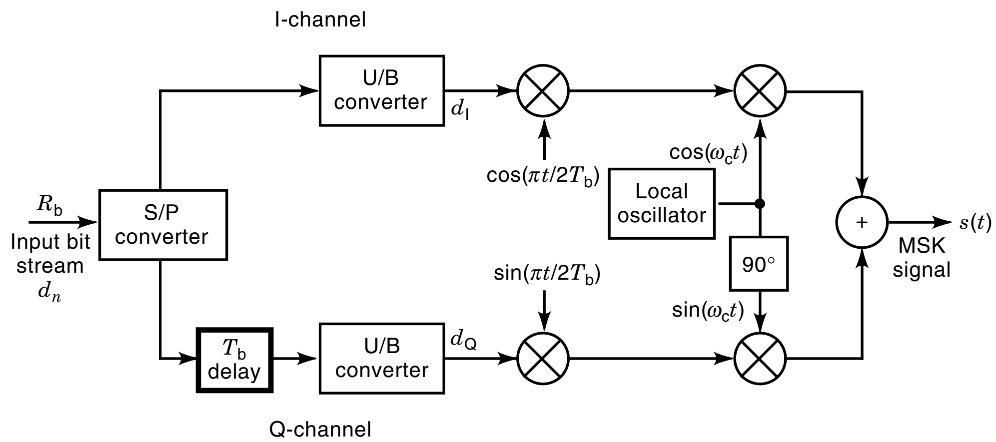

Figure 21. Block diagram of MSK modulator.

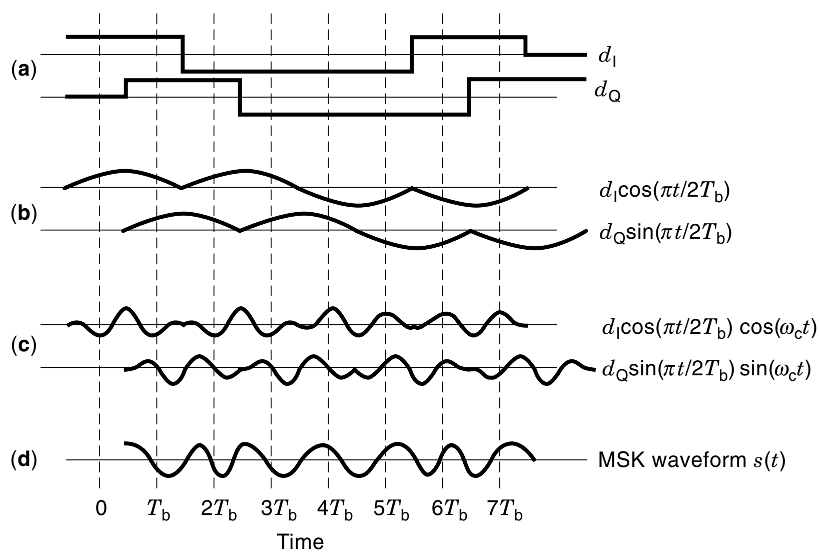

Figure 22. MSK waveforms for the bit stream 11000011. (a) in-phase bipolar data, d_I, and bipolar quadrature data, d_Q; (b) I data stream multiplied by $\cos(\pi t/2T_b)$ and Q data stream multiplied by $\sin(\pi t/2T_b)$; (c) Modified I data stream multiplied by the in-phase component of the carrier and modified Q data stream multiplied by the quadrature component of the carrier; (d) resulting MSK waveform.

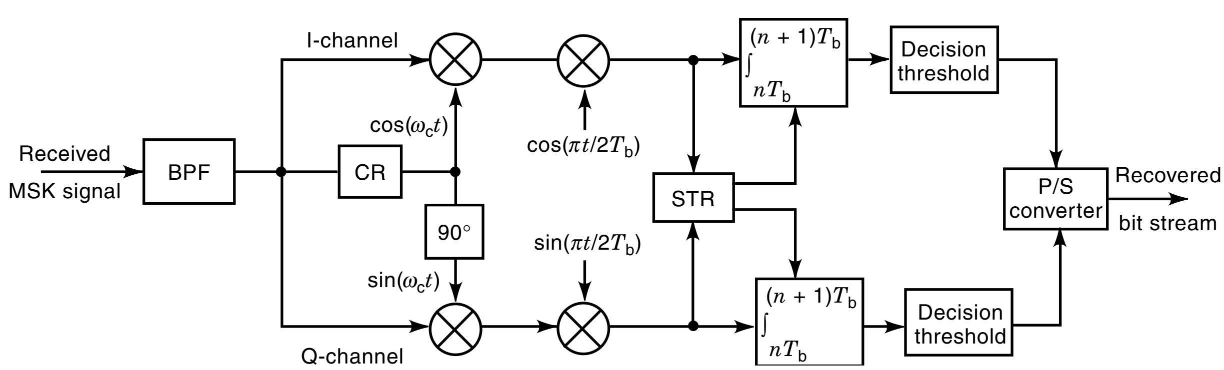

Figure 23. Block diagram of MSK detector.

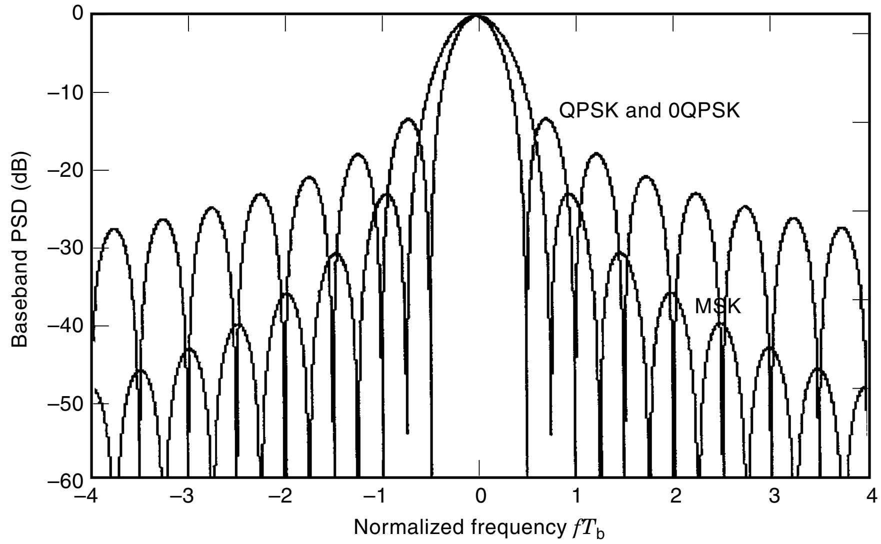

Figure 24. Power spectral density of QPSK, OQPSK, and MSK.

Performance of MSK

The baseband power spectral density (11–13) can be shown to be given by

$$S(f) = \frac{16P_c T_b[1 + \cos(4\pi f T_b)]}{\pi^2(1 - 16T_b^2 f^2)^2} \tag{23}$$

This power spectral density is generated and is shown in Fig. 24. It can be seen from this figure that the main lobe null-to-null bandwidth is $1.5/T_b$, which is 1.5 times larger than that of OQPSK. However, its sidelobes are much lower than OQPSK. Integration of this power spectral density to determine the fraction of the total transmitted power lying outside a given bandwidth produces the results of Table 7. The −20 dB point corresponding to the FCC criterion for the bandwidth containing 99% of the signal energy is now only $1.18/T_b$. This is a major improvement compared with the $8/T_b$ of OQPSK. Finally, the probability of bit error performance for ideal MSK detection is identical to that of QPSK and OQPSK, since orthogonality between I- and Q-channels is preserved (11–13).

CONCLUSIONS AND FURTHER TECHNIQUES

There is ever-increasing demand for bandwidth-efficient constant envelope digital radio transmission techniques with good power efficiency. With *M-ary* phase shift keying increasing the value of *M* has the advantage of increasing the bandwidth efficiency, but at the cost of requiring higher signal-to-noise ratios for the same error rate. Although the probability of error performance is acceptable for BPSK or QPSK, the discontinuous phase of these signals results in an envelope fluctuation, which gives rise to undesirable out-of-band radiation. The MSK technique, with its phase continuity and constant envelope, provides an identical probability of error performance as BPSK or QPSK, but with lower sidelobes. However, the MSK spectrum has a wider main lobe bandwidth than QPSK. Therefore, to improve the bandwidth efficiency, the MSK may be filtered before modulation in a way that keeps the signal envelope constant. Using a Gaussian-shaped filter leads to Gaussian-filtered MSK (GMSK). This gives an improvement in bandwidth efficiency over MSK. It also has low-level sidelobes and this, in turn, translates into decreased adjacent channel interference. This has led to its use in the European digital cellular mobile communication system (GSM) (9). However, a technique known as $\pi/4$-shift QPSK is used in the second-generation digital cellular mobile radio systems in North America and Japan (9). This technique is a compromise between QPSK and OQPSK, with a maximum phase change of 135° compared with 180° for QPSK and 90° for OQPSK. $\pi/4$-shift QPSK provides the same bandwidth efficiency as QPSK but with less envelope fluctuation. MSK is considered to be a special case from a class of modulation techniques known as continuous phase modulation (CPM). CPM techniques, in which the phase of the signal is constrained to be continuous (unlike that of PSK signals), has been much researched for satellite and terrestrial digital radio systems to provide a good trade-off between the bandwidth, power, and complexity (14–16).

There is always considerable demand for techniques to improve both the power and bandwidth efficiency of modulation techniques. For example, the use of channel coding techniques for error detection and correction are well known to provide power efficiency but at the cost of bandwidth expansion and system complexity. With the advent of advanced techniques such as trellis coded modulation (17), which combines channel coding and modulation, power efficiency is achieved without bandwidth expansion and with acceptable

Table 7. Fractional Out-of-Band Power for MSK

RF Bandwidth, B (Hz)	Fractional Out-of-Band Power
$1/T_b$	2.2% (−16.5 dB)
$1.18/T_b$	1% (−20 dB)
$1.5/T_b$	0.5% (−23 dB)
$2/T_b$	0.2% (−27 dB)
$3/T_b$	0.063% (−32 dB)
$4/T_b$	0.025% (−36 dB)
$8/T_b$	0.0031% (−45 dB)

level of complexity. Much research work is being carried out in this area (18).

BIBLIOGRAPHY

1. J. B. Anderson, T. Aulin, and C.-E. Sunberg, *Digital Phase Modulation,* New York: Plenum, 1986.
2. B. Sklar, *Digital Communications Fundamental and Applications,* Englewood Cliffs, NJ: Prentice-Hall, 1988.
3. I. A. Glover and P. M. Grant, *Digital Communications,* Piscataway, NJ: Prentice-Hall, 1998.
4. J. G. Proakis, *Digital Communications,* 3rd ed., New York: McGraw-Hill, 1995.
5. R. E. Ziemer and R. L. Peterson, *Introduction to Digital Communication,* Indianapolis, IN: Macmillan, 1992.
6. L. W. Couch II, *Digital and Analog Communication Systems,* 5th ed., Prentice-Hall, 1997.
7. S. Haykin, *Communication Systems,* 3rd ed., New York: Wiley, 1994.
8. F. Amoroso, The bandwidth of digital data signals, *IEEE Commun. Mag.,* **18** (6): 13–24, 1980.
9. A. H. Aghvami, Digital modulation techniques for mobile and personal communication systems, *IEE Electron. Commun. Eng. J.,* **5** (3): 125–132, 1993.
10. M. I. Irshid and I. S. Salous, Bit error probability for coherent M-ary PSK systems, *IEEE Trans. Commun.,* **39**: 349–352, 1991.
11. S. A. Gronemeyer and A. L. McBride, MSK and offset QPSK modulation, *IEEE Trans. Commun.,* **COM-24**: 809–820, 1976.
12. D. H. Morais and K. Feher, Bandwidth efficiency and probability of error performance of MSK and offset QPSK systems, *IEEE Trans. Commun.,* **COM-27**: 1794–1801, 1979.
13. S. Pasupathy, Minimum shift keying: A spectrally efficient modulation, *IEEE Commun. Mag.,* **17** (4): 14–22, 1979.
14. C.-E. Sundberg, Continuous phase modulation, *IEEE Commun. Mag.,* **24** (4): 25–38, 1986.
15. I. Sasase and S. Mori, Multi-h phase-coded modulation, *IEEE Commun. Mag.,* **29** (7): 46–56, 1991.
16. J. B. Anderson and C.-E. Sundberg, Advances in constant envelope coded modulation, *IEEE Commun. Mag.,* **29** (7): 36–45, 1991.
17. G. Ungerboeck, Channel coding with multilevel/phase signals, *IEEE Trans. Inf. Theory,* **IT-28**: 55–67, 1982.
18. E. Biglieri et al., *Introduction to Trellis-Coded Modulation with Applications,* Indianapolis, IN: Macmillan, 1991.

FALAH H. ALI
PHILIP DENBIGH
University of Sussex

PHASE SHIFT MEASUREMENT. See POWER FACTOR MEASUREMENT.