Bandpass modulation of a carrier signal

11.1 Introduction

Modulation, at its most general, refers to the modification of one signal's characteristics in sympathy with another signal. The signal being modified is called the carrier and, in the context of communications, the signal doing the modifying is called the information signal. Chapter 5 described pulse modulation in which the carrier is a rectangular pulse train and the characteristics adjusted are, for example, pulse amplitude, or, in PCM, the coded waveform. In this chapter, intermediate or radio frequency (IF or RF) bandpass modulation is described, in which the carrier is a sinusoid and the characteristics adjusted are amplitude, frequency or phase.

The principal reason for employing IF modulation is to transform information signals (which are usually generated at baseband) into signals with more convenient (bandpass) spectra. This allows:

1. Signals to be matched to the characteristics of transmission lines or channels.
2. Signals to be combined using frequency division multiplexing and subsequently transmitted using a common physical transmission medium.
3. Efficient antennas of reasonable physical size to be constructed for radio communication systems.
4. Radio spectrum to be allocated to services on a rational basis and regulated so that interference between systems is kept to acceptable levels.

11.2 Spectral and power efficiency

Different modulation schemes can be compared on the basis of their spectral [1] and power efficiencies. Spectral efficiency is a measure of information transmission rate per Hz of

[1] See footnote on first page (pp. 260) of Chapter 8.

bandwidth used. A frequent objective of the communications engineer is to transmit a maximum information rate in a minimum possible bandwidth. This is especially true for radio communications in which radio spectrum is is a scarce, and therefore valuable, resource. The appropriate units for spectral efficiency are clearly bit/s/Hz. It would be elegant if an analogous quantity, i.e. power efficiency, could be defined as the information transmission rate per W of received power. This quantity, however, is not useful since the information received (correctly) depends not only on received signal power but also on noise power.

It is thus carrier to noise ratio (CNR) or equivalently the ratio of symbol energy to noise power spectral density (NPSD), E/N_0, which must be used to compare the power efficiencies of different schemes. It is legitimate, however, to make comparisons between different digital communications systems on the basis of the relative signal power needed to support a given received information rate assuming identical noise environments. In practice this usually means comparing the signal power required by different modulation schemes to sustain identical BERs for identical transmitted information rates.

11.3 Binary IF modulation

Binary IF modulation schemes represent the simplest type of bandpass modulation. They are easy to analyse and occur commonly in practice. Each of the basic binary schemes is therefore examined below in detail. The ideal BER performance of these modulation schemes could be found directly using a general result, equation (8.61), from Chapter 8. It is instructive, however, to obtain the probability of error, P_e, formula for each scheme by considering matched filtering or correlation detection (see section 8.3) as an ideal demodulation process followed by an ideal, baseband, sampling and decision process.

11.3.1 Binary amplitude shift keying (and on-off keying)

In binary amplitude shift keyed (BASK) systems the two digital symbols, zero and one, are represented by pulses of a sinusoidal carrier (frequency, f_c) with two different amplitudes A_0 and A_1. In practice, one of the amplitudes, A_0, is invariably chosen to be zero resulting in on-off keyed (OOK) IF modulation, i.e.:

$$f(t) = \begin{cases} A_1 \, \Pi(t/T_o) \cos 2\pi f_c t, & \text{for a digital 1} \\ 0, & \text{for a digital 0} \end{cases} \qquad (11.1)$$

where T_o is the symbol duration (as used in Chapters 6 and 8) and Π is the rectangular pulse function.

An OOK modulator can be implemented either as a simple switch, which keys a carrier on and off, or as a double balanced modulator (or mixer) which is used to multiply the carrier by a baseband unipolar OOK signal. A schematic diagram of the latter type of modulator is shown with rectangular pulse input and output waveforms, spectra and allowed phasor states in Figure 11.1. The modulated signal has a DSB spectrum centred

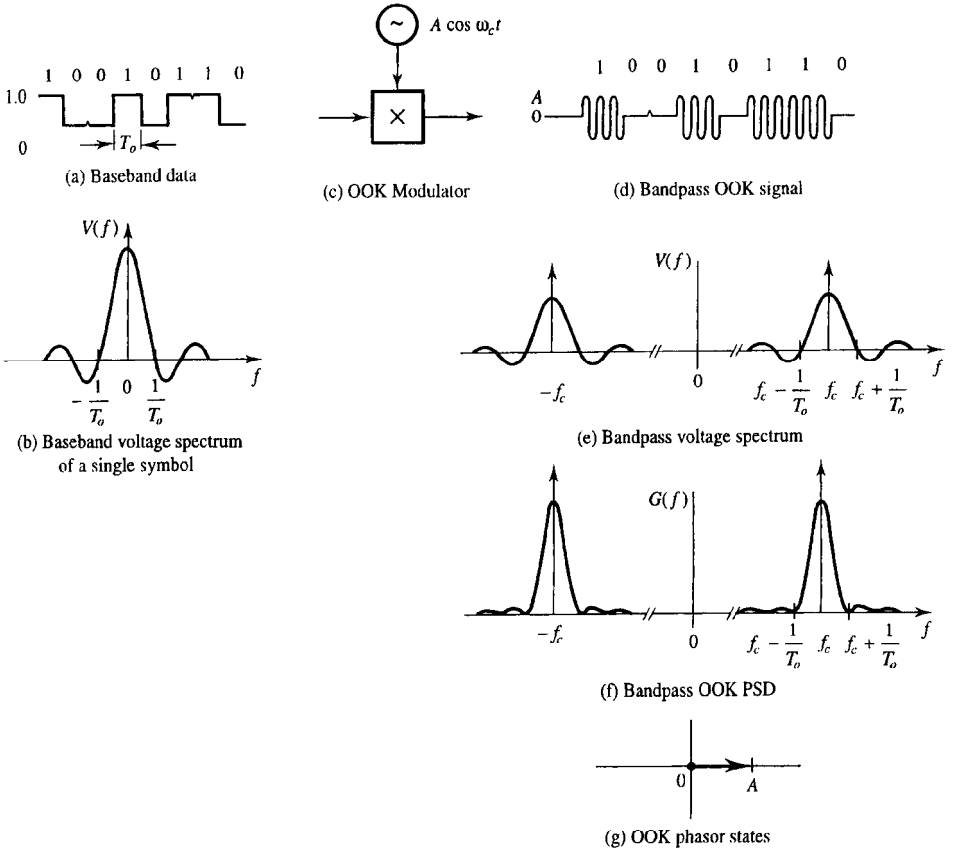
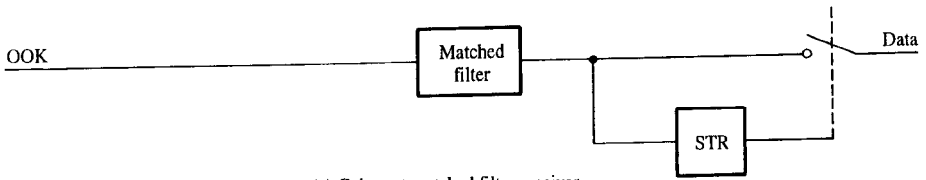
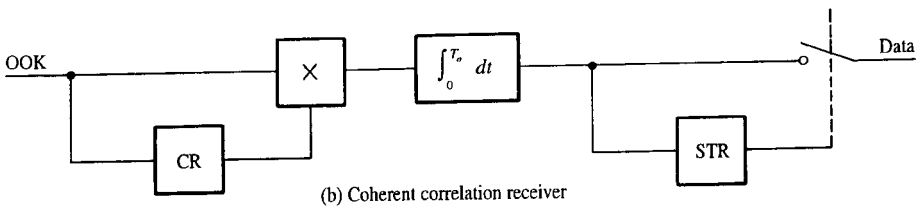


Figure 11.1 *On-off keyed (OOK) modulator, waveforms, spectra and phasor states.*

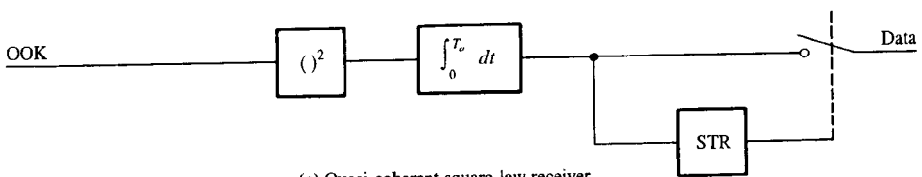
on $\pm f_c$ (Figure 11.1(e) and (f)) and, since a constant carrier waveform is being keyed, the OOK signal has two phasor states, 0 and $A = A_1$ (Figure 11.1(g)). Detection of IF OOK signals can be coherent or incoherent. In the former case a matched filter or correlator is used prior to sampling and decision thresholding (Figure 11.2(a),(b)). In the latter (more common) case envelope detection is used to recover the baseband digital signal followed by centre point sampling or integrate and dump (I + D) detection (Figure 11.2(d)). (The envelope detector would normally be preceded by a bandpass filter to improve the CNR.) Alternatively, an incoherent detector can be constructed using two correlation channels configured to detect inphase (I) and quadrature (Q) components of the signal followed by I and Q squaring operations and a summing device (Figure 11.2(e)). This arrangement overcomes the requirement for precise carrier phase synchronisation (c.f. Figure 11.2(b)). Whilst such a dual channel incoherent detector may seem unnecessarily complicated compared with the other incoherent detectors, recent advances in VLSI technology mean



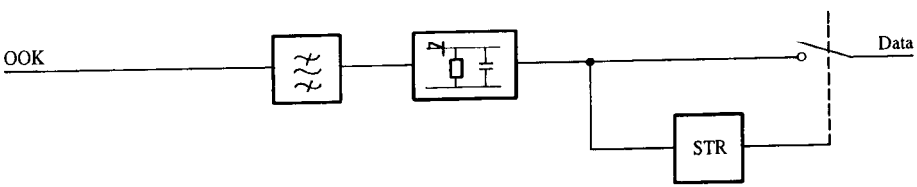
(a) Coherent matched filter receiver



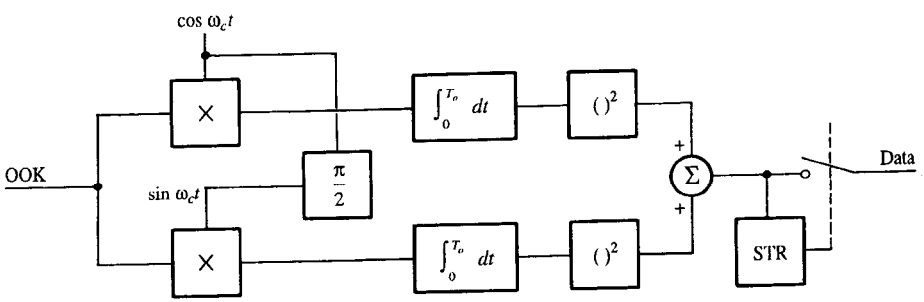
(b) Coherent correlation receiver



(c) Quasi-coherent square-law receiver



(d) Incoherent envelope receiver



(e) Incoherent quadrature receiver

Figure 11.2 *Coherent and incoherent bandpass OOK receivers.*

that they can often be implemented (digitally) as smaller, lighter and cheaper components than the filters and envelope detectors used in more traditional designs. (Symbol timing recovery is discussed in section 6.7 and carrier recovery is discussed later in this chapter.)

The decision instant voltage, $f(nT_o)$, at the output of an OOK matched filter or correlation detector (see equation (8.37)) is:

$$f(nT_o) = \begin{cases} k\,E_1, & \text{digital 1} \\ 0, & \text{digital 0} \end{cases} \tag{11.2}$$

where E_1 (V^2s) is the normalised energy contained in symbol 1 and k has units of Hz/V. The normalised noise power, σ^2 (V^2), at the detector output (see equation (8.49)) is:

$$\sigma^2 = k^2 E_1 N_0/2 \tag{11.3}$$

where N_0 (V^2/Hz) is the normalised one-sided noise power spectral density at the matched filter, or correlator, input. (Note that the constant k is not shown in equations (8.37) and (8.49) since it is assumed, there, to be 1.0 Hz/V.) The post filtered decision process is identical to the baseband binary decision process described in Chapter 6. Equation (6.8) can therefore be used with $\Delta V = k(E_1 - 0)$ and $\sigma^2 = k^2 E_1 N_0/2$ to give the probability of symbol error:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\,\frac{1}{2}\left(\frac{E_1}{N_0}\right)^{1/2}\right] \tag{11.4}$$

Equation (11.4) can be expressed in terms of the time averaged energy per symbol, $\langle E\rangle = \frac{1}{2}(E_1 + E_0)$ where, for OOK, $E_0 = 0$, i.e.:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\,\frac{1}{\sqrt{2}}\left(\frac{\langle E\rangle}{N_0}\right)^{1/2}\right] \tag{11.5}$$

Finally equations (11.4) and (11.5) can be expressed in terms of received carrier to noise ratios (C/N) using the following relationships:

$$C = \langle E\rangle/T_o \quad (V^2) \tag{11.6}$$

$$N = N_0 B \quad (V^2) \tag{11.7}$$

$$\langle E\rangle/N_0 = T_o B\, C/N \tag{11.8}$$

where C is the received carrier power averaged over all symbol periods and N is the normalised noise power in a bandwidth B Hz.† This gives:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\,\frac{(T_o B)^{1/2}}{\sqrt{2}}\left(\frac{C}{N}\right)^{1/2}\right] \tag{11.9}$$

For minimum bandwidth (i.e. Nyquist) pulses $T_o B = 1.0$ and $\langle E\rangle/N_0 = C/N$. Bandlimited signals will result in symbol energy being spread over more than one symbol period, however, resulting (potentially) in ISI. This will degrade P_e with respect to that given in equations (11.4), (11.5) and (11.9) unless proper steps are taken to ensure ISI

free sampling instants at the output of the receiver matched filter (as in section 8.4).

Incoherent detection of OOK is, by definition, insensitive to the phase information contained in the received symbols. This lost information degrades the detector's performance over that given above. The degradation incurred is typically equivalent to a 1 dB penalty in receiver CNR at a P_e level of 10^{-4}. The modest size of this CNR penalty means that incoherent detection of OOK signals is almost always used in practice.

EXAMPLE 11.1

An OOK IF modulated signal is detected by an ideal matched filter receiver. The non-zero symbol at the matched filter input is a rectangular pulse with an amplitude 100 mV and a duration of 10 ms. The noise at this point is known to be white and Gaussian, and has an RMS value of 140 mV when measured in a noise bandwidth of 10 kHz. Calculate the probability of bit error.

Energy per non-zero symbol:

$$E_1 = v_{RMS}^2 T_o$$

$$= \left(\frac{100 \times 10^{-3}}{\sqrt{2}} \right)^2 10 \times 10^{-3} = 5.0 \times 10^{-5} \quad (\mathrm{V^2 s})$$

Average energy per symbol:

$$\langle E \rangle = \frac{E_1 + 0}{2} = 2.5 \times 10^{-5} \quad (\mathrm{V^2 s})$$

Noise power spectral density (from equation (11.7)):

$$N_0 = \frac{N}{B_N} = \frac{n_{RMS}^2}{B_N}$$

$$= \frac{(140 \times 10^{-3})^2}{10 \times 10^3} = 1.96 \times 10^{-6} \quad (\mathrm{V^2/Hz})$$

From equation (11.5):

$$P_e = \frac{1}{2} \left[1 - \mathrm{erf} \frac{1}{\sqrt{2}} \left(\frac{\langle E \rangle}{N_0} \right)^{1/2} \right]$$

$$= \frac{1}{2} \left[1 - \mathrm{erf} \frac{1}{\sqrt{2}} \left(\frac{2.5 \times 10^{-5}}{1.96 \times 10^{-6}} \right)^{1/2} \right]$$

† If pulse shaping, or filtering, has been employed to bandlimit the transmitted signal, the obvious interpretation of B is the predetection bandwidth, equal to the signal bandwidth. Here C/N in equation (11.9) will be the actual CNR measured after a predetection filter with bandwidth just sufficient to pass the signal intact. In the case of rectangular pulse transmission (or any other non-bandlimited transmission scheme), B must be interpreted simply as a convenient bandwidth within which the noise power is measured or specified (typically chosen to be the doublesided Nyquist bandwidth, $B = 1/T_o$). C/N in equation (11.9) does not then correspond to the CNR at the input to the matched filter or correlation receiver since, strictly speaking, this quantity will be zero.

$$= \tfrac{1}{2}[1 - \text{erf}(2.525)]$$

Using error function tables:

$$P_e = \tfrac{1}{2}[1 - 0.999645] = 1.778 \times 10^{-4}$$

11.3.2 Binary phase shift keying (and phase reversal keying)

Binary phase shift keying (BPSK) impresses baseband information onto a carrier, by changing the carrier's phase in sympathy with the baseband digital data, i.e.:

$$f(t) = \begin{cases} A\,\Pi(t/T_o)\cos 2\pi f_c t, & \text{for a digital 1} \\ A\,\Pi(t/T_o)\cos(2\pi f_c t + \phi), & \text{for a digital 0} \end{cases} \tag{11.10}$$

In principle any two phasor states can be used to represent the binary symbols but usually antipodal states are chosen (i.e. states separated by $\phi = 180°$). For obvious reasons this type of modulation is sometimes referred to as phase reversal keying (PRK). A PRK transmitter with typical baseband and IF waveforms, spectra and allowed phasor states is

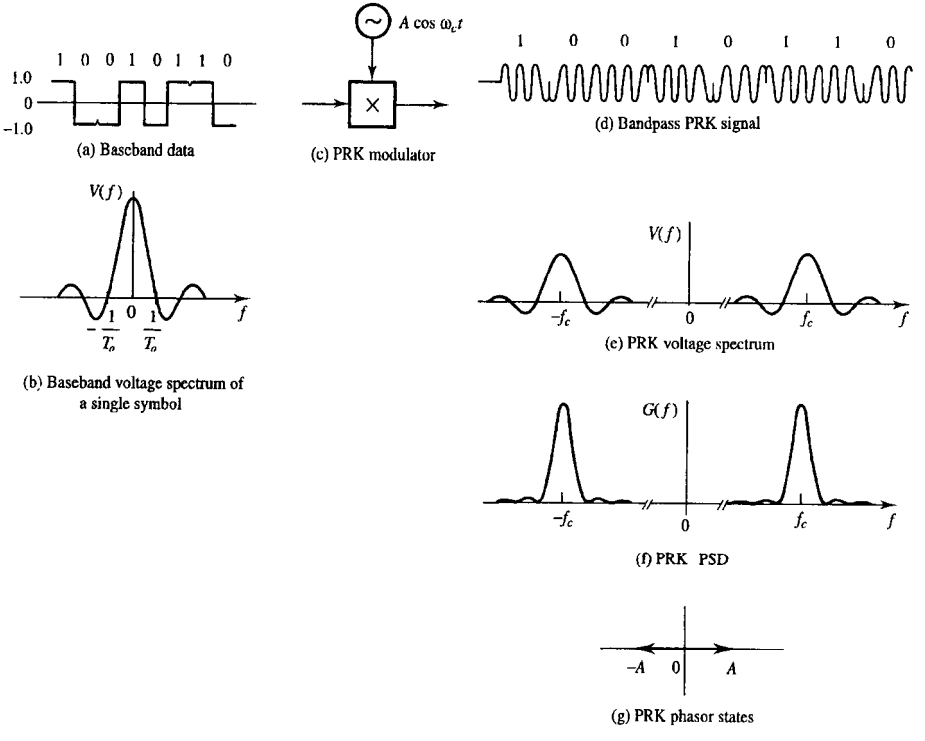


Figure 11.3 *Phase reversal keyed (PRK) modulator waveforms, spectra and phasor states.*

shown in Figure 11.3. PRK systems must obviously employ coherent detectors which can be implemented as either matched filters or correlators. Since the zero and one symbols are antipodal, Figure 11.3(g), only one receiver channel is needed, Figure 11.4. The post-filtered decision instant voltages are $\pm kE$ (V) where E (V^2s) is the normalised energy residing in either symbol. The normalised noise power, σ^2 (V^2), at the filter output is the same as in the BASK case. Substituting $\Delta V = 2kE$ and $\sigma^2 = k^2EN_0/2$ into equation (6.8) gives:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\left(\frac{E}{N_0}\right)^{1/2}\right] \tag{11.11}$$

(Note that in this case $E = \langle E \rangle$.) Using equation (11.8) the PRK probability of symbol error can be expressed as:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\,(T_oB)^{1/2}\left(\frac{C}{N}\right)^{1/2}\right] \tag{11.12}$$

For more general BPSK modulation where the difference between phasor states is less than 180° the P_e performance is most easily derived by resolving the allowed phasor states into a residual carrier and a reduced amplitude PRK signal, Figure 11.5. The residual carrier, which contributes nothing to symbol detection, can be employed as a pilot transmission and used for carrier recovery purposes at the receiver. If the difference between phasor states is $\Delta\theta$ (Figure 11.5) and a BPSK 'modulation index', m, is defined by:

$$m = \sin\left(\frac{\Delta\theta}{2}\right) \tag{11.13}$$

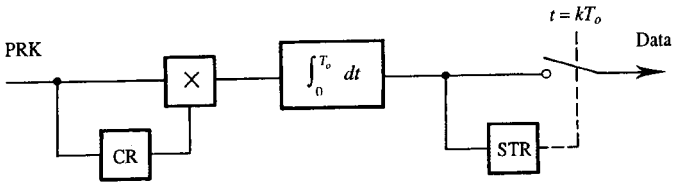


Figure 11.4 *Phase reversal keyed correlation detector.*

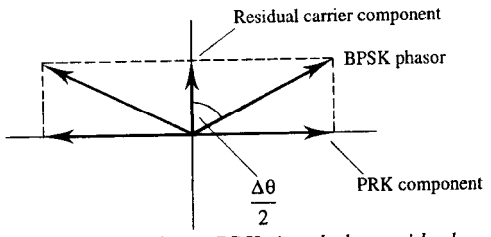


Figure 11.5 *Resolution of BPSK signal into PRK signal plus residual carrier.*

then m is the proportion of the transmitted signal voltage which conveys information and the corresponding proportion of total symbol energy is:

$$m^2 = \sin^2\left(\frac{\Delta\theta}{2}\right) = \frac{1}{2}(1 - \cos\Delta\theta) \tag{11.14}$$

$\cos\Delta\theta$ is the scalar product of the two (unit amplitude) symbol phasors. Denoting this quantity by the normalised correlation coefficient ρ then:

$$m = \sqrt{\tfrac{1}{2}(1-\rho)} \tag{11.15}$$

The BPSK probability of symbol error is found by replacing E in equation (11.11) with m^2E, i.e.:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\, m\left(\frac{E}{N_0}\right)^{1/2}\right] \tag{11.16}$$

and equation (11.16) can be rewritten as:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\,\frac{\sqrt{1-\rho}}{\sqrt{2}}\left(\frac{E}{N_0}\right)^{1/2}\right] \tag{11.17}$$

It is, of course, no coincidence that equation (11.17) is identical to equation (8.61) since, as pointed out both in Chapter 8 and section 11.3, this is a general result for all binary systems having equal energy, equiprobable, symbols. The same caution must be exercised when using equations (11.11), (11.12), (11.16) and (11.17) as when using equations (11.4), (11.5) and (11.9) since all those equations assume ISI free reception. Filtering of PRK signals to limit their bandwidth may, therefore, result in a P_e which is higher than these equations imply.

EXAMPLE 11.2

A 140 Mbit/s ISI free PRK signalling system uses pulse shaping to constrain its transmission to the double sideband Nyquist bandwidth. The received signal power is 10 mW and the one-sided noise power spectral density is 6.0 pW/Hz. Find the BER expected at the output of an ideal matched filter receiver. If the phase angle between symbols is reduced to 165° in order to provide a residual carrier, find the received power in the residual carrier and the new BER.

The double sided Nyquist bandwidth is given by:

$$B = \frac{1}{T_o}$$

i.e.:

$$T_o B = 1.0$$

$$N = N_0 B = N_0/T_o = N_0 R_s$$

$$= 6.0 \times 10^{-12} \times 140 \times 10^{6} = 8.4 \times 10^{-4} \text{ W}$$

Now using equation (11.12):

$$P_e = \frac{1}{2}\left[1 - \text{erf}\ (T_o B)^{1/2}\left(\frac{C}{N}\right)^{1/2}\right] = \frac{1}{2}\left[1 - \text{erf}\left(\frac{10 \times 10^{-3}}{8.4 \times 10^{-4}}\right)^{1/2}\right]$$

$$= \tfrac{1}{2}\,[1 - \text{erf}\,(3.450)] = \tfrac{1}{2}\,[1 - 0.999\,998\,934] = 5.33 \times 10^{-7}$$

$$\text{BER} = P_e R_s$$

$$= 5.33 \times 10^{-7} \times 140 \times 10^{6} = 74.6 \text{ error/s}$$

If $\Delta\theta$ is reduced to 165° then:

$$m = \sin\left(\frac{\Delta\theta}{2}\right) = \sin\left(\frac{165}{2}\right) = 0.9914$$

Proportion of signal power in residual carrier is:

$$1 - m^2 = 1 - 0.9914^2 = 0.0171$$

Power received in residual carrier is therefore:

$$C(1 - m^2) = 10 \times 10^{-3} \times 0.0171 = 1.71 \times 10^{-4} \text{ W}$$

Information bearing carrier power is:

$$C\,m^2 = 10 \times 10^{-3}\,(0.9914)^2 = 9.829 \times 10^{-3} \text{ W}$$

Now from equation (11.12):

$$P_e = \frac{1}{2}\left[1 - \text{erf}\,(T_o B)^{1/2}\left(\frac{C\,m^2}{N}\right)^{1/2}\right] = \frac{1}{2}\left[1 - \text{erf}\left(\frac{9.829 \times 10^{-3}}{8.4 \times 10^{-4}}\right)^{1/2}\right]$$

$$= \tfrac{1}{2}\,[1 - \text{erf}\,(3.421)] = \tfrac{1}{2}\,[1 - 0.999\,998\,688] = 6.55 \times 10^{-7}$$

$$\text{BER} = P_e R_s$$

$$= 6.55 \times 10^{-7} \times 140 \times 10^{6} = 91.7 \text{ error/s}$$

11.3.3 Binary frequency shift keying

Binary frequency shift keying (BFSK) represents digital ones and zeros by carrier pulses with two distinct frequencies, f_1 and f_2, i.e.:

$$f(t) = \begin{cases} A\,\Pi(t/T_o)\cos 2\pi f_1 t, & \text{for a digital 1} \\ A\,\Pi(t/T_o)\cos 2\pi f_2 t, & \text{for a digital 0} \end{cases} \qquad (11.18)$$

Figure 11.6 shows a schematic diagram of a BFSK modulator, signal waveforms and signal spectra. (In practice the BFSK modulator would normally be implemented as a numerically controlled oscillator.) The voltage spectrum of the BFSK signal is the superposition of the two OOK spectra, one representing the baseband data stream

modulated onto a carrier with frequency f_1 and one representing the inverse data stream modulated onto a carrier with frequency f_2 (Figure 11.6(e)). It is important to realise, however, that the PSD of a BFSK signal is not the superposition of two OOK PSDs. This is because in the region where the two OOK spectra overlap the spectral lines must be added with due regard to their phases. (In practice, when the separation of f_1 and f_2 is large, and the overlap correspondingly small, then the BFSK PSD is *approximately* the superposition of two OOK PSDs.)

Detection of BFSK can be coherent or incoherent although the latter is more common. Incoherent detection suffers the same CNR penalty, compared with coherent detection, as is the case for OOK systems. Figure 11.7 shows coherent and incoherent BFSK receivers. The receiver in Figure 11.7(c) is an FSK version of the quadrature receiver shown in Figure 11.2. The coherent receiver (Figure 11.7(a)) is an FSK version of the PSK receiver in Figure 11.4.

FSK does not provide the noise reduction of wideband analogue FM transmissions. If, however, a BFSK carrier frequency, f_c Figure 11.6(e), is defined by:

$$f_c = \frac{f_1 + f_2}{2} \quad \text{(Hz)} \tag{11.19}$$

and a BFSK frequency deviation, Δf, is defined (Figure 11.6(e)) as:

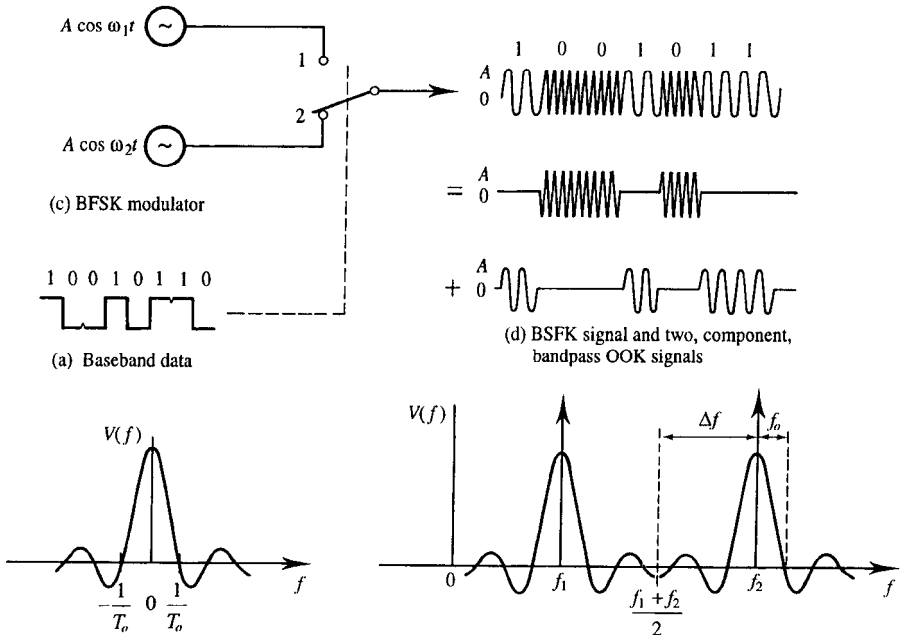


Figure 11.6 *Binary frequency shift keyed (BFSK) modulators, waveforms and spectra.*

$$\Delta f = \frac{f_2 - f_1}{2} \quad \text{(Hz)} \tag{11.20}$$

then using the first zero crossing points in the BFSK voltage spectrum to define its bandwidth, B, gives:

$$B = 2\Delta f + 2f_o \tag{11.21}$$

Here $f_o = 1/T_o$ is both the nominal bandwidth and the baud rate of the baseband data stream. Equation (11.21) is strongly reminiscent of Carson's rule for the bandwidth of an FM signal [Stremler]. If the binary symbols of a BFSK system are orthogonal (see section 2.5.3), i.e.:

$$\int_0^{T_o} \cos(2\pi f_1 t)\cos(2\pi f_2 t)\, dt = 0 \tag{11.22}$$

then, when the output of one channel of a coherent BFSK receiver is a maximum the output of the other channel will be zero. After subtracting the post-filtered signals arising from each receiver channel the orthogonal BFSK decision instant voltage is:

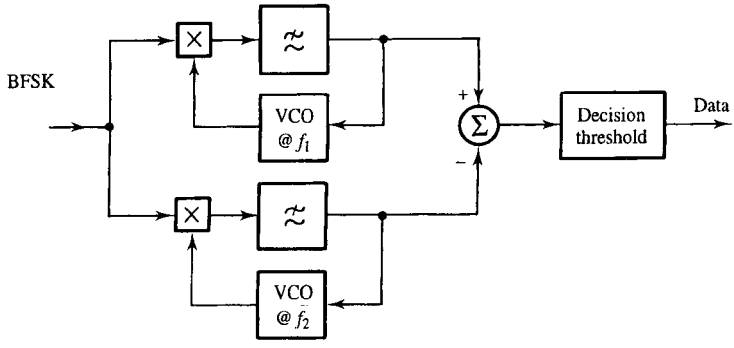
$$f(nT_o) = \begin{cases} kE, & \text{for a digital 1} \\ -kE, & \text{for a digital 0} \end{cases} \tag{11.23}$$

Table 11.1 *P_e formulae for ideal coherent detection of baseband and IF modulated binary signals.*

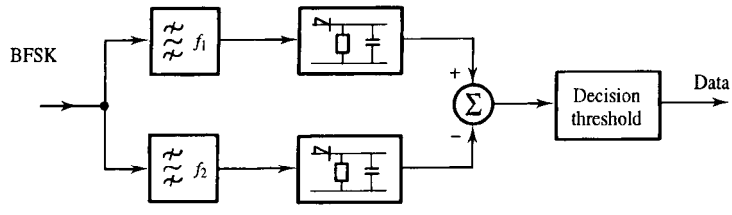
		P_e	
Baseband signalling	Unipolar (OOK)	$\frac{1}{2}\text{erfc}\sqrt{\frac{1}{2}\frac{E}{N_0}}$ (8.61)	$\frac{1}{2}\text{erfc}\sqrt{\frac{1}{4}\frac{S}{N}}$ (6.10(b))
	Polar	$\frac{1}{2}\text{erfc}\sqrt{\frac{E}{N_0}}$ (8.61)	$\frac{1}{2}\text{erfc}\sqrt{\frac{1}{2}\frac{S}{N}}$ (6.12)
IF/RF signalling	OOK	$\frac{1}{2}\text{erfc}\sqrt{\frac{1}{2}\frac{E}{N_0}}$	$\frac{1}{2}\text{erfc}\sqrt{\frac{T_oB}{2}\frac{C}{N}}$
	BFSK (orthogonal)	$\frac{1}{2}\text{erfc}\sqrt{\frac{1}{2}\frac{E}{N_0}}$	$\frac{1}{2}\text{erfc}\sqrt{\frac{T_oB}{2}\frac{C}{N}}$
	PRK	$\frac{1}{2}\text{erfc}\sqrt{\frac{E}{N_0}}$	$\frac{1}{2}\text{erfc}\sqrt{T_oB\frac{C}{N}}$

If the one sided NPSD at the BFSK receiver input is N_0 (V^2Hz^{-1}) then the noise power, $\sigma_1^2 = k^2EN_0/2$, received via channel 1 and the noise power, $\sigma_2^2 = k^2EN_0/2$, received via channel 2 will add power-wise (since the noise processes will be uncorrelated). The total noise power at the receiver output will therefore be:

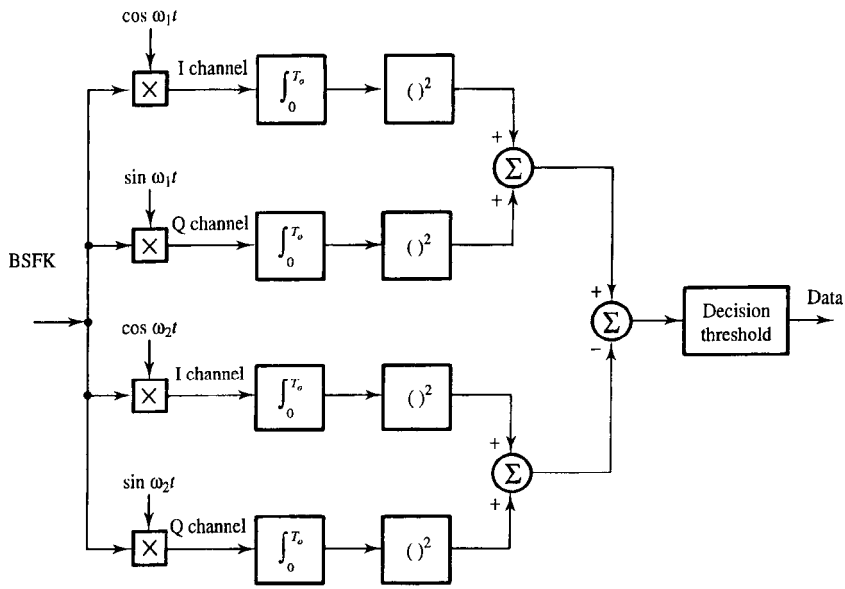
$$\sigma^2 = \sigma_1^2 + \sigma_2^2 = k^2EN_0 \quad (V^2) \tag{11.24}$$



(a) Coherent detection of BFSK using PLLs



(b) Incoherent detection of BFSK



(c) Incoherent detection using I and Q channels

Figure 11.7 *Coherent and incoherent BFSK receivers.*

Substituting $\Delta V = 2kE$ from equation (11.23) and $\sigma = k\sqrt{EN_0}$ from equation (11.24) into equation (6.8) gives the probability of error for coherently detected orthogonal BFSK:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\frac{1}{\sqrt{2}}\left(\frac{E}{N_0}\right)^{1/2}\right] \tag{11.25}$$

Notice that here, as for BPSK signalling, $E = \langle E \rangle$. Since equation (11.25) is identical to equation (11.5) for OOK modulation the BFSK expression for P_e in terms of CNR is identical to equation (11.9).

Table 11.1 compares the various formulae for P_e obtained in section 11.3, Chapter 6 and Chapter 8.

11.3.4 BFSK symbol correlation and Sunde's FSK

Equation (11.25) applies to the coherent detection of orthogonal BFSK. It is, of course, possible to choose symbol frequencies (f_1 and f_2) and a symbol duration (T_o) such that the symbols are not orthogonal, i.e.:

$$\int_0^{T_o} \cos(2\pi f_1 t)\cos(2\pi f_2 t) \neq 0 \tag{11.26}$$

If this is the case then there will be non-zero sampling instant outputs from both BFSK receiver channels when either symbol is present at the receiver input. The key to understanding how this affects the probability of symbol error is to recognise that the common fraction of the two symbols (i.e. the fraction of each symbol which is common to the other) can carry no information. The common fraction is the normalised correlation coefficient, ρ, of the two symbols. Denoting a general pair of symbols by $f_1(t)$ and $f_2(t)$ the normalised correlation coefficient is defined by equation (2.89):

$$\rho = \frac{\langle f_1(t) f_2(t) \rangle}{\sqrt{\langle |f_1(t)|^2 \rangle}\sqrt{\langle |f_2(t)|^2 \rangle}} \tag{11.27}$$

(The normalisation process can be thought of as scaling each symbol to an RMS value of 1.0 before cross correlating.) For practical calculations equation (11.27) can be rewritten as:

$$\begin{aligned} \rho &= \frac{1}{f_{1RMS} f_{2RMS}} \frac{1}{T_o} \int_0^{T_o} f_1(t) f_2(t)\, dt \\ &= \frac{1}{\sqrt{E_1 E_2}} \int_0^{T_o} f_1(t) f_2(t)\, dt \end{aligned} \tag{11.28}$$

and for BFSK systems, in which $E_1 = E_2 = E$, this becomes:

$$\rho = \frac{2}{T_o} \int_0^{T_o} \cos(2\pi f_1 t)\cos(2\pi f_2 t)\, dt \tag{11.29}$$

Figure 11.8 shows BFSK symbol correlation, ρ, plotted against the difference in the number of carrier half cycles contained in the symbols. The zero crossing points represent orthogonal signalling systems and $2(f_2 - f_1)T_o = 1.43$ represents a signalling system somewhere between orthogonal and antipodal, see section 6.4.2. (This is the optimum operating point for BFSK in terms of power efficiency and yields a 0.8 dB CNR saving over orthogonal BFSK. It corresponds to symbols which contain a difference of approximately 3/4 of a carrier cycle.)

Whilst any zero crossing point on the $\rho - T_o$ diagram of Figure 11.8 corresponds to orthogonal BFSK it is not possible to use the first zero (i.e. $2(f_2 - f_1)T_o = 1$) if detection is incoherent. (This can be appreciated if it is remembered that for BFSK operated at this point the two symbols are different by only half a carrier cycle. A difference this small can be detected as a change in phase (of 180°) over the symbol duration but not measured reliably in this time as a change in frequency.) The minimum frequency separation for successful incoherent detection of orthogonal BFSK is therefore given by the second zero crossing point (i.e. $2(f_2 - f_1)T_o = 2$) in Figure 11.8 which corresponds to one carrier cycle difference between the two symbols or $\Delta f = f_o/2$ in Figure 11.6. BFSK operated at the second zero of the $\rho - T_o$ diagram is called Sunde's FSK. The voltage and power spectra for this scheme are shown in Figure 11.9. If the first spectral zero definition of bandwidth (or equivalently Carson's rule) is applied to Sunde's FSK the bandwidth would be given by:

$$B = (f_2 - f_1) + \frac{2}{T_o} = \frac{3}{T_o} \quad \text{(Hz)} \tag{11.30}$$

The overlapping spectral lines of the component OOK signals, however, result in cancellation giving a $1/f^4$ roll-off in the envelope of the power spectral density. This rapid roll-off allows the practical bandwidth of Sunde's FSK to be taken, for most applications, to be:

$$B = f_2 - f_1 = \frac{1}{T_o} \quad \text{(Hz)} \tag{11.31}$$

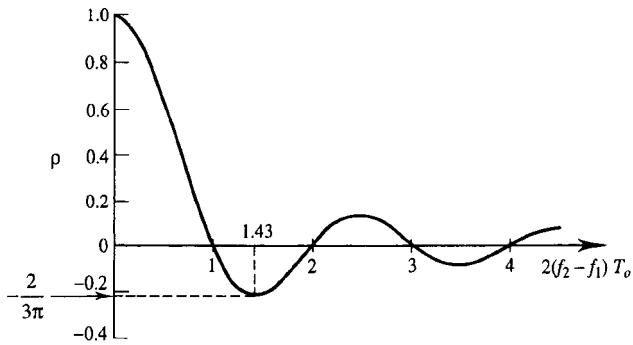


Figure 11.8 *BFSK $\rho - T_o$ diagram,* $\rho = (\sin[\pi(f_2 - f_1)T_o]\cos[\pi(f_2 - f_1)T_o])/([\pi(f_2 - f_1)T_o])$

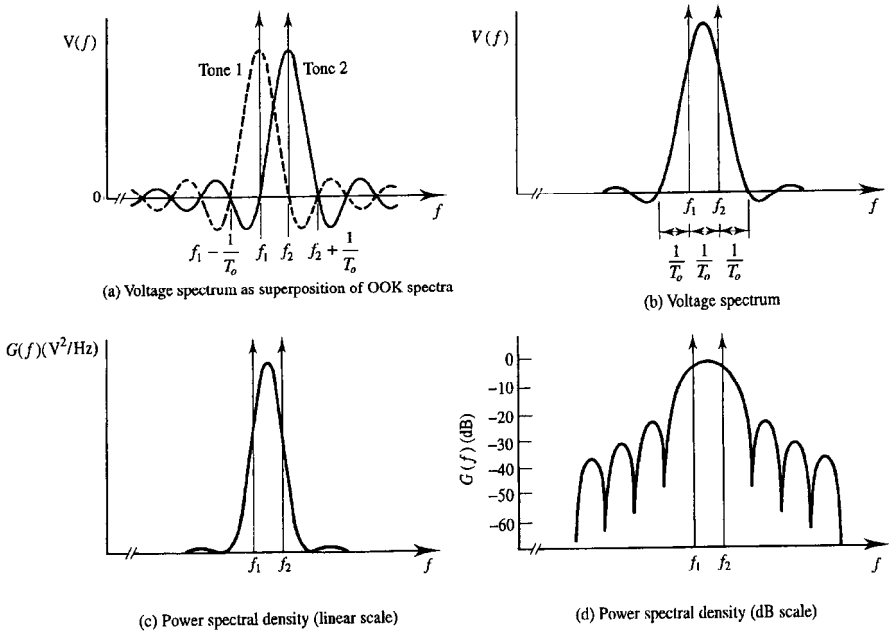


Figure 11.9 *Comparison of spectra for Sunde's FSK.*

EXAMPLE 11.3

A rectangular pulse BFSK system operates at the 3rd zero crossing point of the $\rho - T_o$ diagram of Figure 11.8. The maximum available transmitter power results in a carrier power at the receiver input of 60 mW. The one-sided NPSD referred to the same point is 0.1 nW/Hz. What is the maximum bit rate which the system can support if the probability of bit error is not to fall below 10^{-6}? What is the nominal bandwidth of the BFSK signal if the lower frequency symbol has a frequency of 80 MHz?

$$P_e = \frac{1}{2}\left[1 - \text{erf}\frac{1}{\sqrt{2}}\left(\frac{E}{N_0}\right)^{1/2}\right]$$

i.e.:

$$\frac{E}{N_0} = 2\left[\text{erf}^{-1}(1-2P_e)\right]^2 = 2\left[\text{erf}^{-1}\left(1-2\times10^{-6}\right)\right]^2$$

$$= 2\left[\text{erf}^{-1}(0.999\,998)\right]^2 = 2\,[3.361]^2 = 22.59$$

Therefore:

$$E = 22.59\,N_0 = 22.59 \times 0.1 \times 10^{-9} = 2.259 \times 10^{-9} \text{ J}$$

Now from equation (11.6):

$$T_o = \frac{E}{C} = \frac{2.259 \times 10^{-9}}{60 \times 10^{-3}} = 3.765 \times 10^{-8} \text{ s}$$

$$R_b = R_s = \frac{1}{T_o} = \frac{1}{3.765 \times 10^{-8}} = 2.656 \times 10^{7} \text{ bit/s}$$

Since operation on the diagram of Figure 11.8 is at $n = 3$ then:

$$3 = 2(f_2 - f_1)T_o$$

$$f_2 = \frac{3}{2T_o} + f_1$$

$$= \frac{3}{2 \times 3.765 \times 10^{-8}} + 80 \times 10^6 = 1.198 \times 10^8 \text{ Hz}$$

$$\Delta f = \frac{f_2 - f_1}{2} = \frac{119.8 - 80}{2} = 19.9 \text{ MHz}$$

and from equation (11.21):

$$B = 2(\Delta f + f_o)$$

$$= 2(19.9 + 26.6) = 93.0 \text{ MHz}$$

11.3.5 Comparison of binary shift keying techniques

Table 11.2 *Relative power efficiencies of binary bandpass modulation schemes.*

	Bandpass OOK	*Orthogonal BFSK*	*PRK*
$\frac{E_1}{N_0}$	4	2	1
$\frac{\langle E \rangle}{N_0}$	2	2	1

It is no coincidence that OOK and orthogonal FSK systems are equally power efficient, i.e. have the same probability of symbol error for the same $\langle E \rangle / N_0$, Table 11.1. This is because they are both orthogonal signalling schemes. PRK signalling is antipodal and, therefore, more power efficient, i.e. it has a better P_e performance or, alternatively, a power saving of 3 dB. If comparisons are made on a peak power basis then orthogonal FSK requires 3 dB more power than PRK, as expected, but also requires 3 dB less power than OOK since all the energy of the OOK transmission is squeezed into only one type of symbol. These relative power efficiencies are summarised in Table 11.2.

Although the spectral efficiency for all unfiltered (i.e. rectangular pulse) signalling systems is, strictly, 0 bit/s/Hz (see section 8.2) it is possible to define a nominal spectral efficiency using the signal pulse's main lobe bandwidth, $B = 2/T_o$. (This does not correspond to the theoretical minimum channel bandwidth which is $B = 1/T_o$.) Table 11.3 summarises the nominal spectral efficiencies of unfiltered bandpass BASK, BFSK

and BPSK systems and compares them with the nominal efficiency of unfiltered baseband OOK. It is interesting to note that Sunde's FSK could be regarded as being more spectrally efficient than BASK and BPSK due to the more rapid roll-off of its spectral envelope. (If the null-to-null definition of bandwidth is adhered to strictly then Sunde's FSK has a nominal efficiency of 0.33 bit/s/Hz which is less than BASK and BPSK.) The probability of symbol error versus $\langle E \rangle / N_0$ for orthogonal (OOK and orthogonal BFSK) and antipodal (PRK) signalling is shown in Figure 11.10.

It is useful to develop the generalised expression for P_e which can be applied to all binary (coherently detected, ideal,) modulation schemes. For any two-channel coherent receiver the decision instant voltages for equal energy symbols will be:

$$f(nT_o) = \begin{cases} k\left[E - \int_0^{T_o} f_1(t) f_2(t)\, dt\right], & \text{symbol 0 present} \\ k\left[-E + \int_0^{T_o} f_1(t) f_2(t)\, dt\right], & \text{symbol 1 present} \end{cases}$$

$$= \begin{cases} k\left[E - \rho E\right], & \text{symbol 0 present} \\ k\left[-E + \rho E\right], & \text{symbol 1 present} \end{cases} \qquad (11.32)$$

The decision instant voltage difference between symbols is:

$$\Delta V = k\left[2E - 2\rho E\right] = k2E(1 - \rho) \qquad (11.33(a))$$

Table 11.3 *Relative spectral efficiencies of binary bandpass modulation schemes.*

	Baseband BASK	*Bandpass BASK*	*Orthogonal BFSK ($n \geq 3$)**	*Sunde's BFSK ($n = 2$)**	*BPSK*
Data rate (bit/s)	$1/T_o$	$1/T_o$	$1/T_o$	$1/T_o$	$1/T_o$
Nominal bandwidth (Hz)	$1/T_o$	$2/T_o$	$(n+4)/2T_o$	$1/T_o$† or $3/T_o$	$2/T_o$
Nominal spectral efficiency (bit/s/Hz)	1	1/2	$2/(n+4)$	1† or 1/3	1/2
Minimum ‡ (ISI free) bandwidth	$1/2T_o$	$1/T_o$	$(n+2)/2T_o$	$2/T_o$	$1/T_o$
Maximum ‡ spectral efficiency	2	1	$2/(n+2)$	1/2	1

* n = zero crossing operating point on $\rho - T_o$ diagram of Figure 11.8

† Depends on definition of bandwidth

‡ Based on absolute bandwidth

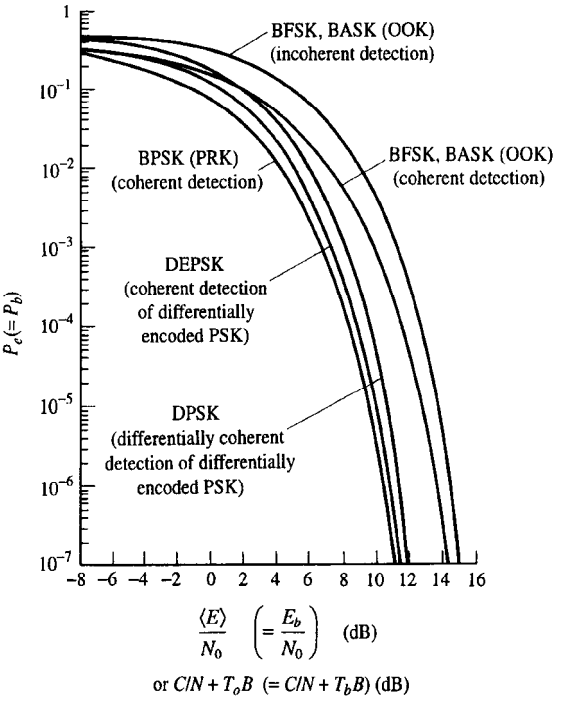


Figure 11.10 *Comparison of binary ASK/PSK/FSK systems performance.*

The two channels result in a total noise power of:

$$\sigma^2 = k^2 E N_0 (1 - \rho) \tag{11.33(b)}$$

which using equation (6.8) gives a probability of error of:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\sqrt{\frac{1-\rho}{2}}\left(\frac{E}{N_0}\right)^{1/2}\right] \tag{11.34(a)}$$

where the correlation coefficient $\rho = 0$ for OOK, $\rho = 0$ for orthogonal BFSK, $\rho = -2/3\pi$ for optimum BFSK and $\rho = -1$ for PRK. (It can now be seen why ρ was chosen to represent $\cos \Delta\theta$ in equation (11.15).) The equivalent expression in terms of C/N is:

$$P_e = \frac{1}{2}\left[1 - \text{erf}\sqrt{\frac{1-\rho}{2}}\,(T_o B)^{1/2}\left(\frac{C}{N}\right)^{1/2}\right] \tag{11.34(b)}$$

11.3.6 Carrier recovery, phase ambiguity and DPSK

Coherent detection of IF modulated signals is required to achieve the lowest error rate, Figure 11.10. This usually requires a reference signal which replicates the phase of the signal carrier. A residual or pilot carrier, if present in the received signal, can be extracted using a filter or phase locked loop or both, then subsequently amplified and employed as a coherent reference. This is possible (though not often implemented) for ASK and FSK signals since both contain discrete lines in their spectra. BPSK signals with phasor states separated by less than 180° also contain a spectral line at the carrier frequency and can therefore be demodulated in the same way. PRK signals, however, contain no such line and carrier recovery must therefore be achieved using alternative methods.

One technique which can be used is to square the received PRK transmission creating a double frequency carrier (with no phase transitions). This occurs because $\sin^2 2\pi f_c t$ and $\sin^2(2\pi f_c t + \pi)$ are both equal to $\sin 2\pi 2 f_c t$. A phase locked loop can then be used as a frequency divider to generate the coherent reference. Such a carrier recovery circuit is shown in Figure 11.11. (Note that the phase locked loop locks to a reference 90° out of phase with the input signal. Thus a 90° phase shifting network is required either within the loop, or between the loop and the demodulator, to obtain the correctly phased reference signal for demodulation.) A second technique is to use a Costas loop [Dunlop and Smith, Lindsey and Simon] in place of a conventional PLL. The Costas loop, Figure 11.12, consists essentially of two PLLs operated in phase quadrature and has the property that it will lock to the suppressed carrier of a PRK signal.

Both the squaring loop and Costas loop solutions to PRK carrier recovery suffer from a 180° phase ambiguity, i.e. the recovered carrier may be either in-phase or in antiphase

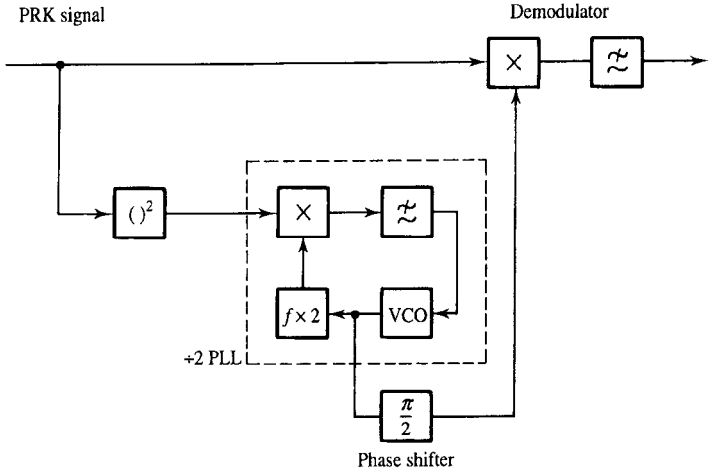


Figure 11.11 *Squaring loop for suppressed carrier recovery.*

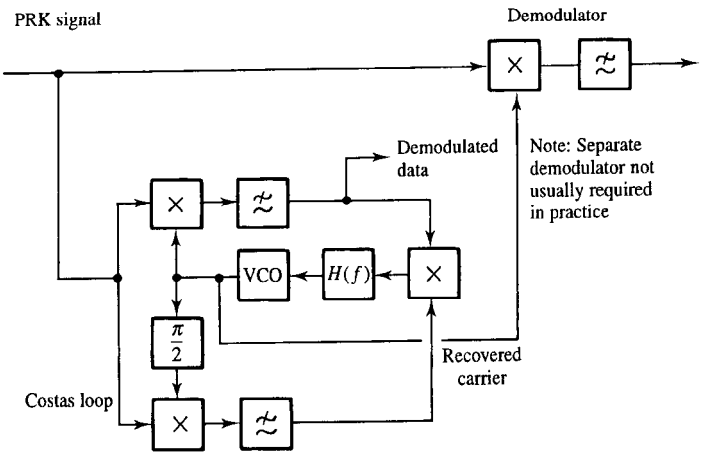


Figure 11.12 *Costas loop for suppressed carrier recovery.*

with the transmitted (suppressed) carrier. This ambiguity can lead to symbol inversion of the demodulated data. Whilst data inversion might be tolerated by some applications, in others it would clearly be disastrous. There are two distinct approaches to resolving phase ambiguities in the recovered carrier.

The first involves periodic transmission of a known data sequence. (This sequence is normally one part of a larger 'preamble' sequence transmitted prior to a block of data, see Chapter 19.) If the received 'training' sequence is inverted this is detected and a second inversion introduced to correct the data. The second approach is to employ differential encoding of the data before PRK modulation (Figure 11.13). This results, for example, in digital ones being represented by a phase transition and digital zeros being represented by no phase transition. The phase ambiguity of the recovered carrier then becomes irrelevant. Demodulation of such encoded PSK signals can be implemented using conventional carrier recovery and detection followed by baseband differential decoding (Figure 11.14). (STR techniques are described in section 6.7.) Systems using this detection scheme are called differentially encoded PSK (DEPSK) and have a probability of symbol error, P'_e, given by:

$$P'_e = P_e(1 - P_e) + P_e(1 - P_e)$$

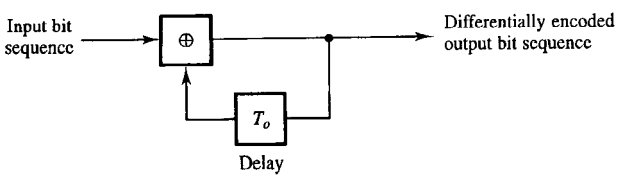


Figure 11.13 *Differential encoding.*

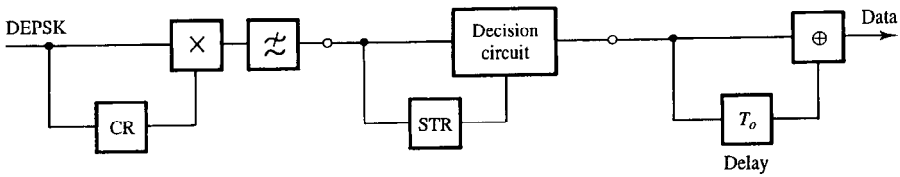


Figure 11.14 *Differentially encoded PSK (DEPSK) detection.*

$$= 2(P_e - P_e^2) \tag{11.35}$$

where P_e is the probability of error for uncoded PSK.

The first term, $P_e(1 - P_e)$, in equation (11.35) is the probability that the current symbol in the decoder is in error and the previous symbol is correct. The second (identical) term is the probability that the current symbol is correct and the previous symbol is in error. (If both current and previous symbols are detected in error the decoded symbol will, of course, be correct.)

An alternative method of demodulating PSK signals with differential coding is to use one symbol as the coherent reference for the next symbol (Figure 11.15). Systems using this detection scheme are called differential PSK (DPSK). They have the advantage of simpler, and therefore cheaper, receivers compared to DEPSK systems. They have the disadvantage, however, of having a noisy reference signal and therefore degraded P'_e performance compared to DEPSK and coherent PSK, Figure 11.10. It might appear that since reference and signal are equally noisy the P'_e degradation in DPSK systems would correspond to a 3 dB penalty in CNR. In practice the degradation is significantly less than this (typically 1 dB) since the noise in the signal and reference channels is correlated. DEPSK and DPSK techniques are easily extended from biphase to multiphase signalling (see section 11.4.2).

11.4 Modulation techniques with increased spectral efficiency

Spectral efficiency, η_s as defined in equation (8.1), depends on symbol (or baud) rate, R_s, signal bandwidth, B, and entropy, H (as defined in equation (9.3)), i.e.:

$$\eta_s = \frac{R_s H}{B} \quad \text{(bit/s/Hz)} \tag{11.36(a)}$$

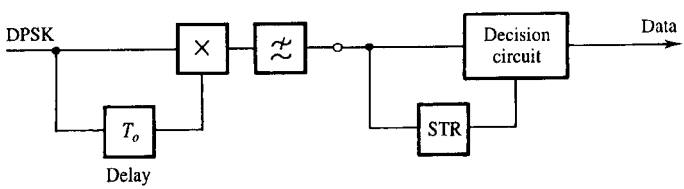


Figure 11.15 *Differential PSK (DPSK) detection.*

Since $R_s = 1/T_o$ and $H = \log_2 M$, for statistically independent, equiprobable symbols (section 9.2.4) then η_s can be expressed as:

$$\eta_s = \frac{\log_2 M}{T_o B} \text{ (bit/s/Hz)} \qquad (11.36(b))$$

It is apparent from equation (11.36(b)) that spectral efficiency is maximised by making the symbol alphabet size, M, large and the $T_o B$ product small. This is exactly the strategy employed by spectrally efficient modulation techniques.

Pulse shaping (or filtering) to decrease B for a given baud rate, $R_s = 1/T_o$, has already been discussed in Chapter 8. There, however, the discussion was in the context of baseband signals where the minimum $T_o B$ product (avoiding ISI) was limited by:

$$T_o B \geq 0.5 \qquad (11.37(a))$$

and the minimum bandwidth $B = 1/(2T_o)$ was called the single sided Nyquist bandwidth. In the context of IF modulation (i.e. bandpass signals) the modulation process results in a double sideband (DSB) signal. The minimum ISI free $T_o B$ product is then given by:

$$T_o B \geq 1.0 \qquad (11.37(b))$$

and the minimum bandwidth is now $B = 1/T_o$, sometimes called the double sided Nyquist bandwidth. Dramatic increases in η_s, however, must usually come from increased alphabet size. Operational systems currently exist with M's of 64, 128, 256 and 1024. In principle such multi-symbol signalling can lead to increased spectral efficiency of MASK, MPSK and MFSK systems (the first letter in each case represents M-symbol). Only MASK, MPSK and hybrid combinations of these two are used in practice, however. This is because MFSK signals are normally designed to retain orthogonality between all symbol pairs. In this case increasing M results in an approximately proportional increase in B which actually results in a decrease in spectral efficiency. MASK and MPSK signals, however, are limited to alphabets of 2 (OOK) and 4 (4-PSK) symbols respectively if orthogonality is to be retained. MASK and MPSK therefore must sacrifice orthogonality to achieve values of M greater than four.

Sub-orthogonal signalling requires greater transmitted power than orthogonal signalling for the same P_e. MASK and MPSK systems are therefore spectrally efficient at the expense of increased transmitted power. Conversely, (orthogonal) MFSK systems are power efficient at the expense of increased bandwidth. This is one manifestation of the general trade-off which can be made between power and bandwidth represented at its most fundamental by the Shannon-Hartley law.

11.4.1 Channel capacity

The Shannon-Hartley channel capacity theorem states that the maximum rate of information transmission, $R_{\max}$, over a channel with bandwidth B and signal to noise ratio S/N is given by:

$$R_{\max} = B \log_2 \left(1 + \frac{S}{N}\right) \text{ bit/s} \qquad (11.38(a))$$

Thus as B increases S/N can be decreased to compensate [Hartley]. Note that in a 3.2 kHz wide audio channel with a SNR of 1000 (30 dB) the theoretical maximum bit rate, $R_{\max}$, is slightly in excess of 30 kbit/s. The corresponding maximum spectral efficiency, $R_{\max}/B$, for a channel with 30 dB SNR is 10 bit/s/Hz. We can determine the value that the channel capacity approaches as the channel bandwidth B tends to infinity:

$$\begin{aligned} R_\infty &= \lim_{B \to \infty} B \log_2 \left(1 + \frac{S}{N_0 B}\right) \\ &= \lim_{B \to \infty} \frac{S}{N_0} \log_2 \left(1 + \frac{S}{N_0 B}\right)^{N_0 B/S} \\ &= \frac{S}{N_0} \log_2 e \;=\; 1.44 \frac{S}{N_0} \end{aligned} \qquad (11.38(b))$$

This gives the maximum possible channel capacity as a function of signal power and noise power spectral density. In an actual system design, the channel capacity might be compared to this value to decide whether a further increase in bandwidth is worth while. In a practical system it is realistic to attempt to achieve a transmission rate of one half R_∞. Some popular modulation schemes which trade an increase in SNR for a decrease in bandwidth, to achieve improved spectral efficiency, are described below.

11.4.2 *M*-symbol phase shift keying

In the context of PSK signalling M-symbol PSK (i.e. MPSK) implies the extension of the number of allowed phasor states from 2 to 4, 8, 16, ... (i.e. 2^n). The phasor diagram (also called the constellation diagram) for 16-PSK is shown in Figure 11.16, where there are now 16 distinct states. Note that as the signal amplitude is constant these states all lie on a circle in the complex plane. 4-phase modulation (4-PSK) with phase states $\pi/4$, $3\pi/4$, $5\pi/4$ and $7\pi/4$ can be considered to be a superposition of two PRK signals using quadrature ($\cos 2\pi f_c t$ and $\sin 2\pi f_c t$) carriers. This type of modulation (usually called quadrature or quaternary phase shift keying, QPSK) is discussed separately in section 11.4.4.

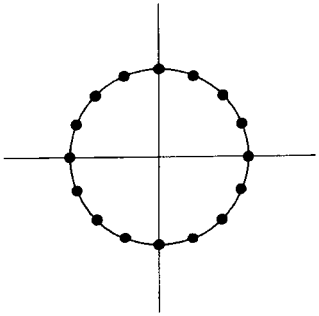


Figure 11.16 *Phasor states (i.e. constellation diagram) for 16-PSK.*

The probability of symbol error for MPSK systems is found by integrating the two-dimensional pdf of the noise centred on the tip of each signal phasor in turn over the corresponding error region and averaging the results. The error region for phasor state 0 is shown as the unhatched region in Figure 11.17. It is possible to derive an exact, but complicated, integral expression for the probability of symbol error in the presence of Gaussian noise. Accurate P_e curves found by numerical evaluation of this expression are shown in Figure 11.18(a). A simple but good approximation [Stein and Jones] which is useful for $M \geq 4$ is:

$$P_e \approx 1 - \text{erf}\left[\sin\left(\frac{\pi}{M} \right) \left(\frac{E}{N_0} \right)^{1/2} \right] \qquad (11.39(a))$$

(Note that when $M = 2$ this expression gives a P_e which is twice the correct result.) The approximation in equation (11.39(a)) becomes better as both M and E/N_0 increase. Equation (11.39(a)) can be rewritten in terms of CNR using $C = E/T_0$ and $N = N_0 B$ from equations (11.6) and (11.7) as usual, i.e.:

$$P_e \approx 1 - \text{erf}\left[(T_0 B)^{1/2} \sin\left(\frac{\pi}{M} \right) \left(\frac{C}{N} \right)^{1/2} \right] \qquad (11.39(b))$$

Multisymbol signalling can be thought of as a coding or bit mapping process in which n binary symbols (bits) are mapped into a single M-ary symbol (as discussed in section 6.4.7), except that here each symbol is an IF pulse. A detection error in a single symbol can therefore translate into several errors in the corresponding decoded bit sequence. The probability of bit error, P_b, therefore depends not only on the probability of symbol error, P_e, and the symbol entropy, $H = \log_2 M$, but also on the code or bit mapping used and the types of error which occur. For example, in the 16-PSK scheme shown in Figure 11.17 the most probable type of error involves a given phasor state being detected as an adjacent state. If a Gray code is used to map binary symbols to phasor states this type of error results in only a single decoded bit error. In this case, providing the probability of errors other than this type is negligible, then the bit error probability is:

$$P_b = \frac{P_e}{\log_2 M} \qquad (11.40(a))$$

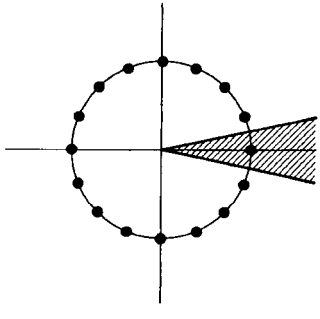


Figure 11.17 *Error region (unhatched) for $\phi = 0°$ state of a 16-PSK signal.*

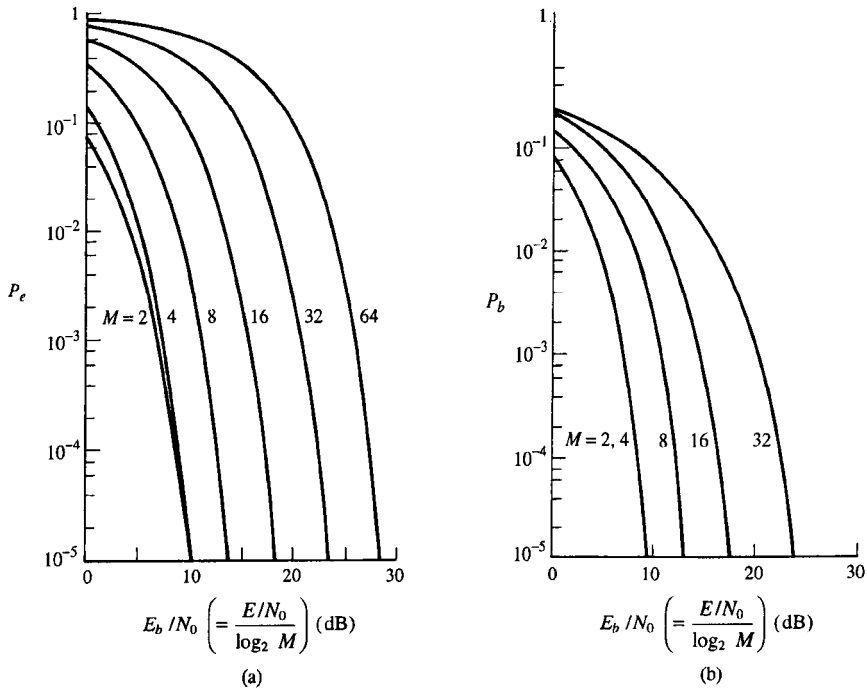


Figure 11.18 *MPSK symbol and bit error probabilities: (a) probability of symbol error, P_e against E_b/N_0; and (b) probability of bit error, P_b against E_b/N_0.*

In order to compare the performance of different modulation schemes on an equitable basis it is useful to express performance in terms of P_b as a function of average energy per bit, E_b. Since the energy, E, of all symbols in an MPSK system are identical the average energy per bit is:

$$E_b = \frac{E}{\log_2 M} \tag{11.40(b)}$$

Substituting equations (11.40) into (11.39):

$$P_b = \frac{1}{\log_2 M}\left\{1 - \text{erf}\left[\sin\left(\frac{\pi}{M}\right)\sqrt{\log_2 M}\left(\frac{E_b}{N_0}\right)^{1/2}\right]\right\} \tag{11.41(a)}$$

In terms of CNR this becomes:

$$P_b = \frac{1}{\log_2 M}\left\{1 - \text{erf}\left[(T_o\, B)^{1/2}\sin\left(\frac{\pi}{M}\right)\left(\frac{C}{N}\right)^{1/2}\right]\right\} \tag{11.41(b)}$$

More accurate closed form expressions for Gray coded and natural binary bit mapped MPSK P_b performance have been derived [Irshid and Salous]. Figure 11.18(a) shows the

probability of symbol error, P_e, against E_b/N_0, and Figure 11.18(b) shows P_b against E_b/N_0. (Figure 13.41 clarifies the distinction between P_e and P_b).

Differential MPSK (DMPSK) can be used to simplify MPSK receiver design and circumvent the phase ambiguity normally present in the recovered carrier. Binary digits are mapped to phase difference between adjacent symbols and each symbol is detected at the receiver using the previous symbol as a coherent reference. The P_e performance of DMPSK is degraded over that for MPSK since the reference and received symbol are now equally noisy. As for DPSK the degradation is not equivalent to a CNR reduction of 3 dB because the noise present in the signal and reference channels is correlated and therefore, to some extent, cancels. Phase noise correlation makes increasingly little difference, however, to the performance degradation as M increases and the phasor states become crowded together. In the limit of large M the degradation approaches the expected 3 dB limit.

Since each symbol of an MPSK signal has an identical amplitude spectrum to all the other symbols the spectral occupancy of MPSK depends only on baud rate and pulse shaping and is independent of M. For unfiltered MPSK (i.e. MPSK with rectangular pulses) the nominal (i.e. main lobe null-to-null) bandwidth is $2/T_0$ Hz. In this case the spectral efficiency (given by equation (11.36(b)) is:

$$\eta_s = 0.5 \log_2 M \quad \text{(bit/s/Hz)} \qquad (11.42(a))$$

The maximum possible, ISI free, spectral efficiency occurs when pulse shaping is such that signalling takes place in the double sided Nyquist bandwidth $B = 1/T_0$ Hz, i.e.:

$$\eta_s = \log_2 M \quad \text{(bit/s/Hz)} \qquad (11.42(b))$$

Thus we usually say BPSK has an efficiency of 1 bit/s/Hz and 16-PSK 4 bit/s/Hz. Table 11.4 compares the performance of several PSK systems.

Table 11.4 *Comparison of several PSK modulation techniques.*

	Required E_b/N_0 for $P_b = 10^{-6}$	*Minimum channel bandwidth for ISI free signalling (R_b = bit rate)*	*Max spectral efficiency (bit/s/Hz)*	*Required CNR in minimum channel bandwidth*
PRK	10.6 dB	R_b	1	10.6 dB
QPSK	10.6 dB	$0.5R_b$	2	13.6 dB
8-PSK	14.0 dB	$0.33R_b$	3	18.8 dB
16-PSK	18.3 dB	$0.25R_b$	4	24.3 dB

EXAMPLE 11.4

An MSPK, ISI free, system is to operate with 2^N PSK symbols over a 120 kHz channel. The minimum required bit rate is 900 kbit/s. What minimum CNR is required to maintain reception with a P_b no worse than 10^{-6}?

Maximum (ISI free) baud rate:

$$R_s = 1/T_o = B$$

Therefore $R_s \leq 120$ kbaud (or k symbol/s). Minimum required entropy is therefore given by:

$$H \geq \frac{R_b}{R_s} = \frac{900 \times 10^3}{120 \times 10^3} = 7.5 \quad \text{bit/symbol}$$

Minimum number of symbols required is given by:

$$H \leq \log_2 M$$

$$M \geq 2^H = 2^{7.5}$$

Since M must be an integer power of 2:

$$M = 2^8 = 256 \quad \text{and } H = \log_2 M = 8$$

For Gray coding of bits to PSK symbols:

$$P_e = P_b \log_2 M = 10^{-6} \log_2 256 = 8 \times 10^{-6}$$

$$P_e = 1 - \text{erf}\left[(T_o B)^{1/2} \sin\left(\frac{\pi}{M} \right) \left(\frac{C}{N} \right)^{1/2} \right]$$

Now:

$$R_s = \frac{R_b}{\log_2 M} = \frac{900 \times 10^3}{8} = 112.5 \times 10^3 \text{ baud}$$

$$T_o = \frac{1}{R_s} = 8.889 \times 10^{-6} \text{ s}$$

$$T_o B = 8.889 \times 10^{-6} \times 120 \times 10^3 = 1.067$$

Thus:

$$\frac{C}{N} = \left[\frac{\text{erf}^{-1}(1 - P_e)}{(T_o B)^{1/2} \sin\left(\frac{\pi}{M} \right)} \right]^2 = \left[\frac{\text{erf}^{-1}(1 - 8 \times 10^{-6})}{(1.067)^{1/2} \sin\left(\frac{\pi}{256} \right)} \right]^2$$

$$= \left[\frac{\text{erf}^{-1}(0.999\,992)}{1.033 \sin\left(\frac{\pi}{256} \right)} \right]^2 = \left(\frac{3.157}{0.01268} \right)^2 = 61988 = 47.9 \text{ dB}$$

11.4.3 Amplitude phase keyed and quadrature amplitude modulation

For an unsaturated transmitter operating over a linear channel it is possible to introduce amplitude as well as phase modulation to give an improved distribution of signal states in the signal constellation. The first such proposal (C. 1960) introduced a constellation with

two amplitude rings and eight phase states on each ring, Figure 11.19(a). For obvious reasons modulation schemes of this type are called amplitude/phase keying (APK). Subsequently it was observed that with half the number of points on the inner ring, Figure 11.19(b), a 3 dB performance improvement could be gained, as the constellation points are more evenly spaced over 12 distinct phases. The square constellation (Figure 11.19(c)), introduced C. 1962, is easier to implement and has a slightly better P_e performance yet. Since a square constellation APK signal can be interpreted as a pair of multilevel ASK (MASK) signals modulated onto quadrature carriers it is normally called quadrature amplitude modulation (QAM). (The terms APK and QAM are sometimes used interchangeably – here, however, we always use the term QAM to represent the square constellation subset of APK.)

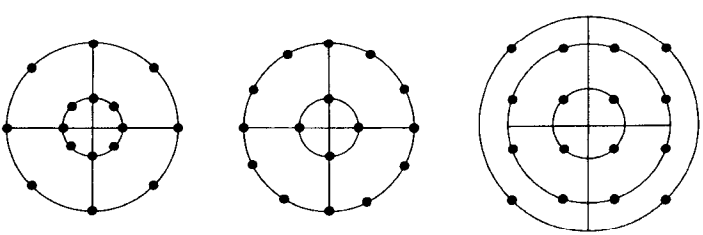
Figure 11.20 shows the time domain waveforms for a 16-state QAM constellation which are obtained by encoding binary data in 4-bit sequences. 2-bit sequences are each encoded in the I and Q channels into 4-level signals, as shown in Figure 11.20(a)/(c) and (b)/(d). These are combined to yield the full 16-state QAM, complex, signal, Figure (11.20(e)). Ideally all of the 16 states are equiprobable and statistically independent. Note that at the symbol sampling times there are 4 possible amplitudes in the inphase and quadrature channels reflecting the constellation's 4×4 structure, Figure 11.19(c). In fact the complex signal has three possible amplitude levels with the intermediate value being the most probable. A simple approximation for the probability of symbol error for MQAM (M even) signalling in Gaussian noise is [Carlson]:

$$P_e = 2\left\{\frac{M^{1/2}-1}{M^{1/2}}\right\}\left[1-\text{erf}\sqrt{\frac{3}{2(M-1)}\left(\frac{\langle E\rangle}{N_0}\right)^{1/2}}\right] \tag{11.43}$$

Where $\langle E\rangle$ is the average energy per QAM symbol. For equiprobable rectangular pulse symbols $\langle E\rangle$ is given by:

$$\langle E\rangle = \frac{1}{3}\left(\frac{\Delta V}{2}\right)^2 (M-1)\,T_o \tag{11.44}$$

where ΔV is the voltage separation between adjacent inphase or quadrature MASK levels and T_o is the symbol duration. Using $C = \langle E\rangle/T_o$ and $N = N_0 B$, equation (11.43) can be



(a) 16 APK (8,8) constellation

(b) 16 APK (4,12) constellation

(c) 16 APK (4,8,4) constellation (or 16 QAM)

Figure 11.19 *16-state quadrature amplitude modulated (QAM) signal constellations.*

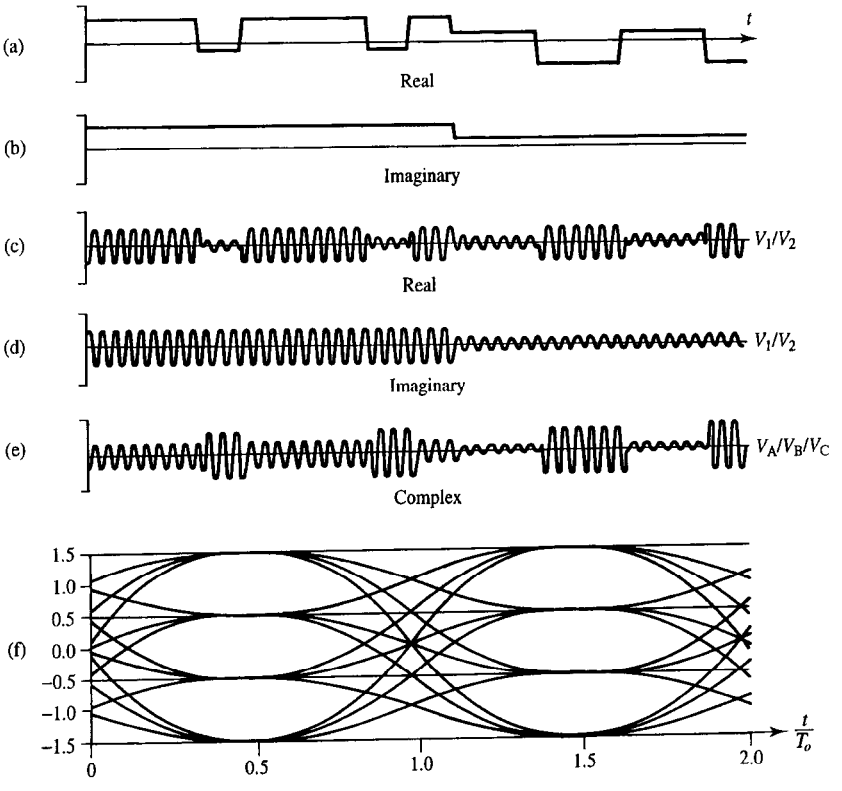


Figure 11.20 *16-state QAM signal: (a) 4-level baseband signals in the inphase and (b) quadrature branches; (c), (d) corresponding 4-level modulated complex signals; (e) resulting combined complex (3-level) QAM signal; (f) demodulated signal eye diagram over 2 symbols (in the I or Q channels).*

rewritten as:

$$P_e = 2\left\{\frac{M^{1/2}-1}{M^{1/2}}\right\}\left[1-\text{erf}\sqrt{\frac{3\,T_o B}{2(M-1)}\left(\frac{C}{N}\right)}^{1/2}\right] \tag{11.45}$$

For Gray code mapping of bits along the inphase and quadrature axes of the QAM constellation the probability of bit error P_b is given approximately by equation (11.40(a)). Denoting the average energy per bit by $E_b = \langle E \rangle / \log_2 M$, equation (11.43) can be written as:

$$P_b = \frac{2}{\log_2 M}\left\{\frac{M^{1/2}-1}{M^{1/2}}\right\}\left[1-\text{erf}\sqrt{\frac{3\log_2 M}{2(M-1)}\left(\frac{E_b}{N_0}\right)}^{1/2}\right] \tag{11.46(a)}$$

and equation (11.45) as:

$$P_b = \frac{2}{\log_2 M}\left\{\frac{M^{1/2}-1}{M^{1/2}}\right\}\left[1-\text{erf}\sqrt{\frac{3\,T_oB}{2(M-1)}\left(\frac{C}{N}\right)}^{1/2}\right] \qquad (11.46(b))$$

Like MPSK all the symbols in a QAM (or APK) signal occupy the same spectral space. The spectral efficiency is therefore identical to MPSK and is given (for statistically independent equiprobable symbols) by equation (11.36b). Unfiltered and Nyquist filtered APK signals therefore have (nominal and maximum) spectral efficiencies given by equations (11.42(a)) and (11.42(b)) respectively. Figure 11.21(a) compares the bit error probability of MPSK and MQAM systems. Note that the superior constellation packing in QAM over MPSK gives a lower required E_b/N_0 for the same P_b value. Figure 11.21(b) shows the spectral efficiencies of Nyquist filtered ($T_oB = 1$) MQAM and MPSK systems plotted against the CNR required for a bit error probability of 10^{-6}. Table 11.5 compares the C/N and E_b/N_0 required for a selection of MPSK and QAM schemes to yield a probability of bit error of 10^{-6}.

Table 11.5 *Comparison of various digital modulation schemes ($P_b = 10^{-6}$, $T_oB = 1.0$).*

Modulation	*C/N ratio (dB)*	E_b/N_0 *(dB)*
PRK	10.6	10.6
QPSK	13.6	10.6
4-QAM	13.6	10.6
8-PSK	18.8	14.0
16-PSK	24.3	18.3
16-QAM	20.5	14.5
32-QAM	24.4	17.4
64-QAM	26.6	18.8

EXAMPLE 11.5

Find the maximum spectral efficiency of ISI free 16-QAM. What is the noise induced probability of symbol error in this scheme for a received CNR of 24.0 dB if the maximum spectral efficiency requirement is retained? What is the Gray coded probability of bit error?

$$\eta_s = \frac{R_s\,H}{B} = \frac{H}{T_oB} \quad \text{(bit/s/Hz)}$$

For maximum (ISI free) spectral efficiency $T_oB = 1$ and $H = \log_2 M$, i.e.:

$$\eta_s = \log_2 M = \log_2 16 = 4 \quad \text{bit/s/Hz}$$

Probability of symbol error is given by equation (11.45) with $T_oB = 1$ and $C/N = 10^{2.4} = 251.2$:

$$P_e = 2\left\{\frac{4-1}{4}\right\}\left[1-\text{erf}\sqrt{\frac{3}{2(16-1)}(251.2)}^{1/2}\right]$$

$$= 1.5[1-\text{erf}(5.012)]$$

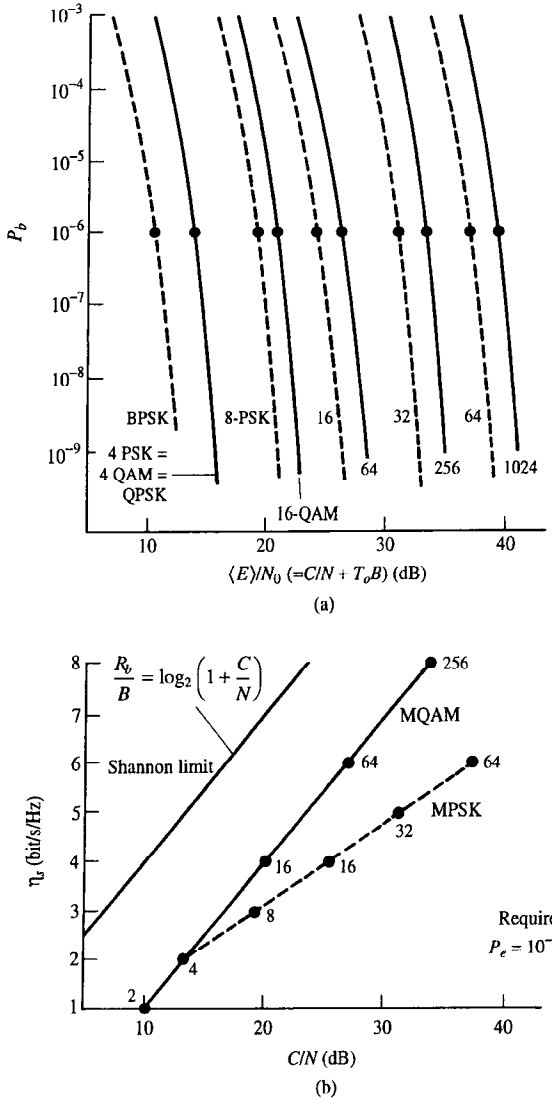


Figure 11.21 *P_e and spectral efficiency for multi-phase PSK and M-QAM modulation: (a) bit error probability against CNR with PSK shown as dashed and QAM as solid curves; (b) comparison of the spectral efficiency of these modulation schemes.*

For $x \geq 4.0$ the error function can be approximated by:

$$\text{erf}(x) = 1 - \frac{e^{-x^2}}{(\sqrt{\pi}\, x)}$$

Therefore:

$$P_e = \frac{1.5\, e^{-5.012^2}}{(\sqrt{\pi}\ 5.012)} = 2.08 \times 10^{-12}$$

For Gray coding:

$$P_b = \frac{P_e}{\log_2 M} = \frac{2.08 \times 10^{-12}}{\log_2 16} = 5.2 \times 10^{-13}$$

11.4.4 Quadrature phase shift keying (QPSK) and offset QPSK

Quadrature phase shift keying (QPSK) [Aghvami] can be interpreted either as 4-PSK with carrier amplitude A (i.e. quaternary PSK) or as a superposition of two (polar) BASK signals with identical 'amplitudes' $\pm A/\sqrt{2}$ and quadrature carriers, $\cos 2\pi f_c t$ and $\sin 2\pi f_c t$, i.e. 4-QAM. The constellation diagram of a QPSK signal is shown in Figure 11.22. The accompanying arrows show that all transitions between the four states are possible. Figure 11.23 shows a schematic diagram of a QPSK transmitter and Figure 11.24 shows the receiver.

The transmitter and receiver are effectively two PRK transmitters and receivers arranged in phase quadrature, the inphase (I) and quadrature (Q) channels each operating at half the bit rate of the QPSK system as a whole. If pulse shaping and filtering are absent the signal is said to be unfiltered, or rectangular pulse, QPSK. Figure 11.25 shows an example sequence of unfiltered QPSK data and Figure 11.26(a) and (b) shows the corresponding power spectral density with $T_o = 2T_b$.

The spectral efficiency of QPSK is twice that for BPSK. This is because the symbols in each quadrature channel occupy the same spectral space and have half the spectral width of a BPSK signal with the same data rate as the QPSK signal. This is illustrated in the form of orthogonal voltage spectra in Figure 11.27.

The P_e performance of QPSK systems will clearly be worse than that of PRK systems, Figure 11.18, since the decision regions on the constellation diagram, Figure

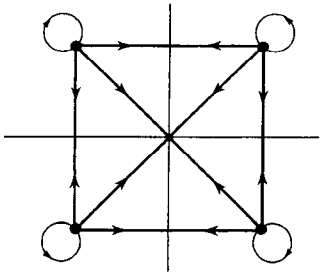


Figure 11.22 *Quadrature phase shift keyed (QPSK) signal constellation showing allowed state transitions.*

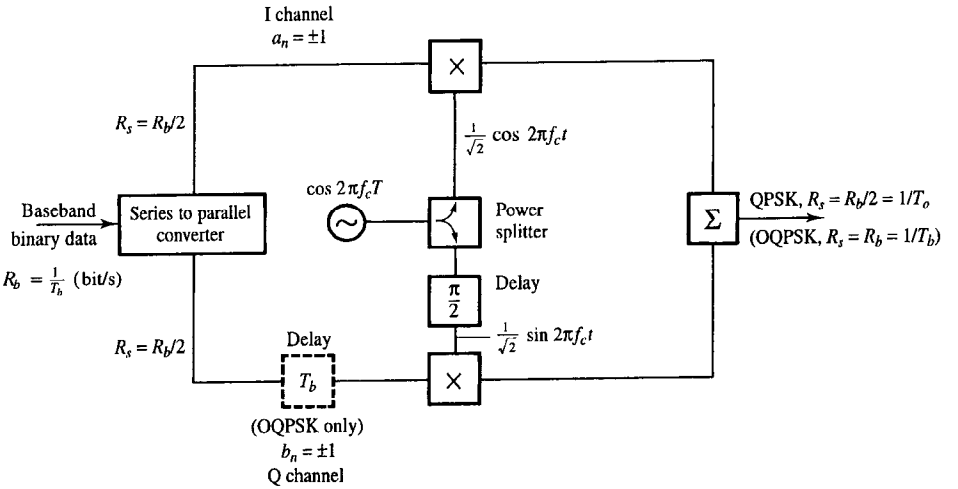


Figure 11.23 *Schematic for QPSK (OQPSK) modulator.*

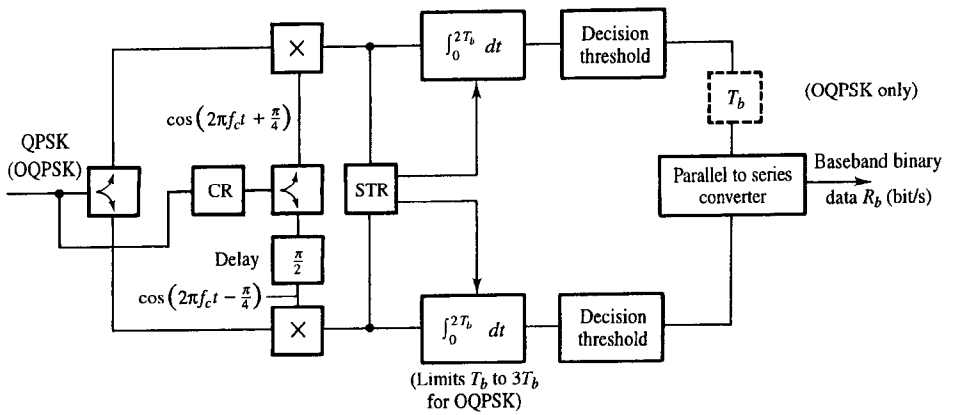


Figure 11.24 *Schematic for QPSK (OQPSK) demodulator.*

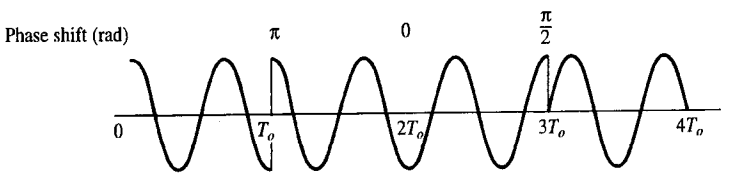


Figure 11.25 *Unfiltered (constant envelope) QPSK signal ($T_o = 2T_b$).*

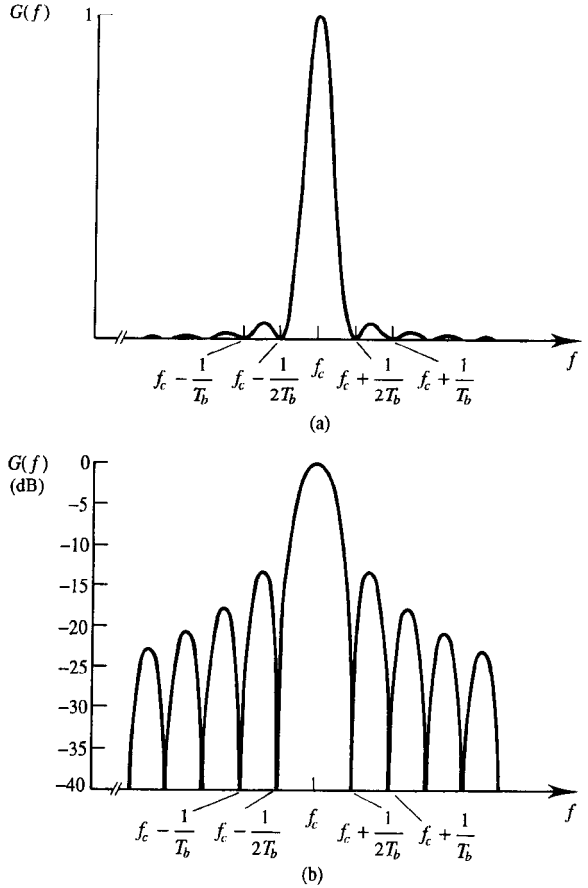


Figure 11.26 *Power spectral density of QPSK/OQPSK signals.* $G(f) \propto [(\sin(4\pi fT_b))/(4\pi fT_b)]^2$ *on (a) linear; and (b) dB scales.*

11.22, are reduced from half spaces to quadrants. The P_b performances of these modulation schemes are, however, identical. This is easily appreciated by recognising that each channel (I and Q) of the QPSK system is independent of (orthogonal to) the other. In principle, therefore, the I channel (binary) signal could be transmitted first followed by the Q channel signal. The total message would take twice as long to transmit in this form but, because each QPSK I or Q channel symbol is twice the duration and half the power of the equivalent PRK symbols, the total message energy (and therefore the energy per bit, E_b) is the same in both the QPSK and PRK cases, Figure 11.28. The P_b performance of ideal QPSK signalling is therefore given by equation (11.11) with E interpreted as the energy per bit, E_b, i.e.:

$$P_b = \frac{1}{2}\left[1 - \text{erf}\left(\frac{E_b}{N_0}\right)^{1/2}\right] \qquad (11.47(a))$$

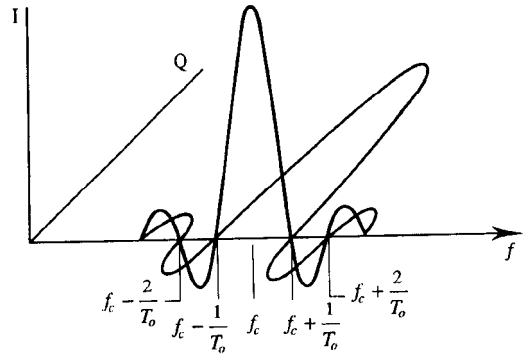


Figure 11.27 *QPSK orthogonal inphase (I) and quadrature (Q) voltage spectra ($T_o = 2T_b$).*

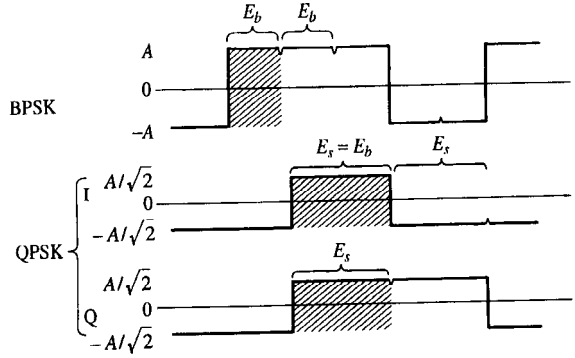


Figure 11.28 *Envelopes of (unfiltered) BPSK and equivalent QPSK signals showing distribution of bit energy in I and Q channels.*

In terms of CNR, therefore:

$$P_b = \frac{1}{2}\left[1 - \text{erf}\,(T_b B)^{1/2}\left(\frac{C}{N}\right)^{1/2}\right] \qquad (11.47(b))$$

where the bit period T_b is half the QPSK symbol period, T_o. This is the origin of the statement that the BER performances of QPSK and PRK systems are identical. (Note that the minimum $T_o B$ product for ISI free QPSK signalling is 1.0, corresponding to $T_b B = 0.5$.) The probability of symbol error, P_e, is given by the probability that the I channel bit is detected in error, or the Q channel bit is detected in error, or both channel bits are detected in error, i.e.:

$$P_e = P_b(1 - P_b) + (1 - P_b)P_b + P_b P_b = 2P_b - P_b^2 \qquad (11.48)$$

Offset QPSK (OQPSK), also sometimes called staggered QPSK, is identical to QPSK

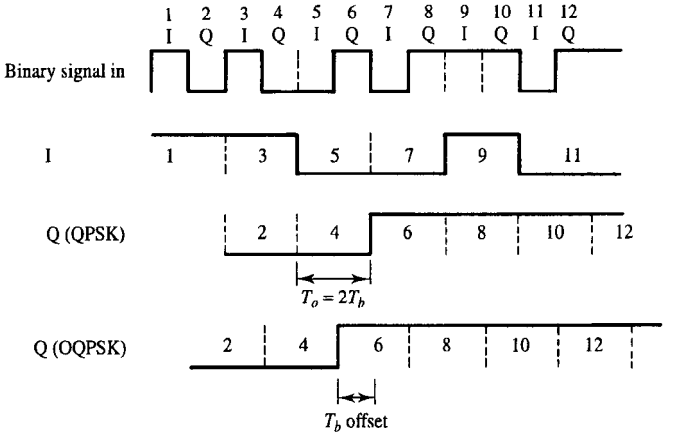


Figure 11.29 *Input bit stream and I and Q channel bit-streams for QPSK and OQPSK systems.*

except that, immediately prior to multiplication by the carrier, the Q channel symbol stream is offset with respect to the I channel symbol stream by half a QPSK symbol period (i.e. one bit period), Figure 11.23. The OQPSK relationships between the input binary data stream and the I and Q channel bit streams are shown in Figure 11.29, and the OQPSK constellation and state transition diagram is shown in Figure 11.30. The lines on Figure 11.30 represent the allowed types of symbol transition and show that transitions across the origin (i.e. phase changes of 180°) are, unlike QPSK, prohibited. The offset between I and Q channels means that OQPSK symbol transitions occur, potentially, every T_b seconds (i.e. $T_o = T_b$).

The spectrum, and therefore spectral efficiency, of OQPSK is thus identical to that of QPSK. (Transitions occur more often but, on average, are less severe.) OQPSK has a potential advantage over QPSK, however, when transmitted over a non-linear channel. This is because envelope variations in the signal at the input of a non-linear device produce distortion of the signal (and therefore the signal spectrum) at the output of the device. This effect leads to spectral spreading which may give rise, ultimately, to adjacent channel interference. To minimise spectral spreading (and some other undesirable effects, see section 14.3.3) in non-linear channels, constant envelope

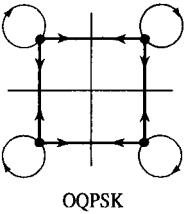


Figure 11.30 *Offset QPSK (OQPSK) constellation diagram showing allowed state transitions.*

modulation schemes are desirable. Since some degree of filtering (i.e. pulse shaping) is always present in a modulated RF signal (even if this is unintentional, and due only to the finite bandwidth of the devices used in the modulator) then the 180° phase transitions present in a QPSK signal will result in severe (∞ dB) envelope fluctuation. In contrast the envelope fluctuation of filtered OQPSK is limited to 3 dB since only one quadrature component of the signal can reverse at any transition instant, Figure 11.31. The spectral spreading suffered by an OQPSK signal will therefore be less than that suffered by the equivalent QPSK signal for the same non-linear channel.

Quadrature partial response systems

Quadrature partial response (QPR) modulation is a quadrature modulated form of partial response signalling (sections 8.2.7 and 8.2.8). A block diagram showing the structure of a duobinary QPR transmitter is shown in Figure 11.32(a) and the resulting duobinary constellation diagram is shown in Figure 11.32(b). The (absolute) bandwidth of the (DSB) QPR signal is $1/T_0$ (Hz) or $1/2T_b$ (Hz) and the spectral efficiency is 2 bit/s/Hz.

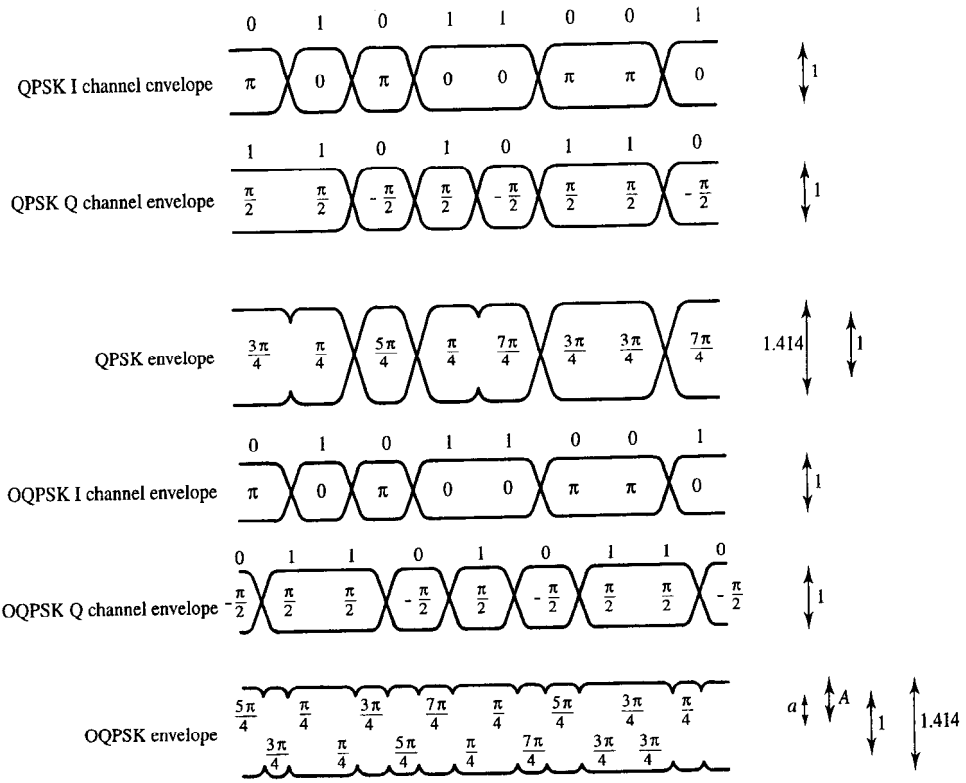


Figure 11.31 *Origin and comparison of QPSK and OQPSK envelope fluctuations ($a/A = 1/\sqrt{2}$).*

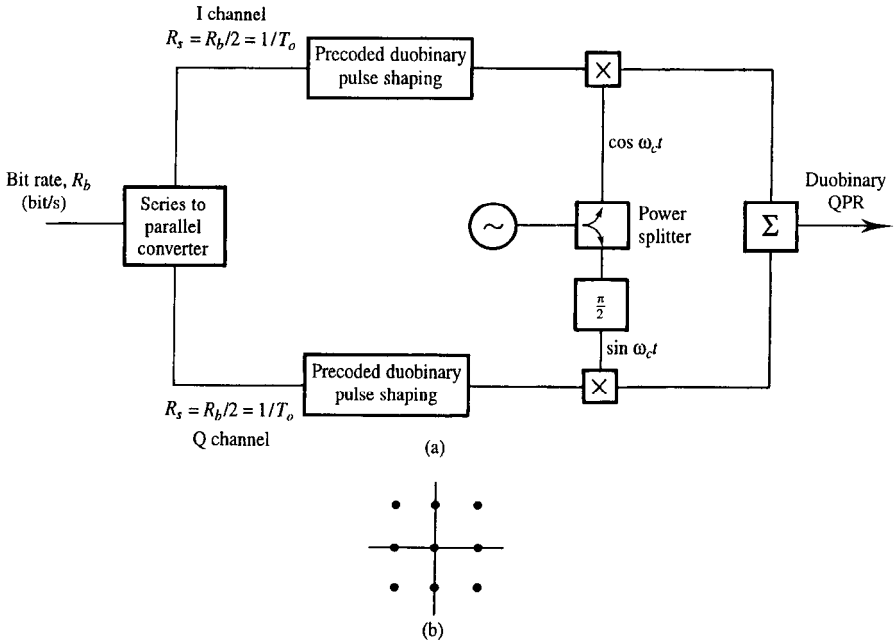


Figure 11.32 *(a) Duobinary QPR transmitter (b) constellation diagram.*

This is the same spectral efficiency as baseband duobinary signalling since a factor of 2 is gained by using quadrature carriers but a factor of 2 is lost by using DSB modulation. The duobinary QPR probability of bit error in the presence of white Gaussian noise [Taub and Schilling] corresponds to a CNR of approximately 3 dB more than that required for the same P_e performance in QPSK and 2 dB less than that required in 8-PSK.

11.4.5 Minimum shift keying

Minimum shift keying (MSK) is a modified form of OQPSK in that I and Q channel sinusoidal pulse shaping is employed prior to multiplication by the carrier, Figure 11.33 [Pasapathy]. The transmitted MSK signal can be represented by:

$$f(t) = a_n \sin\left(\frac{2\pi t}{4T_b}\right)\cos 2\pi f_c t + b_n \cos\left(\frac{2\pi t}{4T_b}\right)\sin 2\pi f_c t \qquad (11.49)$$

where a_n and b_n are the nth I and Q channel symbols. Figure 11.34(a) to (c) shows an example sequence of binary data, the corresponding, sinusoidally shaped, I and Q symbols and the resulting MSK signal. MSK signalling is an example from a class of modulation techniques called continuous phase modulation (CPM). It has a symbol constellation which must now be interpreted as a time varying phasor diagram, Figure 11.34(d). The phasor rotates at a constant angular velocity from one constellation point to an adjacent point over the duration of one MSK symbol. (Like OQPSK the MSK

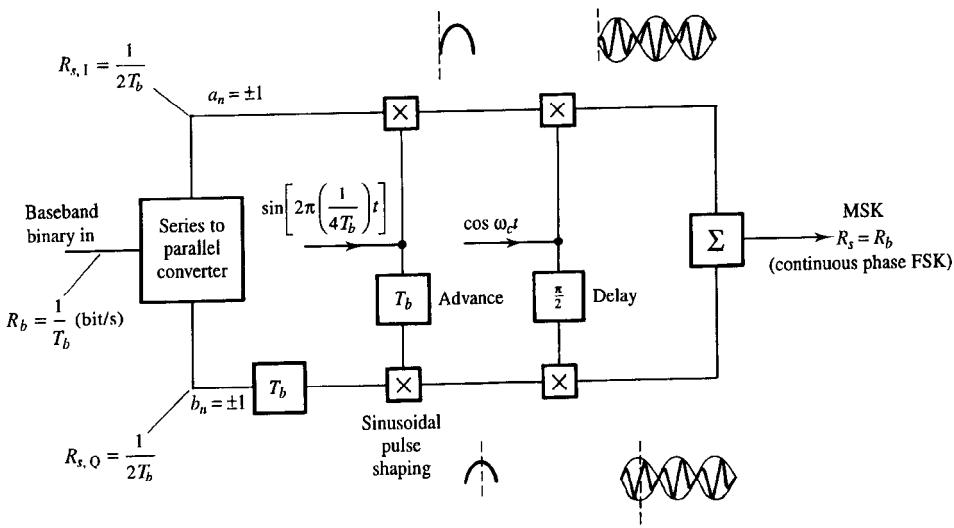


Figure 11.33 *Minimum shift keyed (MSK) transmitter.*

symbol period is the same as the bit period of the unmodulated binary data.) When $a_n = b_n$ the phasor rotates clockwise and when $a_n \neq b_n$ the phasor rotates anticlockwise. A consequence of MSK modulation is that one to one correspondence between constellation points and the original binary data is lost.

An alternative interpretation of MSK signalling is possible in that it can be viewed as a special case of BFSK modulation (and is therefore sometimes called continuous phase FSK). This is obvious, in that when the phasor in Figure 11.34(d) is rotating anticlockwise the MSK symbol has a constant frequency of $f_c + 1/(4T_b)$ Hz and when rotating clockwise it has a frequency of $f_c - 1/(4T_b)$. This corresponds to BFSK operated at the first zero crossing point of the $\rho - T_o$ diagram (see section 11.3.4). When viewed this way MSK can be seen to be identical to BFSK with inherent differential coding (since frequency f_1 is transmitted when $a_n = b_n$ and frequency f_2 is transmitted when $a_n \neq b_n$). A one to one relationship between bits and frequencies can be re-established, however, by differentially precoding the serial input bit stream prior to MSK modulation. The appropriate precoder is identical to that used for differential BPSK, Figure 11.13. When precoding is used the modulation is called fast frequency shift keying (FFSK) although this term is sometimes also used indiscriminately for MSK without precoding.

Since MSK operates at the first zero on the $\rho - T_o$ diagram of Figure 11.8 one of the BFSK signalling frequencies has an integer number of cycles in the symbol period and the other has either one half cycle less or one half cycle more.

The normalised power spectral density of MSK/FFSK is shown in Figure 11.35. The spectra of MSK/FFSK and QPSK/OQPSK are compared, on a dB scale, in Figure 11.36. As can be seen the MSK spectrum has a broader main lobe but more rapidly decaying sidelobes, which is particularly attractive for FDM systems in order to achieve reduced

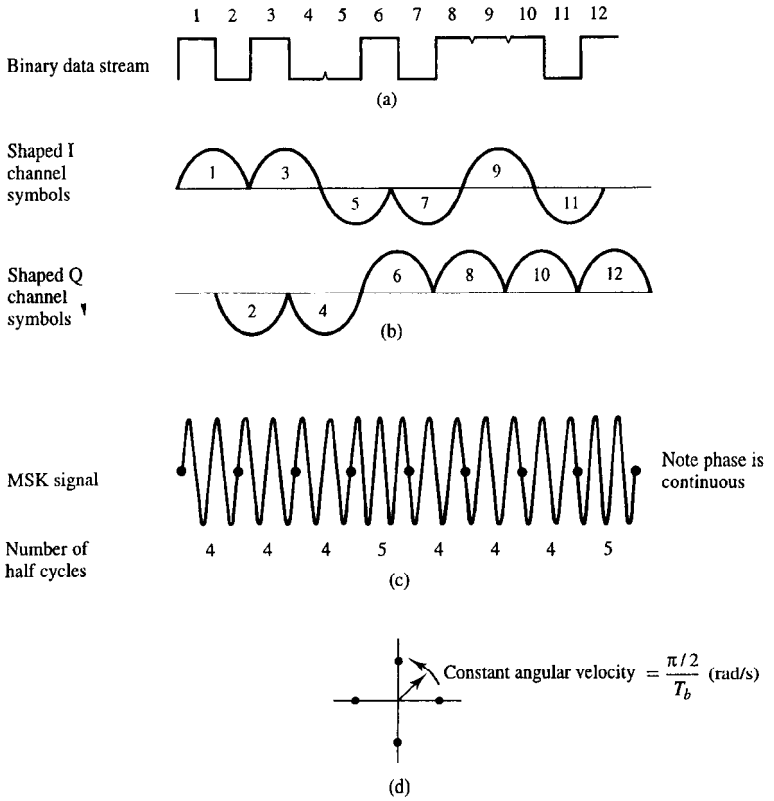


Figure 11.34 *MSK transmitter waveforms and phasor diagrams: (a) binary data; (b) sinusoidally shaped channel symbols; (c) transmitted MSK signal; (d) MSK phasor diagram.*

adjacent channel interference. The probability of bit error for ideal MSK detection, Figure 11.37, is identical to that for QPSK (equations (11.47)) systems since orthogonality between I and Q channels is preserved. (MSK with differentially precoded data, i.e. FFSK, suffers the same degradation in P_b performance as differentially encoded PSK.)

EXAMPLE 11.6

Find the probability of bit error for a 1.0 Mbit/s MSK transmission with a received carrier power of −130 dBW and a NPSD, measured at the same point, of −200 dBW/Hz.

$$C = 10^{\frac{-130}{10}} = 10^{-13} \text{ W}$$

$$E_b = C\,T_b = 10^{-13} \times 10^{-6} = 10^{-19} \text{ J}$$

$$N_0 = 10^{\frac{-200}{10}} = 10^{-20} \text{ W/Hz}$$

$$P_b = \frac{1}{2}\left[1 - \text{erf}\left(\frac{E_b}{N_0}\right)^{1/2}\right] = \frac{1}{2}\left[1 - \text{erf}\left(\frac{10^{-19}}{10^{-20}}\right)^{1/2}\right]$$

$$= \tfrac{1}{2}\left[1 - \text{erf}\,(3.162)\right] = \tfrac{1}{2}\left[1 - 0.999\,999\,24\right] = 3.88 \times 10^{-6}$$

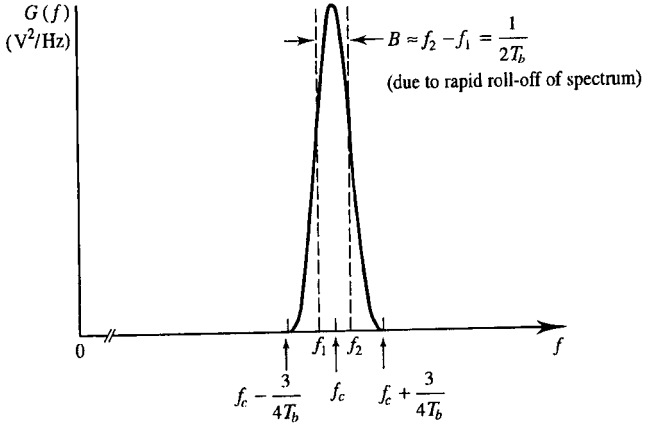


Figure 11.35 *Power spectral density of MSK/FFSK signal.* $G(f) \propto [(\cos(2\pi f T_b))/(1 - 16f^2T_b^2)]^2$. *(Note: no spectral lines are present.)*

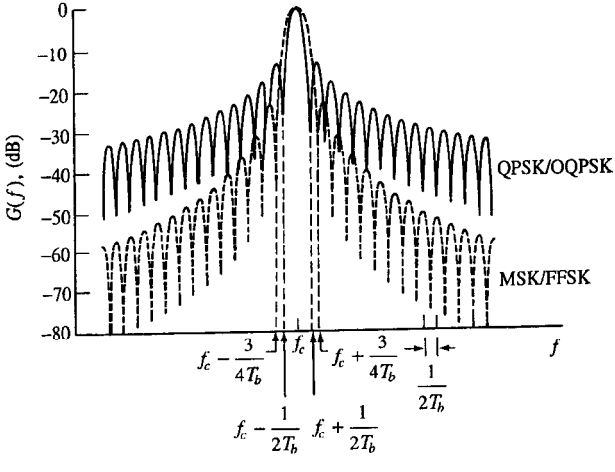


Figure 11.36 *Comparison of QPSK/OQPSK and MSK/FFSK spectra (spectral envelopes roll-off with* $(f - f_c)^{-2}$, *i.e. −6 dB/octave and* $(f - f_c)^{-4}$, *i.e. −12 dB/octave).*

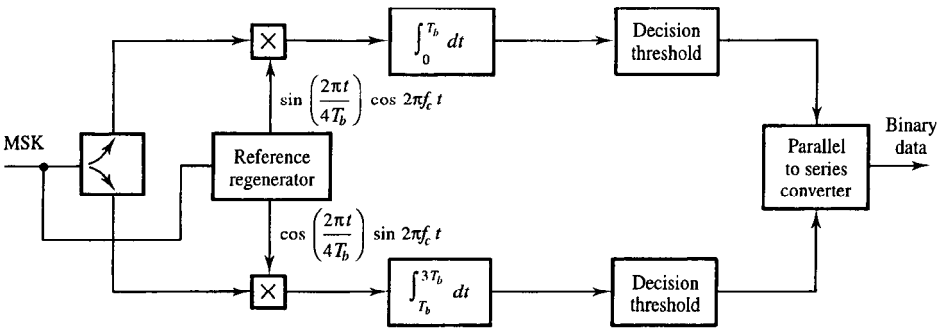


Figure 11.37 *Schematic of MSK demodulator.*

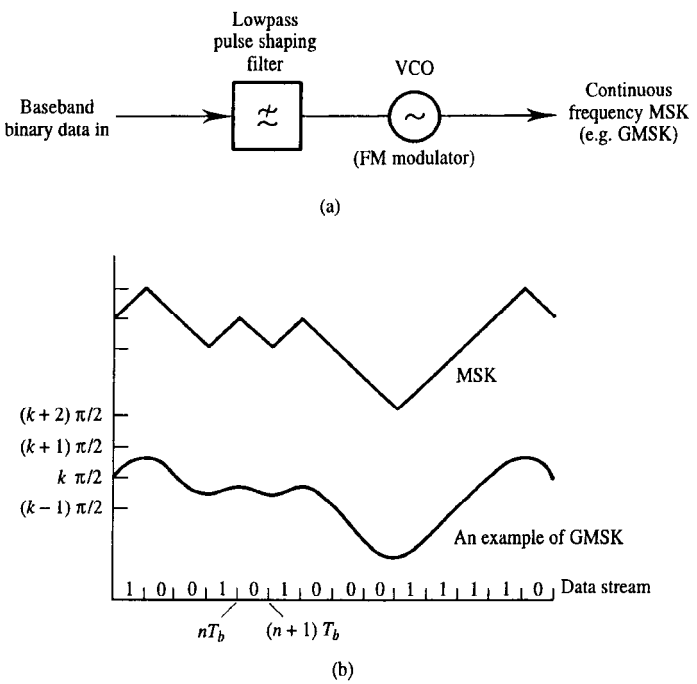


Figure 11.38 *(a) Generation of GMSK; (b) MSK and GMSK phase trajectories for typical bit sequences with the MSK trajectory displaced vertically for clarity (source: Hirade and Murota, 1979, reproduced with the permission of the IEEE).*

11.4.6 Gaussian MSK

Although the phase of an MSK signal is continuous, its first derivative (i.e. frequency) is discontinuous. A smoother modulated signal (with correspondingly narrower spectrum) can be generated by reducing this frequency discontinuity. The conceptually simplest way of achieving this is to generate the MSK signal directly as a BFSK signal (with $(f_2 - f_1)T_o = ½$) using a voltage controlled oscillator, and to employ premodulation pulse shaping of the baseband binary data, Figure 11.38(a). If the pulse shape adopted is Gaussian then the resulting modulation is called Gaussian MSK (GMSK). Figure 11.38(b) compares the phase trajectory of MSK and GMSK signals and Figure 11.39 shows the corresponding power spectra. The parameter B_bT_o in Figure 11.39 is the normalised bandwidth of the premodulation lowpass filter in Figure 11.38(a). (B_b is the baseband bandwidth and T_o is the symbol, or bit, period.) $B_bT_o = \infty$ corresponds to no filtering and therefore MSK signalling. Figure 11.40 shows the impulse response of the premodulation filter for various values of B_bT_o. The baseband Gaussian pulses essentially occupy $1/(B_bT_o)$ bit periods which results in significant sampling instant ISI for $B_bT_o < 0.5$. The improvement in spectral efficiency for $B_bT_o < 0.5$ outweighs the resulting degradation in BER performance, however, in some applications. Figure 11.41 shows measured P_e performance curves for GMSK modulation. GMSK with $B_bT_o = 0.3$ is the modulation scheme adopted by the CEPT's Groupe Spéciale Mobile (Global

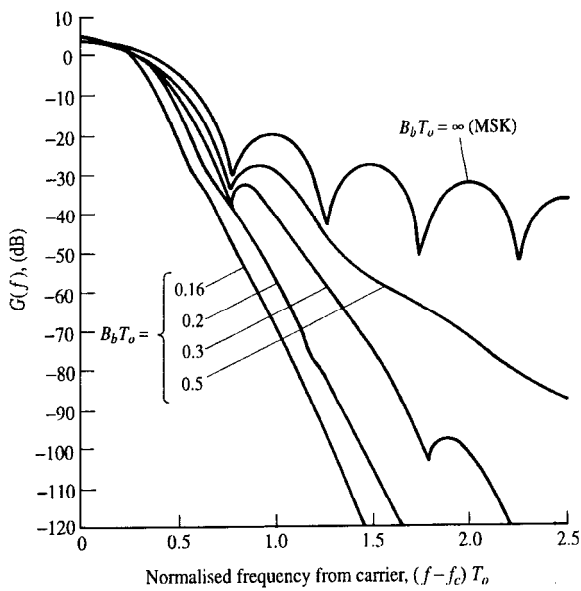


Figure 11.39 *Comparison of GMSK spectra with various values of B_bT_o (B_b is the bandwidth of the premodulation pulse shaping filter, $T_o(= T_b)$ is the symbol period (source: Parsons and Gardiner, 1989, reproduced with the permission of Chapman & Hall).*

System for Mobile), GSM, for the implementation of digital cellular radio systems in Europe, see Chapter 15.

Although Figure 11.38(a) implies a simple implementation for GMSK, in practice the implementation is significantly more complex.

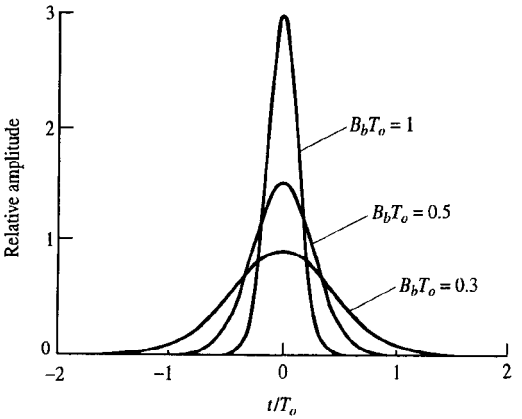


Figure 11.40 *Gaussian filter impulse response for different B_bT_o products.*

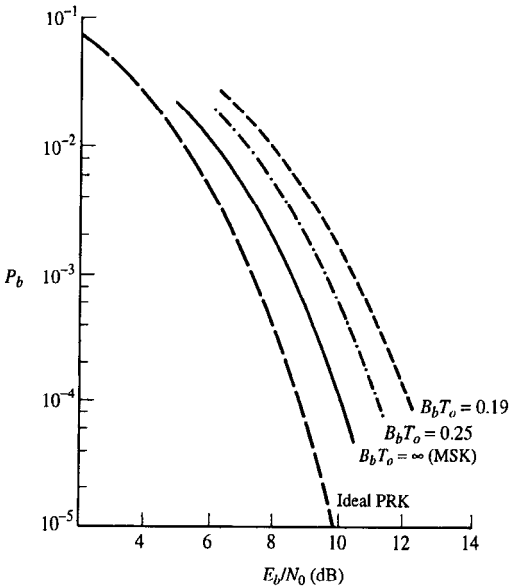


Figure 11.41 *Measured P_b versus E_b/N_0 for MSK with various values of B_bT_o after bandlimiting by an ideal bandpass channel with bandwidth $B = 0.75/T_o$. ($T_o = T_b$) (Source: Hirade and Murota, 1979, reproduced with the permission of the IEEE.)*

11.4.7 Trellis coded modulation

Trellis coded modulation (TCM) [Ungerboeck] is a combined coding and modulation technique for digital transmission over band-limited channels. In TCM there is a restriction on which states may occur so that for a given received symbol sequence the number of possible transmitted symbol sequences is reduced. This increases the decision space across the constellation. Its main attraction is that it allows significant coding gains over the previously described (uncoded) multi-level modulation schemes, without compromising the spectral efficiency.

TCM employs redundant non-binary modulation in combination with a finite-state encoder which governs the selection of transmitted symbols to generate the coded symbol sequences. In the receiver, the noisy symbols are decoded by a soft-decision maximum-likelihood sequence (Viterbi) decoder as described for FEC decoding in Chapter 10. Simple four-state TCM schemes can improve the robustness of digital transmission against additive noise by 3 dB, compared to conventional uncoded modulation. With more complex TCM schemes, the coding gain can reach 6 dB or more. These gains are obtained without bandwidth expansion, or reduction of the effective information rate, as required by traditional error-correction schemes. Figure 11.42 depicts symbol sets and Figure 11.43 the state-transition (trellis) diagrams for uncoded 4-PSK modulation and coded 8-PSK modulation with four trellis states. The trivial one-state diagram in Figure 11.43(a) is shown only to illustrate uncoded 4-PSK from the viewpoint of TCM. Each one of the four connected paths labelled 0, 2, 4, 6 through the trellis in Figure 11.43(a) represents an allowed symbol sequence based on two binary bits (since $2^2 = 4$). In both

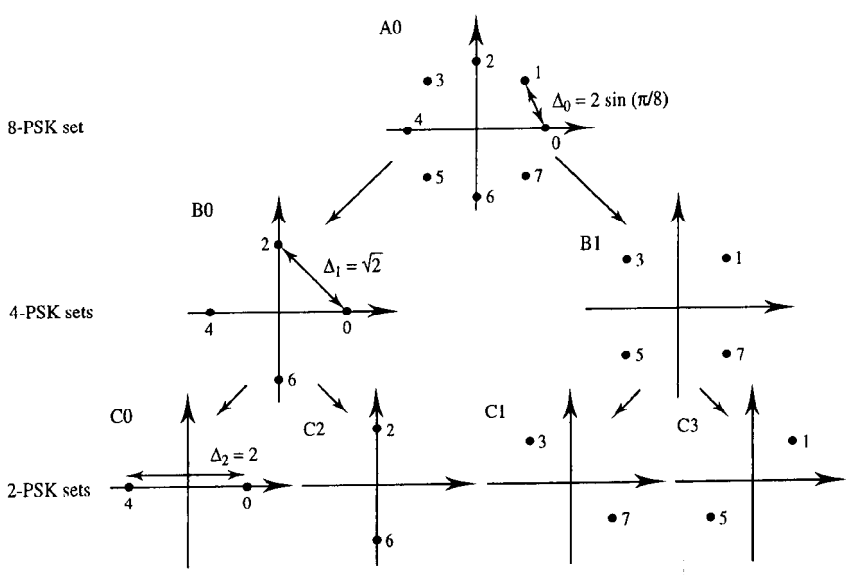
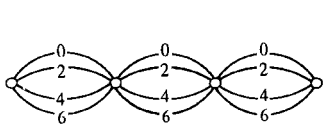
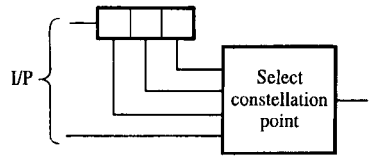


Figure 11.42 *Partitioning of 8-PSK into 4-phase (B0/B1) and 2-phase (C0 ··· C3) subsets.*

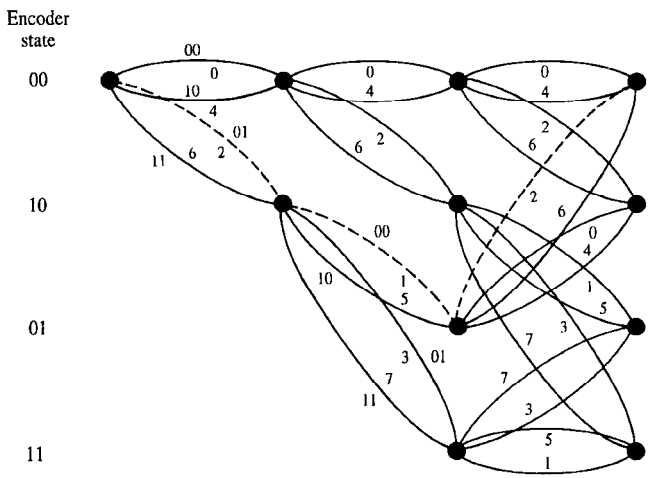


One-state trellis for B0 4-PSK set

(a)



(b)



(c)

Figure 11.43 *Trellis coded modulation (TCM): (a) trellis diagram for 4-PSK uncoded transmission where there is no restriction in the choice of constellation point; (b) four-state TCM encoder; (c) resulting four-state TCM 8-PSK trellis diagram.*

systems, starting from any state, four transitions can occur, as required to encode two information bits per modulation interval (and obtain a spectral efficiency of 2 bit/s/Hz).

The four 'parallel' transition paths in the one-state trellis diagram of Figure 11.43(a) for uncoded 4-PSK do not restrict the sequence of 4-PSK symbols that can be transmitted as there is no sequence coding. For unit transmitted amplitude the smallest distance between the 4-PSK signals is $\sqrt{2}$, denoted by Δ_1 in Figure 11.42. 4-PSK at 2 bit/s/Hz has identical spectral efficiency to rate 2/3 FECC data transmitted over an uncoded 8-PSK system.

In order to understand the ideas behind TCM the 8-PSK signal set can be partitioned as shown in Figure 11.42 into B and C subsets with increasing distance or separation between the constellation states [Burr]. The encoder, like the convolutional coder of Chapter 10, consists of a shift register which contains the current data and a number of previous data bits, so that the state of the encoder depends on the previous data samples,

Figure 11.43(b). The nodes in Figure 11.43(c) represent the state of the encoder at a given time, while the branches leaving each node represent the different input data which arrives in that symbol period. Each branch is labelled with the input data bit pairs and the number of the constellation point which is transmitted for that encoder state.

Any pair of paths between the same two nodes represents a pair of codewords which could be confused. In TCM we choose a symbol from the different subsets of the constellation for each branch in the trellis to maximise the distance between such pairs and, therefore, minimise the error probability. This means using points from the same C subset, Figure 11.42, for these parallel branches. Other branches which diverge from, or converge to, the same node require a smaller minimum distance and are chosen from the same B subset.

Consider the trellis diagram of Figure 11.43(c). The closest paths (in terms of Euclidean distance) to the all-zeros symbol path, which is used as a reference, are shown dashed. There are four branches in the trellis from each node, corresponding to two bits per transmitted constellation point (or channel symbol). Thus the scheme can be compared directly with uncoded QPSK, which also transmits two bits per symbol. But in uncoded QPSK the minimum Euclidean distance between transmitted signals is that between two adjacent constellation points, i.e. $\Delta_1 = \sqrt{2}$, Figure 11.42. The TCM scheme achieves $\sqrt{2}$ times this distance, i.e. $\Delta_2 = 2$. As squared distances correspond to signal powers, the squared Euclidean distance between two points determines the noise tolerance between them. The TCM code has twice the square distance of uncoded QPSK, thus it will tolerate twice, or 3 dB, more noise power.

In general TCM offers coding gains of 3 to 6 dB at spectral efficiencies of 2 bit/s/Hz or greater. An 8-state coder (with a 16-state constellation) gives 4 dB of gain while 128 states are required for 6 dB gain. TCM is thus a variation on FECC, Chapter 10, where the redundancy is obtained by increasing the number of vectors in the signal constellation rather than by explicitly adding redundant digits, thus increasing bandwidth. Further design of the TCM signal constellation to cluster the points into *clouds* of closely spaced points with larger gaps between the individual clouds can obtain error control where there are varying degrees of protection on different information bits. Such systems, which offer unequal error protection, are being actively considered for video signal coding where the priority bits reconstruct the basic picture and the less well protected bits add in more detail, to avoid picture loss in poor channels.

Table 11.6 *Commercially available TCM coder chipset examples.*

Data rate (Mbit/s)		75
Rate, *R*	2/3	3/4
Modulation	8-PSK	16-PSK
Coding gain (dB) at 10^{-5}	3.2	3.0

EXAMPLE 11.7

Compare the carrier to noise ratio required for a P_b of 10^{-6} with: (a) 4-PSK; and (b) TCM derived 8-PSK, using Figure 11.43. How does the performance of the TCM system compare with PRK operation? Is it possible to make any enhancements to the TCM coder to achieve further reductions in the CNR requirement?

For uncoded 4-PSK the distance between states of amplitude 1 is $\sqrt{2}$ and the CNR for $P_b = 10^{-6}$ is given in Figure 11.21 and Table 11.5. The CNR requirement (for $T_o B = 1$) is 13.6 dB.

For TCM based 8-PSK the minimum distance is increased from $\sqrt{2}$ to 2 and the CNR requirement reduces by 3 dB. Thus the minimum CNR will be 10.6 dB.

TCM 8-PSK therefore operates at the same CNR as PRK but achieves twice the spectral efficiency of 2 bit/s/Hz.

The TCM encoder performance can be improved further by extending the constraint length, but this increases the encoder and decoder complexity. (Blahut (1990) shows, in his Table 6.1, that increasing constraint length to 8 increases the coding gain to 5.7 dB.)

11.5 Power efficient modulation techniques

Some communications systems operate in environments where large bandwidths are available but signal power is limited. Such power limited systems rely on power efficient modulation schemes to achieve acceptable bit error, and data, rates. In general data rate can be improved by increasing the number of symbols (i.e. the alphabet size) at the transmitter. If this is to be done without degrading P_e then the enlarged alphabet of symbols must remain at least as widely spaced in the constellation as the original symbol set. This can be achieved without increasing transmitted power by adding orthogonal axes to the constellation space – a technique which results in multidimensional signalling. Power can also be conserved by carefully optimising the arrangement of points in the constellation space. The most significant power saving comes from optimising the lattice pattern of constellation points. This is called symbol packing and results in an increased constellation point density without decreasing point separation. An additional (small) power saving can be obtained by optimising the boundary of the symbol constellation.

11.5.1 Multidimensional signalling and MFSK

Multi-frequency shift keying (MFSK) is a good example of a power efficient, multidimensional, modulation scheme if, as is usually the case, its symbols are designed to be mutually orthogonal, section 2.5.3. Figure 11.44(a), developed from Figure 11.9, shows the voltage spectrum of an orthogonal MFSK signal as a superposition of OOK signals and Figure 11.44(b) shows the power spectrum plotted in dB. The increased data rate realised by MFSK signalling is achieved entirely at the expense of increased bandwidth. Since each symbol (for equiprobable, independent, symbol systems) conveys $H = \log_2 M$ bits of information then the nominal spectral efficiency of orthogonal MFSK is given by:

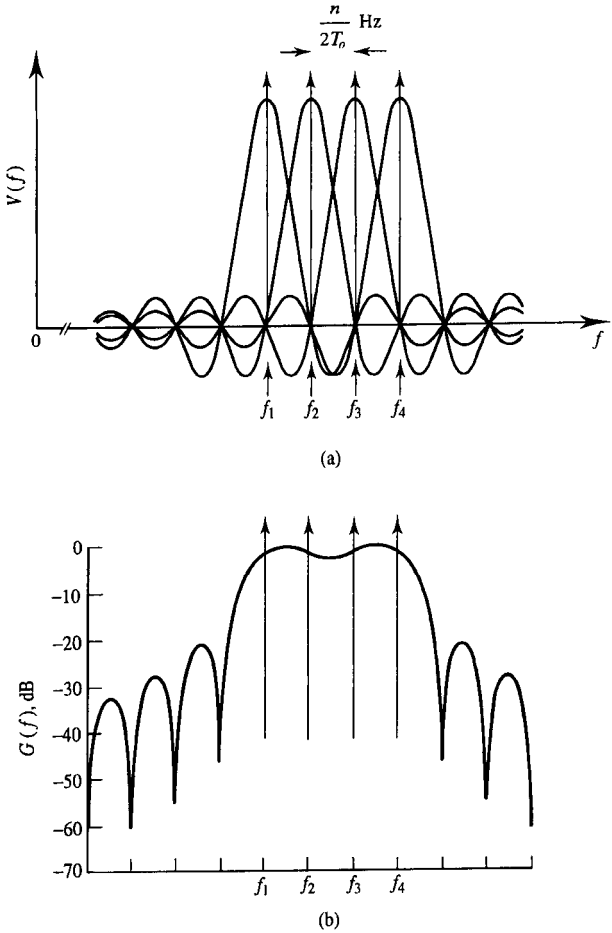


Figure 11.44 *Spectrum of orthogonal MFSK signal ($M = 4$): (a) as a superposition of OOK signal voltage spectra and (b) combined power spectrum on a dB scale. Tone spacing corresponds to second zero crossing point ($n = 2$) on $\rho - T_o$ diagram.*

$$\eta_s = \frac{\log_2 M}{(n/2)(M-1)+2} \quad \text{(bit/s/Hz)} \tag{11.50}$$

where n is the selected zero crossing point on the $\rho - T_o$ diagram of Figure 11.8 and the (nominal) signal bandwidth is defined by the first spectral nulls above and below the highest and lowest frequency symbols respectively. For incoherent detection ($n \geq 2$) the maximum spectral efficiency ($n = 2$) in Figure 11.44(b) is given by:

$$\eta_s = \frac{\log_2 M}{M+1} \quad \text{(bit/s/Hz)} \tag{11.51}$$

EXAMPLE 11.8

Find the maximum spectral efficiency of an 80 kbaud, 8-FSK, orthogonal signalling system which operates with a frequency separation between adjacent tones corresponding to the third zero crossing point of the $\rho - T_o$ diagram of Figure 11.8. Compare this with the nominal spectral efficiency as given by equation (11.50).

The frequency separation between orthogonal tones satisfies:

$$n = 2\,(f_i - f_{i-1})\,T_o$$

In this case, therefore:

$$(f_i - f_{i-1}) = \frac{3}{2\,T_o}$$

$$= \frac{3 \times 80 \times 10^3}{2} = 120 \times 10^3 \quad \text{Hz}$$

The maximum (ISI free) spectral efficiency is realised when $T_o B_{symbol} = 1.0$ and $H = \log_2 M$. The bandwidth of each FSK symbol is therefore $B_{symbol} = 1/T_o$ Hz and for the MFSK signal is:

$$B_{opt\ MFSK} = (M-1)(f_i - f_{i-1}) + 2\,(\tfrac{1}{2} \times B_{symbol})$$

$$= (8-1)\,120 \times 10^3 + 80 \times 10^3 = 920 \times 10^3 \quad \text{Hz}$$

The maximum spectral efficiency is therefore given by:

$$\eta_s = \frac{\log_2 M}{T_o B}$$

$$= \frac{(\log_2 8) \times 80 \times 10^3}{920 \times 10^3} = 0.261 \quad \text{bit/s/Hz}$$

The nominal spectral efficiency given by equation (11.50) is:

$$\eta_s = \frac{\log_2 8}{(3/2)\,(8-1)+2} = 0.240 \quad \text{bit/s/Hz}$$

It can be seen that the difference between the nominal and maximum spectral efficiencies (and signal bandwidths) becomes small for MFSK signals as M (and/or tone separation) becomes large.

Multidimensional signalling can also be achieved using sets of orthogonally coded bit patterns, Figure 11.45 [developed from Sklar, p. 253]. (See also equations (15.17) to (15.20).) For an M-symbol alphabet each coded symbol (ideally) carries $H = \log_2 M$ bits of information but, to be mutually orthogonal, symbols must be M binary digits (or 'chips') long, Figure 11.45. The nominal bandwidth of such symbols is M/T_o Hz for baseband signalling and $2M/T_o$ Hz for bandpass signalling where T_o/M is the duration of the binary chips which make up the orthogonal symbols. For equiprobable symbols the nominal spectral efficiency for bandpass orthogonal code signalling is therefore:

$$\eta_s = \frac{\log_2 M}{2M} \quad \text{(bit/s/Hz)} \tag{11.52}$$

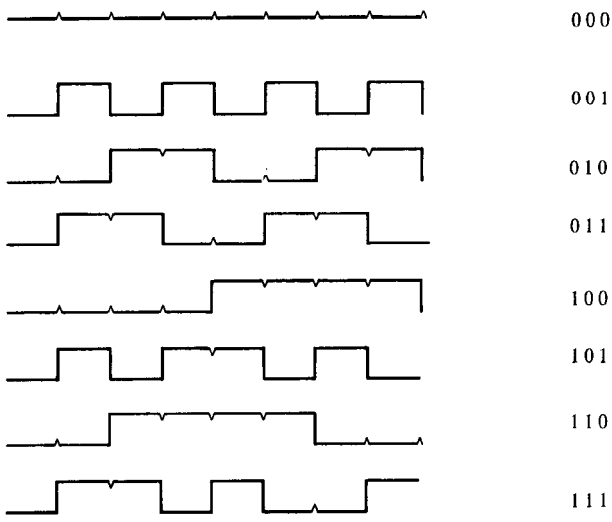


Figure 11.45 *Orthogonal code set comprising eight symbols.*

and the maximum spectral efficiency (with $T_o B = M$) is twice this. For optimum (coherent) detection of any equal energy, equiprobable, M-ary orthogonal symbol set (including MFSK), in the presence of white Gaussian noise, the probability of symbol error is traditionally bounded by the (union bound) formula:

$$P_e \leq \frac{1}{2}(M-1)\left[1 - \text{erf}\sqrt{\frac{E}{2N_0}}\right] \qquad (11.53(a))$$

A tighter bound, however, given by [Hughes] is:

$$P_e \leq 1 - \left\{\frac{1}{2}\left[1 + \text{erf}\sqrt{\frac{E}{2N_0}}\right]\right\}^{M-1} \qquad (11.53(b))$$

where E, as usual, is the symbol energy. The energy per information bit is $E_b = E/\log_2 M$. Figure 11.46 shows P_e against E_b/N_0. The parameter $H = \log_2 M$ is the number of information bits associated with each (equiprobable) symbol. Since the symbols are mutually orthogonal the distance between any pair is a constant. It follows that all types of symbol error are equiprobable, i.e.:

$$P(i_{RX}|j_{TX}) = P_e \quad \text{for all } i \neq j \qquad (11.54)$$

From Figure 11.45(b) it can be seen that for any particular information bit associated with any given symbol only $M/2$ symbol errors out of a total of $M - 1$ possible symbol errors will result in that bit being in error. The probability of bit error for orthogonal signalling

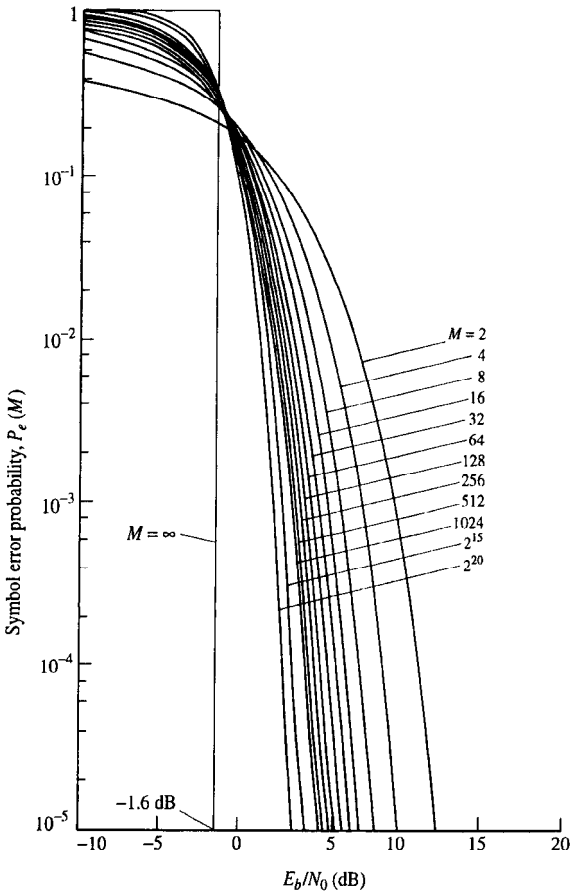


Figure 11.46 *Symbol error probability for coherently detected M-ary orthogonal signalling (source: Lindsey and Simon, 1973, reproduced with permission of Prentice-Hall).*

(including MFSK and orthogonal code signalling) is therefore:

$$P_b = \frac{(M/2)}{M-1} P_e \tag{11.55}$$

Figure 11.47 shows the probability of bit error against E_b/N_0 for orthogonal M-ary systems.

If antipodal symbols are added to an orthogonal symbol set the resulting set is said to be *biorthogonal*. The P_e (and therefore P_b) performance is improved since the average symbol separation in the constellation space is increased for a given average signal power. Figure 11.48 shows the biorthogonal P_b performance. The spectral efficiency of biorthogonal systems is improved by a factor of two over orthogonal systems since pairs of symbols now occupy the same spectral space. Biorthogonal signals must, of course, be

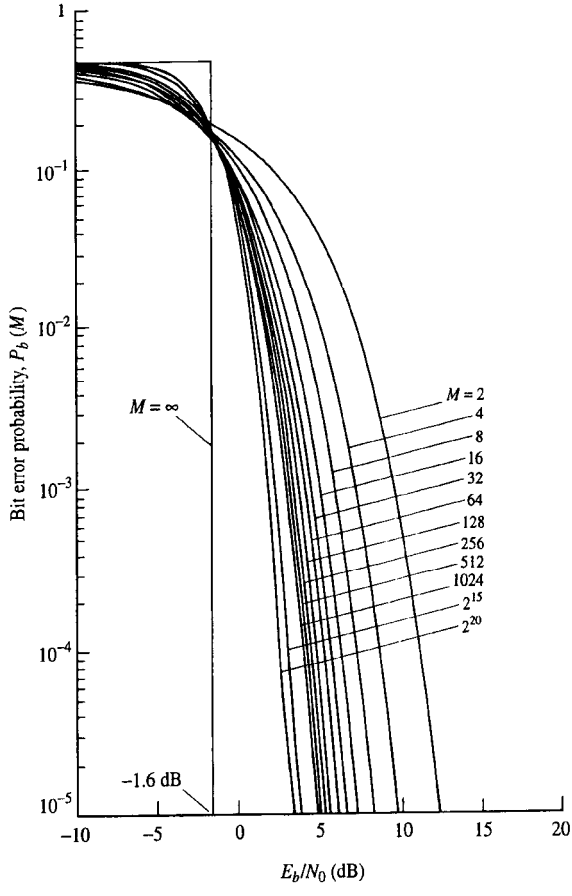


Figure 11.47 *Coherent detection of orthogonally coded transmission (source: Lindsey and Simon, 1973, reproduced with permission of Prentice-Hall).*

detected coherently.

Inspection of Figure 11.45 reveals that for an orthogonal code word set one binary digit (or chip) is identical for all code words. This binary digit therefore carries no information and need not be transmitted. An alphabet of code words formed by deleting the redundant chip from an orthogonal set is said to be *transorthogonal.* It has a correlation between symbol (i.e. code word) pairs given by:

$$\rho = -\frac{1}{M-1} \tag{11.56}$$

and therefore has slightly better P_e (and P_b) performance than orthogonal signalling. Since one binary digit from a total of M has been deleted it has a spectral efficiency which is $M/(M-1)$ times better than the parent orthogonal scheme.

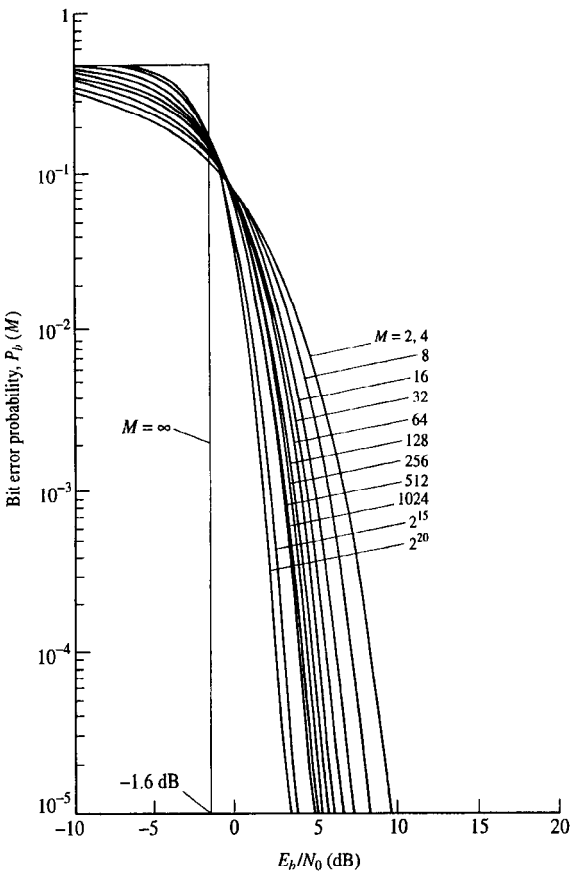


Figure 11.48 *Coherent detection of biorthogonally coded transmission (source: Lindsey and Simon, 1973, reproduced with permission of Prentice-Hall).*

The major application of MFSK type signalling systems is the telegraphy system used for teleprinter messages. This system, called Piccolo [Ralphs], has 32 tones and sends data at only 10 symbol/s with 100 ms duration symbols. (This was chosen to overcome the 0 to 8 ms multipath effects in over-the-horizon HF radio links.)

Another system is currently being developed for multicarrier broadcast to mobiles. It uses *simultaneous* MFSK or orthogonal FDM (OFDM) transmissions [Alard and Lassalle] to reduce a single high data rate signal into many parallel low data rate signals, again with lengthened symbol periods. By suitable FEC coding (Chapter 10) of the 'orthogonal carrier' channels, errors introduced by multipath effects can be tolerated in some of the channels without degrading significantly the overall system error rate. This system is now being developed separately for audio and TV broadcast applications.

11.5.2 Optimum constellation point packing

For a given alphabet of symbols distributed over a fixed number of orthogonal axes there are many possible configurations of constellation points. In general these alternative configurations will each result in a different transmitted power. Power efficient systems may minimise transmitted power for a given P_e performance by optimizing the packing of constellation points. This is well illustrated by the alternative constellation patterns possible for the (two dimensional) 16-APK system shown in Figure 11.19. These patterns are sometimes referred to as (8,8), (4,12) and (4,8,4) APK where each digit represents the number of constellation points on each amplitude ring.

The power efficiency of these constellation patterns increases moving from (8,8) to (4,8,4). The QAM rectangular lattice (represented in Figure 11.19(c) by (4,8,4) APK) is already quite power efficient and as such is a popular practical choice for M-ary signalling. A further saving of 0.6 dB is, however, possible by replacing the rectangular lattice pattern with a hexagonal lattice, Figure 11.49. This represents the densest possible packing of constellation points (in 2 dimensions) for fixed separation of adjacent points.

The power saved by optimising symbol packing increases as the number of constellation dimensions is increased. For a two dimensional constellation the saving gained by replacing a rectangular lattice with a hexagonal lattice is modest. The optimum lattice for a 64-dimensional constellation, however, gives a saving over a rectangular lattice of 8 dB [Forney *et al.*].

11.5.3 Optimum constellation point boundaries

Further modest savings in power can be made by optimising constellation boundaries. Consider, for example, Figure 11.50(a) which represents conventional 64-QAM signalling. If the points at the corners of the constellation are moved to the positions shown in Figure 11.50(b) then the voltage representing each of the moved symbols is reduced by a factor of (9.055/9.899) and their energy reduced by a factor of 0.84. Since the energy residing in each of the other symbols is unaffected the average transmitted power is reduced.

Not surprisingly the most efficient QAM constellation point boundary is that which most closely approximates a circle, and for a three-dimensional signalling scheme the most efficient boundary would be a sphere. It is clear that this idea extends to N-dimensional signalling where the most efficient boundary is an N-dimensional sphere.

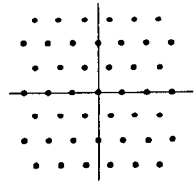
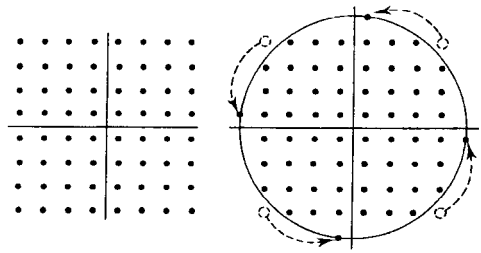


Figure 11.49 *Optimum 2-dimensional lattice pattern.*



(a) Square boundary QAM (b) Reduced power QAM

Figure 11.50 *64-QAM constellations.*

The power savings obtained in moving from N-dimensional cubic boundaries to N-dimensional spherical boundaries increase with N. For $N = 2$ (i.e. QAM) the saving is only 0.2 dB whilst for $N = 64$ the saving has increased to 1.3 dB [Forney *et al.*].

If efficient constellation point boundaries are used with efficient constellation point packing then the power savings from both are, of course, realised. Figure 11.51 shows such a (2-dimensional) 64-APK constellation.

11.6 Data modems

Many different techniques used in voiceband (3.2 kHz bandwidth) data modems are covered by the ITU-T V series of recommendations. The QAM constellation of Figure 11.52(a) is used in V.22 data modems. Low bit rate, telephone line based, data modems employ a wide variety of multi-amplitude, multi-phase schemes. The V.32 recommendation covers three different modulation options, 4800 bit/s and 9600 bit/s using conventional modulation techniques and 9600 bit/s TCM. At these low data rates the complexity of TCM is not a problem as current receiver processors have sufficient time to make all the required computations during the symbol period. A standard 9600 bit/s modem (V.29, Figure 11.52(b)) needs a four-wire connection to operate. Leased

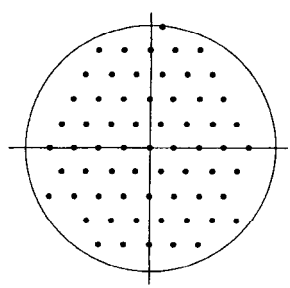


Figure 11.51 *Lattice and boundary efficient 64-APK.*

circuits are available with four-wire connections, but each public switched telephone network (PSTN) connection provides only two wires.

Two PSTN lines are therefore needed to back up a single leased circuit. V.32, modems however, provide major cost savings and other benefits. Firstly, instead of a four-wire leased circuit, only a two-wire circuit is required. In addition only a single PSTN connection per end is needed. Several manufacturers offer products with data rates from 9600 bit/s to 19.2 kbit/s for operation over dial-up networks. With V.33 modems operating at 14.4 kbit/s TCM with full two-way duplex, and the new V.fast standards one now has data rates of 24 kbit/s, and the possibility of 28.8 kbit/s over voiceband circuits. Given that the Shannon-Hartley law, equation (11.38(a)), predicts 32 kbit/s over a 3.2 kHz bandwidth voiceband circuit at 30 dB SNR, then these sophisticated modems are fast approaching the theoretical limits of performance.

11.7 Summary

IF or RF modulation is used principally to shift the spectrum of a digital information signal into a convenient frequency band. This may be to match the spectral band occupied by a signal to the passband of a transmission line, to allow frequency division multiplexing of signals, or to enable signals to be radiated by antennas of practical size.

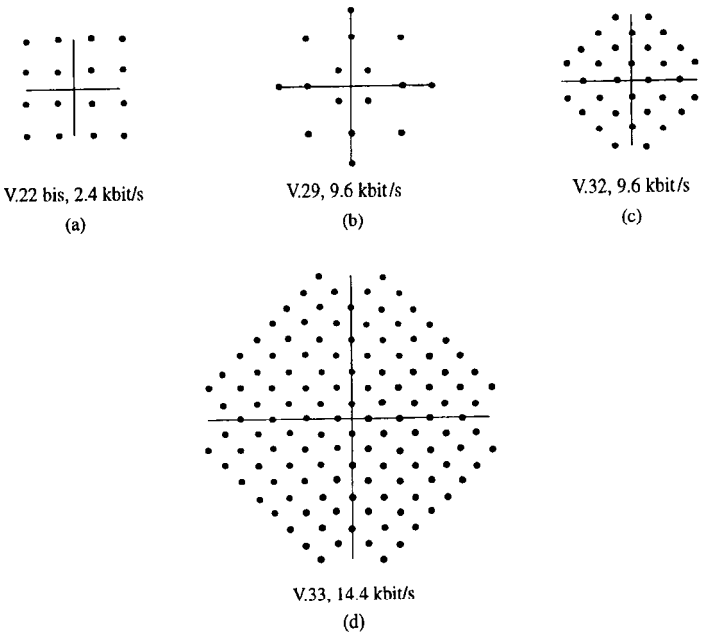


Figure 11.52 *Examples of signal constellations used in speech band data modems: (a) and (c) as used on switched lines; and (b) on leased lines.*

Two performance measures are commonly used to compare different IF modulation techniques – spectral efficiency and power efficiency. The former has units of bit/s/Hz and the latter relates to the required value of E_b/N_0 for a given probability of bit error.

There are three generic IF modulation techniques for digital data. These are ASK, PSK and FSK. Binary ASK and FSK are usually operated as orthogonal signalling schemes. Binary PSK is often operated as an antipodal scheme and therefore requires 3 dB less (average) power. ASK and FSK modulated signals can be detected incoherently which simplifies transmitter and receiver design. A small P_e penalty is incurred in this case over matched filter, or correlation, detection. PSK systems may suffer from phase ambiguity. This can be avoided, however, using either a known phase training sequence, transmitted in a preamble to the information data blocks, or differential coding.

MPSK, APK and QAM are spectrally efficient modulation schemes. QPSK and its variants are also reasonably spectrally efficient. Sophisticated 64-QAM and 256-QAM systems are now being applied in digital microwave long haul radio communications systems (see Chapter 14), but the power amplifier linearity requirements are severe for high level QAM systems. TCM is a technique which increases power efficiency without compromising spectral efficiency by combining modulation with error control coding.

MFSK is usually operated as an orthogonal modulation scheme and is therefore power efficient (as is orthogonal code signalling). Non-orthogonal M-ary schemes can be made more power efficient by distributing the symbols (constellation points) over an increased number of orthogonal axes. Further improvements in power efficiency can be made by choosing the optimum symbol packing arrangement for the number of dimensions used and adopting an (N-dimensional) spherical boundary for the set of constellation points.

Many of these techniques are widely applied in speech band data modems (described in section 11.6) and also in wideband high speed microwave communication systems (described in Chapter 14).

11.8 Problems

11.1. A rectangular pulse OOK signal has an average carrier power, at the input to an ideal correlation receiver, of 8.0 nW. The (one sided) noise power spectral density, measured at the same point, is 2.0×10^{-14} W/Hz. What maximum bit rate can this system support whilst maintaining a P_e of 10^{-6}? [17.7 kbit/s]

11.2. A BPSK, 1.0 Mbaud, communication system is to operate with a P_e of 8×10^{-7}. The signal amplitude at the input to the ideal correlation receiver is 150 mV and the one sided noise power spectral density at the same point is 15 pW/Hz. The impedance level at the correlation receiver input is 50 Ω. What minimum phase shift between phasor states must the system employ and what residual carrier power is therefore available for carrier recovery purposes? [122°, 5.23×10^{-5} W]

11.3. Define on-off keying (OOK), frequency shift keying (FSK), and phase shift keying (PSK) as used in binary signalling. Compare their respective advantages and disadvantages.

In a Datel 600 modem FSK tone frequencies of 1300 Hz and 1700 Hz are used for a signalling rate of 600 bit/s. Comment on the consistency between these tone frequencies, the $\rho - T_o$ diagram, and the signalling rate. What value would the upper tone frequency need to have, to handle a 900

bit/s signalling rate? [1900 Hz]

11.4. A binary, rectangular pulse, BFSK modulation system employs two signalling frequencies f_1 and f_2 and operates at the second zero of the $\rho - T_o$ diagram. The lower frequency f_1 is 1200 Hz and the signalling rate is 500 baud. (a) Calculate f_2. (b) Sketch the PSD of the FSK signal. (c) Calculate the channel bandwidth which would be required to transmit the FSK signal without 'significant' distortion. Indicate in terms of the PSD those frequencies which you assume should be passed to keep the distortion 'insignificant'. [1700 Hz, B = 1500 Hz]

11.5. If the CNR of the system described in problem 11.4 is 4 dB at the input to an ideal (matched filter) receiver and the channel has a low-pass rectangular frequency characteristic with a cut-off frequency of 3.6 kHz, estimate the probability of bit error. [1.06×10^{-5}]

11.6. (a) Define the correlation coefficient of two siganls: $s_1(t)$ and $s_2(t)$. (b) Sketch the correlation coefficient of two FSK rectangular RF signalling pulses, such as might be used in an FSK system, as a function of tone spacing. (c) Find the P_e of an orthogonal FSK binary signalling system using ideal matched filter detection if the signalling pulses at the matched filter input have rectangular envelopes of 10 μs duration and 1.0 V peak to peak amplitude. The one sided NPSD at the input to the filter is 5.5 nW/Hz and the impedance level at the filter input is 50 Ω. [1.653×10^{-2}]

11.7. If the spectrum of each signalling pulse described in problem 11.6(c) is assumed to have significant spectral components only within the points defined by its first zeros, what is the minimum bandwidth of the FSK signal? [0.25 MHz]

11.8. A receiver has a mean input power of 25 pW and is used to receive binary FSK data. The carrier frequencies used are 5 MHz and 5.015 MHz. The noise spectral density at the receiver has a value of 2.0×10^{-16} W/Hz. If the error rate is fixed at 2×10^{-4} find the maximum data rate possible, justifying all assumptions. [10 kbits/s] If the transmitter is switched to PRK what new bit rate can be accommodated for the same error rate? [20 kbit/s]

11.9. What is meant by orthogonal and antipodal signalling? Can the error performance of an FSK system ever be better than that given by the orthogonal case and, if so, how is the performance improved, and at what cost?

11.10. A DEPSK, rectangular pulse, RZ signal, with 50% duty cycle, has a separation of 165° between phasor states. The baud rate is 50 kHz and the peak received power (excluding noise) at the input of an ideal correlation receiver is 100 μW. The (one sided) noise power spectral density measured at the same point is 160 pW/Hz. What is the probability of symbol error? [4.56×10^{-4}]

11.11. Find the maximum spectral efficiency of ISI free 16-PSK. What is the probability of symbol error of this scheme for a received CNR of 24 dB if the maximum spectral efficiency requirement is retained? What is the Gray coded probability of bit error? [4 bit/s/Hz, 1.227×10^{-5}, 3.068×10^{-6}]

11.12. Sketch the constellation diagram of 64-QAM. What is the spectral efficiency of this scheme if pulse shaping is employed such that $BT_o = 2$? What is the best possible spectral efficency whilst maintaining zero ISI? Why is QAM (with the exception of QPSK) usually restricted to use in linear (or approximately linear) channel applications? [3 bit/s/Hz, 6 bit/s/Hz]

11.13. (a) Draw a block diagram of a QPSK modulator. Show clearly on this diagram or elsewhere the relationship in time between the binary signals in I and Q channels at the points immediately preceeding multiplication by the I and Q carriers. How does this relationship differ in OQPSK systems? What advantage do OQPSK systems have over QPSK systems? Sketch the phasor diagrams for both systems showing which signal state transitions are allowed in each.
(b) Calculate the probability of bit error for an ideal OQPSK system if the single sided NPSD and bit energy are 1.0 pW/Hz and 10.0 pJ respectively. [3.88×10^{-6}]

11.14. Justify the statement: 'The performance of BPSK and QPSK systems are identical'.

11.15. Two engineers are arguing about MSK modulation. One says it is a special case of BFSK and the other asserts it is a form of OQPSK. Resolve their argument. Is incoherent detection of MSK possible? (Explain your answer.) Draw a block diagram of an MSK transmitter and sketch a typical segment of its output.

11.16. A power limited digital communication system is to employ orthogonal MFSK modulation and ideal matched filter detection. P_b is not to exceed 10^{-5} and the available receiver carrier power is limited to 1.6×10^{-8} V^2. Find the minimum required size of the MFSK alphabet which will achieve a bit rate of 1.0 Mbit/s at the specified P_b if the two sided noise power spectral density at the receiver input is 10^{-15} V^2/Hz. (The use of a computer running an interactive maths package such as MATLAB, or a good programmable calculator, will take some of the tedium out of this problem.) [M=6]

11.17. A 3-APK modulation scheme contains the constellation points 0, $e^{j\pi/3}$, $e^{j2\pi/3}$. What is the power saving (in dB) if this scheme is replaced with a minimum power 3-PSK scheme having the same P_e performance? [3 dB]