Section 7: Digital Modulation Systems

Contents

7. Digital Modulation Systems

These are systems in which the transmitted signal consists of a sequence of ***symbols*** (pulses), each of which represents binary data 0 or 1, or some combination of 0's and 1's.

If the symbols are transmitted at a rate f_s symbols/sec, then the ***symbol period*** T is $1/f_s$. In many cases the symbol pulse is of length T or less, but as we shall see later, this is not absolutely necessary.

Binary Systems

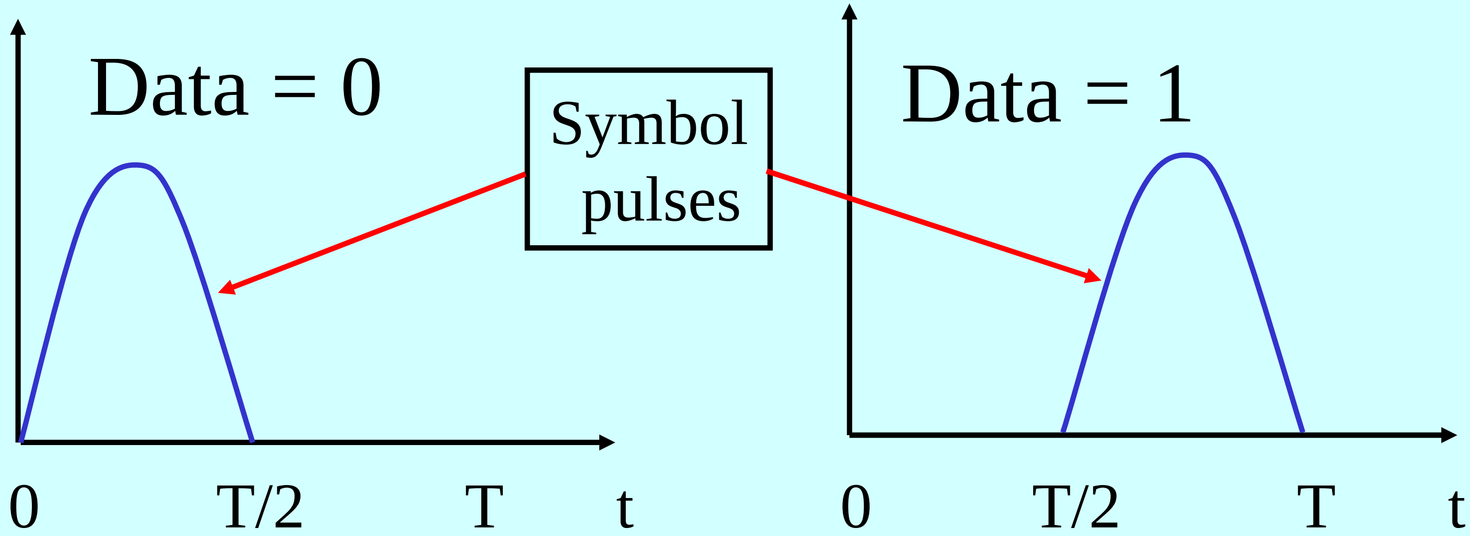
The simplest situation is where we only transmit two symbols, one representing a "0" and the other representing a "1". ***Bit rate*** (bits/sec) = Symbol rate.

M-ary Systems

In this case each symbol represents K bits, so we require $M = 2^K$ symbols. For example, with M = 4, (K = 2) the symbols would represent 00, 01, 10, 11 respectively. ***Bit rate*** (bits/sec) = K × Symbol rate.

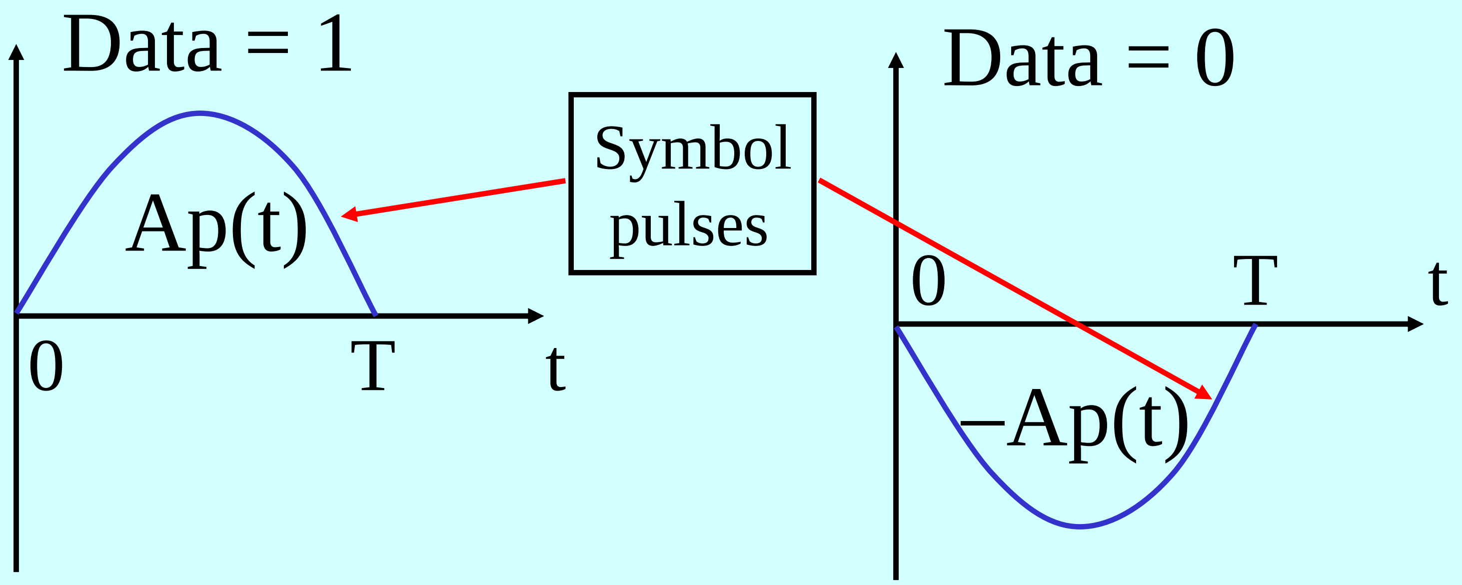
7.1 Pulse Position Modulation

In **baseband** pulse position modulation (PPM), the information is carried by the position of the pulse. This is not very bandwidth efficient, but is simple. Larger values of M require more bandwidth.



7.2 Pulse Amplitude Modulation

In **baseband** pulse amplitude modulation (PAM), the information is carried as the amplitude of the pulse. In the binary case we use ***antipodal*** signals, since this gives the best performance.

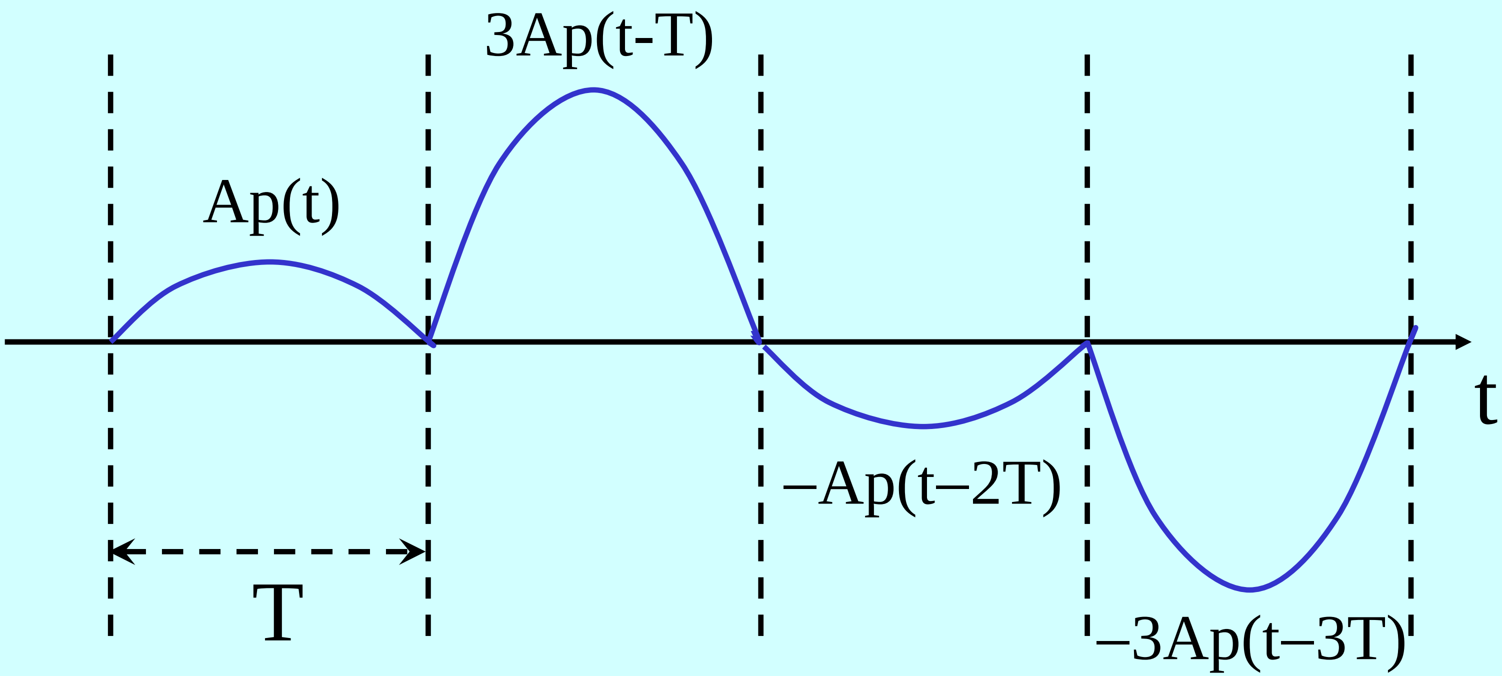


An M-ary **baseband** PAM signal is of the form:

$$s(t) = \sum_{k=-\infty}^{\infty} A\, a_k\, p(t - kT)$$

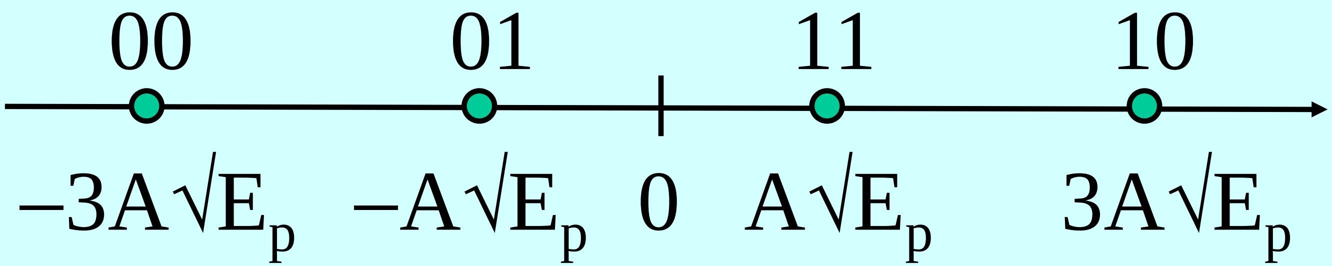
where a_k can take any of M possible values. These values are usually equally spaced and symmetrical about 0, since this gives the best performance in the presence of noise. With M = 4 we would use $a_k = \pm 1$ or ± 3. We note that the same pulse shape p(t) is used, all that is varied is the amplitude.

A typical 4PAM signal might look like this:



Data	11	10	01	00
	$a_0 = 1$	$a_1 = 3$	$a_2 = -1$	$a_3 = -3$

For each symbol $A\, a_k\, p(t-kT)$, if we plot the possible values of the square root of the symbol energy we obtain the ***signal constellation***. For the 4PAM signal on the previous page, these points would be as shown below.

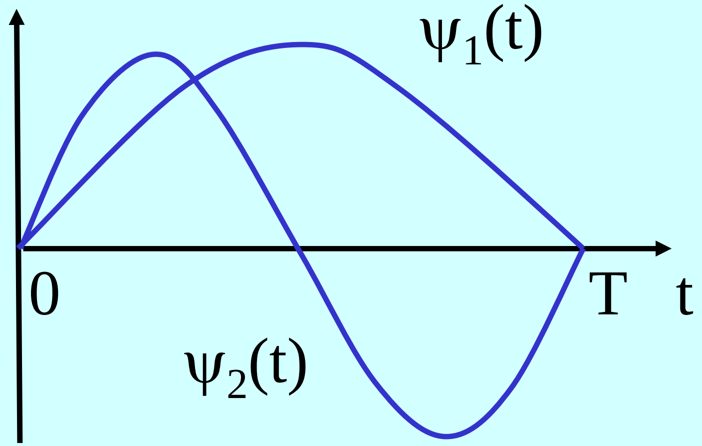


$$\text{where } E_p = \int_{-\infty}^{\infty} p^2(t)\,dt$$

7.3 Two Dimensional Signals

Consider two orthogonal signals $\psi_1(t)$ and $\psi_2(t)$ such as those shown below. Orthogonality means that:

$$\int_{-\infty}^{\infty} \psi_1(t)\,\psi_2(t)\,dt = 0$$



The pulses shown span the interval (0,T) but in general may span $(-\infty,\infty)$.

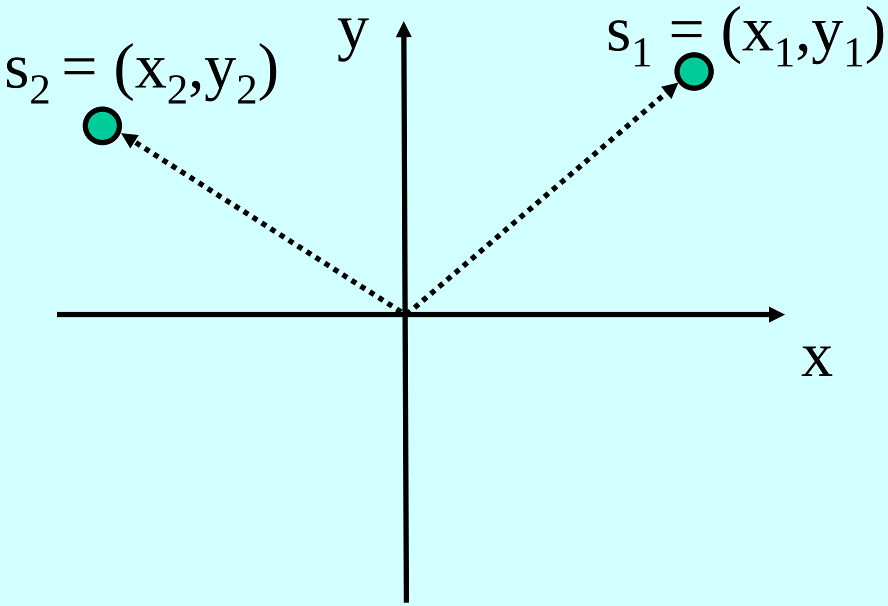
These are called ***basis functions***, and are chosen so that they have ***unit energy***.

$$\int_{-\infty}^{\infty} \psi_1^2(t)\,dt = \int_{-\infty}^{\infty} \psi_2^2(t)\,dt = 1$$

Suppose we transmit signals

$$s_1(t) = x_1\psi_1(t)+y_1\psi_2(t)$$
$$s_2(t) = x_2\psi_1(t)+y_2\psi_2(t)$$

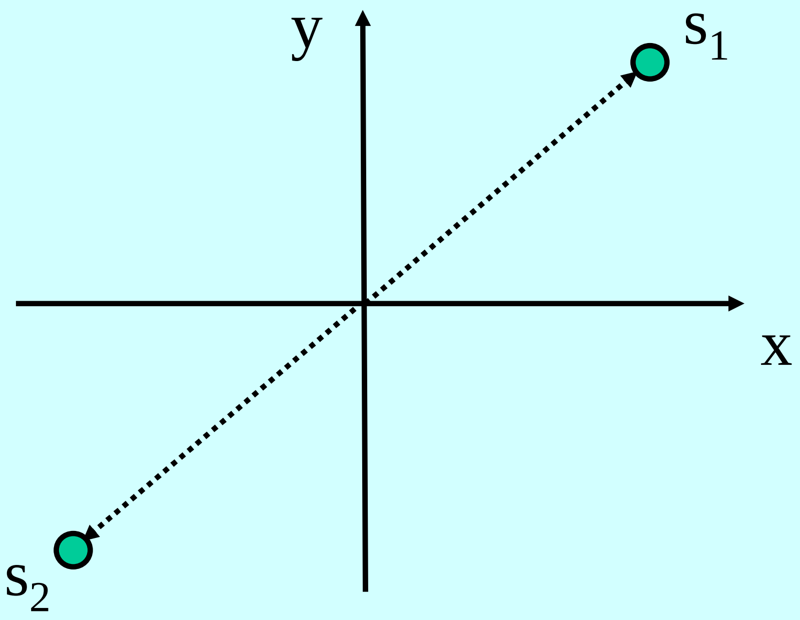
These can be represented geometrically.



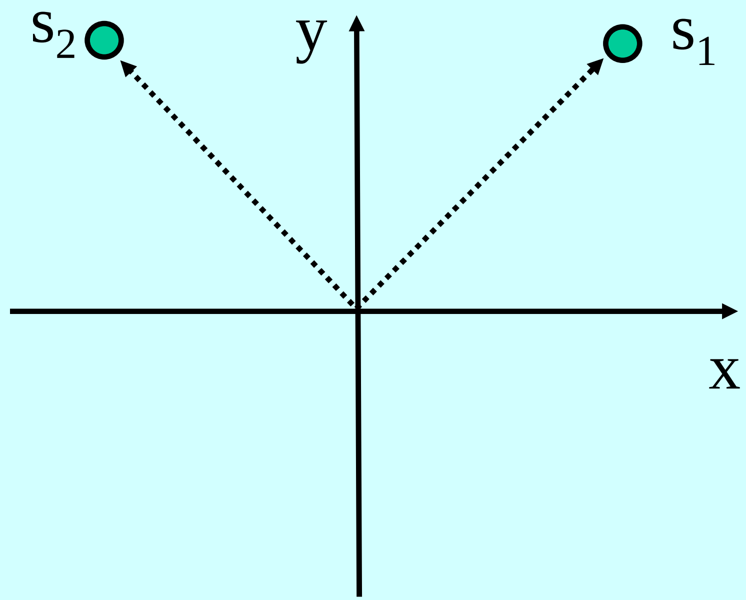
Each signal is called a ***symbol***, and the square of the distance from the origin represents the ***signal energy*** E.

$$E_1 = \int_0^T s_1^2(t)\,dt = x_1^2 + y_1^2 = r_1^2$$

If $x_2 = -x_1$ and $y_2 = -y_1$ then the signals $s_1(t)$ and $s_2(t)$ are ***antipodal***. If the angle between the vectors is 90^o, then they are ***orthogonal***.

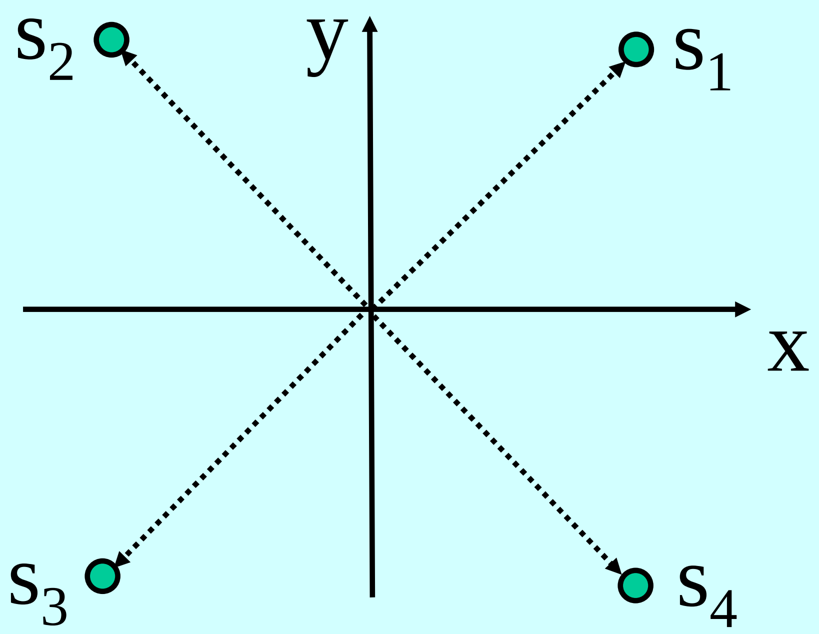


Antipodal



Orthogonal

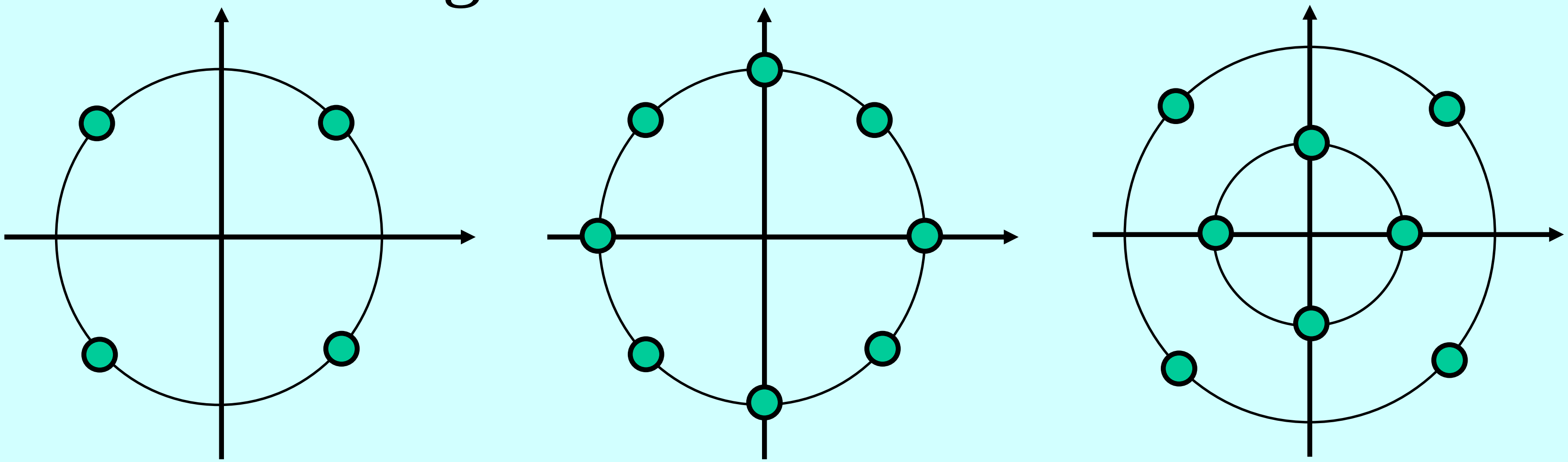
We can generate larger sets of signal waveforms by adding various weights of $\psi_1(t)$ and $\psi_2(t)$. For example the signal set below is called a ***bi-orthogonal*** set.



Bi-orthogonal

THE UNIVERSITY OF ADELAIDE AUSTRALIA

The diagram showing the points corresponding to the various symbols is called the ***signal constellation.*** Various possible constellations are shown below. Note that the symbols may have different energies.



M = 4 M = 8 M = 8

With a symbol set of size $M = 2^K$, we can transmit K bits of information in each symbol interval T. If E_s is the average energy per symbol, then the (average) transmitter power is $P_{av} = E_s/T$.

E_s = average energy per symbol.
E_b = average energy per bit = E_s/K.

The energy per bit E_b is a useful way of comparing modulation schemes of different sizes.

7.4 Carrier Modulation

We can form **carrier modulated** signals by choosing the basis functions (with unit energy) to be:

$$\psi_1(t) = C\, p(t) \cos(2\pi f_o t)$$

$$\psi_2(t) = C\, p(t) \sin(2\pi f_o t)$$

$$C = \sqrt{\frac{2}{E_p}}, \quad E_p = \int_{-\infty}^{\infty} p^2(t)\, dt$$

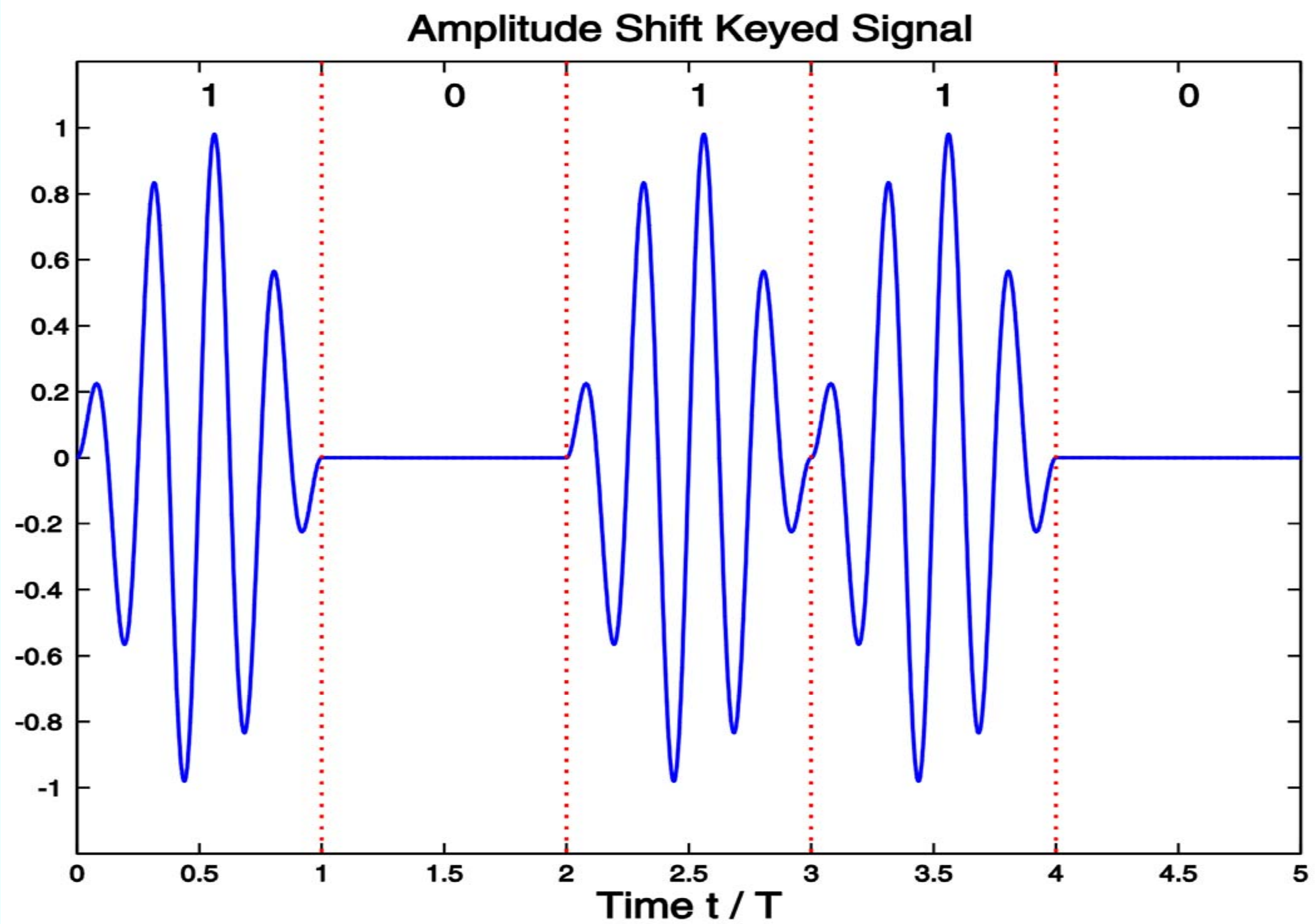
where p(t) is the pulse shape. The required bandwidth is usually **twice** that required for p(t).

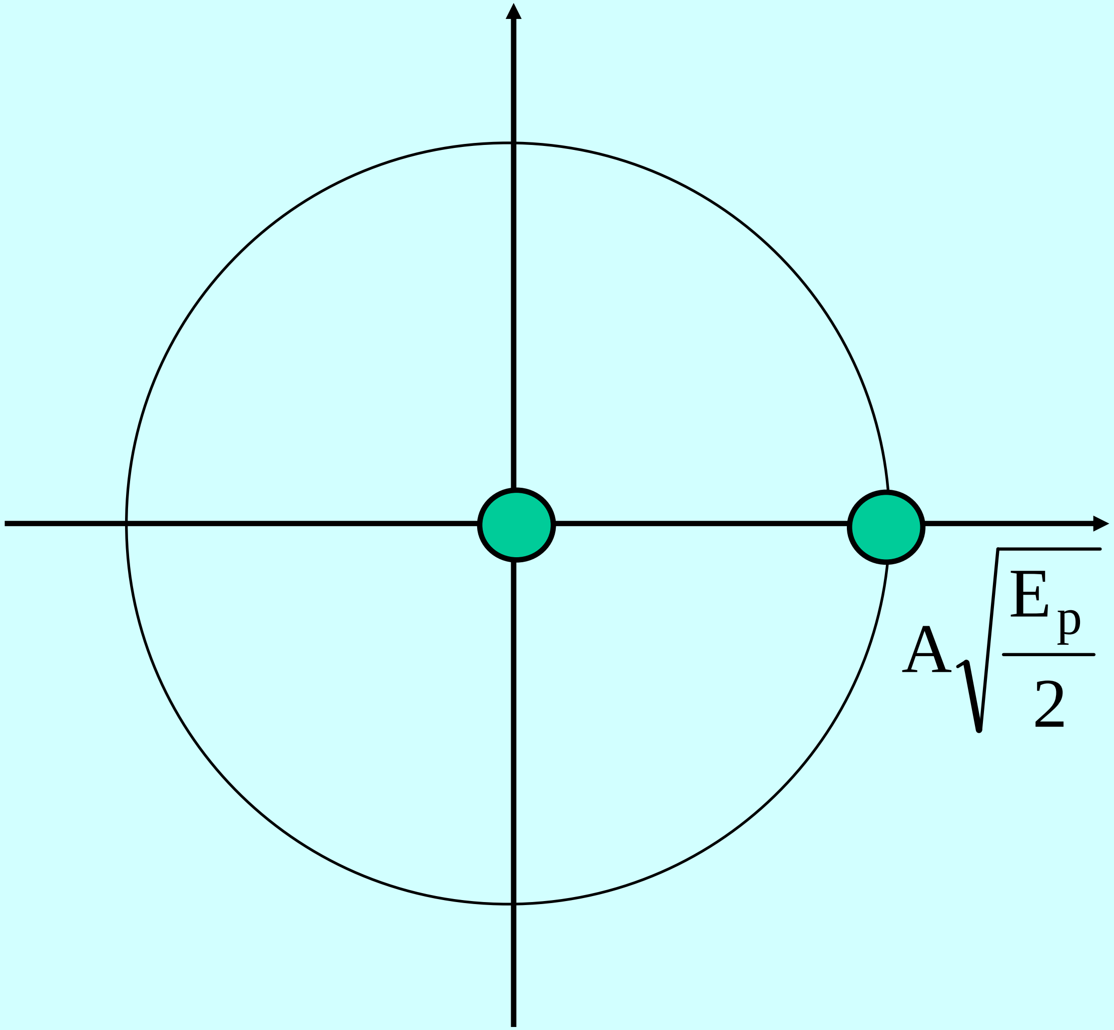
7.5 Amplitude Shift Keying

The simplest form of carrier modulation is ***amplitude shift keying*** (ASK) which is the digital equivalent of AM. (The energy of the "1" symbol is $0.5A^2E_p$, that of a "0" symbol is 0, so the average symbol energy $E_s = 0.25A^2E_p$).

$$s(t) = \sum_{k=-\infty}^{\infty} s_k(t)$$

$$s_k(t) = \begin{cases} A\,p(t-kT)\cos(2\pi f_o t) & \text{for a "1"} \\ 0 & \text{for a "0"} \end{cases}$$


Amplitude Shift Keyed Signal
1
0
1
1
0
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Time t / T

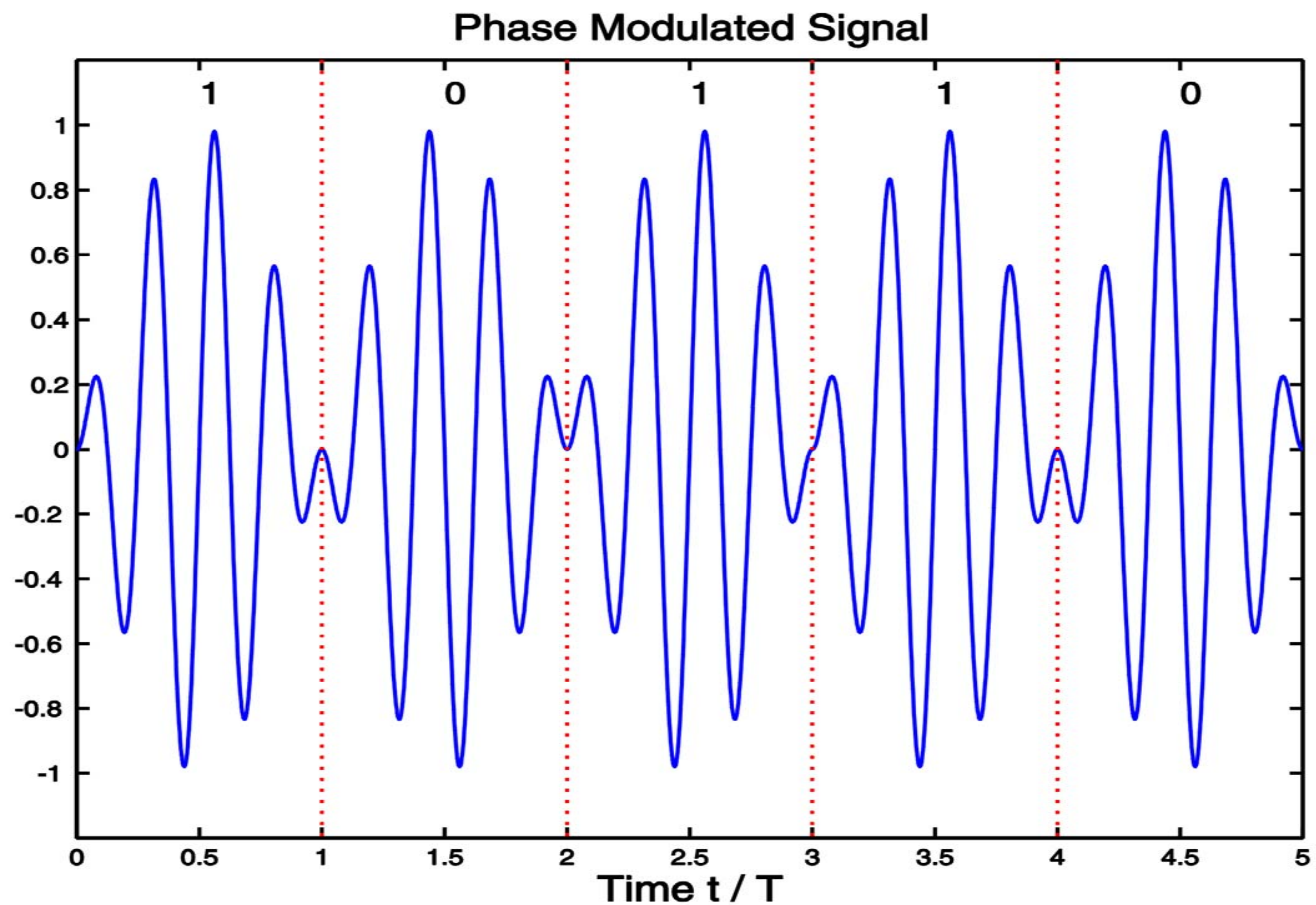

$A\sqrt{\frac{E_p}{2}}$

7.6 Phase Shift Keying

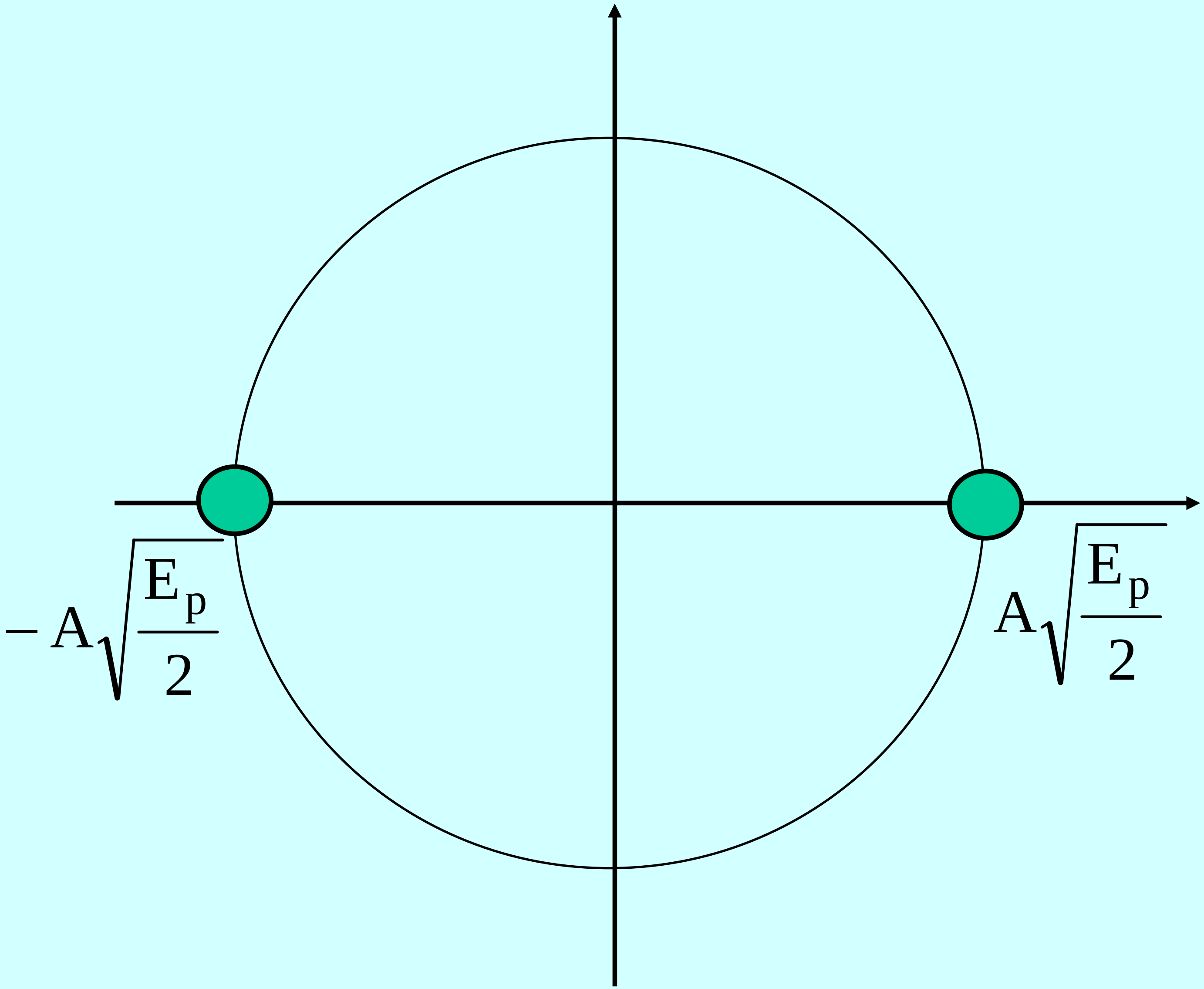
The simplest form of this is ***binary phase shift keying*** (BPSK) where the signal s(t) is:

$$s(t) = A \sum_{k=-\infty}^{\infty} a_k \, p(t - kT) \cos(2\pi f_o t)$$

where p(t) is the pulse shape, $a_k = \pm 1$ is the digital data and $E_s = 0.5A^2|a_k|^2E_p = 0.5A^2E_p$ is the energy per symbol. Note that the signal s(t) is a superposition of all the individual signals for each symbol interval.



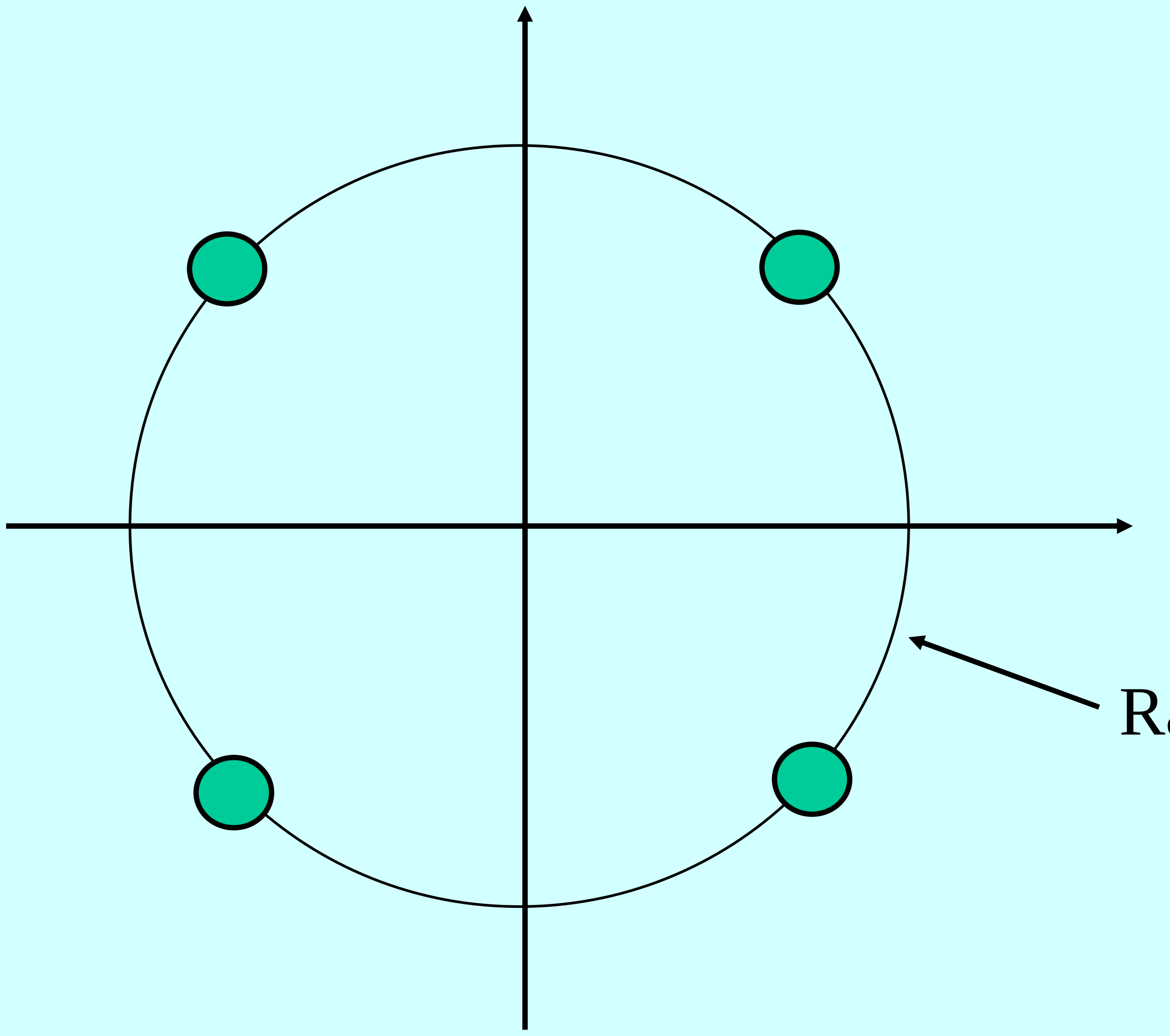
Phase Modulated Signal
1
0
1
1
0
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Time t / T


$-A\sqrt{\frac{E_p}{2}}$
$A\sqrt{\frac{E_p}{2}}$

The M = 4 version is called ***quaternary phase shift keying*** (QPSK) which can transmit two bits of information per symbol.

$$s(t) = \sum_{k=-\infty}^{\infty} A\, p(t - kT) \cos(2\pi f_o t + \theta_k)$$

Where θ_k takes values $\pm 45^o$ and $\pm 135^o$. In this case $E_s = 0.5A^2E_p$ and $E_b = E_s/2$. This form of modulation is very popular in satellite systems and HF communication systems.



$$\text{Radius} = A\sqrt{\frac{E_p}{2}}$$

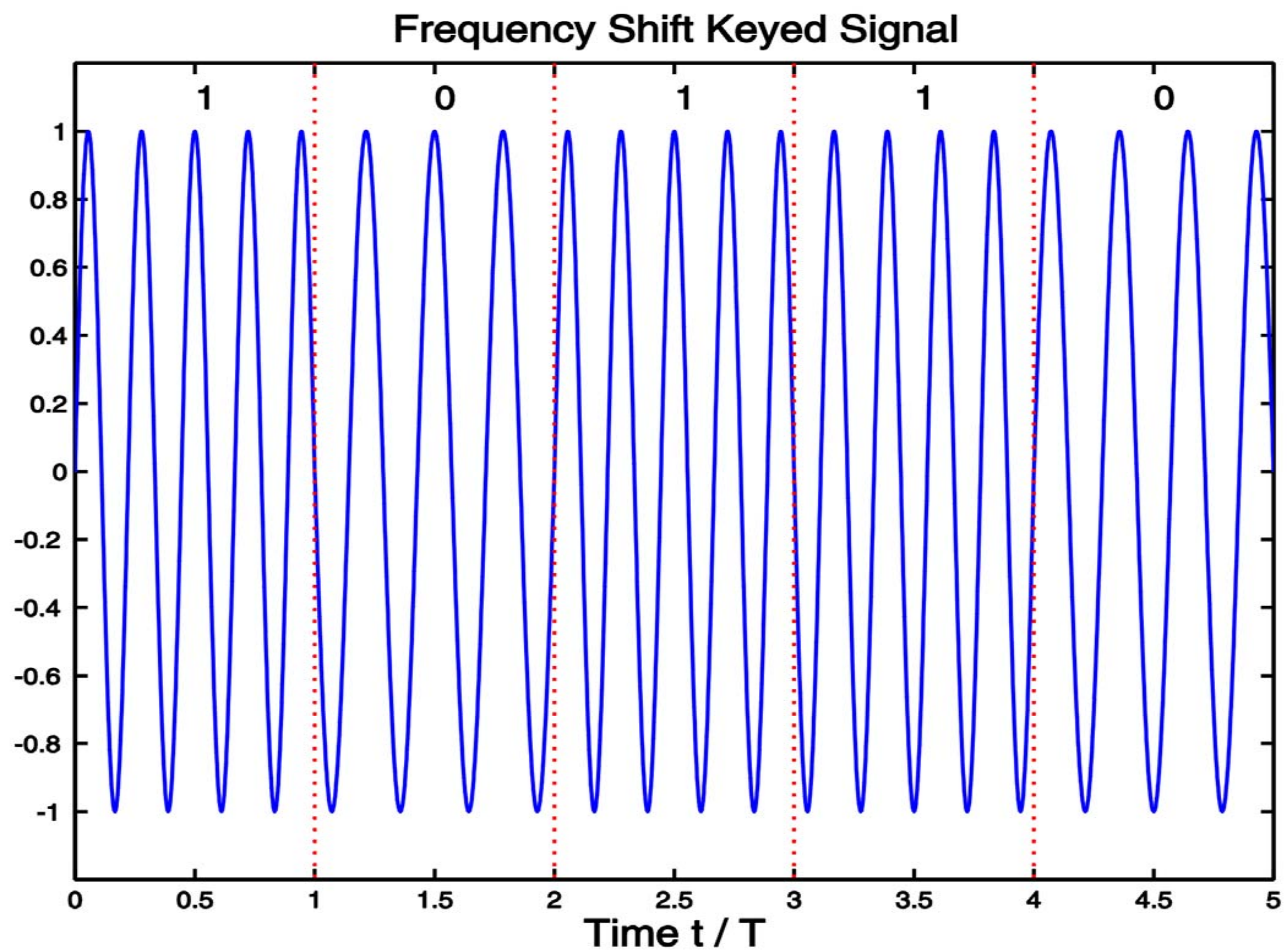
7.7 Frequency Shift Keying

In ***binary frequency shift keying*** (BFSK) we have:

$$s(t) = \sum_{k=-\infty}^{\infty} s_k(t)$$

$$s_k(t) = \begin{cases} A\cos(2\pi f_1 t + \theta_{1k}) & \text{for a "1"} \\ A\cos(2\pi f_2 t + \theta_{2k}) & \text{for a "0"} \end{cases}$$

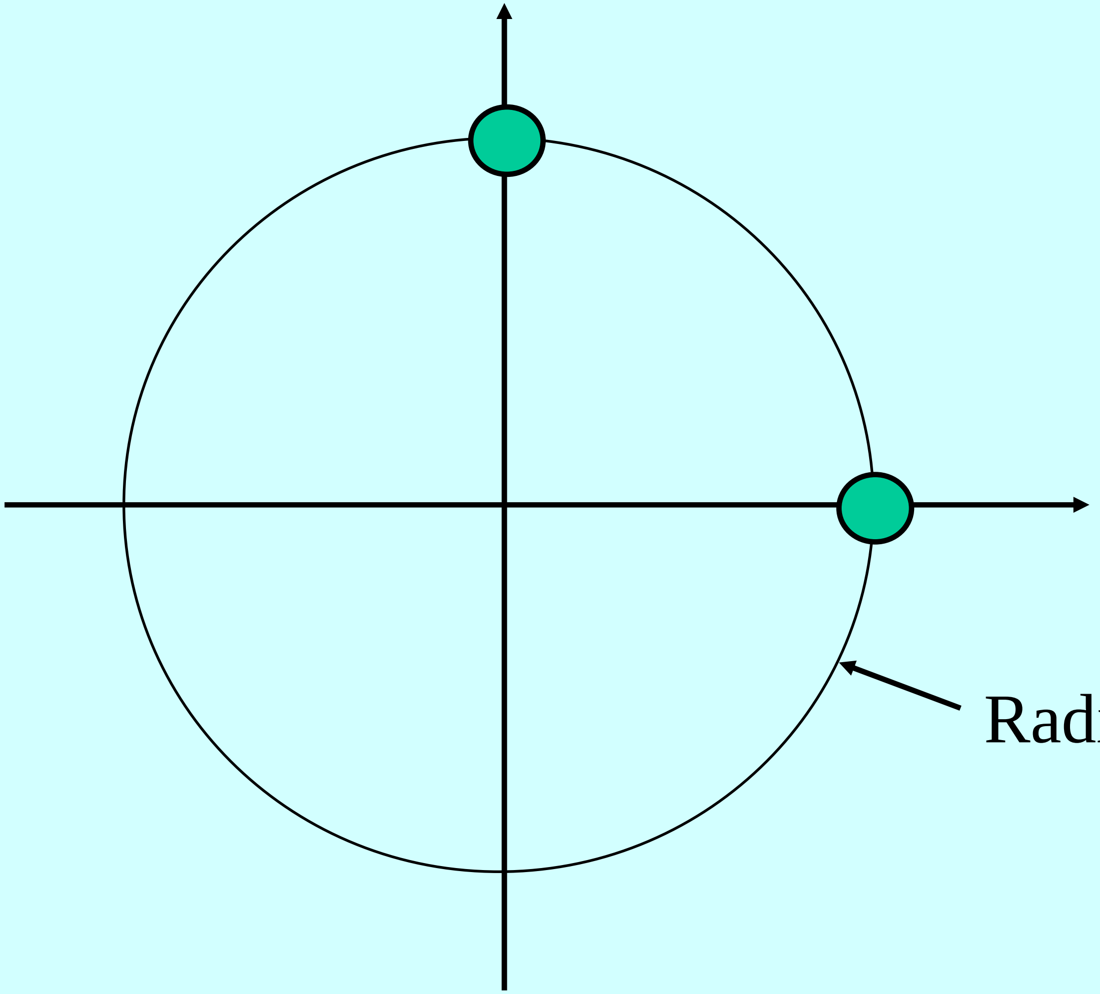
The phases of the two sinewaves are usually adjusted so that the signal s(t) is continuous when switching from a “0” to “1” and vice versa.


Frequency Shift Keyed Signal
1
0
1
1
0
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Time t / T

FSK has the advantage that simple detectors are possible, but because FSK does not use antipodal signals, it requires 3 dB more power than BPSK for the same error rate.

The ***modulation index*** **$h = (f_1 – f_2)T$** is usually about unity, but special forms with h = 1/2 exist.

Demodulation for h = 1 is achieved by filtering the received signal with filters tuned to the frequencies f_1 and f_2 and selecting whichever output is greater, or alternatively a correlation detector can be used.



$$\text{Radius} = \sqrt{\frac{A^2 T}{2}}$$

We can also express the FSK signal as:

$$s(t) = A\cos[2\pi f_o t + \theta(t)]$$

$$\frac{d\theta(t)}{dt} = \frac{\pi h a_k}{T} \quad ; kT \le t < (k+1)T$$

$$f_o = \frac{f_1 + f_2}{2}$$

where $a_k = \pm 1$ is the digital data and the instantaneous frequency of s(t) is either f_1 or f_2.

FSK is digital frequency modulation, since the information is carried in the frequency of the signal.

In some FSK systems the frequency change is smoothed by using a pulse shape p(t) with unit area

$$\frac{d\theta(t)}{dt} = \sum_{k=-\infty}^{\infty} \pi h a_k p(t-kT)$$

With binary data the phase changes by πh in each symbol interval.

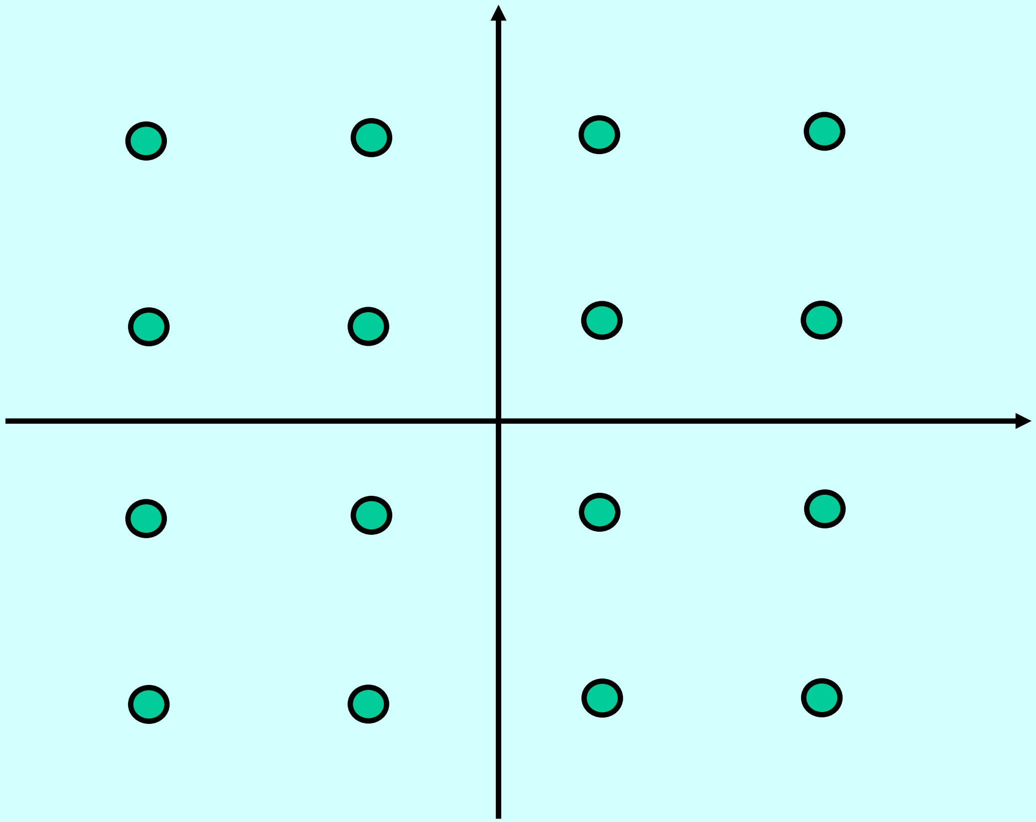
7.8 Quadrature Amplitude Modulation

The two simplest versions of this are equivalent to BPSK and QPSK. For larger symbol sets we have:

$$s(t) = \sum_{k=-\infty}^{\infty} A\,p(t-kT)\left[a_k \cos(2\pi f_o t) - b_k \sin(2\pi f_o t)\right]$$

where a_k and b_k are multilevel signals. The symbol energy is $E_k = 0.5A^2E_p(a_k{}^2+b_k{}^2)$ so E_s is the average over all possible values of a_k and b_k. For instance, with 16QAM, we would have a_k and b_k chosen from the values ±1 and ±3.

The constellation would appear as:



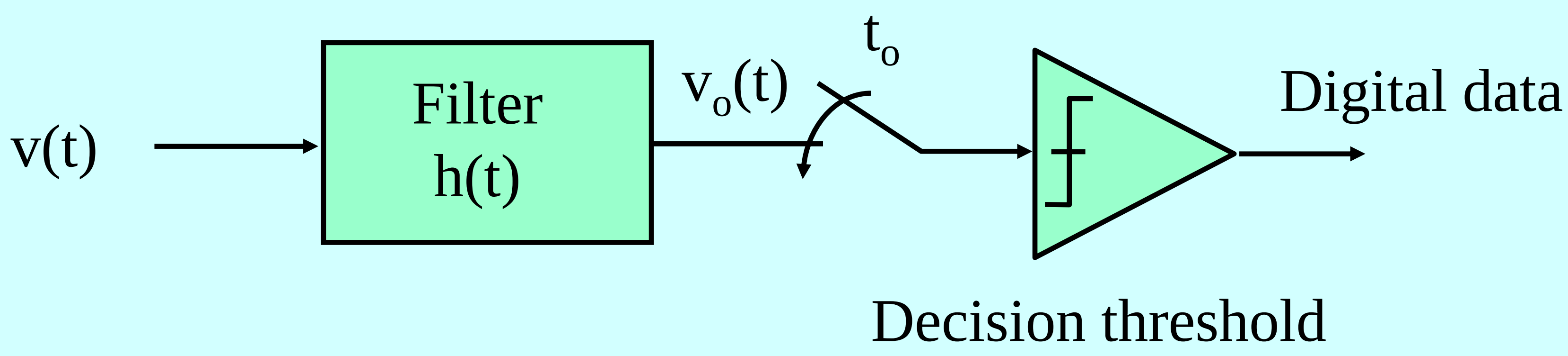
7.9 The Matched Filter

We will derive the optimum receiver for a baseband binary PAM signal and extend the result to other situations. We will assume we have a symbol which is received accompanied by additive white Gaussian noise (AWGN).

$$v(t) = \pm s(t) + n(t)$$

where s(t) is the received signal of energy E_s and n(t) is white Gaussian noise of power spectral density $S_{nn}(f) = \alpha = N_o/2$.

The receiver will be assumed to consist of a filter (to reduce the noise) and whose output is sampled at some time t_o. On the basis of this sample we we decide whether the symbol was +1 or –1.



The decision threshold will be zero volts.

The signal to noise ratio at the sampling instant is:

$$\Gamma = \frac{s_o^2(t_o)}{\left\langle n_o^2(t) \right\rangle}$$

$$s_o(t_o) = \int_{-\infty}^{\infty} h(\lambda)\, s(t_o - \lambda)\, d\lambda$$

$$\left\langle n_o^2(t) \right\rangle = \alpha \int_{-\infty}^{\infty} \left| H(f) \right|^2 df = \alpha \int_{-\infty}^{\infty} h^2(\lambda)\, d\lambda$$

To maximise Γ we will maximise $s_o(t_o)$ while keeping $\langle n_o^2(t)\rangle$ constant. For a perturbation $\delta h(t)$:

$$\delta s_o(t_o) = \int_{-\infty}^{\infty} \delta h(\lambda) s(t_o - \lambda) \, d\lambda = 0 \quad \text{(for a maximum)}$$

$$\delta \left\langle n_o^2(t) \right\rangle = 2\alpha \int_{-\infty}^{\infty} \delta h(\lambda) h(\lambda) \, d\lambda = 0 \quad \text{(constant output noise)}$$

Since $\delta h(t)$ is arbitrary except that the second relation must be zero, we must have $h(t) = c\, s(t_o - t)$.

THE UNIVERSITY OF ADELAIDE AUSTRALIA

$$h(t) = c\,s(t_o - t)$$

$$s_o(t_o) = c\int_{-\infty}^{\infty} s^2(t_o - t)\,dt = c\int_{-\infty}^{\infty} s^2(t)\,dt = c\,E_s$$

$$\left\langle n_o^2(t)\right\rangle = \alpha c^2 \int_{-\infty}^{\infty} s^2(t_o - t)\,dt = \alpha\, c^2 E_s$$

$$\Gamma_{max} = \frac{E_s}{\alpha} = \frac{E_s}{\frac{1}{2}N_o}$$

In the frequency domain we have

$$H(f) = cS^*(f)e^{-j2\pi f t_o}$$

where S(f) is the Fourier transform of s(t). This is called a ***matched filter***. It is equivalent to a ***correlator***, where the received signal is multiplied by a replica of the known signal s(t) and integrated.

$$v_o(t_o) = \int_{-\infty}^{\infty} v(\lambda)h(t_o - \lambda)d\lambda = \int_{-\infty}^{\infty} v(\lambda)s(\lambda)d\lambda$$

7.10 Receiver for Carrier Systems

For carrier systems, the received signal is synchronously demodulated using $\cos(2\pi f_o t)$ and $\sin(2\pi f_o t)$ to obtain baseband signals which can then be applied to baseband receivers of the type discussed in the previous section.

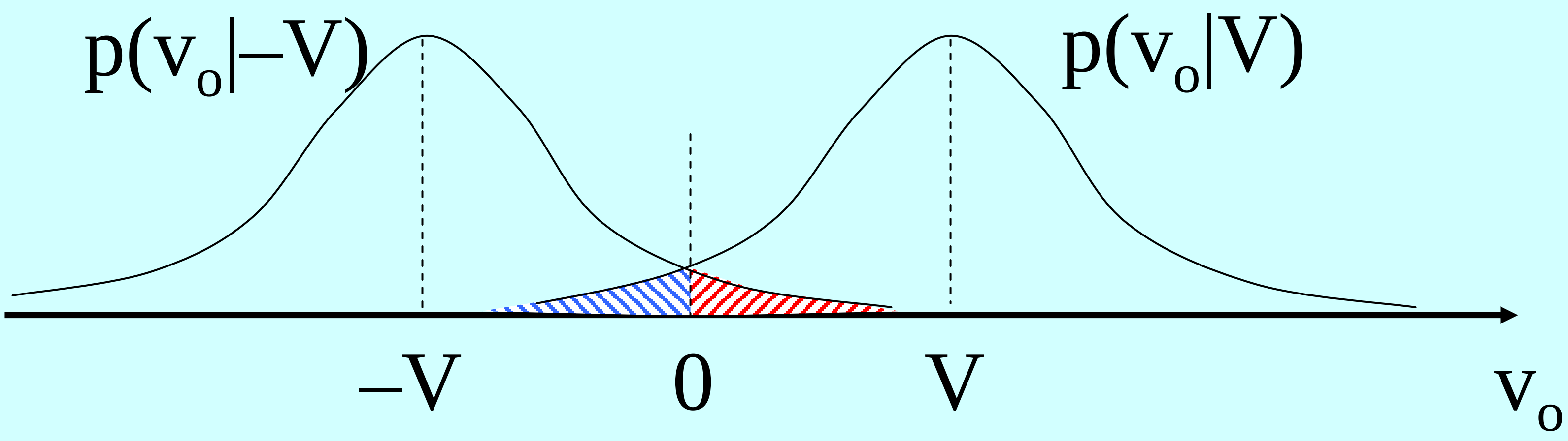
In some cases, such as FSK for example, simple non-optimal receivers can be used. Alternatively, the matched filter can be realised by correlation applied directly to the RF signal.

7.11 Probability of Error

We will derive the result for a ***baseband binary PAM signal*** and extend the result to other situations. We will assume we have a symbol which is received accompanied by additive white Gaussian noise (AWGN).

We will assume the sample value is $v_o = s_o(t_o) + n_o(t_o) = \pm V + n_o$. The noise sample n_o is of variance $\sigma^2 = \langle n_o^2(t) \rangle$.

The conditional probability density functions of v_o are called a ***likelihood functions***. They are simply the pdf of the noise n_o, shifted by the signal component ±V.



An error occurs if $n_o > V$ when $s_o = -V$, or if $n_o < -V$ when $s_o = V$. The probabilities of these are equal, so the probability of a bit error is:

$$P_b = \int_V^\infty p(n_o)\,dn_o = \frac{1}{\sigma\sqrt{2\pi}}\int_V^\infty e^{-n_o^2/2\sigma^2}\,dn_o$$

$$= \frac{1}{\sqrt{2\pi}}\int_{V/\sigma}^\infty e^{-t^2/2}\,dt = Q\left(\frac{V}{\sigma}\right)$$

$$P_b = Q\left(\sqrt{\Gamma}\right)$$ (always true)

If the filter is a matched filter, then:

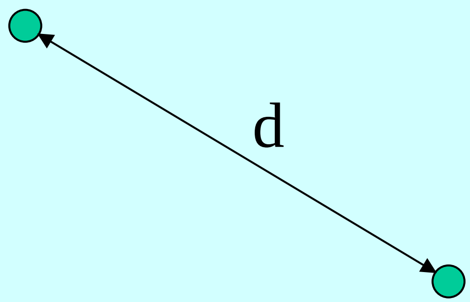
$$P_b = Q\left(\sqrt{\frac{E_s}{\alpha}}\right) = Q\left(\sqrt{\frac{E_s}{\frac{1}{2}N_o}}\right)$$

Note that this is ***only true for a matched filter*** receiver.

7.12 Probability of Error (General Constellation)

The square of the distance of a constellation point from the origin is the ***symbol energy***. If two symbols in a constellation are spaced by a distance "d" on the diagram, then with a ***matched filter*** receiver and a noise spectral density $S_{nn}(f) = N_o/2$, the probability of selecting the wrong symbol is:

$$P_{sym} \approx Q\left\{\sqrt{\frac{d^2}{2N_o}}\right\}$$

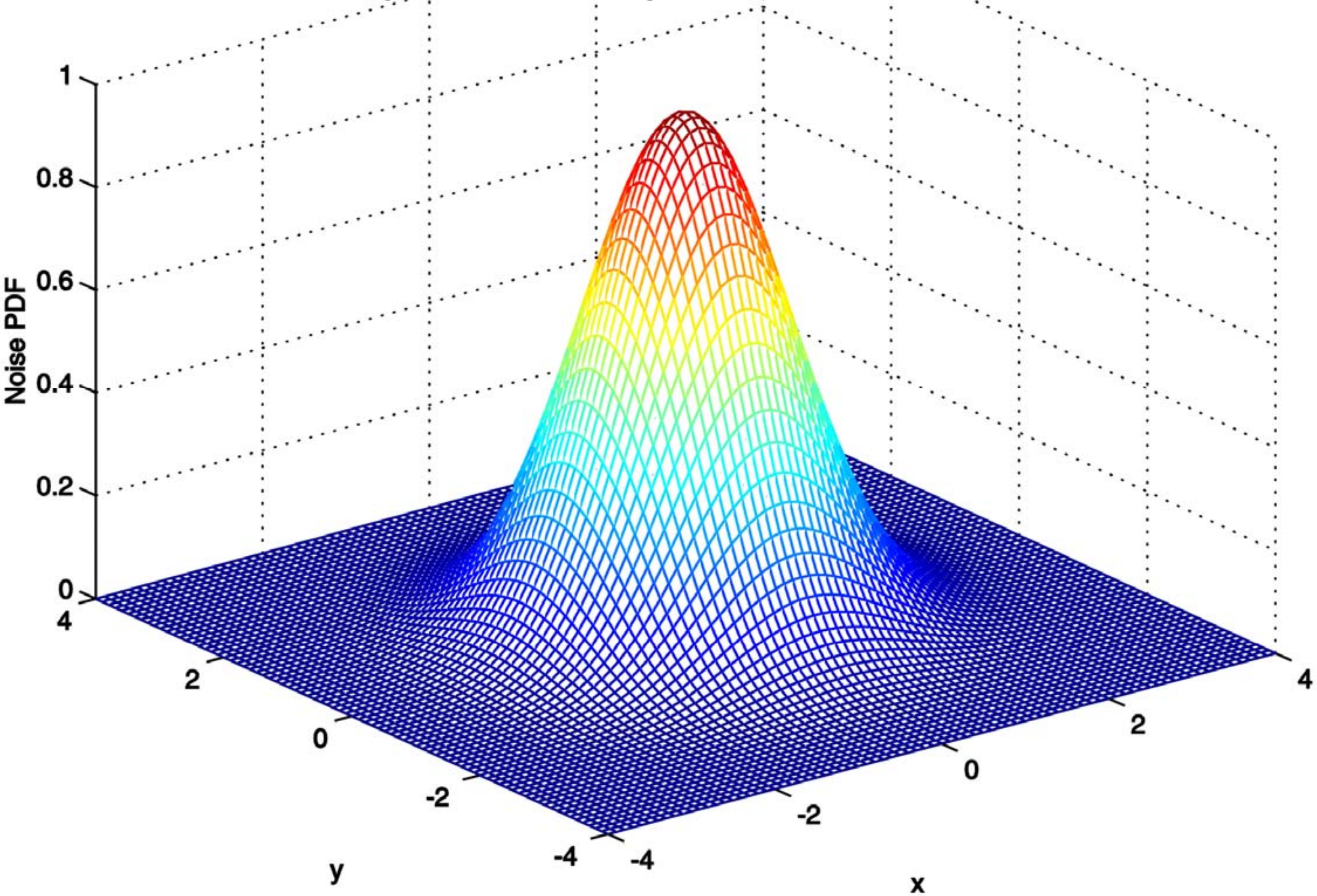


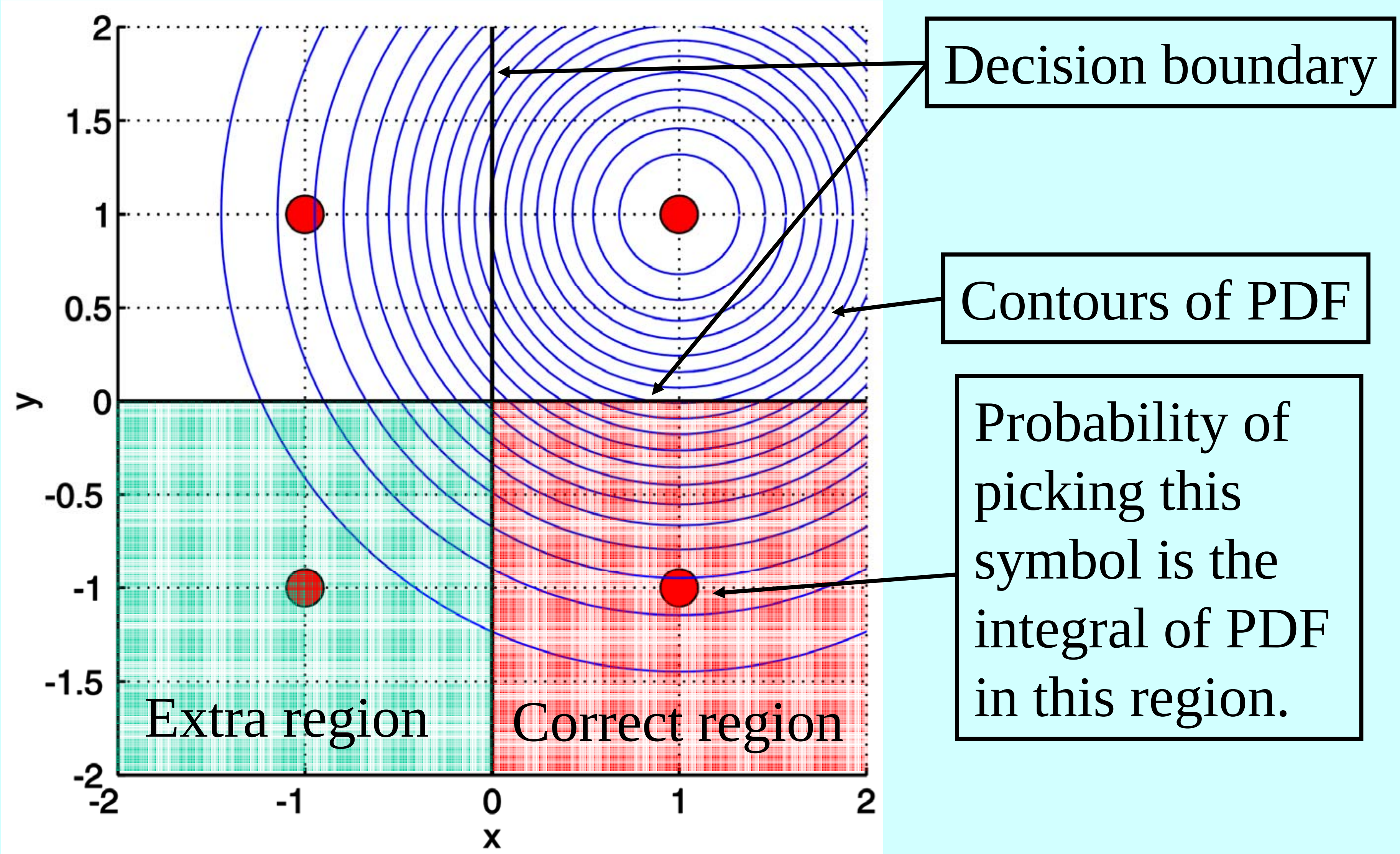
This formula is approximate as it calculates the probability of a received symbol being in the half plane containing the other symbol, and not the actual ***decision region*** for the other symbol.

The ***decision regions*** are defined by ***decision boundaries*** which are equidistant from each pair of symbols.

Hence the result will be greater than the true value, but in most cases the difference is negligible.

Probability Density Function of Noise





2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
y
-2
-1
0
1
2
x
Extra region
Correct region
Decision boundary
Contours of PDF
Probability of picking this symbol is the integral of PDF in this region.

When we compare digital modulation systems, we do it on the basis of ***energy per bit* (E_b)**, and not ***energy per symbol* (E_s)**. For a system which transmits K bits per symbol (which requires $M = 2^K$ symbols) we have:

$$E_b = \frac{E_s}{K} = \frac{P_s(av)T}{K}$$

Where $P_s(av)$ is the average power of each symbol, and T is the symbol interval.

To calculate the energy per symbol E_s or the energy per bit E_b for an arbitrary constellation of $M = 2^K$ symbols represented by phasors $Z_{k.}$:

$$E_s = \mathrm{av}\left\{ |Z_k|^2 \right\} \quad \text{; average energy per symbol}$$

$$E_b = \frac{E_s}{K} \quad \text{; average energy per bit}$$

$$P_s = E_s / T \quad \text{; average power}$$

Example: For BPSK, $E_b = E_s$, and $d = 2\sqrt{E_s}$:

$$P_b = P_{sym} = Q\left\{\sqrt{\frac{4E_b}{2N_o}}\right\} = Q\left\{\sqrt{\frac{E_b}{\frac{1}{2}N_o}}\right\}$$

Example: For QPSK, $E_s = 2E_b$ and $d = \sqrt{(2E_s)}$ and hence:

$$P_{sym} = Q\left\{\sqrt{\frac{2E_s}{2N_o}}\right\} = Q\left\{\sqrt{\frac{E_s}{N_o}}\right\} = Q\left\{\sqrt{\frac{E_b}{\frac{1}{2}N_o}}\right\}$$

If there are a number of symbols near a particular symbol, then the ***symbol error probability*** is the sum of the probabilities of picking each of the wrong ones. We can ignore symbols far away.

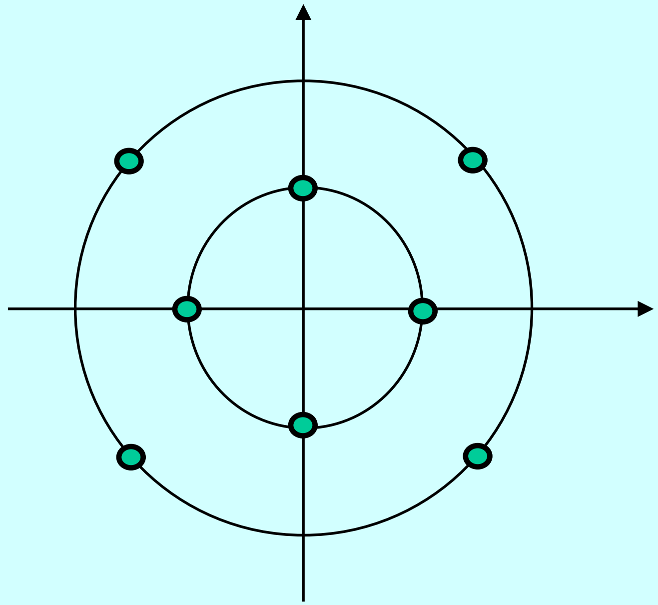
For instance, with QPSK the probability of picking an adjacent symbol is actually $2 * Q\{\sqrt{(E_s/N_o)}\}$ because there are two adjacent symbols. However, this will usually cause only one of the two bits to be in error, so the ***bit error probability*** is $Q\{\sqrt{(E_s/N_o)}\} = Q\{\sqrt{(2E_b/N_o)}\}$.

Exercise: In M-ary PSK, ($M=2^K$), show that the probability of choosing an adjacent symbol (the most likely error event) is given by:

$$P_{sym} = Q\left(\sin\left(\frac{\pi}{M}\right)\sqrt{\frac{E_s}{\frac{1}{2}N_o}}\right) = Q\left(\sin\left(\frac{\pi}{M}\right)\sqrt{\frac{KE_b}{\frac{1}{2}N_o}}\right)$$

Hence determine how much transmitter power is required for 8-PSK compared with 4-PSK for the same probability of error for the same a) symbol rate, b) bit rate. **Ans:** 5.3dB, 3.6dB more.

Exercise: For the 8QAM constellation shown calculate the radii such that the nearest symbol is distance d in all cases. Hence find the probability of a symbol error if E_b/N_o = 13 dB, assuming a matched filter receiver.



7.13 Non-Matched Filters

When the receiver filter is matched to p(t), we have the useful result that in a binary PAM or BPSK system, the signal to noise ratio Γ is:

$$\Gamma_{\text{matched}} = \frac{E_b}{\frac{1}{2}N_o}$$

If the receiver filter is not a matched filter, the analysis becomes more complicated in that the signal and noise components at the receiver output must be calculated individually.

7.14 Carrier and Clock Recovery

In digital communication systems neither the ***carrier*** nor the symbol ***clock*** is transmitted, and these must be recovered from the received signal.

The carrier is required for demodulation of the RF signal (except for FSK) and the clock is required to time the digital data signal. The recovery circuits usually involve some sort of non-linear operation on the received signal, followed by a phase locked loop to extract the required signal.

Exercises: You are expected to attempt the following exercises in Proakis & Salehi. Completion of these exercises is part of the course. Solutions will be available later.

7.10
7.15
7.17
7.34
7.43