Lecture Notes 6: Basic Modulation Schemes

In this lecture we examine a number of different simple modulation schemes. We examine the implementation of the optimum receiver, the error probability and the bandwidth occupancy. We would like the simplest possible receiver, with the lowest error probability and smallest bandwidth for a given data rate.

Binary Phase Shift Keying (BPSK)

The first modulation considered is binary phase shift keying. In this scheme during every bit duration, denoted by T, one of two phases of the carrier is transmitted. These two phases are 180 degrees apart. This makes these two waveforms antipodal. Any binary modulation where the two signals are antipodal gives the minimum error probability (for fixed energy) over any other set of binary signals. The error probability can only be made smaller (for fixed energy per bit) by allowing more than two waveforms for transmitting information.

BPSK Modulator

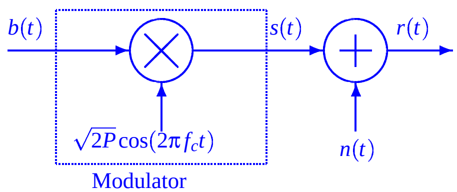
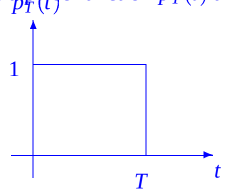


Figure 33: Modulator for BPSK

To mathematically described the transmitted signal we define a pulse function $p_T(t)$ as

$$p_T(t) = \begin{cases} 1, & 0 \le t \le T \\ 0, & \text{otherwise.} \end{cases}$$



Let $b(t)$ denote the data waveform consisting of an infinite sequence of pulses of duration T

and height ± 1.

$$b(t) = \sum_{l=-\infty}^{\infty} b_l \, p_T(t - lT), \quad b_l \in \{+1, -1\}.$$

The transmitted signal then is given by

$$\begin{aligned} s(t) &= \sqrt{2P} \sum_{l=-\infty}^{\infty} b_l \cos(2\pi f_c t) p_T(t - lT) \\ &= \sqrt{2P}\, b(t) \cos(2\pi f_c t) = \sqrt{2P} \cos(2\pi f_c t + \phi(t)) \end{aligned}$$

where $\phi(t)$ is the phase waveform. The signal power is P. The energy of each transmitted bit is $E = PT$.

The phase of a BPSK signal can take on one of two values as shown in Figure VI-3.

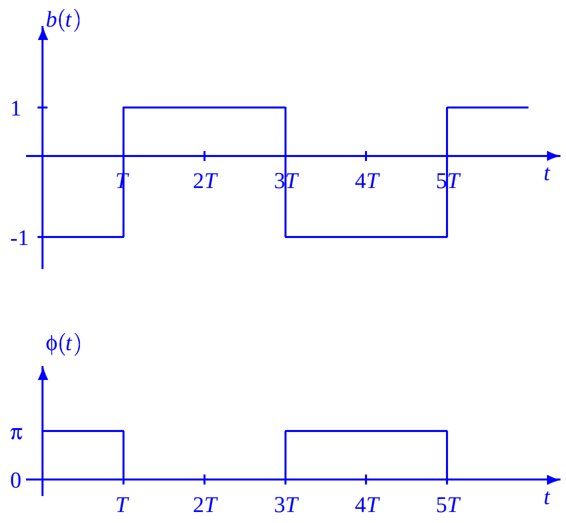


Figure 34: Data and Phase waveforms for BPSK

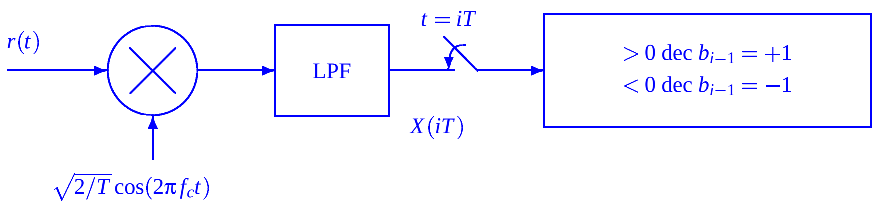


Figure 35: Demodulator for BPSK

The optimum receiver for BPSK in the presence of additive white Gaussian noise is shown in Figure VI-3. The low pass filter (LPF) is a filter “matched” to the baseband signal being transmitted. For BPSK this is just a rectangular pulse of duration T. The impulse response is $h(t) = p_T(t)$. The output of the low pass filter is

$$X(t) = \int_{-\infty}^{\infty} \sqrt{2/T}\cos(2\pi f_c \tau) h(t-\tau) r(\tau) d\tau.$$

The sampled version of the output is given by

$$\begin{aligned} X(iT) &= \int_{-\infty}^{\infty} \sqrt{2/T}\cos(2\pi f_c \tau) p_T(iT-\tau) r(\tau) d\tau \\ &= \int_{(i-1)T}^{iT} \sqrt{2/T}\cos(2\pi f_c \tau) \Big[\sqrt{2P}\, b(\tau)\cos(2\pi f_c \tau) \\ &\qquad + n(\tau)\Big] d\tau \\ &= \int_{(i-1)T}^{iT} 2\sqrt{P/T} b_{i-1} \cos(2\pi f_c \tau)\cos(2\pi f_c \tau) d\tau + \eta_i. \end{aligned}$$

η_i is Gaussian random variable, mean 0 variance $N_0/2$. Assuming $2\pi f_c T = 2\pi n$ for some integer n (or that $f_c T >> 1$)

$$\boxed{X(iT) = \sqrt{PT}\, b_{i-1} + \eta_i = \sqrt{E}\, b_{i-1} + \eta_i.}$$

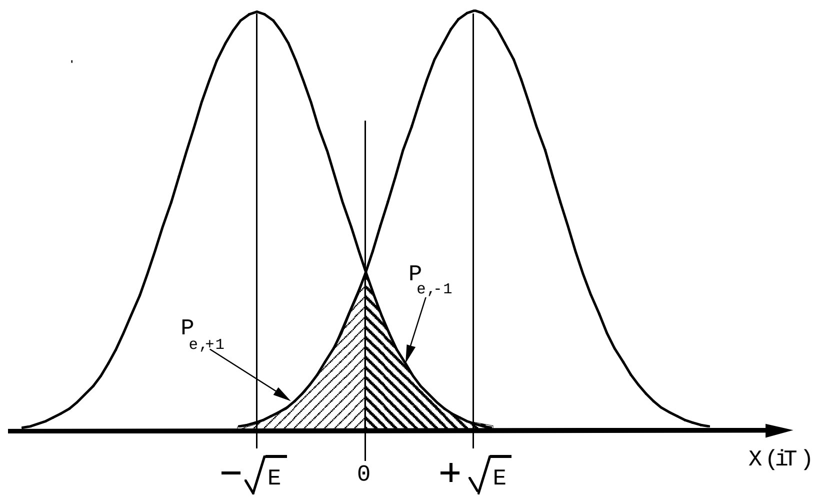


Figure 36: Probability Density of Decision Statistic for Binary Phase Shift Keying

Bit Error Probability of BPSK

$$P_{e,b} = Q\left(\sqrt{\frac{2E}{N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

where

$$Q(x) = \int_x^{\infty} \frac{1}{2\pi} e^{-u^2/2} du$$

For binary signals this is the smallest bit error probability, i.e. BPSK are optimal signals and the receiver shown above is optimum (in additive white Gaussian noise). For binary signals the energy transmitted per information bit E_b is equal to the energy per signal E. For $P_{e,b} = 10^{-5}$ we need a bit-energy, E_b to noise density N_0 ratio of $E_b/N_0 = 9.6$dB. **Note:** $Q(x)$ is a decreasing function which is 1/2 at $x = 0$. There are efficient algorithms (based on Taylor series expansions) to calculate $Q(x)$. Since $Q(x) \leq e^{\{-x^2/2\}}/2$ the error probability can be upper bounded by

$$P_{e,b} \leq \frac{1}{2} e^{\{-E_b/N_0\}}$$

which decreases exponentially with signal-to-noise ratio.

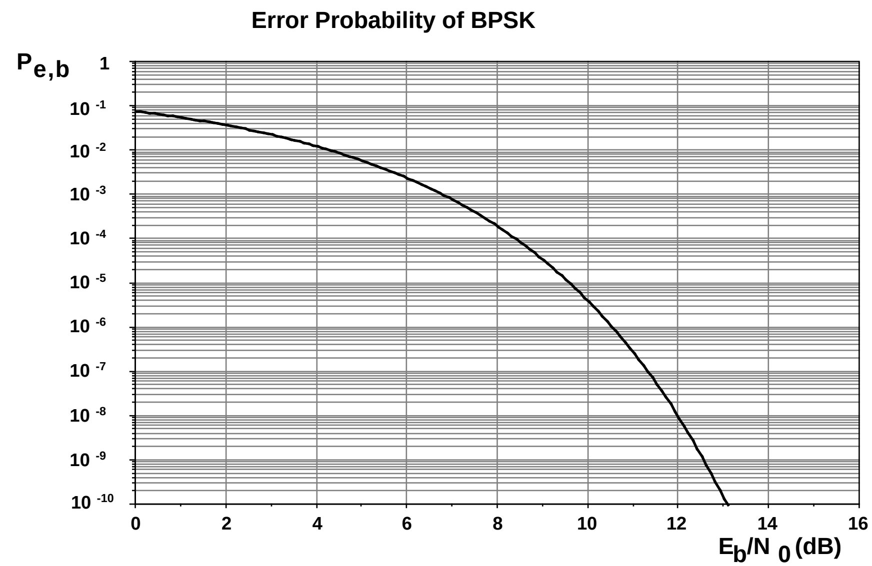


Figure 37: Error Probability of BPSK.

Bandwidth of BPSK

The power spectral density is a measure of the distribution of power with respect to frequency. The power spectral density for BPSK has the form

$$S(f) = \frac{PT}{2}\left[\text{sinc}^2((f-f_c)T) + \text{sinc}^2((f+f_c)T)\right]$$

where

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}.$$

Notice that

$$\int_{-\infty}^{\infty} S(f)df = P.$$

The power spectrum has zeros or nulls at $f - f_c = i/T$ except for $i = 0$; that is there is a null at $f - f_c = \pm 1/T$ called the first null; a null at $f - f_c = \pm 2/T$ called the second null; etc. The bandwidth between the first nulls is called the null-to-null bandwidth. For BPSK the null-to-null bandwidth is $2/T$. Notice that the spectrum falls off as $(f - f_c)^2$ as f moves away from f_c. (The spectrum of MSK falls off as the fourth power, versus the second power for BPSK).

It is possible to reduce the bandwidth of a BPSK signal by filtering. If the filtering is done properly the (absolute) bandwidth of the signal can be reduced to $1/T$ without causing any intersymbol interference; that is all the power is concentrated in the frequency range

$-1/(2T) \leq |f - f_c| \leq 1/(2T)$. The drawbacks are that the signal loses its constant envelope property (useful for nonlinear amplifiers) and the sensitivity to timing errors is greatly increased. The timing sensitivity problem can be greatly alleviated by filtering to a slightly larger bandwidth $-(1+\alpha)/(2T) \leq |f - f_c| \leq (1+\alpha)/(2T)$.

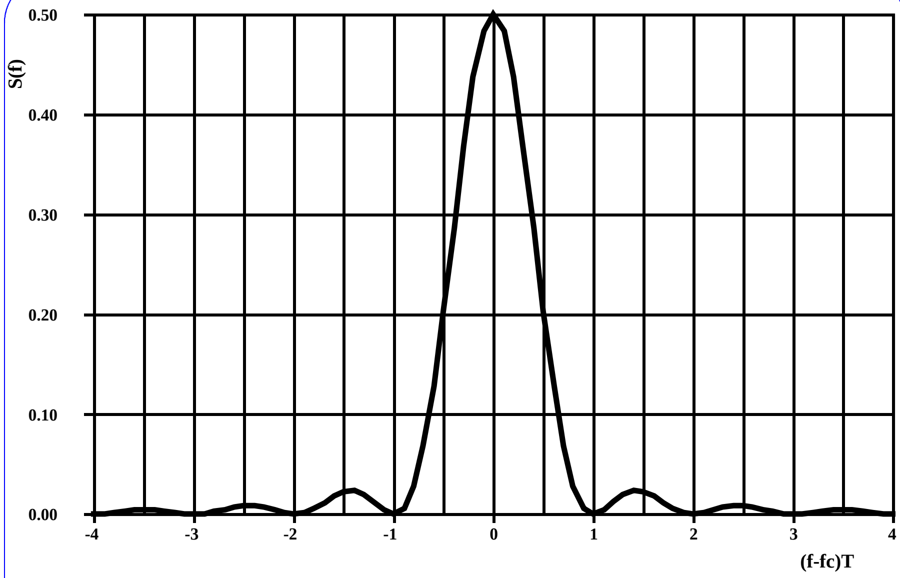


Figure 38: Spectrum of BPSK

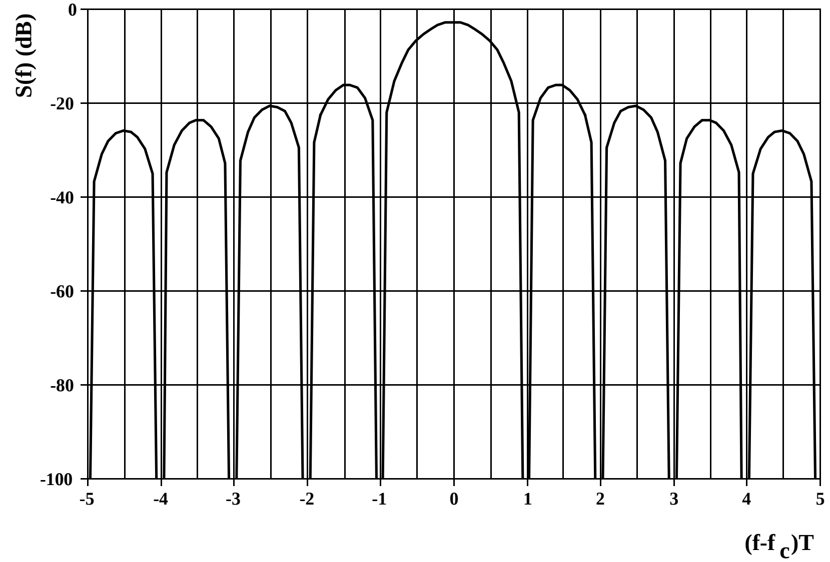


Figure 39: Spectrum of BPSK

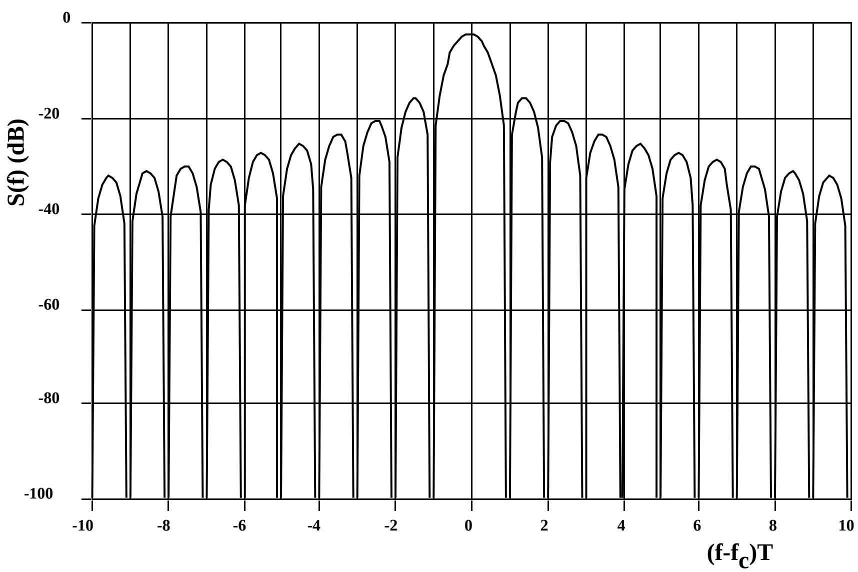


Figure 40: Spectrum of BPSK

Example

Given:

- Noise power spectral density of $N_0/2 = -180$ dBm/Hz $=10^{-21}$ Watts/Hz.
- $P_r = 3 \times 10^{-13}$ Watts
- Desired $P_e = 10^{-7}$.

Find: The data rate that can be used and the bandwidth that is needed.

Solution: Need $Q(\sqrt{2E_b/N_0}) = 10^{-7}$ or $E_b/N_0 = 11.3$dB or $E_b/N_0 = 13.52$. But $E_b/N_0 = P_r T/N_0 = 13.52$. Thus the data bit must be at least $T = 9.0 \times 10^{-8}$ seconds long, i.e. the data rate $1/T$ must be less than 11 Mbits/second. Clearly we also need a (null-to-null) bandwidth of 22 MHz.

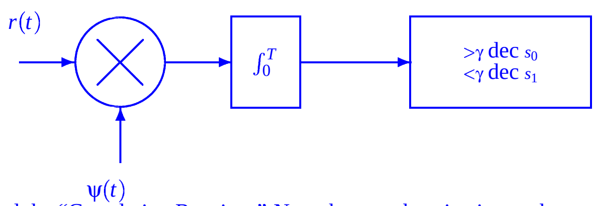
An alternative view of BPSK is that of two antipodal signals; that is

$$s_0(t) = \sqrt{E}\psi(t), \quad 0 \le t \le T$$

and

$$s_1(t) = -\sqrt{E}\psi(t), \quad 0 \le t \le T$$

where $\psi(t) = \sqrt{2/T}\cos(2\pi f_c t)$, $0 \le t \le T$ is a unit energy waveform. The above describes the signals transmitted only during the interval $[0,T]$. Obviously this is repeated for other intervals. The receiver correlates with $\psi(t)$ over the interval $[0,T]$ and compares with a threshold (usually 0) to make a decision. The correlation receiver is shown below.



This is called the "Correlation Receiver." Note that synchronization to the symbol timing and oscillator phase are required.

Effect of Filtering and Nonlinear Amplification on a BPSK waveform

In this section we illustrate one main drawback to BPSK. The fact that the signal amplitude has discontinuities causes the spectrum to have fairly large sidelobes. For a system that has a constraint on the bandwidth this can be a problem. A possible solution is to filter the signal. A bandpas filter centered at the carrier frequency which removes the sidbands can be inserted after mixing to the carrier frequency. Alternatly we can filter the data signal at baseband before mixing to the carrier frequency.

Below we simulate this type of system to illustrate the effect of filtering and nonlinear amplification. The data waveform $b(t)$ is mixed onto a carrier. This modulated waveform is denoted by

$$s_1(t) = \sqrt{2P}\cos(2\pi f_c t)$$

The signal $s_1(t)$ is filtered by a fourth order bandpass Butterworth filter with passband from $f_c - 4Rb$ to $f_c + 4Rb$ The filtered signal is denoted by $s_2(t)$. The signal $s_2(t)$ is then amplified.

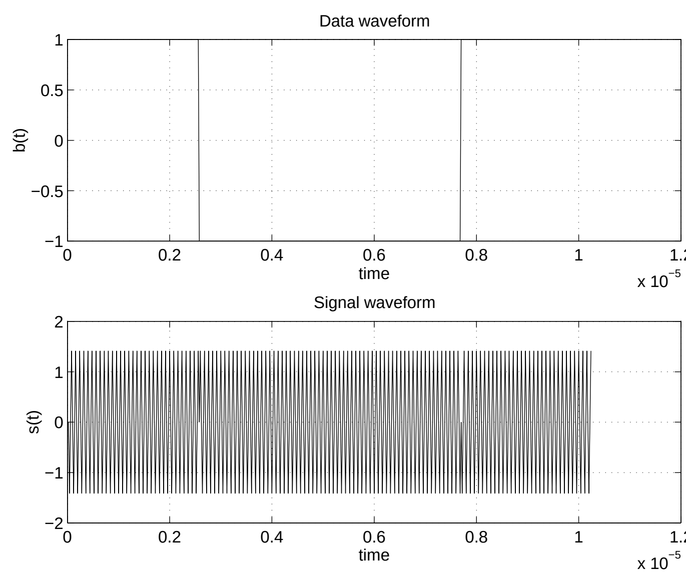
The input-output characteristics of the amplifier are

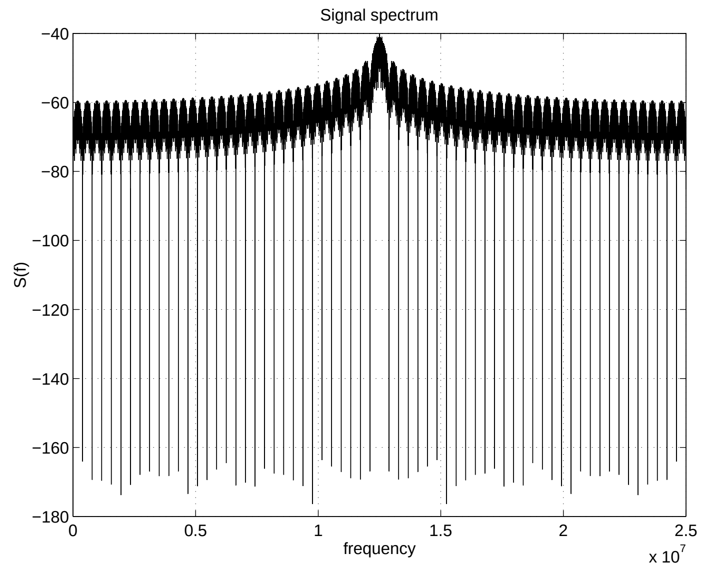
$$s_3(t) = 100\tanh(2s_1(t))$$

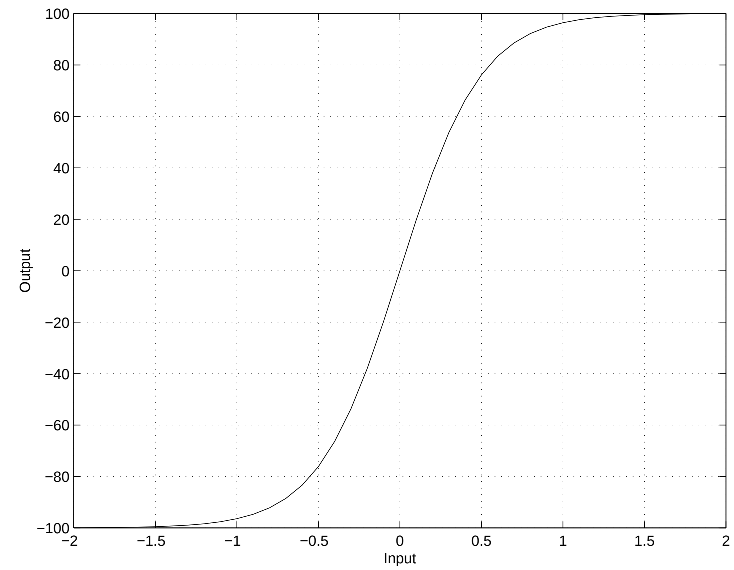
This amplifier is fairly close to a hard limiter in which every input greater than zero is mapped to 100 and every input less than zero is mapped to -100.

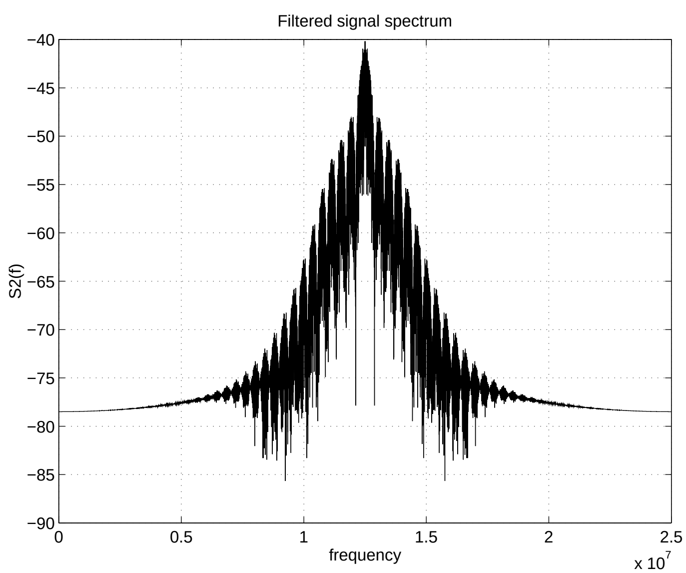
Simulation Parameters

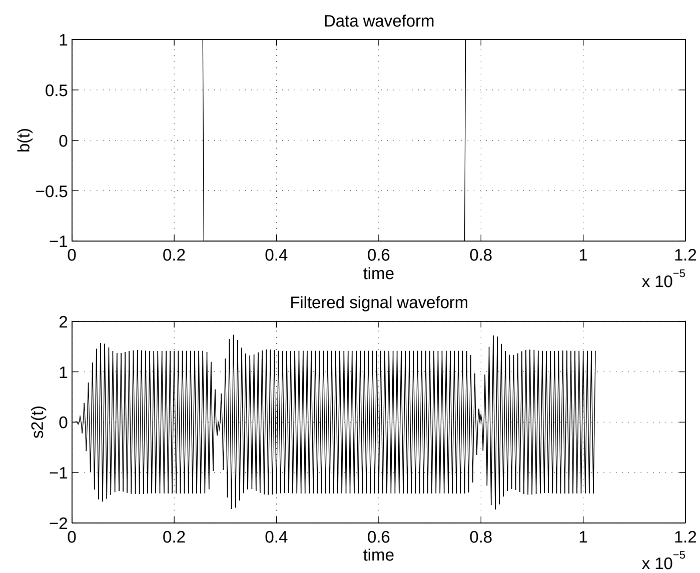
Sampling Frequency= 50MHz
Sampling Time =20nseconds
Center Frequency= 12.5MHz
Data Rate=390.125kbps
Simulation Time= 1.31072 m s

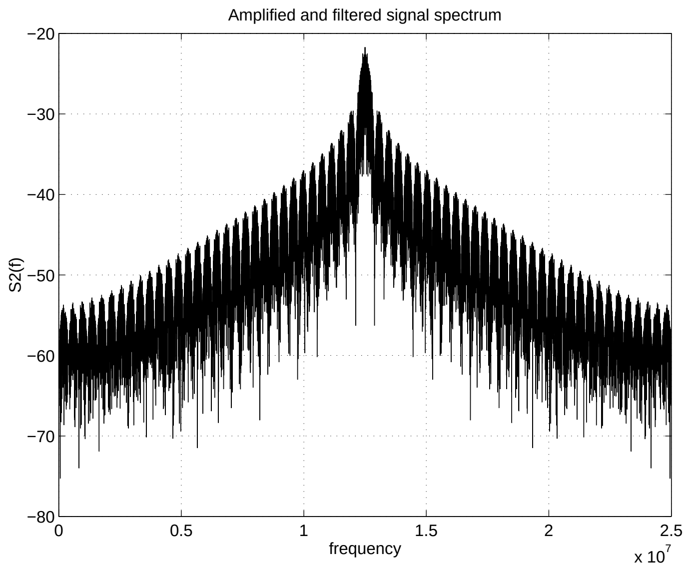


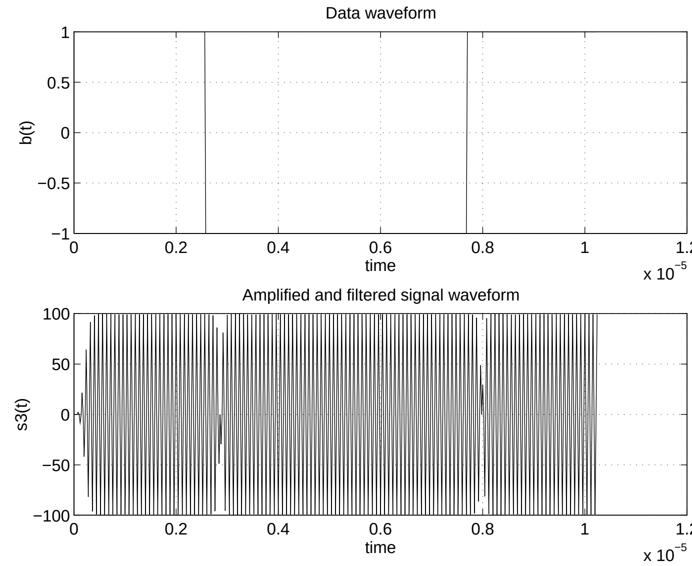

Signal spectrum
S(f)
frequency
x 10^7


Output
Input


Filtered signal spectrum
S2(f)
frequency
x 10^7


Data waveform
b(t)
time
x 10^−5
Filtered signal waveform
s2(t)
time
x 10^−5





Quaternary Phase Shift Keying (QPSK)

The next modulation technique we consider is QPSK. In this modulation technique one of four phases of the carrier is transmitted in a symbol duration denoted by T_s. Since one of four waveforms is transmitted there are two bits of information transmitted during each symbol duration. An alternative way of describing QPSK is that of two carriers offset in phase by 90 degrees. Each of these carriers is then modulated using BPSK. These two carriers are called the inphase and quadrature carriers. Because the carriers are 90 degrees offset, at the output of the correlation receiver they do not interfer with each other (assuming perfect phase synchronization). The advantage of QPSK over BPSK is that the the data rate is twice as high for the same bandwidth. Alternatively single-sideband BPSK would have the same rate in bits per second per hertz but would have a more difficult job of recovering the carrier frequency and phase.

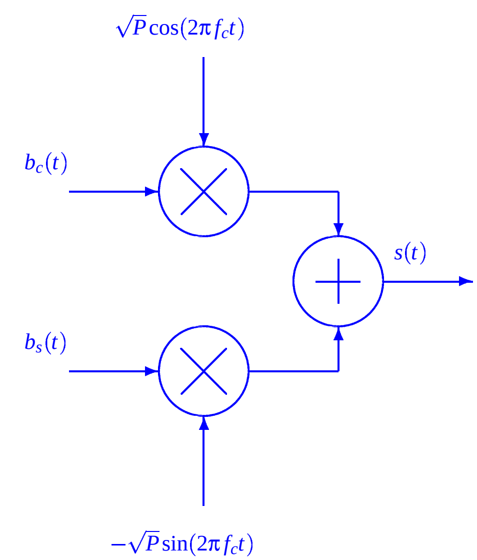


Figure 41: Modulator for QPSK

$$
\begin{aligned}
b_c(t) &= \sum_{l=-\infty}^{\infty} b_{c,l}\; p_{T_s}(t-lT_s), \quad b_{c,l} \in \{+1,-1\} \\
b_s(t) &= \sum_{l=-\infty}^{\infty} b_{s,l}\; p_{T_s}(t-lT_s), \quad b_{s,l} \in \{+1,-1\}
\end{aligned}
$$

$$
\begin{aligned}
s(t) &= \sqrt{P}[b_c(t)\cos(2\pi f_c t) - b_s(t)\sin(2\pi f_c t)] \\
&= \sqrt{2P}\cos(2\pi f_c t + \phi(t))
\end{aligned}
$$

The transmitted power is still P. The symbol duration is T_s seconds. The data rate is $R_b = 2/T_s$ bits seconds.

The phase $\phi(t)$, of the transmitted signal is related to the data waveform as follows.

$$
\phi(t) = \sum_{l=-\infty}^{\infty} \phi_l\; p_{T_s}(t-lT_s), \qquad \phi_l \in \{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}
$$

The relation between ϕ_l and $b_{c,l}, b_{s,l}$ is shown in the following table

$b_{c,l}$	$b_{s,l}$	ϕ_l
+1	+1	$\pi/4$
-1	+1	$3\pi/4$
-1	-1	$5\pi/4$
+1	-1	$7\pi/4$

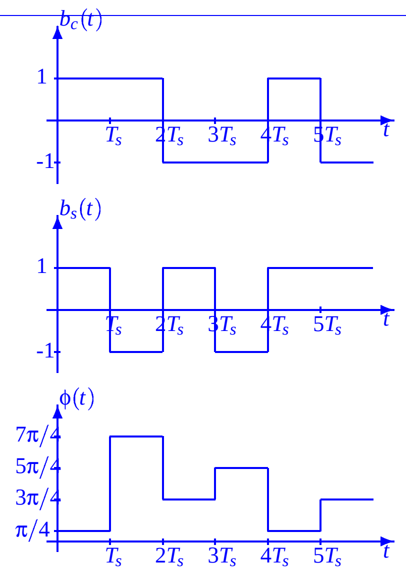


Figure 42: Timing and Phase of QPSK

The constellation of QPSK is shown below. The phase of the overall carrier can be on of four values. Transitions between any of the four values may occur at any symbol transition. Because of this, it is possible that the transition is to the 180 degree opposite phase. When this happens the amplitude of the signal goes through zero. In theory this is an instantaneous transition. In practice, when the signal has been filtered to remove out-of-band components this transition is slowed down. During this transition the amplitude of the carrier goes throguh zero. This can be undesireable for various reasons. One reason is that nonlinear amplifiers with a non constant envelope signal will regenerate the out-of-band spectral components. Another reason is that at the receiver, certain synchronization circuits need constant envelope to maintain their tracking capability.

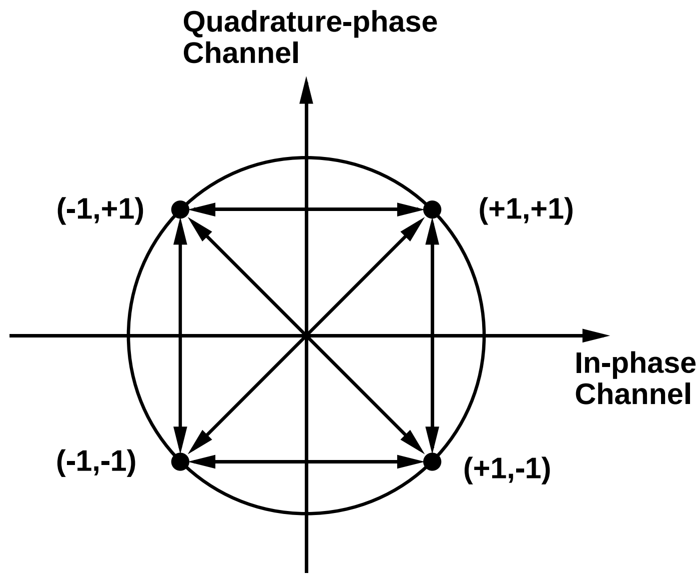


Figure 43: Constellation of QPSK

The bandwidth of QPSK is given by

$$
\begin{aligned}
S(f) &= PT_s/2\,[\mathrm{sinc}^2((f-f_c)T_s)+\mathrm{sinc}^2((f+f_c)T_s)] \\
&= PT_b\,[\mathrm{sinc}^2(2(f-f_c)T_b)+\mathrm{sinc}^2(2(f+f_c)T_b)]
\end{aligned}
$$

since $T_s = T_b/2$. Thus while the spectrum is compressed by a factor of 2 relative to BPSK with the same bit rate, the center lobe is also 3dB higher, that is the peak power density is higher for QPSK than BPSK. The null-to-null bandwidth is $2/T_s = R_b$.

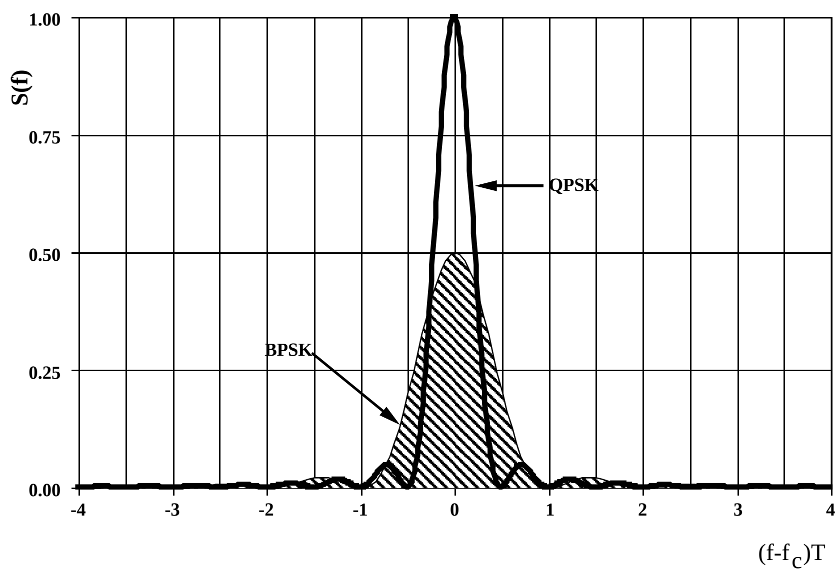


Figure 44: Spectrum of QPSK

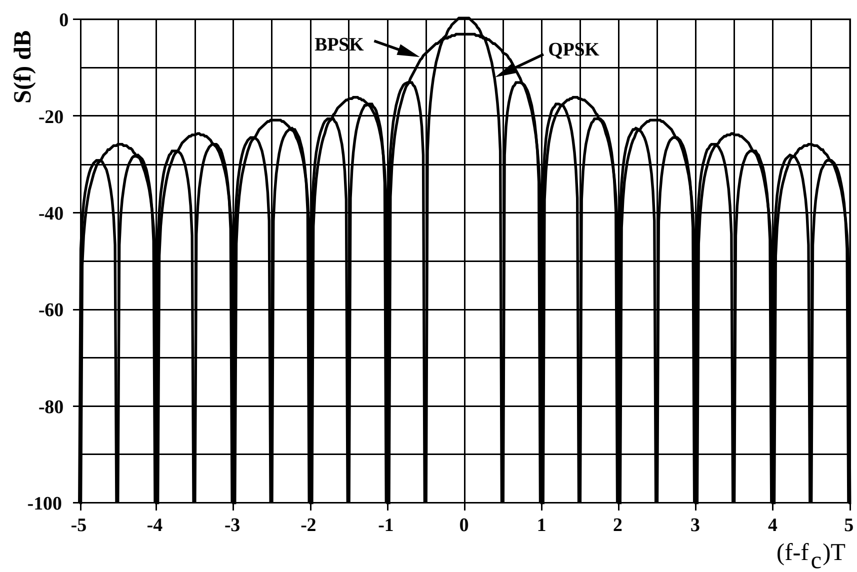


Figure 45: Spectrum of QPSK

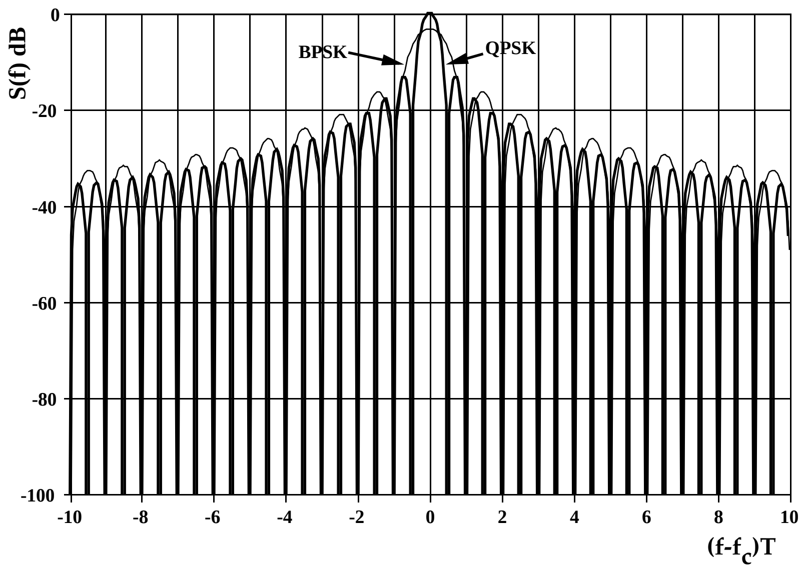


Figure 46: Spectrum of QPSK

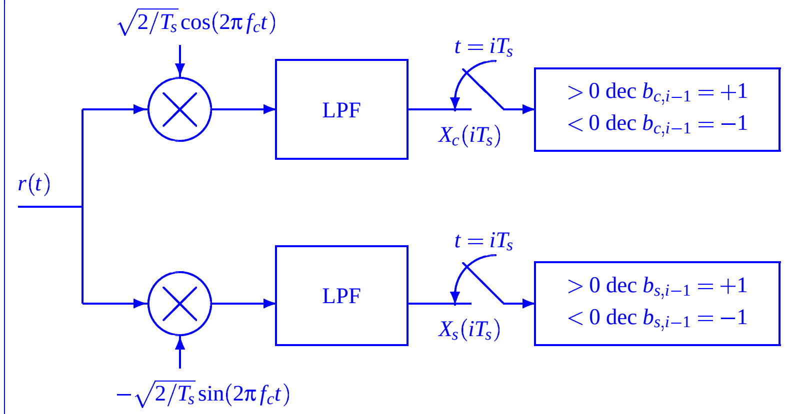


Figure 47: QPSK Demodulator

Assuming $2\pi f_c T_s = 2\pi n$ or $2\pi f_c T_s \gg 1$

$$X_c(iT_s) = \sqrt{PT_s/2}\, b_{c,i-1} + \eta_{c,i} = \sqrt{E_b}\, b_{c,i-1} + \eta_{c,i}$$

$$X_s(iT_s) = \sqrt{PT_s/2}\, b_{s,i-1} + \eta_{s,i} = \sqrt{E_b}\, b_{s,i-1} + \eta_{s,i}$$

where $E_b = PT_s/2$ is the energy per transmitted bit. Also $\eta_{c,i}$ and $\eta_{s,i}$ are Gaussian random variables, with mean 0 and variance $N_0/2$.

Bit Error Probability of QPSK

$$P_{e,b} = Q(\sqrt{\frac{2E_b}{N_0}})$$

The probability that a symbol error is made is

$$P_{e,s} = 1 - (1 - P_{e,b})^2 = 2P_{e,b} - P_{e,b}^2$$

Thus for the same data rate, transmitted power, and bit error rate (probability of error), QPSK has half the (null-to-null) bandwidth of BPSK.

Example

Given:

- Noise power spectral density of $N_0/2 = -110$ dBm/Hz $=10^{-14}$ Watts/Hz.
- $P_r = 3 \times 10^{-6}$ Watts
- Desired $P_e = 10^{-7}$.

Find: The data rate that can be used and the bandwidth that is needed for QPSK.

Solution: Need $Q(\sqrt{2E_b/N_0}) = 10^{-7}$ or $E_b/N_0 = 11.3$dB or $E_b/N_0 = 13.52$. But

$$\begin{aligned} E_b/N_0 &= \frac{\frac{P_r}{2}(T_s)}{N_0} \\ &= P_r T / N_0 = 13.52 \end{aligned}$$

since $T_s = 2T$. Thus the data bit must be at least $T = 9.0 \times 10^{-8}$ seconds long, i.e. the data rate $1/T$ must be less than 11 Mbits/second. Clearly we also need a (null-to-null) bandwidth of 11 MHz.

Offset Quaternary Phase Shift Keying (OQPSK)

The disadvantages of QPSK can be fixed by offsetting one of the data streams by a fraction (usually 1/2) of a symbol duration. By doing this we only allow one data bit to change at a time. When this is done the possible phase transitions are ± 90 deg. In this way the transitions through the origin are illiminated. Offset QPSK then gives the same performance as QPSK but will have less distorition when there is filtering and nonlinearities.

h

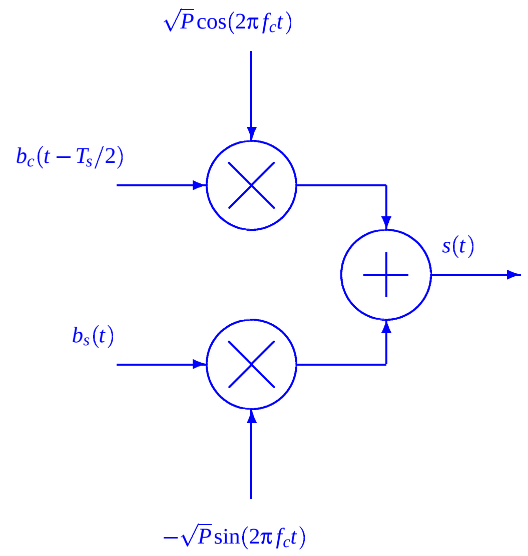


Figure 48: Modulator for OQPSK

$$
\begin{aligned}
b_c(t) &= \sum_{l=-\infty}^{\infty} b_{c,l}\; p_{T_s}(t-lT_s), \quad b_{c,l} \in \{+1,-1\} \\
b_s(t) &= \sum_{l=-\infty}^{\infty} b_{s,l}\; p_{T_s}(t-lT_s), \quad b_{s,l} \in \{+1,-1\} \\
s(t) &= \sqrt{P}[b_c(t-T_s/2)\cos(2\pi f_c t) - b_s(t)\sin(2\pi f_c t)] \\
s(t) &= \sqrt{2P}\cos(2\pi f_c t + \phi(t))
\end{aligned}
$$

The transmitted power is still P. The symbols duration is T_s seconds. The data rate is $R_b = 2/T_s$ bits seconds. The bandwidth (null-to-null) is $2/T_s = R_b$. This modification of QPSK removes the possibility of both data bits changing simultaneously. However, one of the data bits may change every $T_s/2$ seconds but 180 degree changes are not allowed. The bandwidth of OQPKS is the same as QPSK. OQPSK has advantage over QPSK when passed through nonlinearities (such as in a satellite) in that the out of band interference generated by first bandlimiting and then hard limiting is less with OQPSK than QPSK.

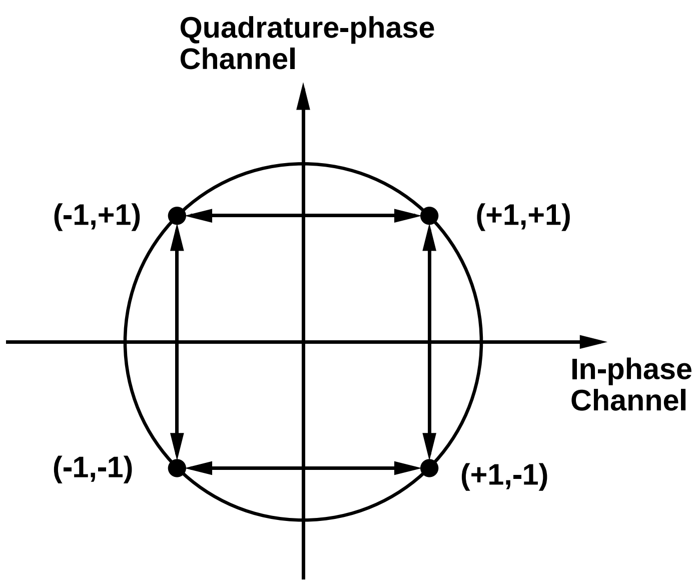


Figure 49: Constellation of QPSK

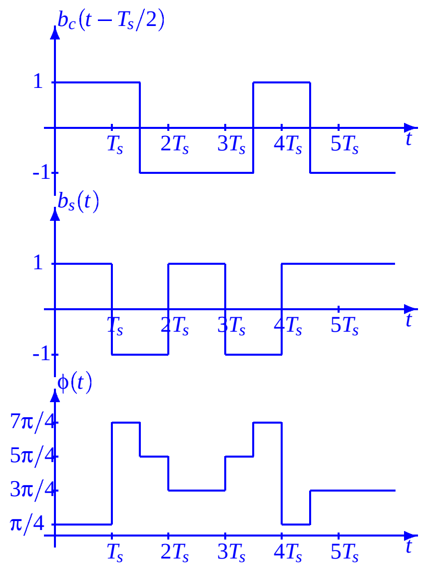


Figure 50: Data and Phase Waveforms for OQPSK

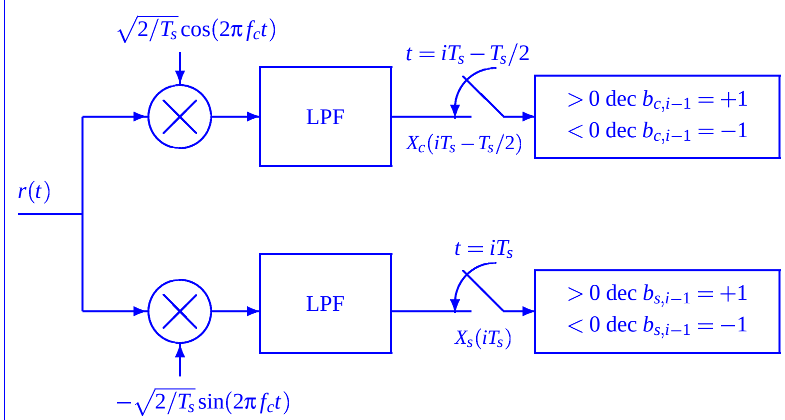


Figure 51: Demodulator for OQPSK

Assuming $2\pi f_c T_s = 2\pi n$ or $2\pi f_c T_s \gg 1$

$$\begin{aligned} X_c(iT_s - T_s/2) &= \sqrt{PT_s/2}\, b_{c,i-1} + \eta_{c,i} = \sqrt{E_b}\, b_{c,i-1} + \eta_{c,i} \\ X_s(iT_s) &= \sqrt{PT_s/2}\, b_{s,i-1} + \eta_{s,i} = \sqrt{E_b}\, b_{s,i-1} + \eta_{s,i} \end{aligned}$$

where $E_b = PT_s/2$ is the energy per transmitted bit. Also $\eta_{c,i}$ and $\eta_{s,i}$ are Gaussian random variables, with mean 0 variance $N_0/2$.

Bit Error Probability of OQPSK

$$P_{e,b} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

The probability that a symbol error is made is

$$P_{e,s} = 1 - (1 - P_{e,b})^2 = 2P_{e,b} - P_{e,b}^2$$

This is the same as QPSK.

Minimum Shift Keying (MSK)

Minimum shift keying can be viewed in several different ways and has a number of significant advantages over the previously considered modulation schemes. MSK can be thought of as a variant of OQPSK where the data pulse waveforms are shaped to allow smooth transition between phases. It can also be thought of a a form of frequency shift keying where the two frequencies are separated by the minimum amount to maintain orthogonality and have continuous phase when switching from one frequency to another (hence the name minimum shift keying). The advantages of MSK include a better spectral efficiency in most cases. In fact the spectrum of MSK falls off at a faster rate than BPSK, QPSK and OQPSK. In addition there is an easier implementation than OQPSK (called serial MSK) that avoids the problem of having a precisely controlled time offset between the two data streams. An additional advantage is that MSK can be demodulator noncoherently as well as coherently. So for applications requiring a low cost receiver MSK may be a good choice.

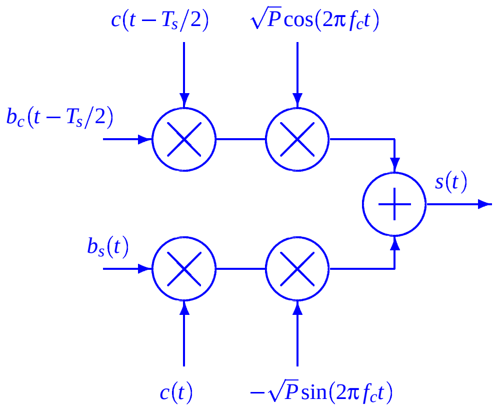


Figure 52: Modulator for MSK

$$
\begin{aligned}
b_c(t) &= \sum_{l=-\infty}^{\infty} b_{c,l}\, p_{T_s}(t-lT_s), \quad b_{c,l} \in \{+1,-1\} \\
b_s(t) &= \sum_{l=-\infty}^{\infty} b_{s,l}\, p_{T_s}(t-lT_s), \quad b_{s,l} \in \{+1,-1\} \\
c(t) &= \sqrt{2}\sin(\pi t/T_s) \quad c(t-T_s/2) = -\sqrt{2}\cos(\pi t/T_s)
\end{aligned}
$$

$$
\begin{aligned}
s(t) &= \sqrt{P}[b_c(t-T_s/2)c(t-T_s/2)\cos(2\pi f_c t) - b_s(t)c(t)\sin(2\pi f_c t)] \\
s(t) &= \sqrt{2P}\,[\{-b_c(t-T_s/2)\cos(\pi t/T_s)\}\cos(2\pi f_c t) \\
&\qquad -\{b_s(t)\sin(\pi t/T_s)\}\sin(2\pi f_c t)] \\
&= \sqrt{2P}\cos(2\pi f_c t + \phi(t))
\end{aligned}
$$

where

$$
\cos(\phi(t)) = -b_c(t-T_s/2)\cos(\pi t/T_s)
$$

$$
\sin(\phi(t)) = b_s(t)\sin(\pi t/T_s)
$$

$$
\phi(t) = \tan^{-1}\left(\frac{b_s(t)\sin(\pi t/T_s)}{-b_c(t-T_s/2)\cos(\pi t/T_s)}\right)
$$

$b_c(t-T_s/2)$	$b_s(t)$	$\phi(t)$
+1	+1	$\pi - \frac{\pi t}{T_s}$
+1	-1	$\pi + \frac{\pi t}{T_s}$
-1	+1	$\frac{\pi t}{T_s}$
-1	-1	$-\frac{\pi t}{T_s}$

In the above table, because of the delay of the bit stream corresponding to the cosine branch, only one bit is allowed to change at a time. During each time interval of duration $T_s/2$ during which the data bits remain constant there is a phase shift of $\pm\pi/2$. Because the phase changes linearly with time MSK can also be viewed as frequency shift keying. The two different frequencies are $f_c + \frac{1}{2T_s}$ and $f_c - \frac{1}{2T_s}$. The change in frequency is $\Delta f = \frac{1}{T_s} = \frac{1}{2T_b}$ where $T_b^{-1} = 2/T_s$ is the data bit rate. The transmitted power is still P. The symbols duration is T_s seconds. The data rate is $R_b = 2/T_s$ bits seconds. The signal has constant envelope which is useful for nonlinear amplifiers. The bandwidth is different because of the pulse shaping waveforms.

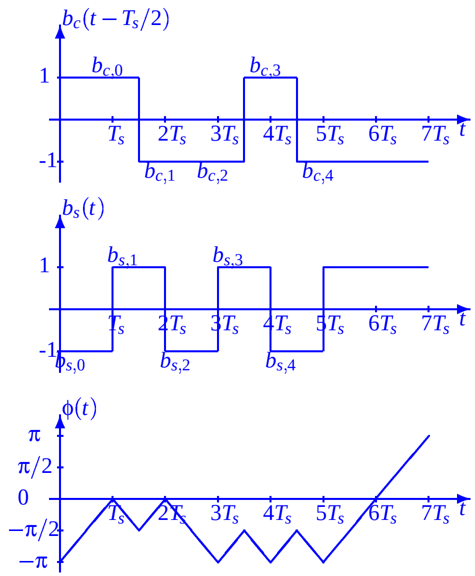


Figure 53: Data and phase waveforms for MSK

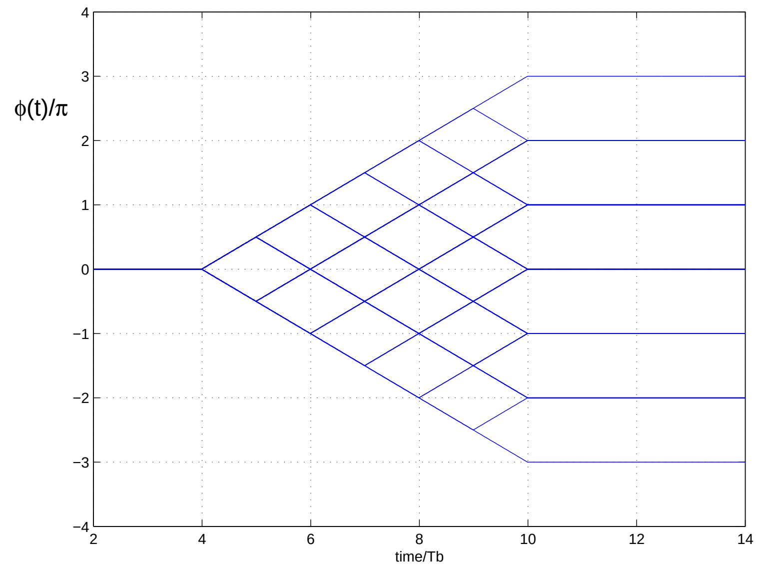


Figure 54: Phase of MSK signals

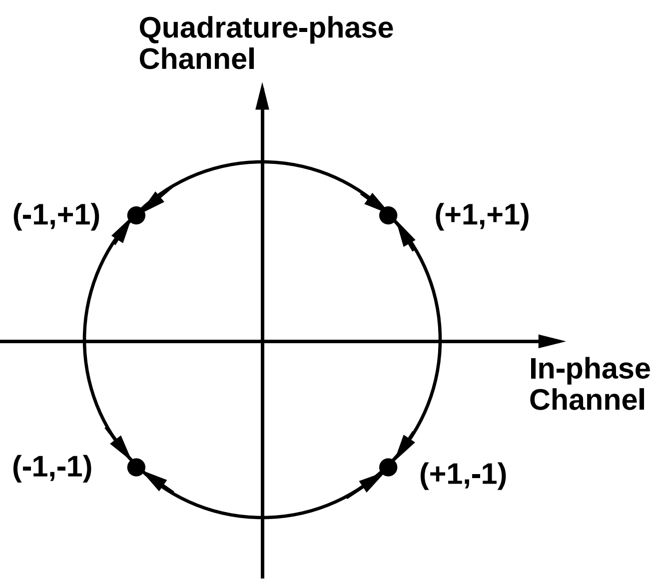


Figure 55: Constellation of MSK

The spectrum of MSK is given by

$$S(f) = \frac{8PT_b}{\pi^2}\left\{\frac{\cos^2(2\pi T_b(f-f_c))}{[1-(4T_b(f-f_c))^2]^2} + \frac{\cos^2(2\pi T_b(f+f_c))}{[1-(4T_b(f+f_c))^2]^2}\right\}$$

The nulls in the spectrum are at $(f-f_c)T_b = 0.75, 1.25, 1.75,\ldots$. Because we force the signal to be continuous in phase MSK has significantly faster decay of the power spectrum as the frequency from the carrier becomes larger. MSK decays as $1/f^4$ while QPSK, OQPSK, and BPSK decay as $1/f^2$ as the frequency differs more and more from the center frequency.

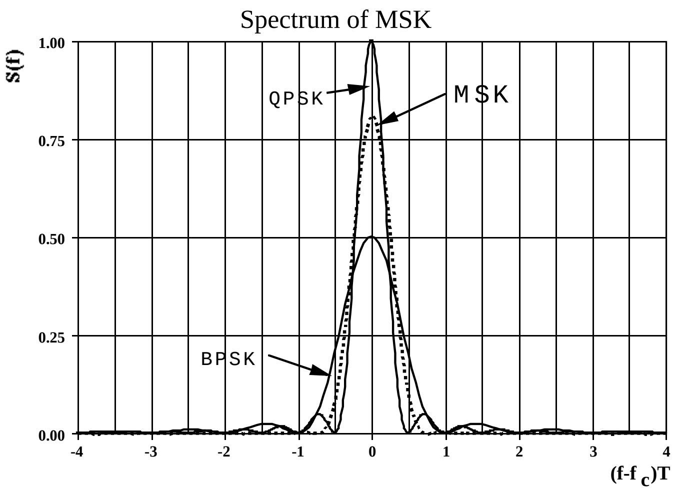


Figure 56: Spectrum of MSK

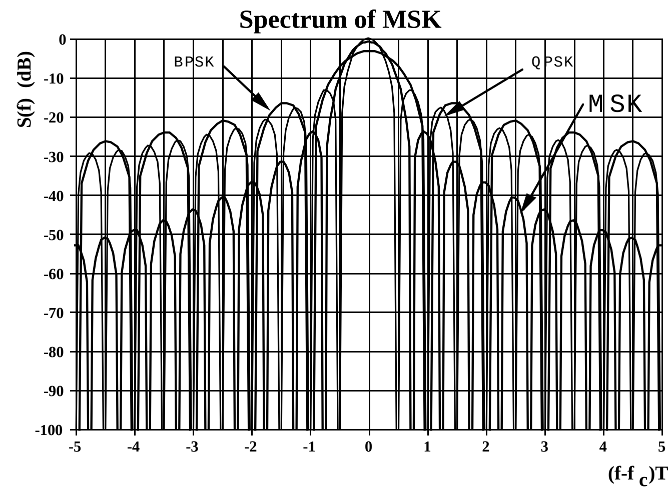


Figure 57: Spectrum of MSK

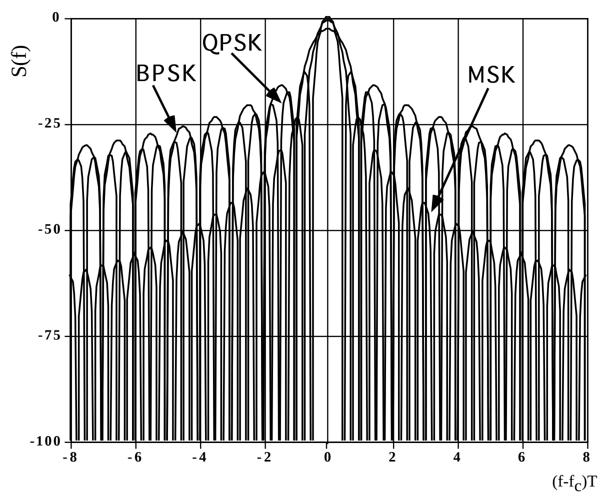


Figure 58: Spectrum of MSK

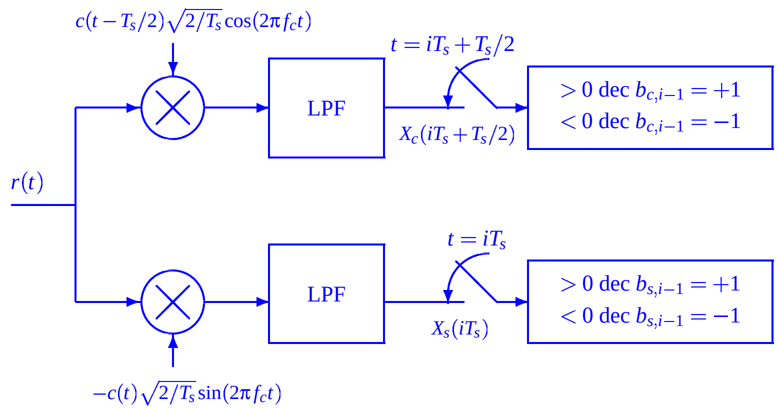


Figure 59: Coherent Demodulator for MSK

Assuming $2\pi f_c T_s = 2\pi n$ or $2\pi f_c T_s \gg 1$

$$\begin{aligned} X_c(iT_s + T_s/2) &= \sqrt{PT_s/2}\, b_{c,i-1} + \eta_{c,i} = \sqrt{E_b}\, b_{c,i-1} + \eta_{c,i} \\ X_s(iT_s) &= \sqrt{PT_s/2}\, b_{s,i-1} + \eta_{s,i} = \sqrt{E_b}\, b_{s,i-1} + \eta_{s,i} \end{aligned}$$

where $E_b = PT_s/2$ is the energy per transmitted bit. Also $\eta_{c,i}$ and $\eta_{s,i}$ are Gaussian random variables, with mean 0 variance $N_0/2$.

Bit Error Probability of MSK with Coherent Demodulation

Since the signals are still antipodal

$$P_{e,b} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

The probability that a symbol error is made is

$$P_{e,s} = 1 - (1 - P_{e,b})^2 = 2P_{e,b} - P_{e,b}^2$$

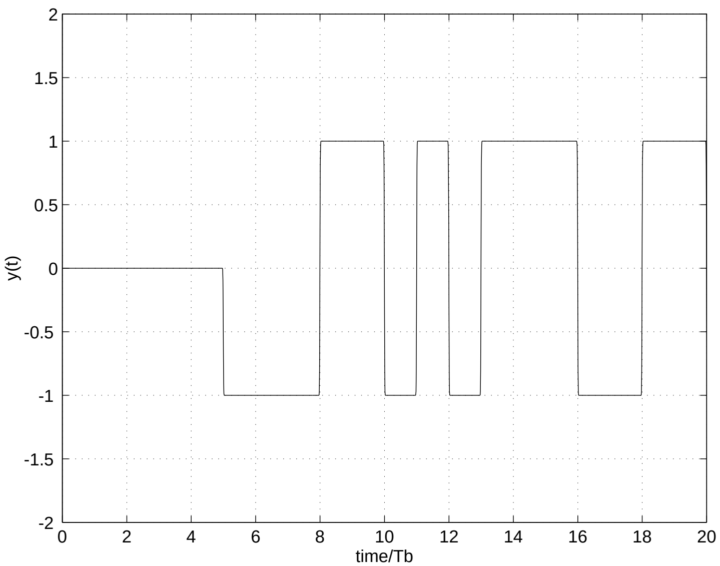


Figure 60: Waveform for Minimum Shift Keying

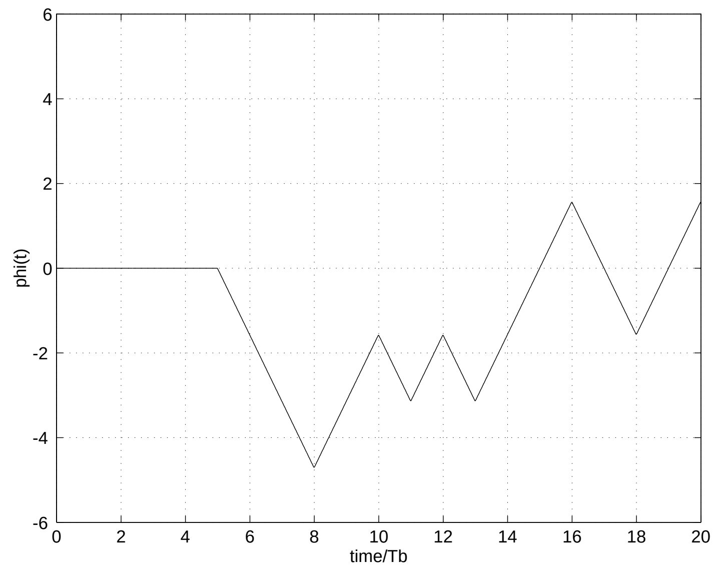


Figure 61: Phase Waveform for Minimum Shift Keying

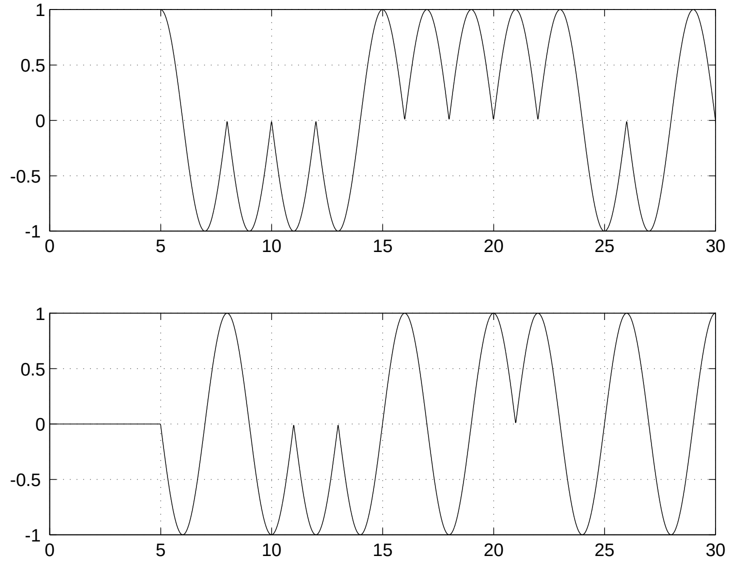


Figure 62: Quadrature Waveforms for Minimum Shift Keying

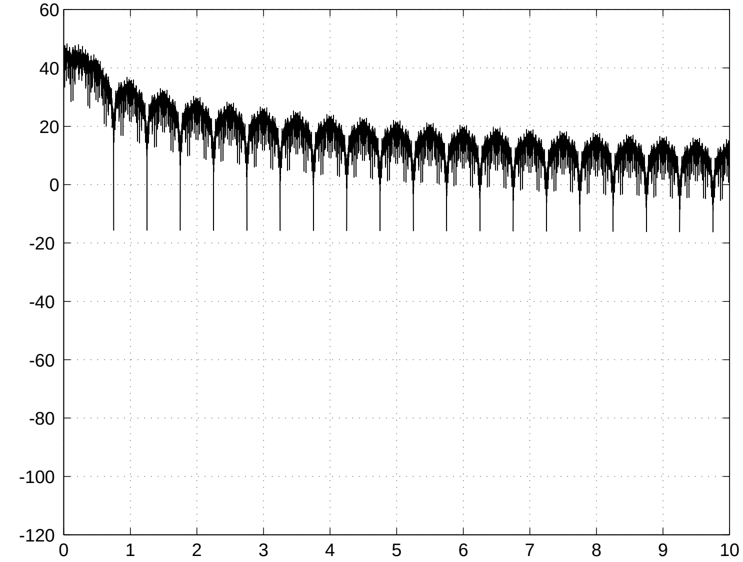


Figure 63: Spectrum for Minimum Shift Keying

Noncoherent Demodulation of MSK

Because MSK can be viewed as a form of Frequency Shift Keying it can also be demodulated noncoherently. For the same sequence of data bits the frequency is $f_c - 1/2T_s$ if $b_c(t - T_s/2) = b_s(t)$ and is $f_c + 1/2T_s$ if $b_c(t - T_s/2) \neq b_s(t)$.

Consider determining $b_{s,i-1}$ at time $(i-1/2)T_s$. Assume we have already determined $b_{c,i-2}$ at time $(i-1)T_s$. If we estimate which of two frequencies is sent over the interval $[(i-1)T_s, (i-1/2)T_s)$ the decision rule is to decide that $b_{s,i-1} = b_{c,i-2}$ if the frequency detected is $f_c - 1/(2T_s)$ and to decide that $b_{s,i-1} = -b_{c,i-2}$ if the frequency detected is $f_c + 1/(2T_s)$.

Consider determining $b_{c,i-1}$ at time iT_s. Assume we have already determined $b_{s,i-1}$ at time $(i-1/2)T_s$. If we estimate which of two frequencies is sent over the interval $[(i-1/2)T_s, iT_s)$ the decision rule is to decide that $b_{c,i-1} = b_{s,i-1}$ if the frequency detected is $f_c - 1/(2T_s)$ and to decide that $b_{c,i-1} = -b_{s,i-1}$ if the frequency detected is $f_c + 1/(2T_s)$.

The method to detect which of the two frequencies is transmitted is identical to that of Frequency Shift Keying which will be considered later.

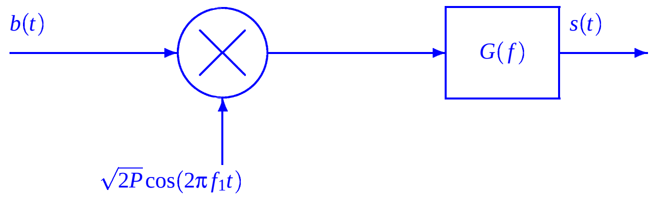
For the example phase waveform shown previously we have that

Time Interval	$[0, T_s/2)$	$[T_s/2, T_s)$	$[T_s, 3T_s/2)$	$[3T_s/2, 2T_s)$	$[2T_s, 5T_s/2)$
Frequency	$+$	$+$	$-$	$+$	$-$
Previous Data	$b_{c,-1} = +1$	$b_{s,0} = -1$	$b_{c,0} = 1$	$b_{s,1} = +1$	$b_{c,1} = -1$
Detected Data	$b_{s,0} = -1$	$b_{c,0} = +1$	$b_{s,1} = +1$	$b_{c,1} = -1$	$b_{s,2} = -1$

So detecting the frequency can also be used to detect the data.

Serial Modulation and Demodulation

The implementation of MSK as parallel branches suffers from significant sensitivity to precise timing of the data (exact shift by T for the inphase component) and the exact balance between the inphase and quadriphase carrier signals. An alternative implementation of MSK that is less complex and does not have these draw backs is known as serial MSK. Serial MSK does have an additional restriction that $f_c = (2n+1)/4T$ which may be important when f_c is about the same as $1/T$ but for $f_c \gg 1/T$ it is not important. The block diagram for serial MSK modulator and demodulator is shown below.



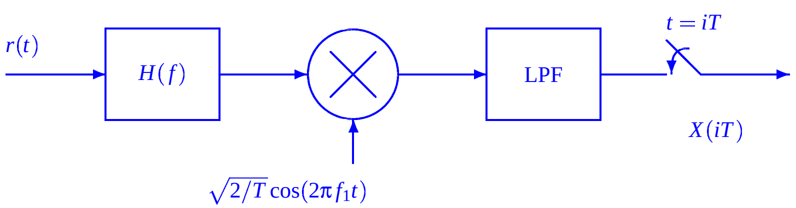
The filter $G(f)$ is given by filter

$$G(f) = T\,\mathrm{sinc}[(f-f_1)T]e^{-j\pi(f-f_1)T} + T\,\mathrm{sinc}[(f+f_1)T]e^{-j\pi(f+f_1)T}$$

$$g(t) = 2\sin(2\pi f_1 t)p_T(t)$$

where $f_1 = f_c - \frac{1}{4T}$ and $f_2 = f_c + \frac{1}{4T}$. (For serial MSK we require $f_c = (2n+1)/4T$ for some integer n. Otherwise the implementation does not give constant envelope).

Demodulator



The filter $H(f)$ is given by

$$H(f) = \frac{4T}{\pi}\frac{\cos[2\pi(f-f_1)T - 0.25]}{1-16[(f-f_1)T-0.25]^2}e^{-j2\pi(f-f_1)T}$$

The low pass filter (LPF) removes double frequency components. Serial MSK is can also be viewed as a filtered form of BPSK where the BPSK signal center frequency is f_1 but the filter is not symmetric with respect to f_1. The receiver is a filter matched to the transmitted signal (and hence optimal). The output is then mixed down to baseband where it is filtered (to remove the double frequency terms) and sampled.

Continuous Phase Modulation

MSK is a special case of a more general form of modulation known as continuous phase modulation where the phase is continuous. The general form of CPM is given by

$$s(t) = \sqrt{2P}\cos(2\pi f_c t + \phi(t))$$

where the phase waveform has the form

$$\begin{aligned}\phi(t) &= 2\pi h \int_0^t \sum_{i=0}^{k} b_i g(\tau - iT)d\tau + \phi_0 \quad kT \le t \le (k+1)T \\ &= 2\pi h \sum_{i=0}^{k} b_i q(t-iT) + \phi_0 \quad kT \le t \le (k+1)T\end{aligned}$$

The function $g(\cdot)$ is the (instantaneous) frequency function, h is called the modulation index and b_i is the data. The function $q(t) = \int_0^t g(\tau)d\tau$ is the phase waveform. The function $g(t) = \frac{dg(t)}{dt}$ is the frequence waveform.

For example if CPM has $h = 1/2$ and

$$q(t) = \begin{cases} 0, & t < 0 \\ t/2, & 0 \le t < T \\ 1/2, & t > T. \end{cases}$$

then the modulation is the same as MSK. Continuous Phase Modulation Techniques have constant envelope which make them useful for systems involving nonlinear amplifiers which also must have very narrow spectral widths.

Example

Given:

- Noise power spectral density of $N_0/2 = -110$ dBm/Hz $=10^{-14}$ Watts/Hz.
- $P_r = 3 \times 10^{-6}$ Watts
- Desired $P_e = 10^{-7}$.
- Bandwidth available=26MHz (at the 902-928MHz band). The peak power outside must be 20dB below the peak power inside the band.

Find: The data rate that can be used for MSK.

Solution: Need $Q(\sqrt{2E_b/N_0}) = 10^{-7}$ or $E_b/N_0 = 11.3$dB or $E_b/N_0 = 13.52$. But $E_b/N_0 = P_r T/N_0 = 13.52$. Thus the data bit must be at least $T = 9.0 \times 10^{-8}$ seconds long, i.e. the data rate $1/T$ must be less than 11 Mbits/second.

Gaussian Minimum Shift Keying

Gaussian minimum shift keying is a special case of continuous phase modulation discussed in the previous section. For GMSK the pulse waveforms are given by

$$g(t) = Q(\frac{t-T}{\sigma}) - Q(\frac{t}{\sigma})$$

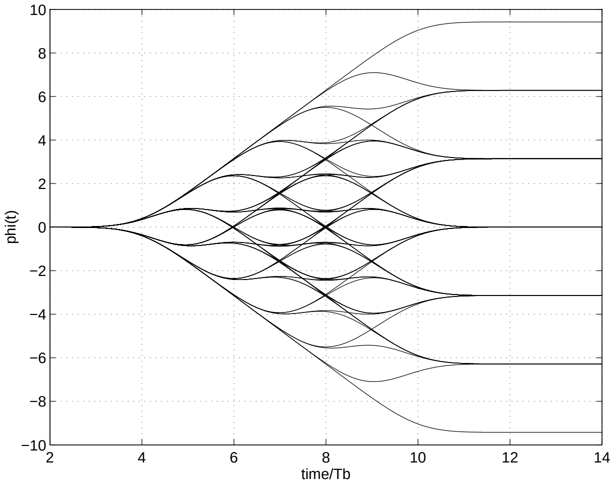


Figure 64: Phase Waveform for Gaussian Minimum Shift Keying (BT=0.3)

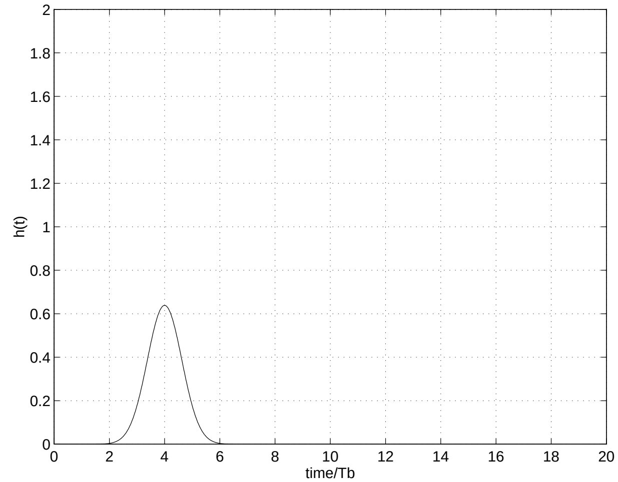


Figure 65: Data Waveform for Gaussian Minimum Shift Keying (BT=0.3)

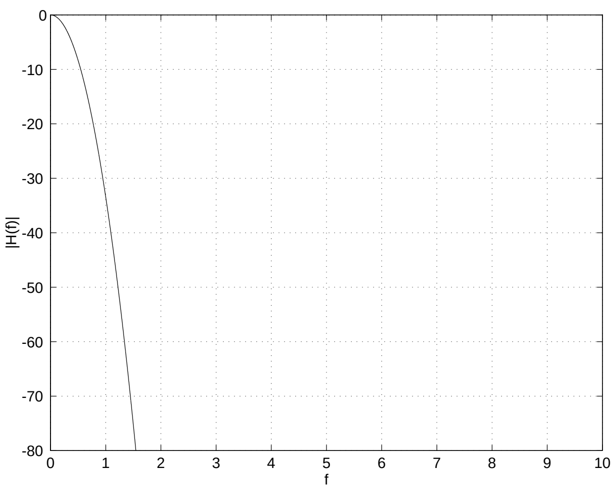


Figure 66: Waveform for Gaussian Minimum Shift Keying (BT=0.3)

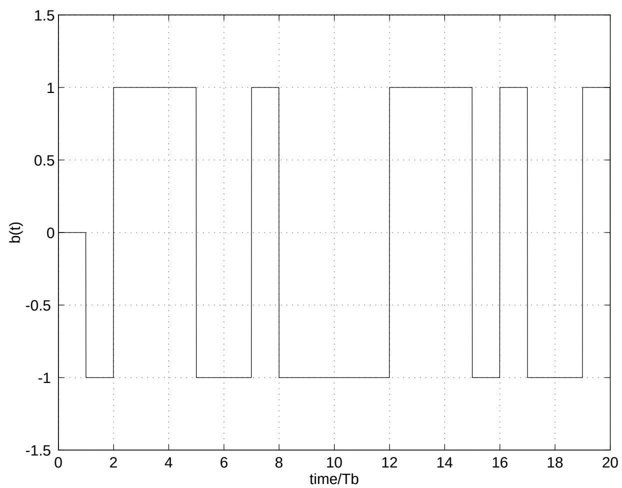


Figure 67: Waveform for Gaussian Minimum Shift Keying (BT=0.3)

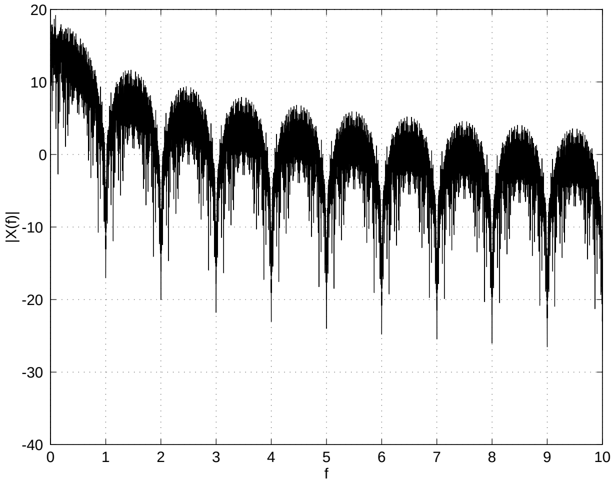


Figure 68: Waveform for Gaussian Minimum Shift Keying (BT=0.3)

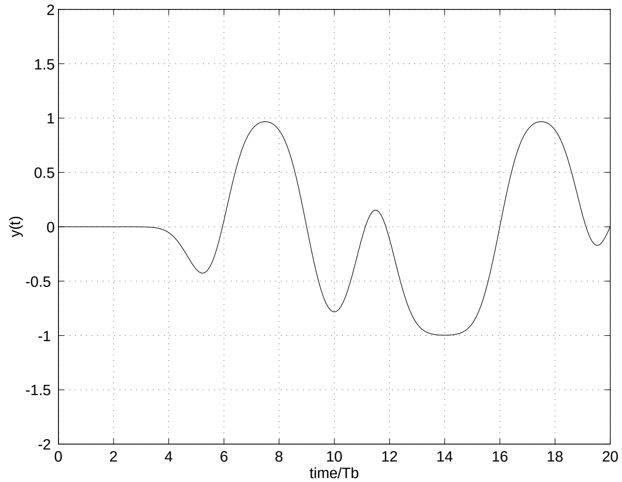


Figure 69: Waveform for Gaussian Minimum Shift Keying (BT=0.3)

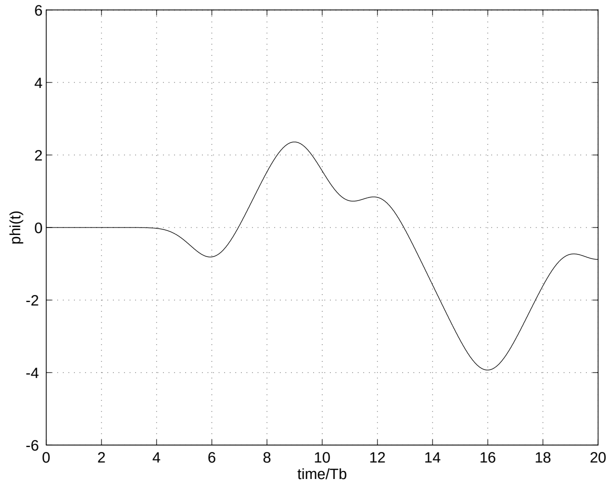


Figure 70: Waveform for Gaussian Minimum Shift Keying (BT=0.3)

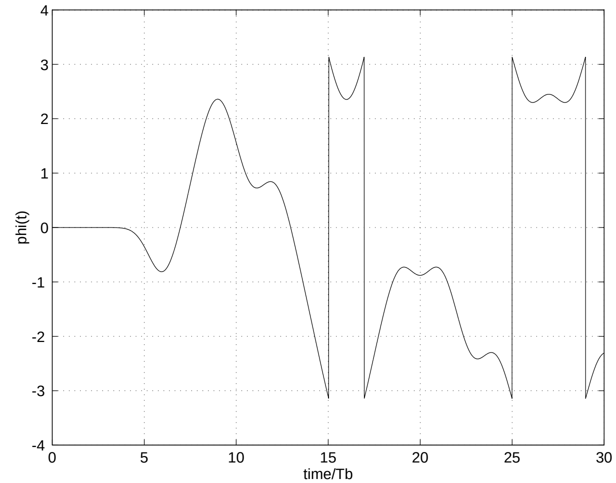


Figure 71: Waveform for Gaussian Minimum Shift Keying (BT=0.3)

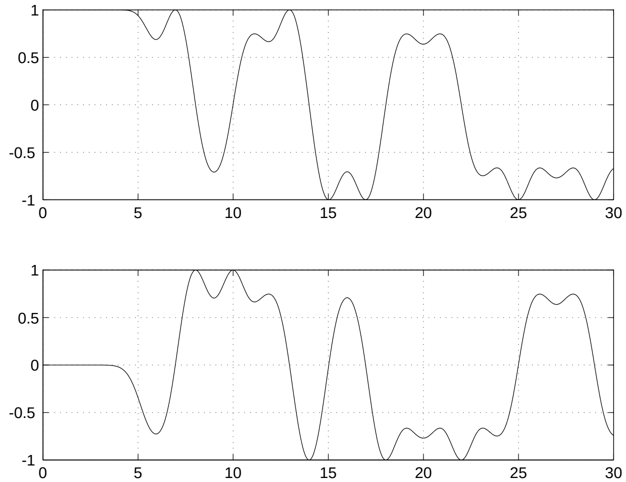


Figure 72: Waveform for Gaussian Minimum Shift Keying (BT=0.3)

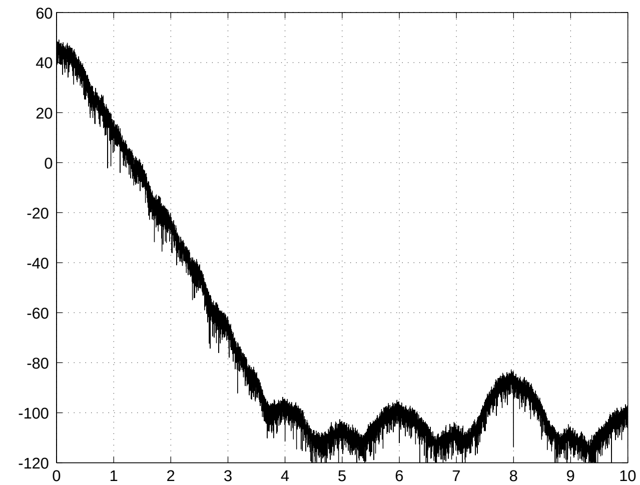


Figure 73: Waveform for Gaussian Minimum Shift Keying (BT=0.3)

$\pi/4$ QPSK

As mentioned earlier the effect of filtering and nonlinearly amplifying a QPSK waveform causes distortion when the signal amplitude fluctuates significantly. Another modulation scheme that has less fluctuation that QPSK is $\pi/4$ QPSK. In this modulation scheme every other symbol is sent using a rotated (by 45 degrees) constellation. Thus the transitions from one phase to the next are still instantaneous (without any filtering) but the signal never makes a transition through the origin. Only ± 45 and ± 135 degree transitions are possible. This is shown in the constellation below where a little bit of filtering was done.

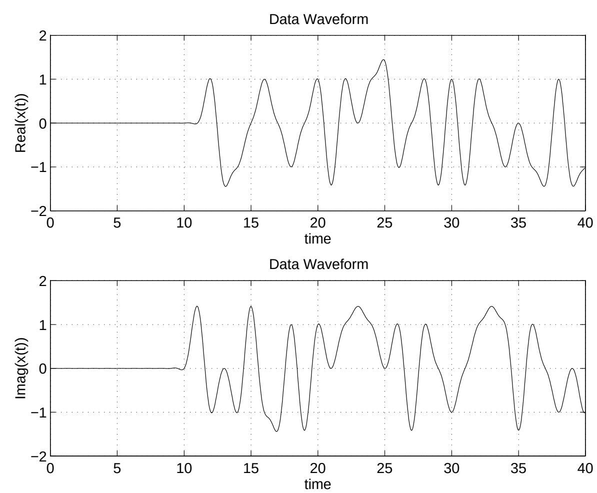


Figure 74: Data Waveforms for $\pi/4$ QPSK

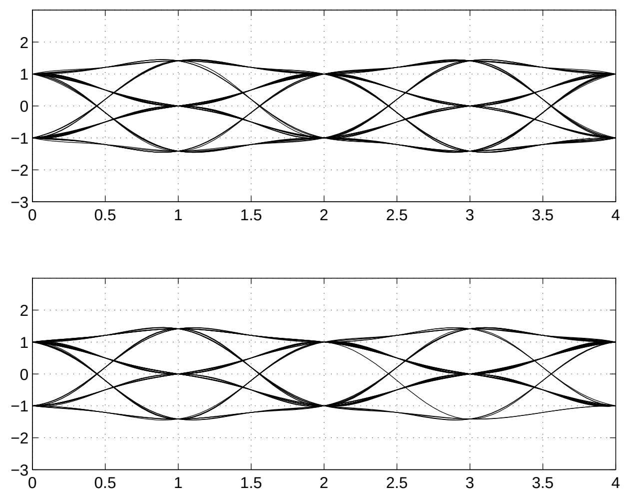


Figure 75: Eye Diagram for $\pi/4$ QPSK

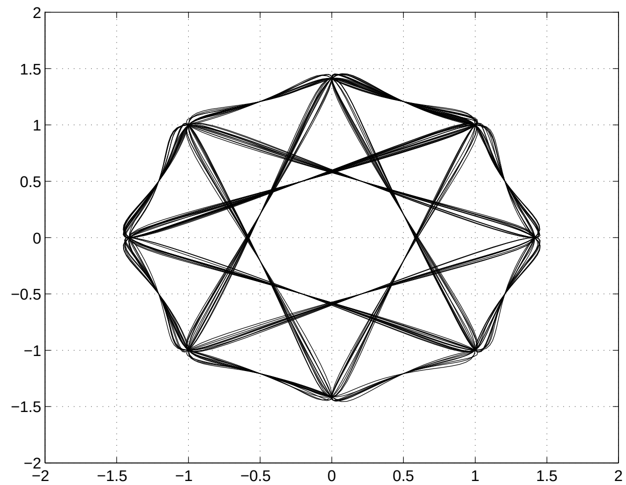


Figure 76: Constellation for $\pi/4$ QPSK.

Lecture 6b: Other Modulation Techniques

Orthogonal Signals

A set of signals $\{\psi_i(t) : 0 \le t \le T, 0 \le i \le M-1\}$ are said to be orthogonal (over the interval $[0,T]$) if

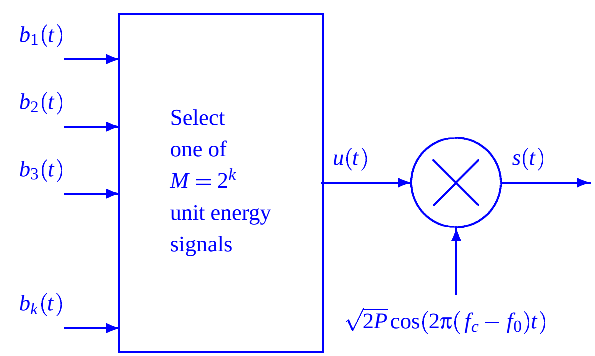
$$\int_0^T \psi_i(t)\psi_j(t)dt = 0, \quad i \neq j.$$

In most cases the signals will have the same energy and it is convenient to normalize the signals to unit energy. A set of signals $\{\psi_i(t) : 0 \le t \le T, 1 \le i \le M\}$ are said to be orthonormal (over the interval $[0,T]$) if

$$\int_0^T \psi_i(t)\psi_j(t)dt = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

Many signal sets can be described as linear combinations of orthonormal signal sets as we will show later. Below we describe a number of different orthonormal signal sets. The signal sets will all be described at some intermediate frequency f_0 but are typically modulated up to the carrier frequency f_c.

General Modulator



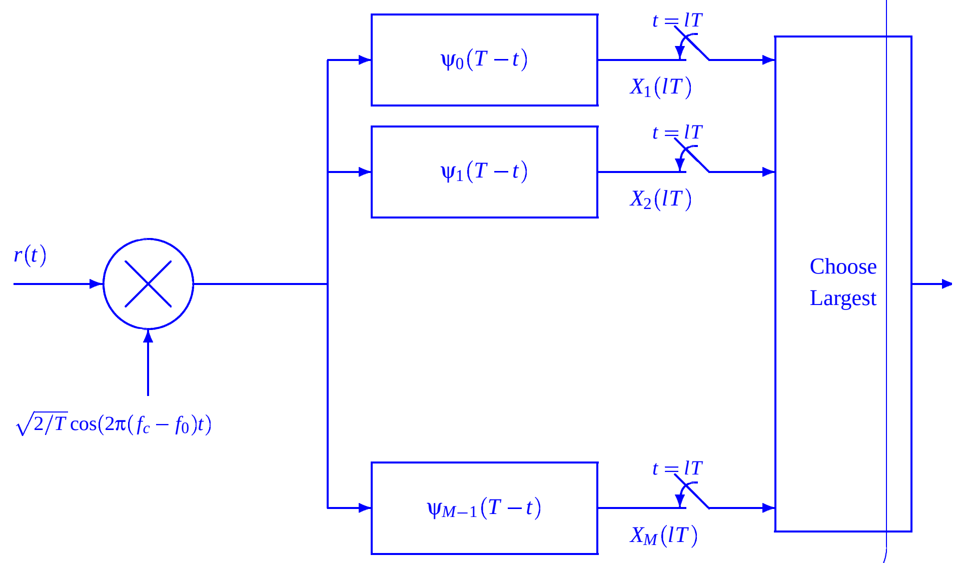
$$b_i(t) = \sum_{l=-\infty}^{\infty} b_l p_T(t - lT), i = 1, 2, \ldots, k$$

$$u(t) = \sum_{l=-\infty}^{\infty} \psi_{i_l}(t - lT)$$

where for $(l-1)T \leq t < T$

$$i_l = \begin{cases} 1, & b_1(t) = b_2(t) = \ldots = b_{k-1}(t) = b_k(t) = +1 \\ 2, & b_1(t) = b_2(t) = \ldots = b_{k-1}(t) = +1, b_k(t) = -1 \\ & \\ M, & b_1(t) = b_2(t) = \ldots = b_{k-1}(t) = b_k(t) = -1 \end{cases}$$

General Coherent Demodulator



$\psi_m(T-t)$ is the impulse response of the m-th matched filter. The output of these filters (assuming that the i_l-th orthogonal signal is transmitted is) given by

$$X_m(lT) = \begin{cases} \eta_m, & m \neq i_l \\ \sqrt{E} + \eta_m, & m = i_l \end{cases}$$

where $\{\eta_m, m = 0, 1, 2, \ldots, M-1\}$ is a sequence of independent, identically distributed Gaussian random variables with mean zero and variance $N_0/2$.

To determine the probability of error we need to determine the probability that the filter output corresponding to the signal present is smaller than one of the other filter outputs.

The symbol error probability of M orthogonal signals with coherent demodulation is given by

$$P_{e,s} = \frac{M-1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Phi(u - \sqrt{\frac{2E}{N_0}}) \Phi^{M-2}(u) e^{-u^2/2} du$$

where $\Phi(u)$ is the distribution function of a zero mean, variance 1, Gaussian random variable given by

$$\Phi(u) = \frac{1}{2\pi} \int_{-\infty}^{u} e^{-x^2/2} dx.$$

The symbol error probability can be upper bounded as

$$P_{e,s} \leq \begin{cases} 1, & \frac{E}{N_0} \leq \ln M \\ \exp\left\{-\left(\sqrt{\frac{E}{N_0}} - \sqrt{\ln M}\right)^2\right\}, & \ln M \leq \frac{E}{N_0} \leq 4 \ln M \\ \exp\{-(\frac{E}{2N_0} - \ln M)\}, & \frac{E}{N_0} \geq 4 \ln M. \end{cases}$$

Normally a communication engineer is more concerned with the energy transmitted per bit rather than the energy transmitted per signal, E. If we let E_b be the energy transmitted per bit then these are related as follows

$$E_b = \frac{E}{\log_2 M}.$$

Thus the bound on the symbol error probability can be expressed in terms of the energy transmitted per bit as

$$P_{e,s} \leq \begin{cases} 1, & \frac{E_b}{N_0} \leq \ln 2 \\ \exp_2\left\{-\log_2 M \left(\sqrt{\frac{E_b}{N_0}} - \sqrt{\ln 2}\right)^2\right\}, & \ln 2 \leq \frac{E_b}{N_0} \leq 4 \ln 2 \\ \exp_2\left\{-\log_2 M \left(\frac{E_b}{2N_0} - \ln 2\right)\right\}, & \frac{E_b}{N_0} \geq 4 \ln 2 \end{cases}$$

where $\exp_2\{x\}$ denotes 2^x. Note that as $M \to \infty$, $P_e \to 0$ if $\frac{E_b}{N_0} > \ln 2$ = -1.59dB.

So far we have examined the symbol error probability for orthogonal signals. Usually the number of such signals is a power of 2, e.g. 4, 8, 16, 32, ... If so then each transmission of a signal is carrying $k = \log_2 M$ bits of information. In this case a communication engineer is usually interested in the bit error probability as opposed to the symbol error probability. These can be related for any equidistant, equienergy signal set (such as orthogonal or simplex signal sets) by

$$P_{e,b} = \frac{2^{k-1}}{2^k - 1} P_{e,s} = \frac{M}{2(M-1)} P_{e,s}.$$

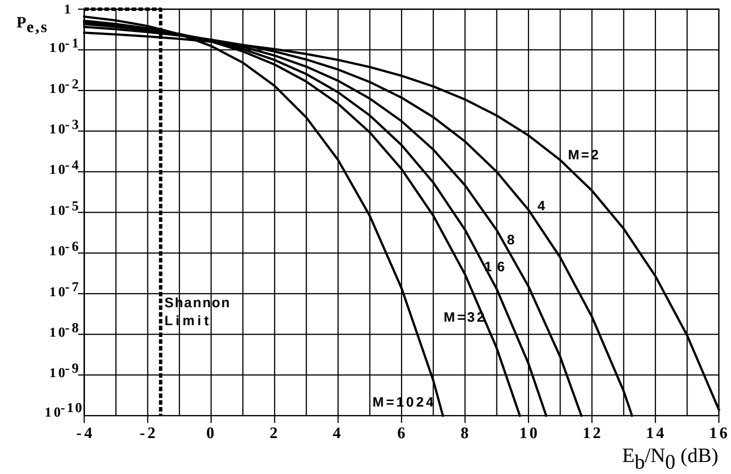


Figure 77: Symbol Error Probability for Coherent Demodulation of Orthogonal Signals

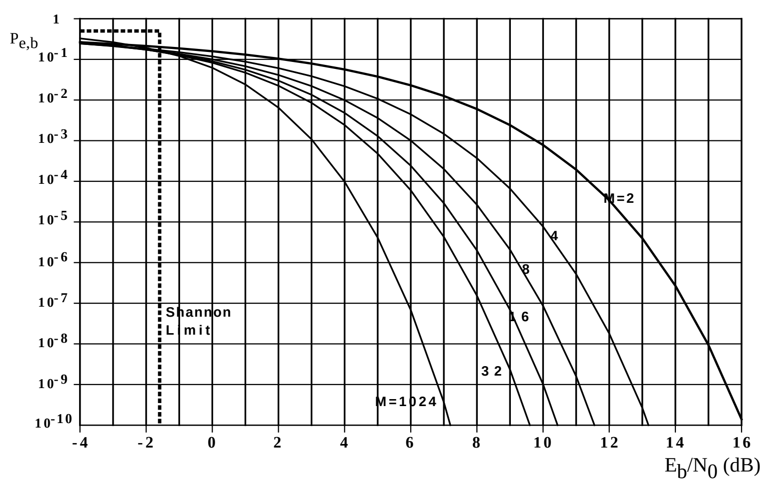
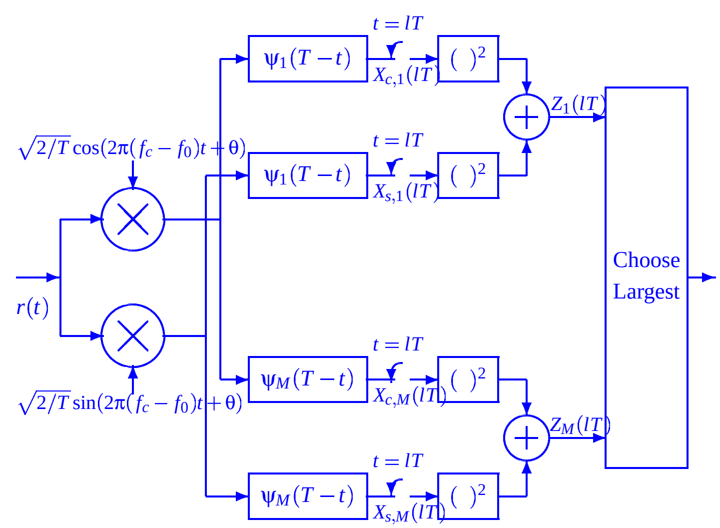


Figure 78: Bit Error Probability for Coherent Demodulation of Orthogonal Signals

General Noncoherent Demodulator



If signal 1 is transmitted during the interval $[(l-1)T, lT)$ then

$$X_{c,m}(lT) = \begin{cases} \sqrt{E}\cos(\theta) + \eta_{c,1}, & m = 1 \\ \eta_{c,m}, & m \neq 1 \end{cases}$$

$$X_{s,m}(lT) = \begin{cases} \sqrt{E}\sin(\theta) + \eta_{s,1}, & m = 1 \\ \eta_{s,m}, & m \neq 1 \end{cases}$$

The decision statistic then (if signal 1 is transmitted) has the form

$$\begin{aligned} Z_1(lT) &= E + 2\sqrt{E}(\eta_{c,1}\cos(\theta) + \eta_{s,1}\sin(\theta)) + \eta_{c,1}^2 + \eta_{s,1}^2, \\ Z_2(lT) &= \eta_{c,2}^2 + \eta_{s,2}^2 \\ Z_3(lT) &= \eta_{c,3}^2 + \eta_{s,3}^2 \\ &\vdots \\ Z_M(lT) &= \eta_{c,M}^2 + \eta_{s,M}^2 \end{aligned}$$

The symbol error probability for noncoherently detection of orthogonal signals is

$$P_{e,s} = \frac{1}{M} e^{\{-E_b/(\log_2 M N_0)\}} \sum_{m=2}^{M} (-1)^m \binom{M}{m} e^{\{E_b/(m \log_2 M N_0)\}}$$

As with coherent demodulation the relation between bit error probability and symbol error probability for noncoherent demodulation of orthogonal signals is

$$P_{e,b} = \frac{2^{k-1}}{2^k - 1} P_{e,s} = \frac{M}{2(M-1)} P_{e,s}.$$

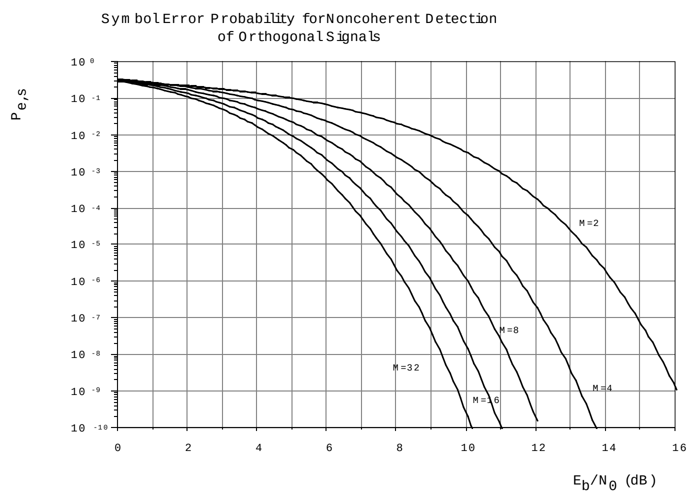


Figure 79: Symbol Error Probability for Nonocherent Detection of Orthogonal Signals.

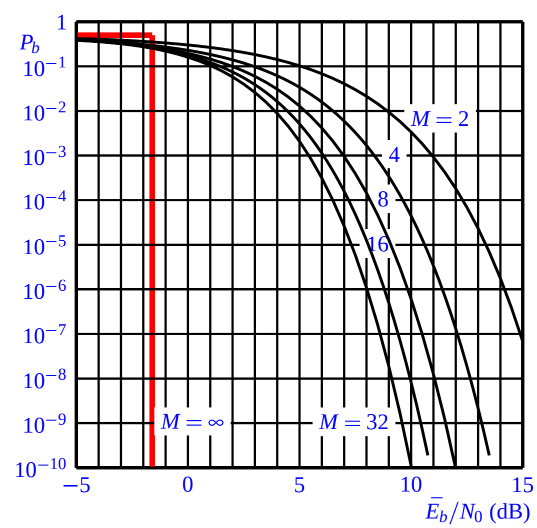


Figure 80: Bit error probability of M-ary orthogonal modulation in an additive white Gaussian noise channel with noncoherent demodulation

A. Time-orthogonal (Pulse position modulation PPM)

$$\psi_i(t) = \begin{cases} \sqrt{\frac{2M}{T}}\sin(2\pi f_0 t), & iT/M \le t < (i+1)T/M \\ 0, & \text{elsewhere} \end{cases}$$

$$i = 0,1,\ldots,M-1, \quad f_0 = n\frac{M}{2T}$$

$$\int_{iT/M}^{(i+1)T/M} \sqrt{\frac{2M}{T}}\sin(2\pi f_0 t)\sqrt{\frac{2M}{T}}\sin(2\pi f_1 t)dt = 0$$

$$f_0 = n\frac{M}{2T}, \quad f_1 = (n+1)\frac{M}{2T}.$$

B. Time-orthogonal quadrature-phase

$$\psi_{2i}(t) = \begin{cases} \sqrt{\frac{2M}{T}}\sin(2\pi f_0 t), & \frac{2iT}{M} \le t < 2(i+1)T/M \\ 0, & \text{elsewhere} \end{cases}$$

$$\psi_{2i+1}(t) = \begin{cases} \sqrt{\frac{2M}{T}}\cos(2\pi f_0 t), & \frac{2iT}{M} \le t < 2(i+1)\frac{T}{M} \\ 0 & \text{elsewhere} \end{cases}$$

$$i = 0,1,\ldots,\frac{M}{2}-1, \quad M \text{ even}, \quad f_0 = n\frac{M}{2T}$$

C. Frequency-orthogonal (Frequency Shift Keying FSK)

$$\psi_i(t) = \sqrt{\frac{2E}{T}}\sin[2\pi(f_0 + \frac{i}{2T}t)], \qquad 0 \le t \le T$$

$$i = 0,1,\ldots,M-1, \quad f_0 = \frac{nM}{2T}.$$

D. Frequency-orthogonal quadrature-phase

$$\psi_{2i}(t) = \sqrt{\frac{2E}{T}}\sin[2\pi(f_0 + \frac{i}{T}),t] \qquad 0 \le t < T$$

$$\psi_{2i+1}(t) = \sqrt{\frac{2E}{T}}\cos[2\pi(f_0 + \frac{i}{T})t], \qquad 0 \le t \le T$$

$$f_0 = \frac{nM}{2T}.$$

E. Hadamard-Walsh Construction

The last construction of orthogonal signals is done via the Hadamard Matrix. The Hadamard matrix is an N by N matrix with components either +1 or -1 such that every pair of distinct rows are orthogonal. We show how to construct a Hadamard when the number of signals is a power of 2 (which is often the case).

Begin with a two by two matrix

$$H_2 = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix}.$$

Then use the recursion

$$H_{2^l} = \begin{bmatrix} +H_{2^{(l-1)}} & +H_{2^{(l-1)}} \\ +H_{2^{(l-1)}} & -H_{2^{(l-1)}} \end{bmatrix}.$$

Now it is easy to check that distinct rows in these matrices are orthogonal. The i-th modulated signal is then obtained by using a single (arbitrary) waveform N times in nonoverlapping time intervals and multiplying by the $j-th$ repetition of the waveform by the jth component of the i-th row of the matrix.

Example ($M = 4$):

$$\begin{aligned} H_4 &= \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} \\ &= \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix}. \end{aligned}$$

Example ($M = 8$):

$$\begin{aligned} H_8 &= \begin{bmatrix} H_4 & H_4 \\ H_4 & -H_4 \end{bmatrix} \\ &= \begin{bmatrix} H_2 & H_2 & H_2 & H_2 \\ H_2 & -H_2 & H_2 & -H_2 \\ H_2 & H_2 & -H_2 & -H_2 \\ H_2 & -H_2 & -H_2 & H_2 \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 & +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 \\ +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ +1 & -1 & -1 & +1 & -1 & +1 & +1 & -1 \end{bmatrix}.$$

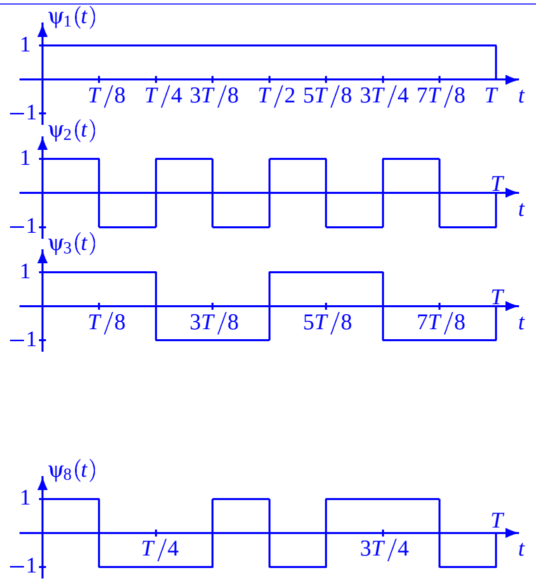


Figure 81: Hadamard-Walsh Orthogonal Signals

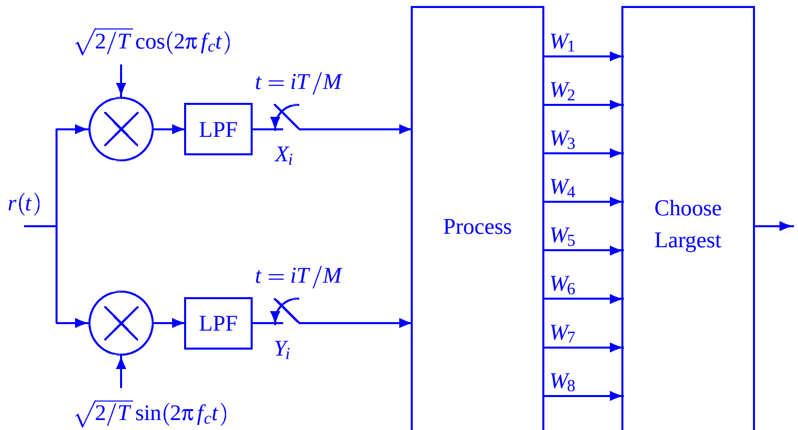


Figure 82: Noncoherent Demodulator

Noncoherent Reception of Hadamard Generated Orthogonal Signals

$$
\begin{aligned}
W_1 &= (X_1+X_2+X_3+X_4+X_5+X_6+X_7+X_8)^2 \\
&\quad +(Y_1+Y_2+Y_3+Y_4+Y_5+Y_6+Y_7+Y_8)^2 \\
W_2 &= (X_1-X_2+X_3-X_4+X_5-X_6+X_7-X_8)^2 \\
&\quad +(Y_1-Y_2+Y_3-Y_4+Y_5-Y_6+Y_7-Y_8)^2 \\
W_3 &= (X_1+X_2-X_3-X_4+X_5+X_6-X_7-X_8)^2 \\
&\quad +(Y_1+Y_2-Y_3-Y_4+Y_5+Y_6-Y_7-Y_8)^2 \\
W_4 &= (X_1-X_2-X_3+X_4+X_5-X_6-X_7+X_8)^2 \\
&\quad +(Y_1-Y_2-Y_3+Y_4+Y_5-Y_6-Y_7+Y_8)^2
\end{aligned}
$$

Noncoherent Reception of Hadamard Generated Orthogonal Signals

$$
\begin{aligned}
W_5 &= (X_1+X_2+X_3+X_4-X_5-X_6-X_7-X_8)^2 \\
&\quad +(Y_1+Y_2+Y_3+Y_4-Y_5-Y_6-Y_7-Y_8)^2 \\
W_6 &= (X_1-X_2+X_3-X_4-X_5+X_6-X_7+X_8)^2 \\
&\quad +(Y_1-Y_2+Y_3-Y_4-Y_5+Y_6-Y_7+Y_8)^2 \\
W_7 &= (X_1+X_2-X_3-X_4-X_5-X_6+X_7+X_8)^2 \\
&\quad +(Y_1+Y_2-Y_3-Y_4-Y_5-Y_6+Y_7+Y_8)^2 \\
W_8 &= (X_1-X_2-X_3+X_4-X_5+X_6+X_7-X_8)^2 \\
&\quad +(Y_1-Y_2-Y_3+Y_4-Y_5+Y_6+Y_7-Y_8)^2
\end{aligned}
$$

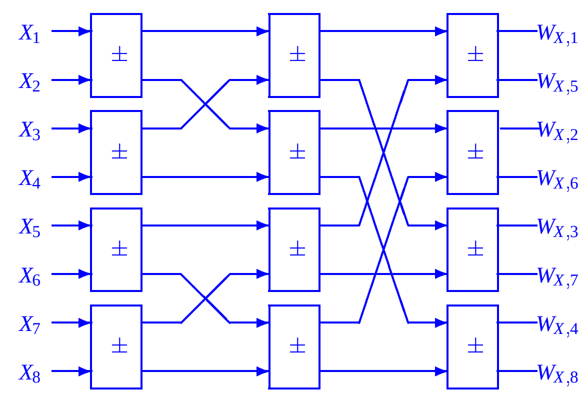


Figure 83: Fast Processing for Hadamard Signals

If we define bandwidth of M signals as minimum frequency separation between two such signal sets such that any signal from one signal set is orthogonal to every signal from a frequency adjacent signal set are orthogonal then for all of these examples of M signals the bandwidth is

$$W = \frac{M}{2T} \Rightarrow M = 2WT!$$

Thus there are $2WT$ orthogonal signals in bandwidth W and time duration T.

Biorthogonal Signal Set

A biorthogonal signal set can be described as

$$\begin{aligned}
s_0(t) &= \sqrt{E}\phi_0(t) \\
s_1(t) &= \sqrt{E}\phi_1(t) \\
&\cdot \quad \cdot \quad \cdot \\
s_{M/2-1}(t) &= \sqrt{E}\phi_{M/2-1}(t) \\
s_{M/2}(t) &= -\sqrt{E}\phi_0(t) \\
&\cdot \quad \cdot \quad \cdot \\
s_{M-1}(t) &= -\sqrt{E}\phi_{M/2-1}(t)
\end{aligned}$$

That is a biorthogonal signal set is the same as orthogonal signal set except that the negative of each orthonormal signal is also allowed.. Thus there are $2N$ signals in N dimensions. We have doubled the number of signals without changing the minimum Euclidean distance of the

signal set. For example:

$$B_8 = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \\ -1 & -1 & -1 & -1 \\ -1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 \\ -1 & +1 & +1 & -1 \end{bmatrix}$$

Let H_j be the hypothesis that signal s_j was sent for $j = 0, \ldots, M-1$. The optimal receiver does a correlation of the received signal with each of the $M/2$ orthonormal signals. Let r_j be the correlation of $r(t)$ with $\phi_j(t)$. The decision rule is to choose hypothesis H_j if r_j is largest in absolute value and is of the appropriate sign. That is, if r_j is larger than $|r_i|$ and is the same sign as the coefficient in the representation of $s_j(t)$.

Symbol Error Probability

Let H_j be the hypothesis that signal s_j was sent for $j = 0, \ldots, M-1$. The probability of correct is (given signal s_0 sent)

$$\begin{aligned}
P_{c,0} &= P\{r_0 > 0, |r_1| < r_0, \ldots, |r_{M/2-1}| < r_0 | H_0\} \\
&= \int_{r_0=0}^{\infty} f_s(r_0)[F_n(r_0) - F_n(-r_0)]^{M/2-1} dr_0
\end{aligned}$$

where $f_s(x)$ is the denisty function of r_0 when H_0 is true and $F_n(x)$ is the distribution of r_1 when H_0 is true.

$$\begin{aligned}
f_s(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\{-\frac{1}{2\sigma^2}(x-\sqrt{E})^2\} \\
F_s(x) &= \Phi(\frac{x-\sqrt{E}}{\sigma}) \\
f_n(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\{-\frac{1}{2\sigma^2}(x)^2\} \\
F_n(x) &= \Phi(\frac{x}{\sigma})
\end{aligned}$$

where $\sigma^2 = N_0/2$. The error probability is then

$$P_{e,0} = 1 - \int_{r_0=0}^{\infty} f_s(r_0)[F_n(r_0) - F_n(-r_0)]^{M/2-1} dr_0$$

Using an integration by parts argument we can write this as

$$P_{e,s} = (M-2)\int_0^{\infty} \Phi(z - \sqrt{\frac{2E}{N_0}})[2\Phi(z)-1]^{M/2-2}\frac{1}{\sqrt{2\pi}}\exp\{-\frac{z^2}{2}\}dz$$

Bit Error Probability

The bit error probabiltiy for birothogonal signals can be determined for the usual mapping of bits to symbols. The mapping is given as

$000000\cdots000$	$s_0(t)$
$000000\cdots001$	$s_1(t)$
$011111\cdots111$	$s_{M/2-1}(t)$
$111111\cdots111$	$s_{M/2}(t) = -s_0(t)$
$111111\cdots110$	$s_{M/2+1}(t) = -s_1(t)$
$100000\cdots000$	$s_{M-1}(t) = -s_{M/2-1}(t)$.

The mapping is such that signals with furthest distance have largest number of bit errors. An error of the first kind is defined to be an error to an orthogonal signal, while an error of the second kind is an error to the antipodal signal. The probability of error of the first kind is the probability that H_j is chosen given that s_0 is transmitted ($j < M/2$) and is given by

$$P_{e,1} = P\{r_j > |r_0|, r_j > |r_1|, \ldots, r_j > |r_{M/2-1}|, r_j > 0 | H_0\}$$

$$= \int_0^{\infty}[F_s(r_j) - F_s(-r_j)][F_n(r_j) - F_n(-r_j)]^{M/2-2} f_n(r_j) dr_j$$

It should be obvious that this is also the error probability to H_j for $j > M/2$. The probability of error of the second kind is the probability that $H_{M/2}$ is chosen given that s_0 is transmitted and is given by

$$\begin{aligned} P_{e,2} &= P\{r_0 < 0, |r_1| < |r_0|, |r_2| < |r_0|, \ldots, |r_{M/2-1}| < |r_0| | H_0\} \\ &= \int_{-\infty}^{0} f_s(r_0)[F_n(r_0) - F_n(-r_0)]^{M/2-1} dr_0 \\ &= (M-2)\int_0^{\infty} F_s(-r_0)[F_n(r_0) - F_n(-r_0)]^{M/2-2} f_n(r_0) dr_0 \end{aligned}$$

The bit error probability is determined by realizing that of the $M-2$ possible errors (all equally likely) of the first kind, $(M-2)/2$ of them result in a particular bit in error while an error of the second kind causes all the bits to be in error. Thus

$$\begin{aligned} P_{e,b} &= \frac{M-2}{2} P_{e,1} + P_{e,2} \\ &= \frac{(M-2)}{2}\int_0^{\infty}[F_s(u) + F_s(-u)][F_n(u) - F_n(-u)]^{M/2-2} f_n(u) du \end{aligned}$$

$$= \frac{(M-2)}{2}\int_0^{\infty}[\Phi(z - \sqrt{\frac{2E}{N_0}}) + \Phi(-z - \sqrt{\frac{2E}{N_0}})][2\Phi(z)-1]^{M/2-2}\frac{1}{\sqrt{2\pi}}\exp\{-\frac{z^2}{2}\}dz$$

Notice that the symbol error probability is $P_{e,s} = (M-2)P_{e,1} + P_{e,2}$.

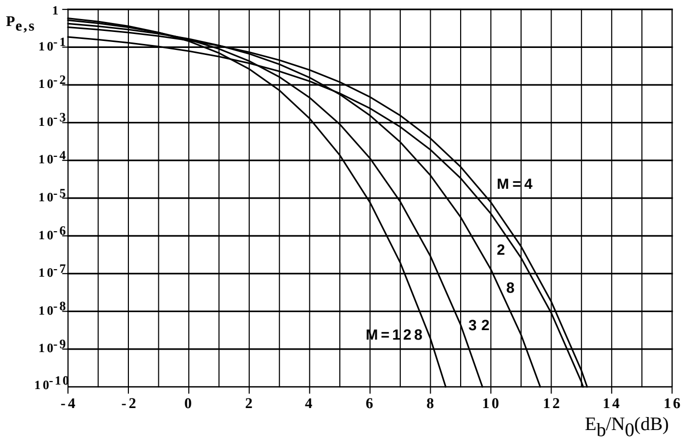


Figure 84: Symbol Error Probability for Coherent Demodulation of Biorthogonal Signals

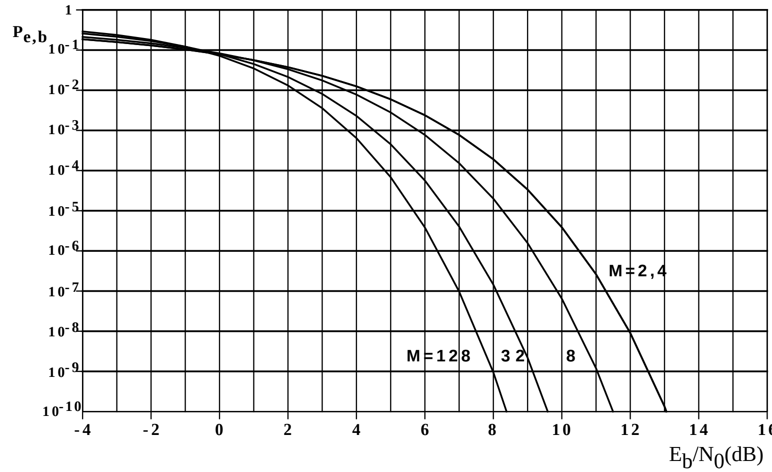


Figure 85: Bit Error Probability for Coherent Demodulation of Biorthogonal Signals

Simplex Signal Set

Same as orthogonal except subtract from each of the signals the average signal of the set, i.e.

$$s'_i(t) = s_i(t) - \frac{1}{M}\sum_{i=0}^{M-1} s_i(t),\ i = 0, 1, \ldots, M-1$$

When the orthogonal set is constructed via a Hadamard matrix this amounts to deleting the first component in the matrix since the other components sum to zero.

For example

$$S_8 = \begin{bmatrix} +1 & +1 & +1 & +1 & +1 & +1 & +1 \\ -1 & +1 & -1 & +1 & -1 & +1 & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 \\ -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ +1 & +1 & +1 & -1 & -1 & -1 & -1 \\ -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ -1 & -1 & +1 & -1 & +1 & +1 & -1 \end{bmatrix}$$

These are slightly more efficient than orthogonal signals.

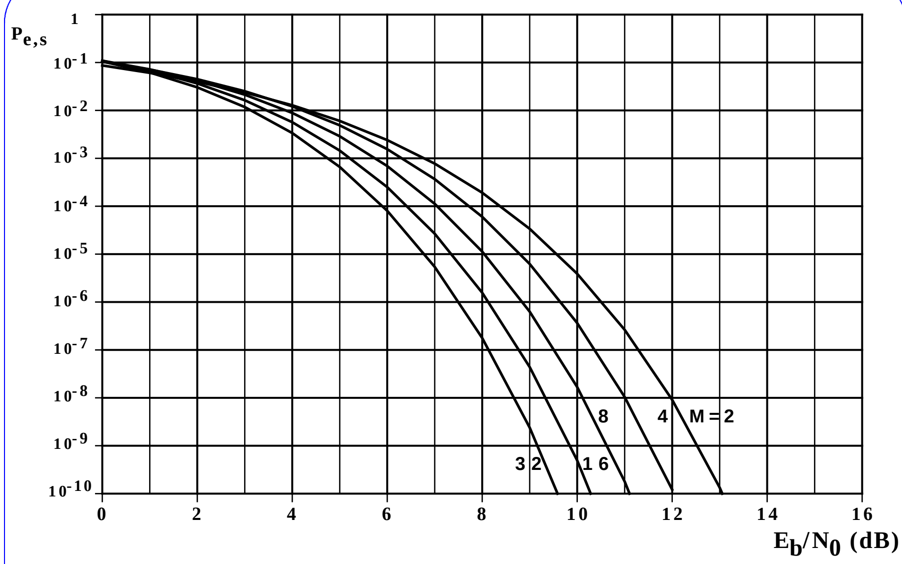


Figure 86: Symbol Error Probability for Simplex Signalling

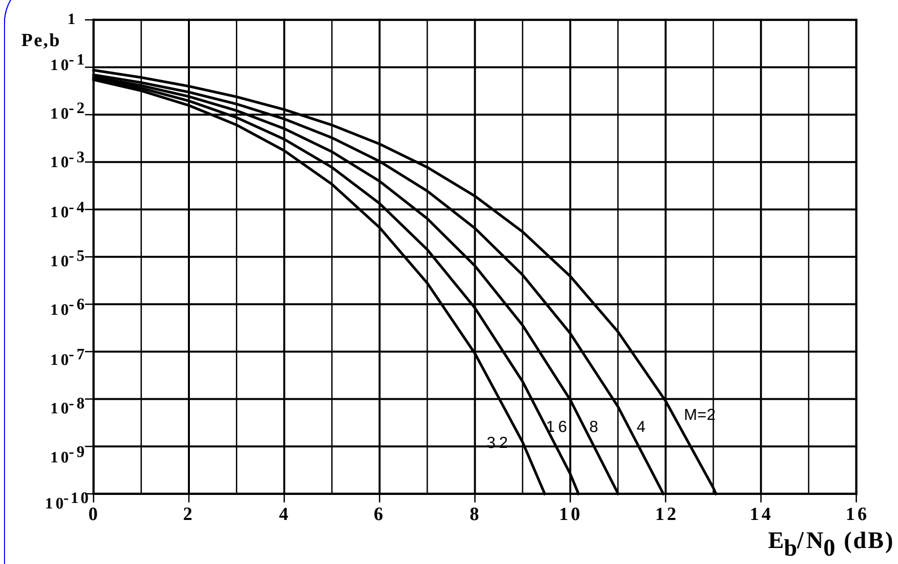


Figure 87: Bit Error Probability for Simplex Signalling

Multiphase Shift Keying (MPSK)

$$\begin{aligned} s_i(t) &= A\cos\left[2\pi f_0 t + \frac{2\pi}{M} i + \lambda\right] \quad 0 \le t \le T \\ &= A_{c,i}\cos 2\pi f_c t - A_{s,i}\sin 2\pi f_c t \end{aligned}$$

where for $i = 0, 1, \ldots, M-1$,

$$\begin{aligned} A_{c,i} &= A\cos\left(\frac{2\pi i}{M} + \lambda\right) \\ A_{s,i} &= A\sin\left(\frac{2\pi i}{M} + \lambda\right) \end{aligned}$$

$$P_{e,s} = 1 - \int_{-\pi/M}^{\pi/M} \frac{e^{-E/N_0}}{2\pi}\left[1 + \sqrt{\frac{4\pi E}{N_0}}\cos\theta e^{\gamma\cos^2\theta}\left(1 - Q\left(\sqrt{\frac{2E}{N_0}}\cos\theta\right)\right)\right] d\theta$$

For this modulation scheme we should use Gray coding to map bits into signals.

$$M = 2 \Rightarrow \text{BPSK} \qquad M = 4 \Rightarrow \text{QPSK}$$

This type of modulation has the properties that all signals have the same power thus the use of nonlinear amplifiers (class C amplifiers) affects each signal in the same manner. Furthermore if we are restricted to two dimensions and every signal must have the same power than this signal set minimizes the error probability of all such signal sets.

(QPSK and BPSK are special cases of this modulation).

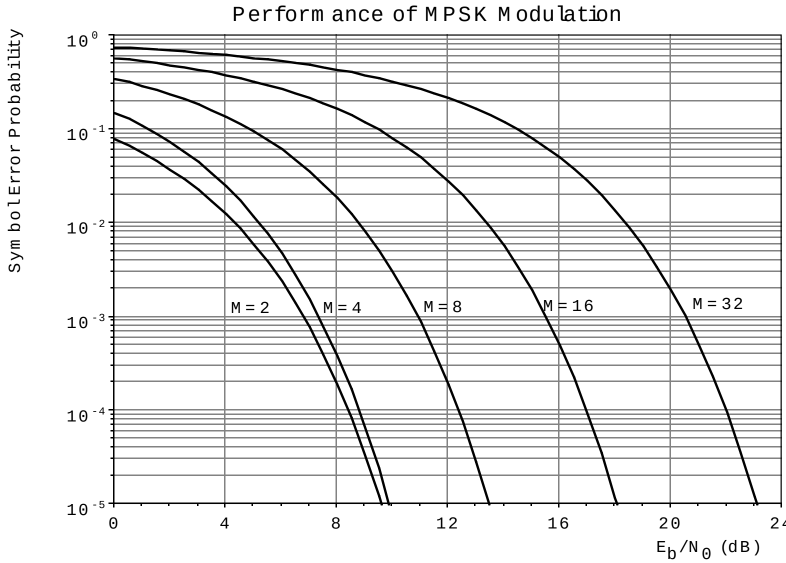


Figure 88: Symbol Error Probability for MPSK Signalling

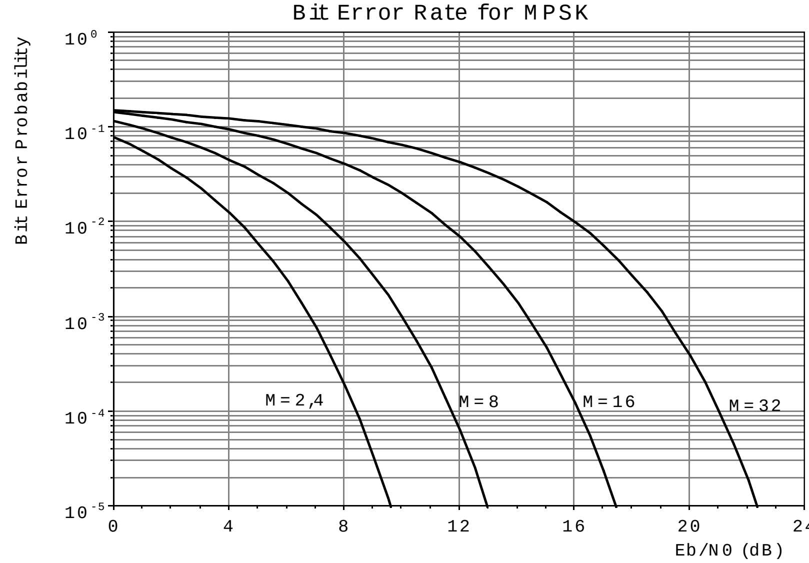


Figure 89: Bit Error Probability for MPSK Signalling

M-ary Pulse Amplitude Modulation (PAM)

$$s_i(t) = A_i\, s(t), \qquad 0 \le t \le T$$

where

$$A_i = (2i+1-M)A \qquad i = 0,1,\ldots,M-1$$

$$E_i = A_i^2$$

$$\begin{aligned}\overline{E} &= \frac{1}{M}\sum_{i=0}^{M-1} E_i = \frac{A^2}{M}\sum_{i=0}^{M-1}(2i+1-M)^2 \\ &= \frac{M^2-1}{3}A^2 \\ P_{e,s} &= \left(\frac{2(M-1)}{M}\right)\, Q\left(\sqrt{\frac{6\overline{E}}{(M^2-1)N_0}}\right)\end{aligned}$$

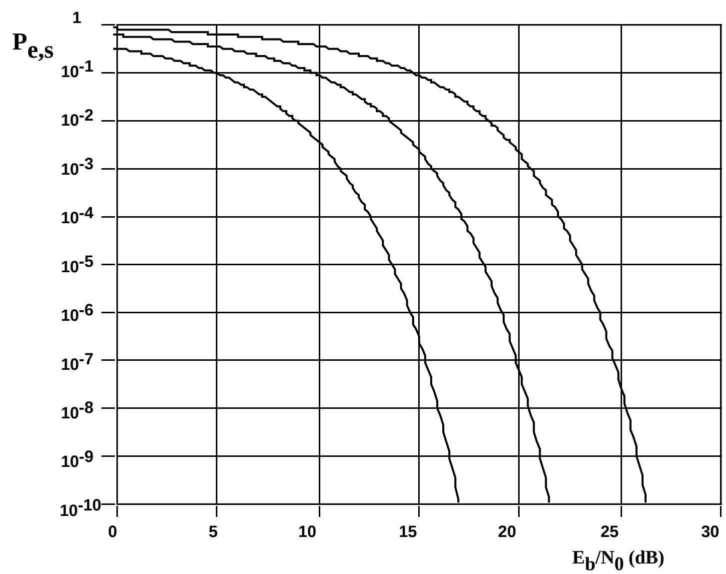


Figure 90: Symbol Error Probability for MPAM Signalling

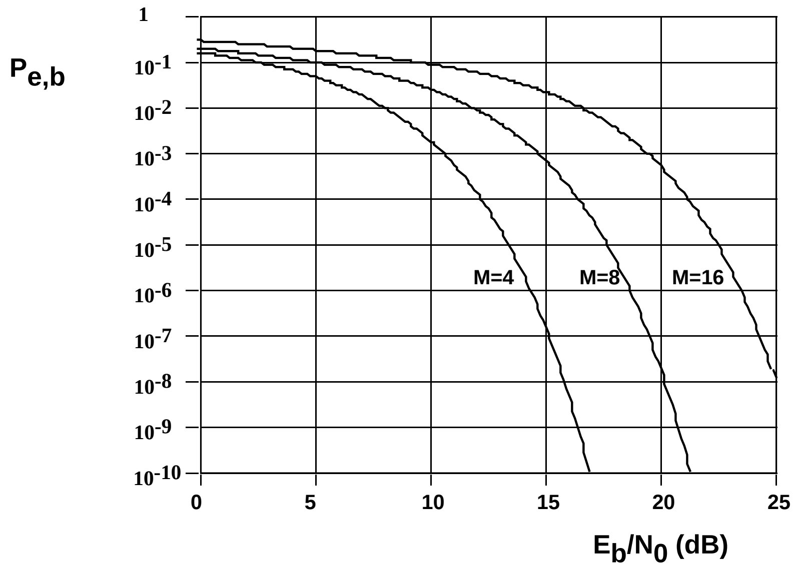


Figure 91: Bit Error Probability for MPAM Signalling

Quadrature Amplitude Modulation

For $i = 0, \ldots, M-1$

$$s_i(t) \quad = \quad A_i \cos 2\pi f_c t + B_i \sin 2\pi f_c t \qquad 0 \le t \le T$$

Since this is two PAM systems in quadrature. $P_{e,2} = 1 - (1 - P_{e,1})^2$ for PAM with $\sqrt{M}$ signals

Bandwidth of Digital Signals:

In practice a set of signals is not used once but in a periodic fashion. If a source produces symbols every T seconds from the alphabet $A = 0, 1, \ldots, M-1$ with b_e representing the l^{th} letter $-\infty \le l < \infty$ then the digital data signal has the form

$$s(t) = \sum_{l=-\infty}^{\infty} s_{b_n}(t - nT)$$

Note: 1) $s_i(t)$ need not be time limited to $[0,T]$. In fact we may design $\{s_i(t)\}_{i=0}^{M-1}$ so that $s_i(t)$ is not time limited to $[0,T]$. If $s_i(t)$ is not time limited to $[0,T]$ then we may have intersymbol interference in the demodulaton. The reason for introducing intersymbol interference is to "shape" the spectral characteristic of the signal (e.g. if ther are nonlinear amplifiers or other nonlinearities in the communication system).

2) The random variables b_n need not be a sequence of i.i.d. random variables. In fact if we are using error-correcting codes there will be some redundancy in b_2 so that it is not a sequence of i.i.d. r.v.

In many of the modulation schemes (the linear ones) considered we can equivalently write the signal as

$$s(t) = \mathrm{Re}[u(t) e^{j\omega_c t}]$$

where $u(t)$ is called the lowpass signal. For general CPM the modulation is nonlinear so that the below does not apply. Also

$$u(t) = \sum_{n=-\infty}^{\infty} I_n g(t - nT)$$

where I_n is possibly complex and $g(t)$ is an arbitrary pulse shape.

Note that while $u(t)$ is a (non stationary) random process $u(t+\hat{\tau})$ where $\hat{\tau}$ is uniform r.v. on $[0,T]$ is stationary.

$$\begin{aligned} \Phi_u(f) &\triangleq \mathcal{F}\{E[u^*(t+\hat{\tau})u(t+\hat{\tau}+\tau)]\} \\ &= \frac{1}{T}\Phi_I(f)|G(f)|^2 \end{aligned}$$

where

$$\begin{aligned} \Phi_I(f) &= \sum_{m=-\infty}^{\infty} E[I_n^* I_{n+m}] e^{-j2\pi f m T} \\ G(f) &= \mathcal{F}[g(t)] = \int_{-\infty}^{\infty} g(t) e^{-j2\pi f t} dt \end{aligned}$$

Example: BPSK $I_n = \pm 1$ (i.i.d).

$$
\begin{aligned}
g(t) &= A\cos\omega_c t\, p_T(t) \quad 0 \le t \le T \\
E[I_n I_{n+m}] &= \delta_{m,0} = \begin{cases} 1, & m = 0 \\ 0, & m \ne 0 \end{cases} \\
\Phi_I(f) &= 1 \\
\Phi_u(f) &= \frac{A^2 T}{4}(\mathrm{sinc}^2(\omega-\omega_c)\frac{T}{2} + \mathrm{sinc}^2(\omega+\omega_c)\frac{T}{2})
\end{aligned}
$$

Definition of Bandwidth

1. Null-to-Null $\triangleq$ bandwidth (in Hz) of main lobe ($=\frac{2}{T}$ for BPSK).

2. 99% containment bandwidth $\triangleq$ bandwith such that $=\frac{1}{2}$% lies above upper bandlimit $\frac{1}{2}$% lies below lower level.

3. x dB bandwidth $W_x \triangleq$ bandwidth such that spectrum is $x\ dB$ below spectrum at center of band (e.g. 3dB bandwidth).

4. Noise bandwidth

$$W_N \triangleq P/S(f_c)$$

where P is total power and $S(f_c)$ is value of spectrum $f = f_c$

$$P = \int_{-\infty}^{\infty} S(f)df$$

5. Gabor bandwidth

$$W_G \triangleq \sigma = \sqrt{\frac{\int_{-\infty}^{\infty}(f-f_c)^2 S(f)df}{\int_{-\infty}^{\infty} S(f)df}}$$

6. Absolute bandwidth

$$W_A \triangleq \min\{W : S(f) = 0 \quad \forall f \ge |W|\}$$

7. half null-to-null $\triangleq \frac{1}{2}$ null-to-null.

	1	2	3 35dB	4	5 and 6	3 3dB
BPSK	2.0	20.56	35.12	1.00	∞	0.88
QPSK	1.0	10.28	17.56	0.50	∞	0.44
MSK	1.5					

QPSK for same date rate T bits/sec

Comparison of Modulation Techniques

BPSK has $P_{e,s} = Q(\sqrt{\frac{2E_b}{N_0}})$

$$W = \frac{1}{T}, \quad R = \frac{1}{T}, \quad \Rightarrow \frac{R}{W} = 1.0$$

$E_b/N_0 = 9.6dB$

QPSK has same P_e but has $\frac{R}{W} = 2.0$

$$R = \frac{2}{T_s} \qquad (R = \frac{1}{T})$$

or

$$W = \frac{1}{T_s} \quad (W = \frac{1}{2T})$$

M-ary PSK has same bandwidth as BPSK but transmits $\log_2 M$ bits/channel use (T sec).

M-ary PSK

$$\begin{aligned} R &= \frac{\log_2 M}{T} \\ &\Rightarrow \frac{R}{W} = \log_2 M \\ W &= \frac{1}{T} \end{aligned}$$

Capacity (Shannon Limits)

$$\begin{aligned} R/W &< \log_2\left(1+\frac{R}{W}\frac{E_b}{N_0}\right) \\ &\text{or} \\ E_b/N_0 &> \frac{(2^{R/W}-1)}{R/W} \end{aligned}$$

We can come close to capacity (at fixed R/W) by use of coding (At $R/W = 1$ there is a possible 9.6 dB ”coding gain”)

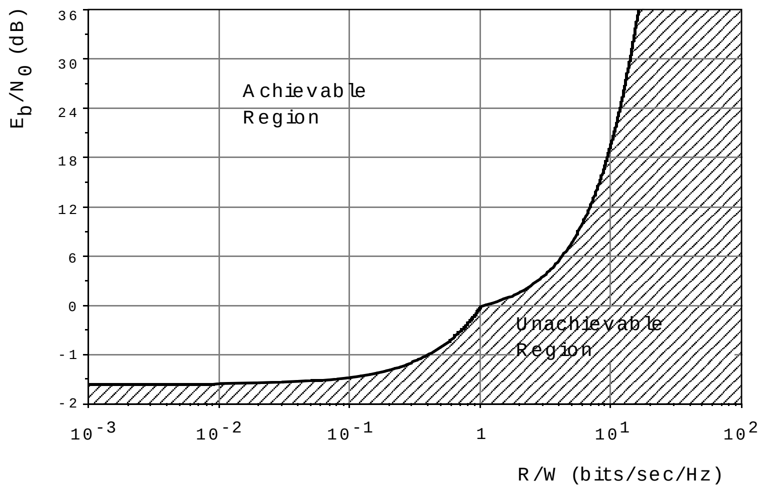


Figure 92: Capacity of Additive White Gaussian Noise Channel.