Optimum Receiver and Digital Binary Transmission

In binary data transmission over a communication channel, logic 1 is represented by a signal $s_1(t)$ and logic 0 by a signal $s_2(t)$. The time allocated for each signal is the symbol duration τ, where τ is related to the data rate by $r_b = 1/\tau$. We have two types of data transmission:

Baseband Data Transmission: Binary data transmission by means of two baseband waveforms (typically, two voltage levels) is referred to as baseband signaling. The spectrum of the transmitted signal occupies the low part of the frequency band (around the zero frequency). No high frequency carrier is used in this mode of transmission.

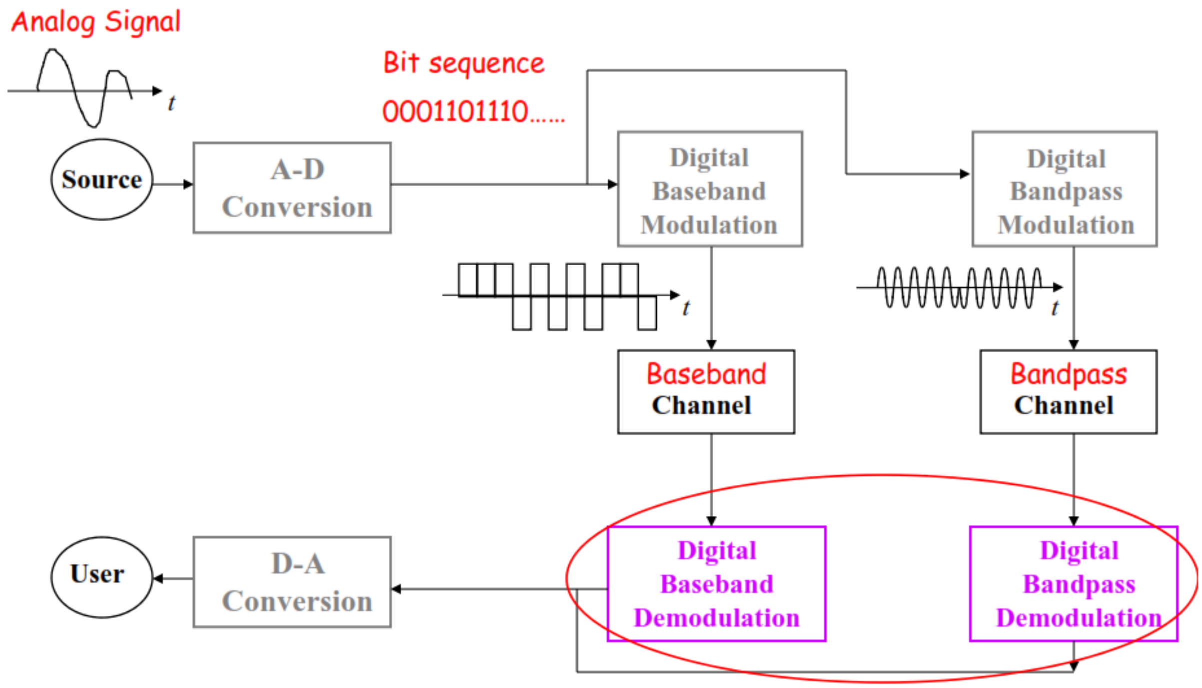
Bandpass Data Transmission: The baseband data modulates a high frequency carrier to produce a modulated signal, whose spectrum is centered around the carrier frequency.

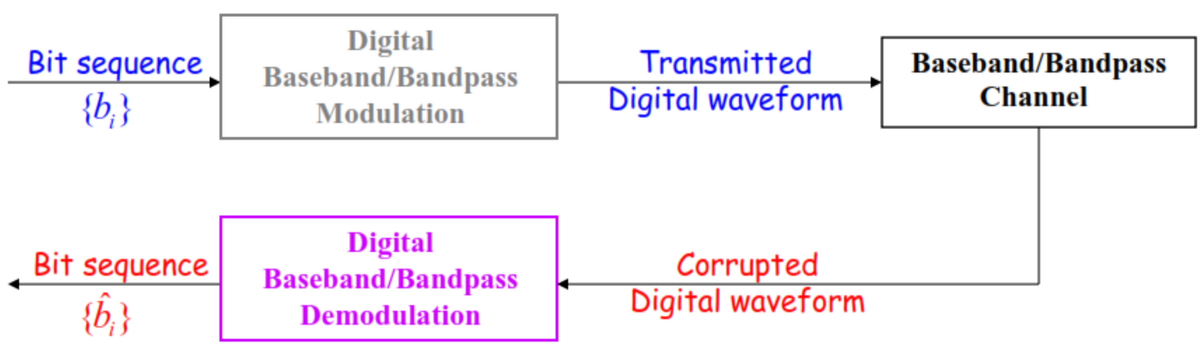
Assumptions:

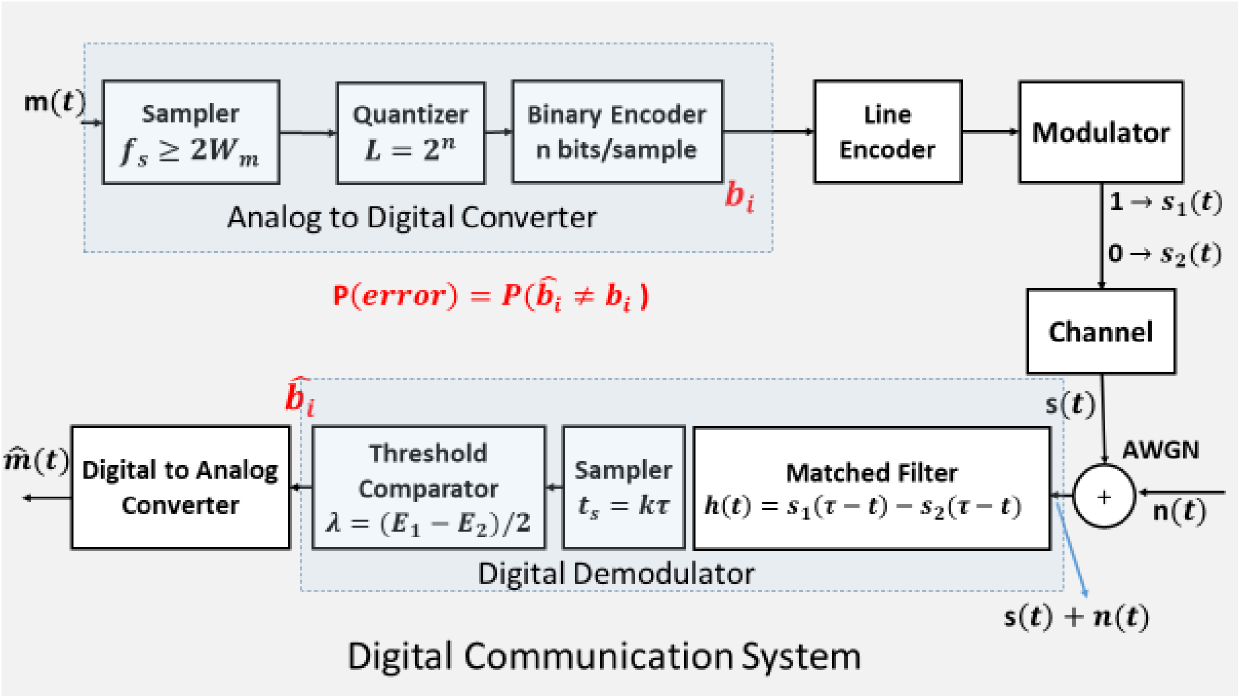
- The channel noise $n(t)$ is additive white Gaussian (AWGN) with a double-sided PSD of $N_0/2$. Noise is assumed to be added at the front end of the receiver.
- The data component at the front end of the receiver is assumed to be an exact replica of the transmitted signal, in the sense that the transmission bandwidth of the medium is wide enough to reproduce the signal without distortion.
- Bits in different time intervals are assumed independent.
- The signal to be processed by the receiver is the noisy signal $\boldsymbol{y(t) = s_i(t) + n(t)}$

Based on $y(t)$, the task of the receiver is to decide whether a 1 or a 0 was transmitted during each transmission slot τ with minimum probability of error.

Digital Communication Block Diagram







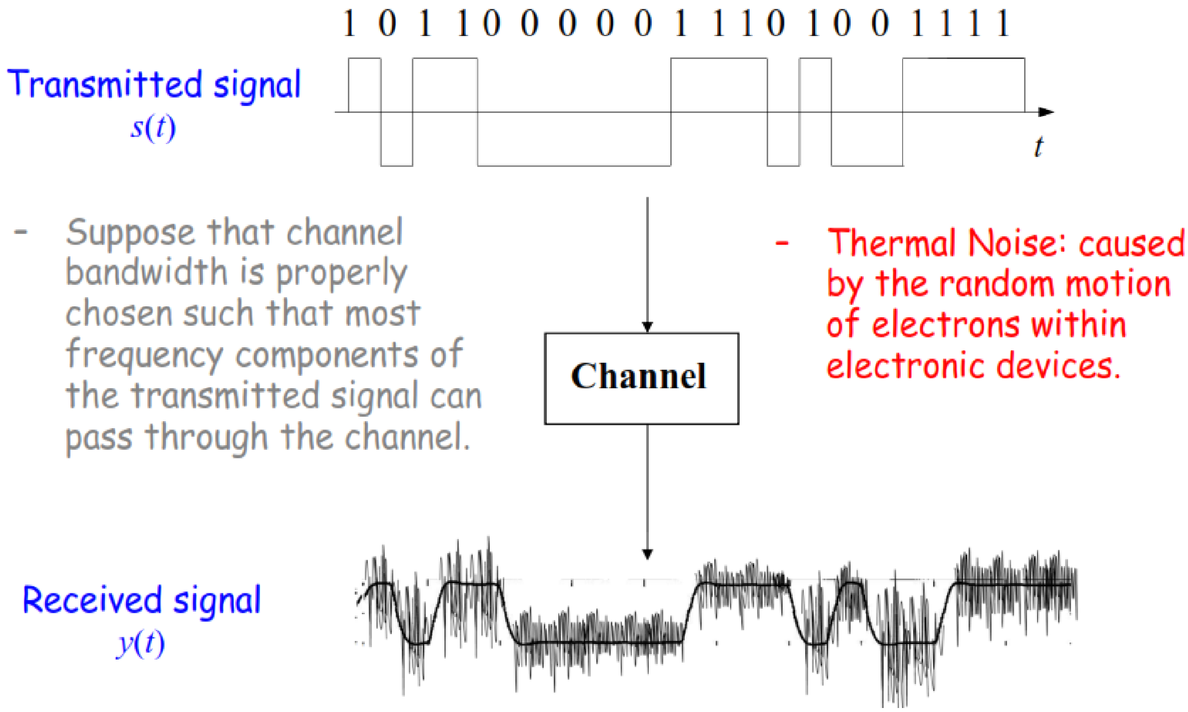
Three Questions:

- What are the sources of signal corruption?
- How to recover the transmitted data (i.e., to obtain the bit sequence $\hat{b}_i$)?
- How to evaluate the receiver performance?

Sources of Signal Corruption

- **Thermal Noise**: caused by the random motion of electrons within electronic devices
- **Communication Channel**: The finite bandwidth of the channel introduces some amount of amplitude distortion, which can be minimized, by increasing the allocated bandwidth. The effect of finite channel bandwidth introduces a type of noise (distortion) called **inter-symbol interference (ISI)**. Here, we confine our discussion to channels with large bandwidth. ISI will be addressed in a later chapter.

The following figure illustrates the effect of the two types of noise on the received signal.

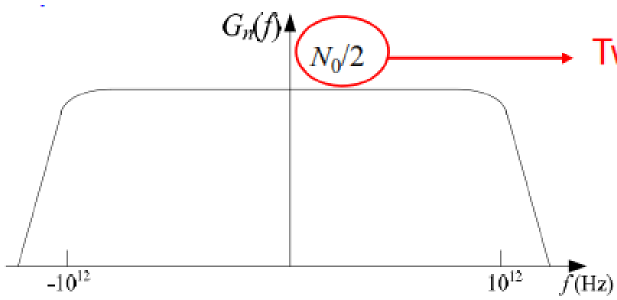


Thermal Noise:

- Thermal noise, $n(t)$, is modeled as a wise sense stationary (WSS) Gaussian random process.
- The thermal noise has a power spectrum that is constant from dc to approximately 10^{12} Hz; hence, $n(t)$ can be approximately regarded as a white process.
- Thermal noise is superimposed (added) on the transmitted signal. The received signal is $y(t) = s(t) + n(t)$.
- The mean value of the thermal noise $n(t)$ is zero.
- At any given time t_0 the probability density function of $n(t_0)$ follows the Gaussian distribution; $N(0, \sigma_0^2\ )$; where $\sigma_0^2 = E(n(t))^2$ is the noise power.

$$f_n(n) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-n^2/2\sigma_0^2}$$

In the following analysis, we will refer to thermal noise as **additive white Gaussian Noise** (**AWGN**), because it is modeled as a white Gaussian WSS process, which is added to the signal.



Basic Elements of the Receiver:

To decide on whether logic 1 or logic 0 was transmitted during a given time slot τ, the received signal (transmitted signal and noise) passes through three basic units.

Filter: The optimum filter, which we will also call the matched filter.

Sampler: Samples the received signal (data component plus noise) at some time $t = t_0 = \tau = symbol\ duration.$

Threshold comparator: If the sampled value is larger than a given threshold, λ, digit 1 is declared true, otherwise digit 0 is declared true.

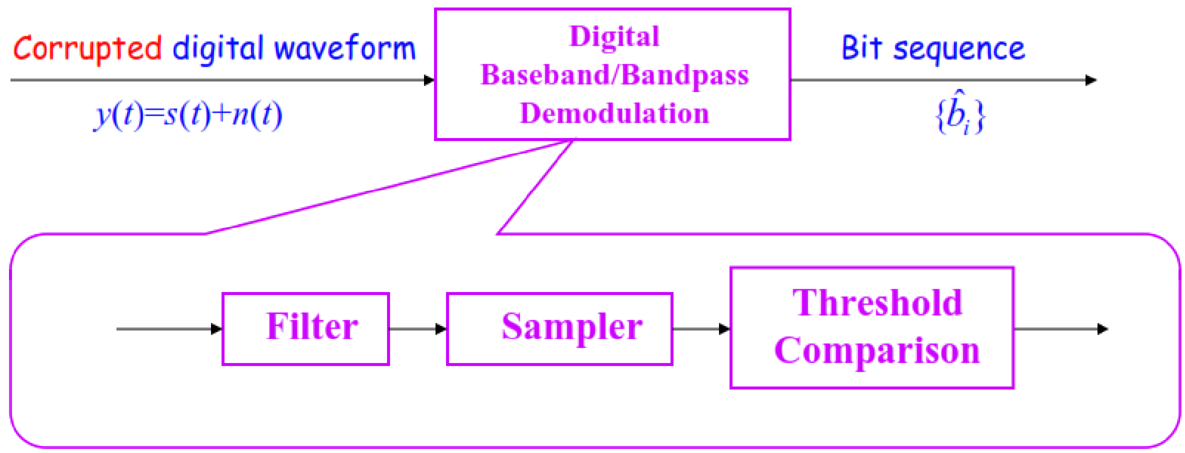
There are three design elements at the receiver

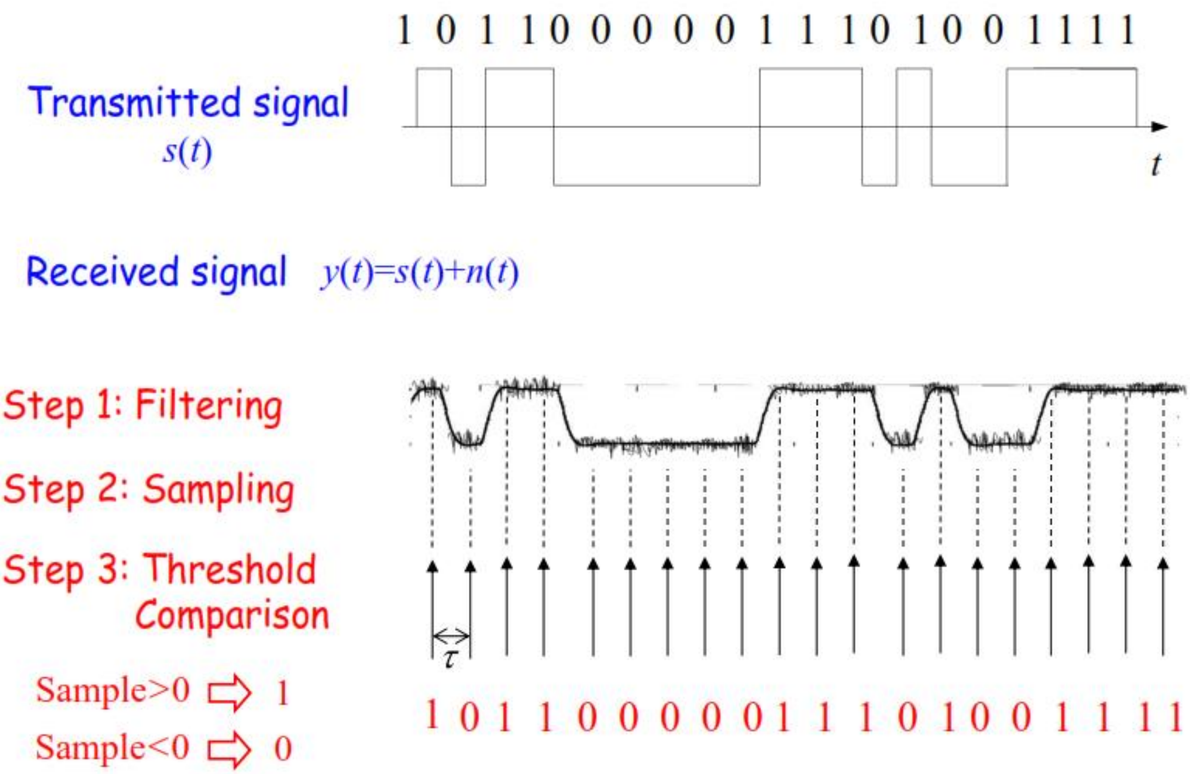
a. The impulse response $h(t)$ of the filter
b. The sampling time t_0
c. The threshold λ

These parameters should be chosen so as to minimize the average probability of error (or bit error rate BER), defined as

$$P_b = P(b_i \neq \widehat{b_i})$$

$$P_b = \Pr\{\hat{b}_i = 1,\ b_i = 0\} + \Pr\{\hat{b}_i = 0,\ b_i = 1\}$$


Corrupted digital waveform
y(t)=s(t)+n(t)
Digital Baseband/Bandpass Demodulation
Bit sequence
{b̂_i}
Filter
Sampler
Threshold Comparison


1 0 1 1 0 0 0 0 0 1 1 1 0 1 0 0 1 1 1 1
Transmitted signal
s(t)
t
Received signal y(t)=s(t)+n(t)
Step 1: Filtering
Step 2: Sampling
Step 3: Threshold Comparison
Sample>0 ⇨ 1
Sample<0 ⇨ 0
τ
1 0 1 1 0 0 0 0 0 1 1 1 0 1 0 0 1 1 1 1

Theorem on the Optimum Binary Receiver

Consider a binary communication system, corrupted by AWGN with power spectral density $N_0/2$, where the equally probable binary digits 1 and 0 are represented by the signals $s_1(t)$ and $s_2(t)$, respectively. The transmission time for each signal is τ sec. The optimum receiver elements, i.e., the elements that minimize the receiver probability of error are given by

Impulse response of the matched filter: $\boldsymbol{h(t) = s_1(\tau - t) - s_2(\tau - t)}$, $0 \le t \le \tau$

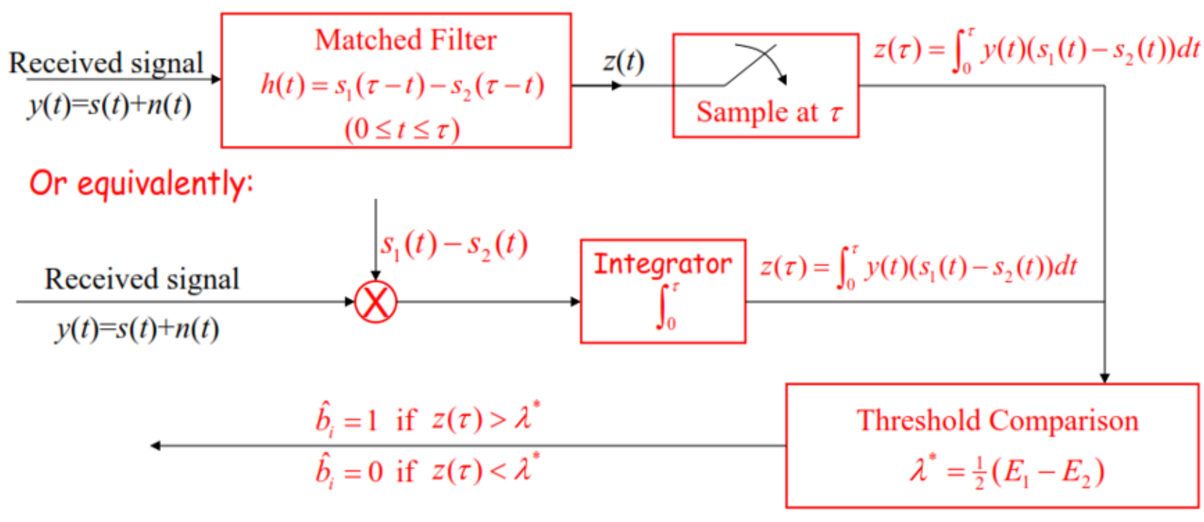
Optimum sampling time: $\boldsymbol{t_s = \tau}$

Optimum threshold of comparator: $\boldsymbol{\lambda^* = \frac{1}{2}(E_1 - E_2)}$, $E_k = \int_0^\tau (s_k(t))^2 dt\,, k = 1,2$

When these elements are used, the system minimum probability of error is

$$\boldsymbol{P_b^* = Q\left(\sqrt{\frac{\int_0^\tau (s_1(t) - s_2(t))^2 dt}{2N_0}}\right)}$$

The structure of the optimum receiver is depicted in the figure below. Note that the receiver can be implemented in terms of the matched filter and, equivalently, in terms of a correlator (a multiplier followed by an integrator).

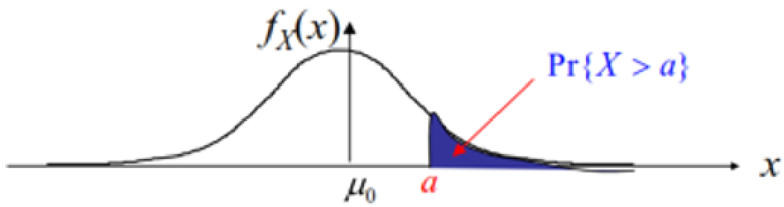


The Q-Function

$$Q(\alpha)=\int_{\alpha}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}dx$$

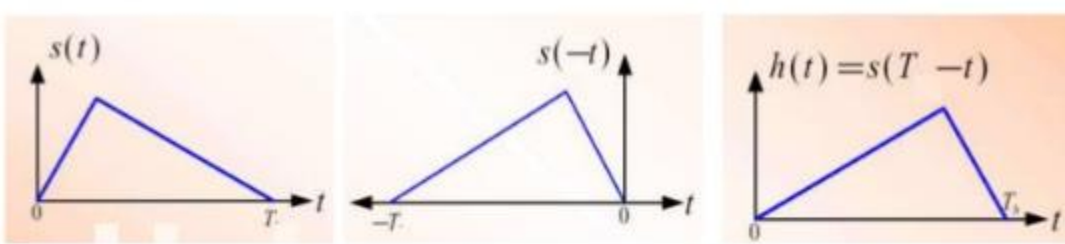
- $Q(\alpha)$ is a decreasing function of α.
- For $X\sim\mathcal{N}(\mu_0,\sigma_0^2)$,

$$\Pr\{X>a\}=\int_a^{\infty}f_X(x)dx=\int_a^{\infty}\frac{1}{\sqrt{2\pi}\sigma_0}\exp\left\{-\frac{(x-\mu_0)^2}{2\sigma_0^2}\right\}dx=Q\left(\frac{a-\mu_0}{\sigma_0}\right)$$



The Matched Filter

Example 1: The figure below shows the steps involved in obtaining the matched filter for the given signal waveform $\boldsymbol{s(t)=s_1(t)-s_2(t)}$

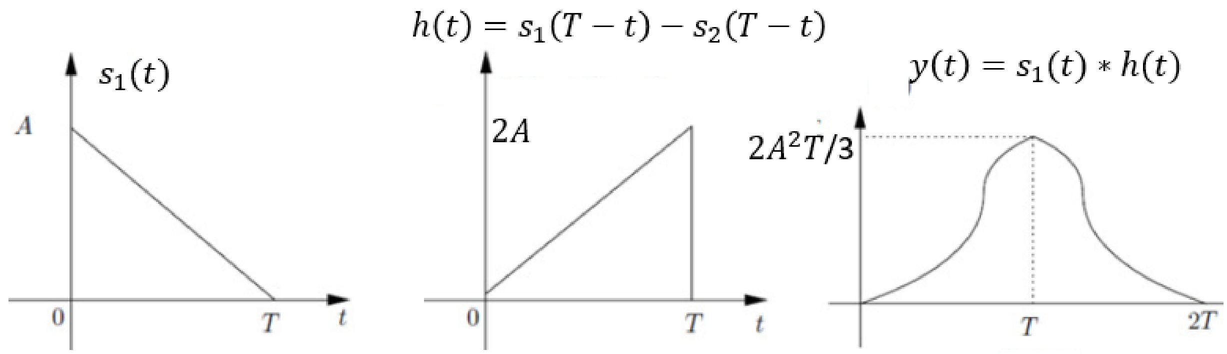


Signal Waveform

Mirror image of signal waveform

Impulse response of matched filter

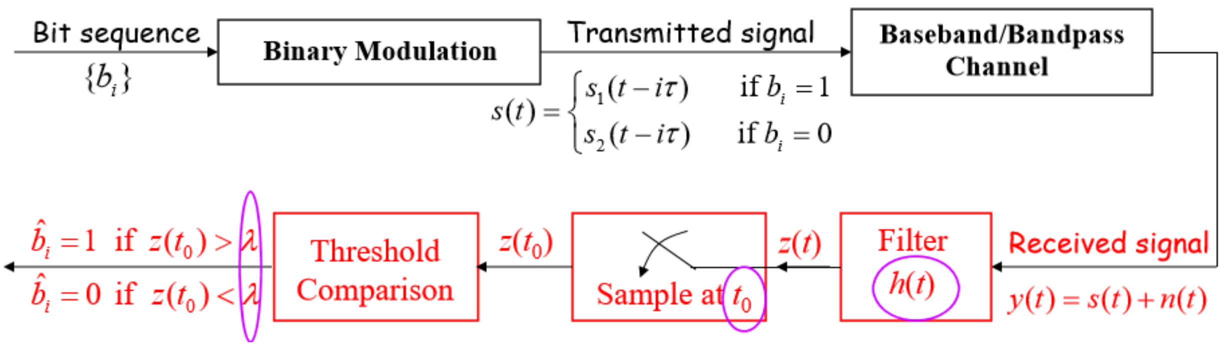
Example 2: The next figure shows a signaling scheme where $s_2(t)=-s_1(t)$. The impulse response of the matched filter is $h(t)=s_1(T-t)-s_2(T-t)$. The figure shows the filter output when $s_1(t)$ is applied to the filter. Note that the output attains its maximum value at time t=T, which is the sampling time chosen to maximize the output SNR.



Next, we prove the theorem.

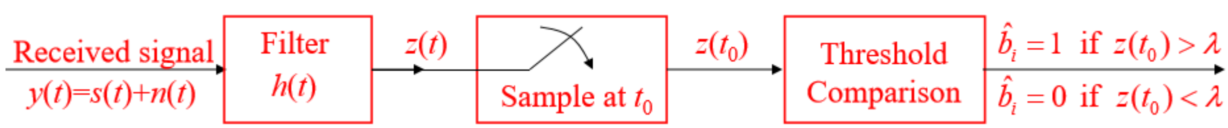
The Optimum (Matched) Filter:

In this section, we will derive the impulse response of the optimum filter, specify the optimum values of the threshold , and the sampling time t_0 that minimize the average probability of error of the system. These three parameters appear in the following block diagram.



The transmitted signal is:

$$s(t)=\begin{cases} s_1(t) & \text{if } b_1=1 \\ s_2(t) & \text{if } b_1=0 \end{cases} \qquad 0\le t\le \tau$$



The received signal is:

$$y(t)=s(t)+n(t)=\begin{cases} s_1(t)+n(t) & \text{if } b_1=1 \\ s_2(t)+n(t) & \text{if } b_1=0 \end{cases}$$

where noise has been added to the transmitted signal. The received signal plus noise $y(t)=s(t)+n(t)$ passes through a filter with an impulse response $h(t)$. Its output, $z(t)$, is the convolution of $h(t)$ with $y(t)$.

$$z(t)=s_o(t)+n_o(t)=\begin{cases} s_{o,1}(t)+n_o(t) & \text{if } b_1=1 \\ s_{o,2}(t)+n_o(t) & \text{if } b_1=0 \end{cases}$$

where,

$$s_{o,i}(t) = \int_0^t s_i(x)h(t-x)dx, \quad i = 1,2.$$

$$n_o(t) = \int_0^t n(x)h(t-x)dx,$$

are the signal and noise components at the filter output, respectively. The filter output, sampled at time t_0 is

$$z(t_0) = s_o(t_0) + n_o(t_0) = \begin{cases} s_{o,1}(t_0) + n_o(t_0) & \text{if } b_1 = 1 \\ s_{o,2}(t_0) + n_o(t_0) & \text{if } b_1 = 0 \end{cases}$$

The sampled output $z(t_0)$ is a random variable and is no more a function of time. The probability density function (pdf) of $z(t_0)$ will be derived next.

Receiver Average Probability of Error

First, we note that $n(t_0)$ is a Gaussian random variable with

mean: $E\big(n(t_0)\big) = \int_0^{t_0} E\big(n(t)\big)h(t_0 - x)dx$=0, since $E\big(n(t)\big) = 0$.

variance: $\sigma_0^2 = E(n(t_0)^2) = \frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2 df = \frac{N_0}{2}\int_0^{t_0} |h(t_0 - t)|^2 dt;$

The last step follows from Parseval's power theorem.

Hence, $n(t_0)$ follows the Gaussian distribution, i.e., $\boldsymbol{n(t_0) \sim N(0, \sigma_0^2)}$.

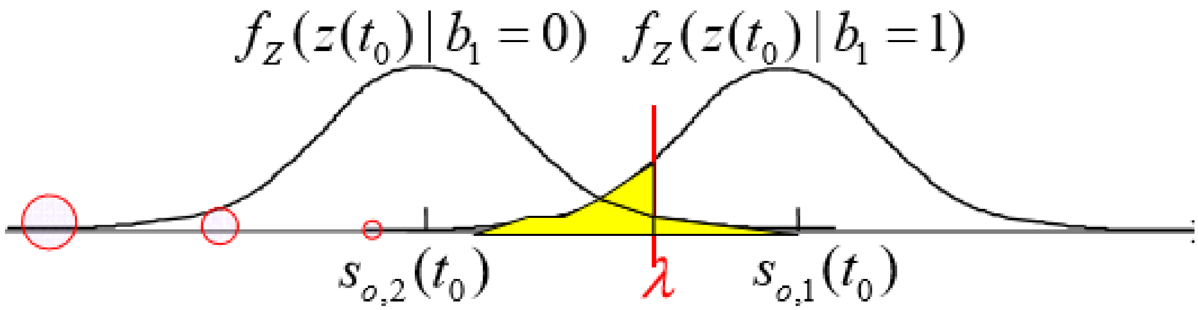
Remark: In ENEE 2307, we came across the following theorem related to the transformation of random variables:

Theorem: If X is a Gaussian random variable with mean μ_X and variance σ_X^2, and Y is another random variable related linearly to X by $Y = aX + b$, then Y is Gaussian with mean $\mu_Y = a\mu_X + b$ and variance $\sigma_Y^2 = a^2\sigma_X^2$.

Based on the above theorem, $z(t_0)$ follows the Gaussian distribution with the following conditional pdf's (depending on the transmitted bit $b_i = 1$ or $b_i = 0$)

$$z(t_0) \mid b_1 = 1 \sim \mathcal{N}(s_{o,1}(t_0), \sigma_0^2)$$
$$z(t_0) \mid b_1 = 0 \sim \mathcal{N}(s_{o,2}(t_0), \sigma_0^2)$$

These pdf's are depicted in this figure



The average probability of error is the sum of the two shaded areas in the figure. To help you understand that, we recall the Theorem of Total Probability, considered in ENEE 2307.

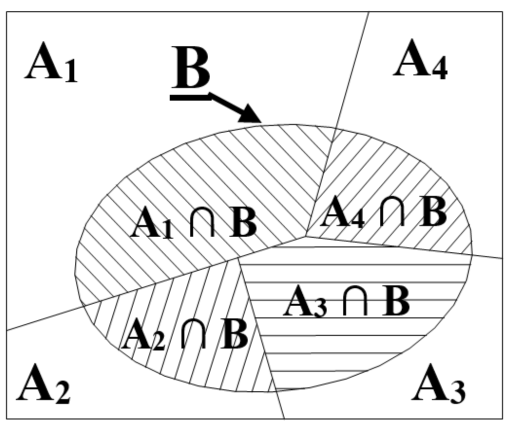
Theorem of Total Probability

Let $A_1, A_2, \ldots, A_n$ be a set of events defined over (S) such that:

$S = A_1 \cup A_2 \cup \ldots \cup A_n$; $A_i \cap A_j = \emptyset$ for $i \neq j$, and $P(A_i) > 0$ for $i = 1, 2, 3, \ldots n$.

For any event (B) defined on (S),

$$P(B) = P(A_1)\, P(B/A_1) + P(A_2)\, P(B/A_2) + \ldots\ldots + P(A_n)\, P(B/A_n)$$



Making use of the theorem of total probability, the probability of error is calculated as:

$$P_b = \Pr\{\hat{b}_1 = 1,\, b_1 = 0\} + \Pr\{\hat{b}_1 = 0,\, b_1 = 1\}$$

$$= \Pr\{z(t_0) > \lambda, b_1 = 0\} + \Pr\{z(t_0) < \lambda, b_1 = 1\}$$

$$= \Pr\{z(t_0) > \lambda \mid b_1 = 0\}\Pr\{b_1 = 0\} + \Pr\{z(t_0) < \lambda \mid b_1 = 1\}\Pr\{b_1 = 1\}$$

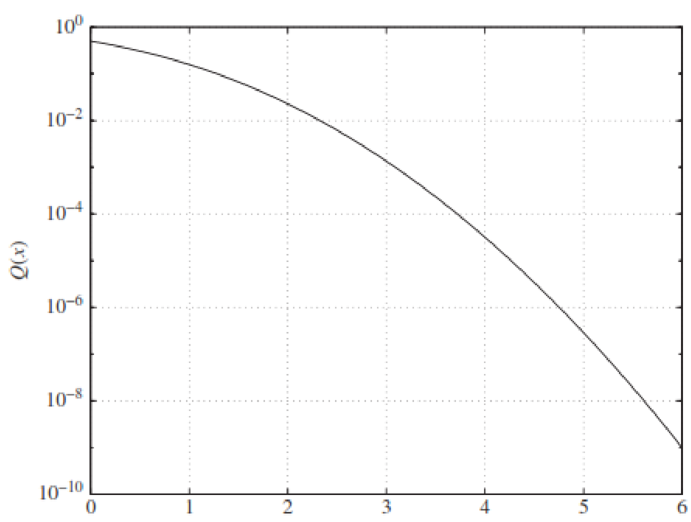
$$= \tfrac{1}{2}\left[\underline{\Pr\{z(t_0) > \lambda \mid b_1 = 0\}} + \underline{\Pr\{z(t_0) < \lambda \mid b_1 = 1\}}\right] \qquad (\Pr\{b_1 = 0\} = \Pr\{b_1 = 1\} = \tfrac{1}{2})$$

$$= \frac{1}{2}\left(Q\left(\frac{\lambda - s_{o,2}(t_0)}{\sigma_0}\right) + Q\left(\frac{s_{o,1}(t_0) - \lambda}{\sigma_0}\right)\right)$$

where, $Q(x)$ is the complementary Gaussian distribution function defined as:

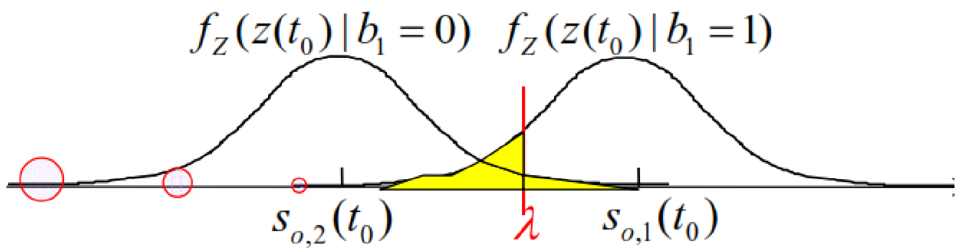
$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

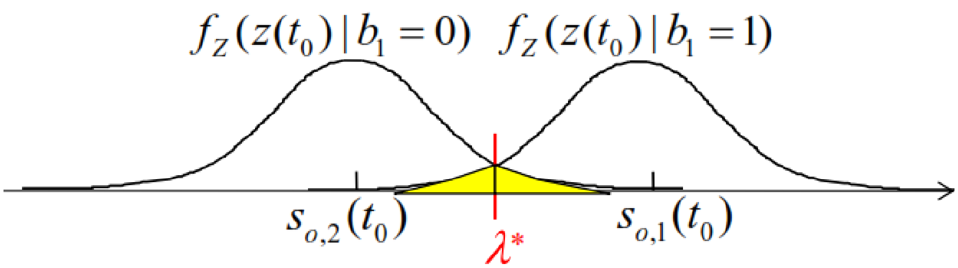
The $Q(x)$ function, plotted below, is a monotonically decreasing function in x, for $x \geq 0$. The Q function is minimized when the argument x is maximized. The results to be derived next on the minimization of the probability of error are based on this observation.



The average probability of error, P_b is a function of three variables, the threshold λ, the sampling time t_0, and the filter impulse response $h(t)$. We will deal with each variable one by one.

First, P_b is minimized by selecting a λ that minimizes the sum of the two shaded areas in the following figure:





The optimum value of λ (for equally probable symbols) is:

$$\lambda^* = \frac{s_{o,1}(t_0) + s_{o,2}(t_0)}{2}$$

With this choice of λ, P_b becomes

$$P_b(\lambda^*) = \frac{1}{2}\left(Q\left(\frac{\lambda^* - s_{o,2}(t_0)}{\sigma_0} \right) + Q\left(\frac{s_{o,1}(t_0) - \lambda^*}{\sigma_0} \right) \right)$$

$$P_b^* = Q\left(\frac{s_{o,1}(t_0) - s_{o,2}(t_0)}{2\sqrt{\sigma_0^2}} \right)$$

$$= Q\left(\frac{1}{2}\sqrt{\frac{\left(\int_0^{t_0}(s_1(x)-s_2(x))h(t_0-x)dx\right)^2}{\frac{N_0}{2}\int_0^{t_0}h^2(t_0-x)dx}}\right)$$

Where,

$$s_{o,1}(t_0)-s_{o,2}(t_0)=\int_0^{t_0}(s_1(x)-s_2(x))h(t_0-x)dx,$$

$$\sigma_0^2=\frac{N_0}{2}\int_0^{t_0}h^2(t_0-x)dx.$$

Next, P_b is minimized when the argument of the Q function is maximized, which by means of Schwartz's inequality is achieved when the sampling time is $t_0=\tau$ and

$$h(t)=k(s_1(\tau-t)-s_2(\tau-t)),\quad 0\le t\le\tau \quad \text{and} \quad t_0=\tau$$

$$\frac{\left[\int_0^{t_0}(s_1(x)-s_2(x))h(t_0-x)dx\right]^2}{\int_0^{t_0}\frac{N_0}{2}h^2(t_0-x)dx}\le\frac{\int_0^{t_0}(s_1(x)-s_2(x))^2dx\int_0^{t_0}h^2(t_0-x)dx}{\frac{N_0}{2}\int_0^{t_0}h^2(t_0-x)dx}$$

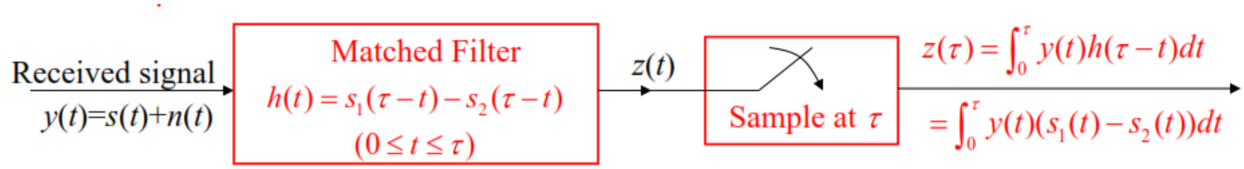
"=" holds when $h(t)=k(s_1(t_0-t)-s_2(t_0-t))$

$$=\frac{\int_0^{t_0}(s_1(x)-s_2(x))^2dx}{N_0/2}\le\frac{\int_0^{\tau}(s_1(x)-s_2(x))^2dx}{N_0/2}$$

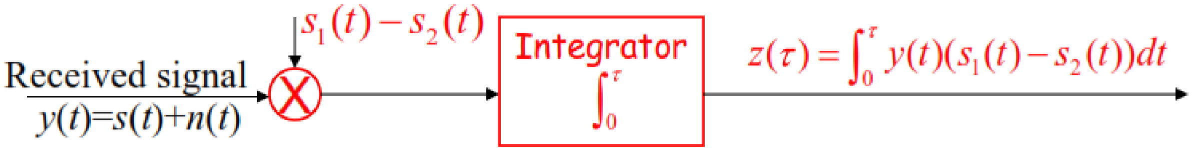
"=" holds when $t_0=\tau$

$h(t)$, as obtained above, is called the optimum or matched filter, since its shape is derived (matched) from the shapes of the transmitted signals.

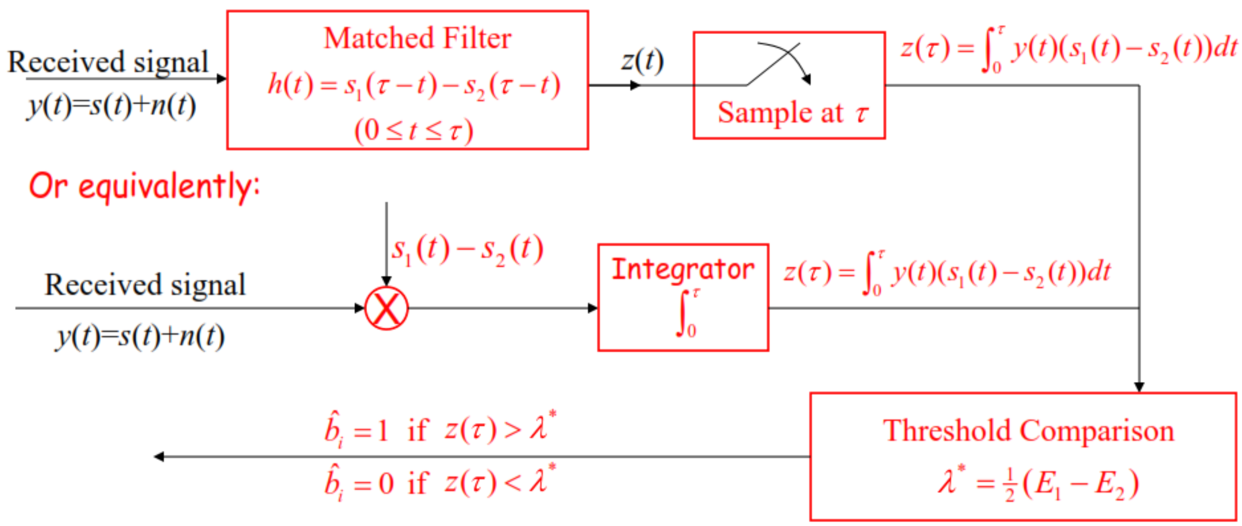
The output of the matched filter at the sampling time may be obtained directly through the convolution integral



Or, equivalently, through correlation. Here, the incoming signal is multiplied by $(s_1(t) - s_2(t))$ and then integrated over the symbol period $(0, \tau)$. The output of the configurations is equal only at $t = \tau$.



The following figure summarizes the above operations and results



$$\lambda^* = \tfrac{1}{2}(s_{o,1}(\tau)+s_{o,2}(\tau)) = \tfrac{1}{2}\int_0^\tau (s_1(x)+s_2(x))h(\tau-x)dx = \tfrac{1}{2}\left(\underbrace{\int_0^\tau s_1^2(t)dt}_{E_1} - \underbrace{\int_0^\tau s_2^2(t)dt}_{E_2}\right)$$

$$= \tfrac{1}{2}(E_1 - E_2) \qquad \text{Energy of } s_i(t)\text{:} \quad E_1 \quad E_2$$

Finally, the BER of the optimal binary detector is:

$$P_b^* = Q\left(\frac{1}{2}\sqrt{\frac{\left(\int_0^\tau (s_1(x)-s_2(x))(s_1(x)-s_2(x))dx\right)^2}{\frac{N_0}{2}\int_0^\tau (s_1(x)-s_2(x))^2 dx}}\right);$$

$$P_b^* = Q\left(\sqrt{\frac{\int_0^\tau (s_1(t)-s_2(t))^2 dt}{2N_0}}\right)$$

The integration on the RHS, can be expanded as

$$\int_0^\tau (s_1(t) - s_2(t))^2 dt = \int_0^\tau (s_1(t))^2 dt + \int_0^\tau (s_2(t))^2 dt - 2\int_0^\tau s_1(t)s_2(t)dt$$

$$\int_0^\tau (s_1(t) - s_2(t))^2 dt = E_1 + E_2 - 2\int_0^\tau s_1(t)s_2(t)dt$$

The correlation coefficient is a measure of similarity between two signals $s_1(t)$ and $s_2(t)$, and is defined as

$$\boldsymbol{\rho = \frac{1}{\sqrt{E_1}\sqrt{E_2}}\int_0^\tau s_1(t)s_2(t)dt}; \quad -1 \le \rho \le 1$$

Therefore,

$$\int_0^\tau (s_1(t) - s_2(t))^2 dt = E_1 + E_2 - 2\rho\sqrt{E_1}\sqrt{E_2}$$

Remark: Please note that $\int_0^\tau (s_1(t) - s_2(t))^2 dt$ represents the square of the distance between the two signals $s_1(t)$ and $s_2(t)$.

In the next section, we will make frequent use of these results to obtain the probability of error for a number of modulation schemes.

Example: Antipodal Binary Transmission

Let us consider a digital binary communication system where bits 1 and 0 are represented by the signals $s_1(t)$ and $-s_1(t)$, respectively.

For this case, $E_1 = E_2 = E = \int_0^{\tau}(s_1(t))^2 dt$

Therefore, the threshold is $\boldsymbol{\lambda^* = (E_1 - E_2) = 0}$

The probability of error is:

$$P_b^* = Q\left(\sqrt{\frac{\int_0^{\tau}(s_1(t) - s_2(t))^2 dt}{2N_0}}\right)$$

$$P_b^* = Q\left(\sqrt{\frac{\int_0^{\tau}(s_1(t) + s_1(t))^2 dt}{2N_0}}\right)$$

$$P_b^* = Q\left(\sqrt{\frac{4\int_0^{\tau}(s_1(t))^2 dt}{2N_0}}\right)$$

$$\boldsymbol{P_b^* = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)}$$

Baseband Data Transmission

Binary data transmission by means of two voltage levels is referred to as baseband signaling. Manchester encoding, for example, is used in the Ethernet local area network as the signaling scheme. Here, we consider polar non-return to zero baseband transmission scheme, in terms of probability of error, optimum receiver structure, power spectral density and bandwidth.

Polar nonreturn to zero (also known as binary pulse amplitude modulation)

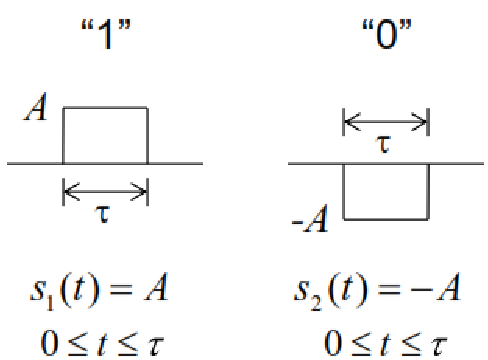
Signal Representation

The baseband signals representing digits 1 and 0 are:

$$s_1(t) = A\,, \qquad 0 \le t \le \tau$$

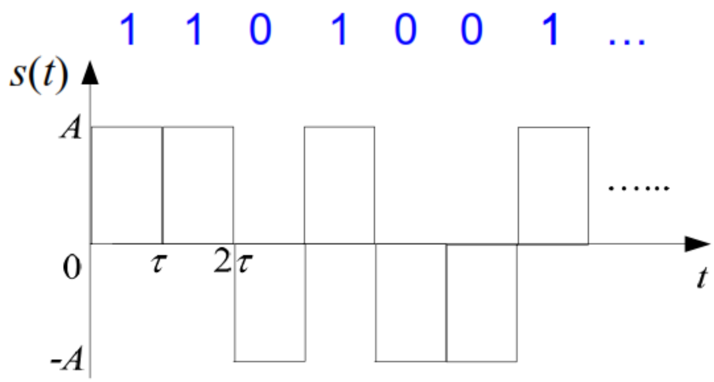
$$s_0(t) = -A\,, \qquad 0 \le t \le \tau$$

where, τ is the symbol duration and $R_b = 1/\tau$ is the data rate in bits/sec.



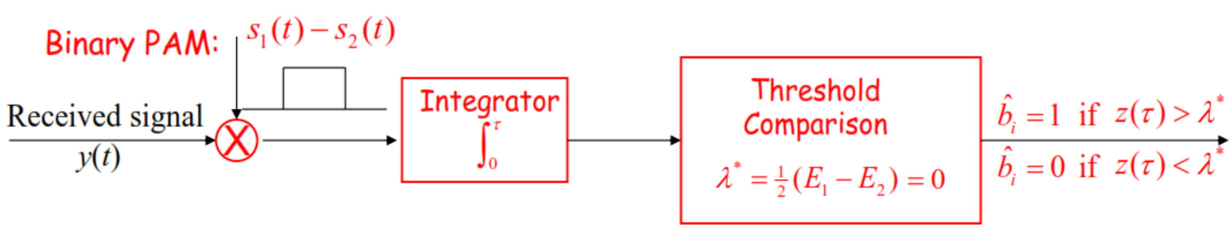
Generation:

The input to the modulator is a sequence of binary digits (0's and 1's). The modulator converts the sequence into a polar nonreturn to zero waveform (also known as binary pulse amplitude modulated waveform BPAM). The pulses are transmitted through the channel to the receiver.



Optimum Receiver

The optimum receiver is, of course, the matched filter, also implemented as a correlator, as shown in this figure.



Probability of Error:

$$P_b^* = Q\left(\sqrt{\frac{\int_0^\tau (s_1(t) - s_2(t))^2 dt}{2N_0}}\right)$$

Note that: $E_1 = E_2 = \int_0^\tau A^2 dt = A^2\tau \Rightarrow \lambda^* = (E_1 - E_2) = 0$

Average Energy per bit: $E_b = \frac{1}{2}(E_1 + E_2) = A^2\tau$

Optimal BER:

$$P_b^* = Q\left(\sqrt{\frac{2A^2\tau}{N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

General Result on the Power Spectral Density of a digital M-ary baseband signal

The time-domain representation of a digital M-ary baseband signal is

$$s(t)=\sum_{n=-\infty}^{\infty} Z_n \cdot v(t-n\tau)$$

where Z_n is a discrete random variable with $\Pr\{Z_n=a_i\}=1/M,\ i=1,\ldots,M,$

$v(t)$ is a unit-baseband signal, and symbols in different time slots are assumed independent. Under these assumptions, the power spectral density of $s(t)$ is given by

$$G_s(f)=\frac{1}{\tau}\left|V(f)\right|^2 \cdot \left(\sigma_Z^2+\frac{\mu_Z^2}{\tau}\sum_{m=-\infty}^{\infty}\delta\left(f-\frac{m}{\tau}\right)\right)$$

Power Spectral Density of the polar non-return to zero baseband signal

The general result stated above for the M-ary baseand signal can be specialized to the polar nonreturn to zero transmission as follows

- The signal amplitude assumes two equally likely values. i.e., $P\{Z_n=\pm 1\}=1/2$
- The basic unit pulse is $v(t)=\begin{cases}A, & 0\le t\le \tau\\ 0, & otherwise\end{cases}$
- The Fourier transform of the basic unit pulse is $V(f)=A\tau sinc(f\tau)$
- The mean and variance of Z are: $\mu_Z=0$, $\sigma_Z^2=1$

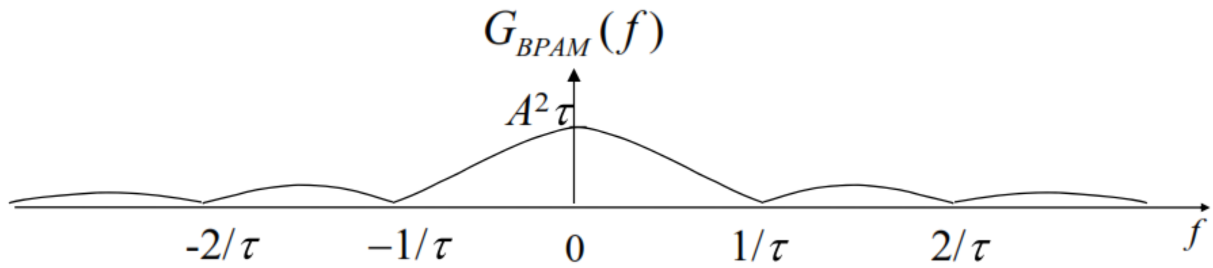
Remark: Recall that for a discrete random variable Z, the mean and variance are defined as

$$E(Z)=\sum_{all\ z_i} z_i P(Z=z_i)$$

$$Var(Z)=\sigma_Z^2=\sum_{all\ z_i}(z_i-E(Z))^2 P(Z=z_i)$$

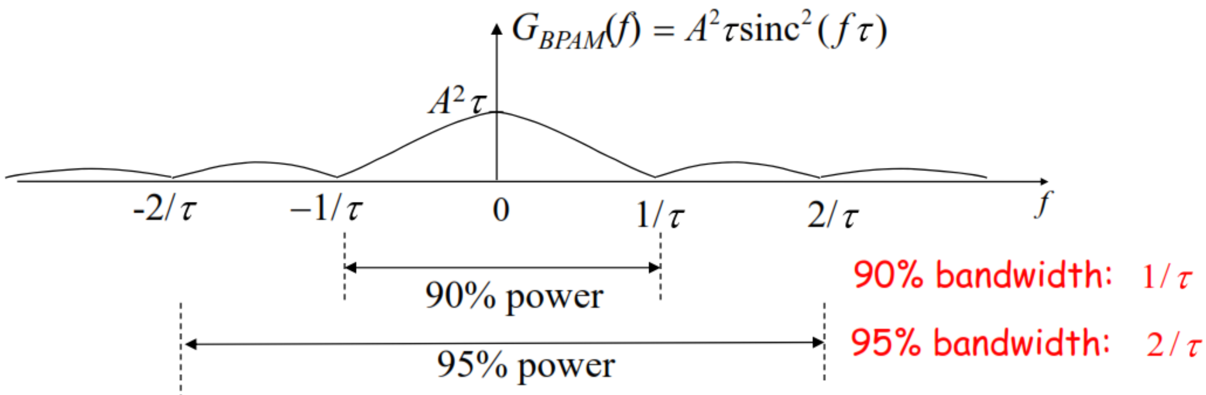
Therefore, the power spectral density of the polar non-return to zero signal is

$$G_{BPAM}(f)=A^2\tau\,\mathrm{sinc}^2(f\tau)$$



Bandwidth

The bandwidth can be obtained from the power spectral density.



The 90% power bandwidth = $\frac{1}{\tau} = R_b$ (data rate)

The 95% power bandwidth = $\frac{2}{\tau} = 2R_b$ (twice the data rate)

Binary Digital Bandpass Modulation

Here, the baseband data modulates a high frequency carrier to produce a modulated signal, whose spectrum is centered on the carrier frequency. We will consider four types of bandpass transmission schemes; Amplitude Shift Keying (ASK), Phase Shift Keying (PSK), Frequency Shift Keying (FSK), and Quadri-phase Shift Keying (QPSK). For each type, we consider the generation, optimum receiver, probability of error, power spectral density, and bandwidth.

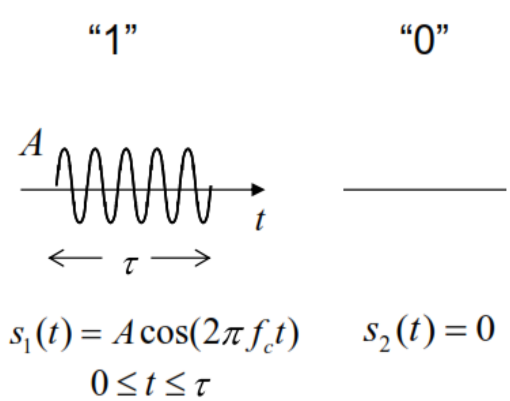
Binary Amplitude Shift Keying (BASK)

Signal Representation

Send: $s_1(t) = Acos(2\pi f_c t)$, if the information bit is "1" $\Rightarrow E_1 = \frac{A^2\tau}{2}$

Send: $s_2(t) = 0$, if the information bit is "0"; $\Rightarrow E_2 = 0$

The average energy per bit $E_b = \frac{1}{2}(E_1 + E_2) = \frac{A^2\tau}{4}$

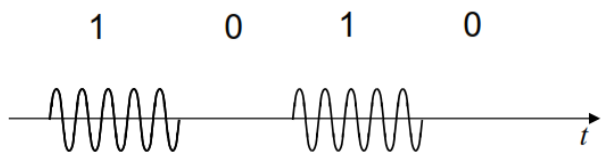


Generation:

The ASK signal is generated by multiplying the unipolar nonreturn to zero waveform (also known as binary On-Off Keying signal) with the sinusoidal high frequency carrier.

$$s_{BASK}(t) = s_{BOOK}(t)cos(2\pi f_c t)$$

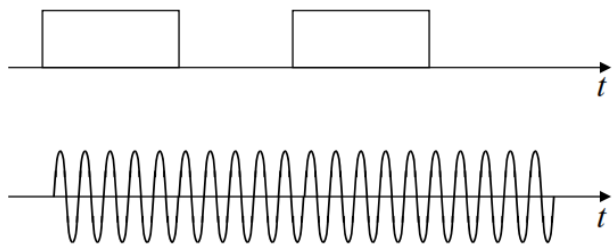
$s_{BASK}(t) = s_{BOOK}(t)\cos(2\pi f_c t)$

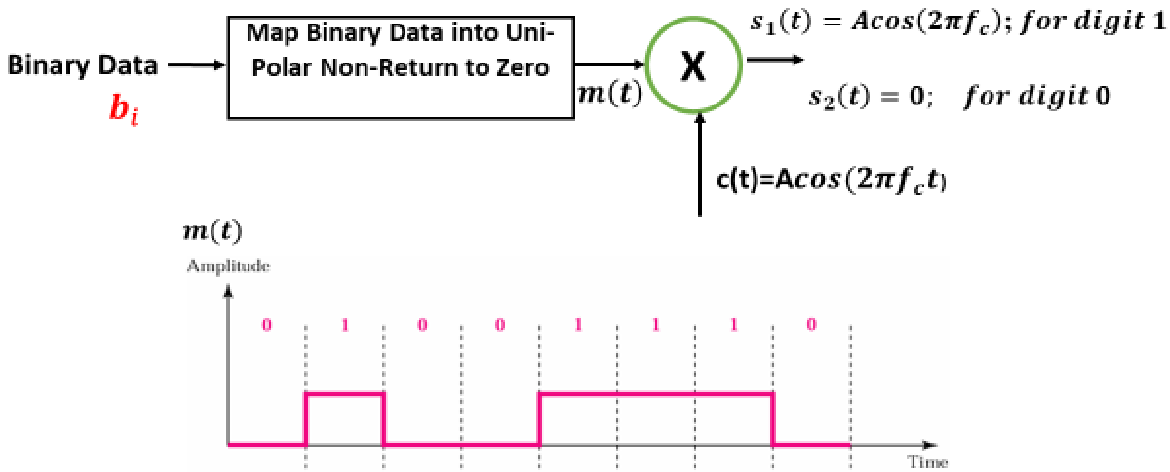


Binary On-Off Keying $s_{BOOK}(t)$

X

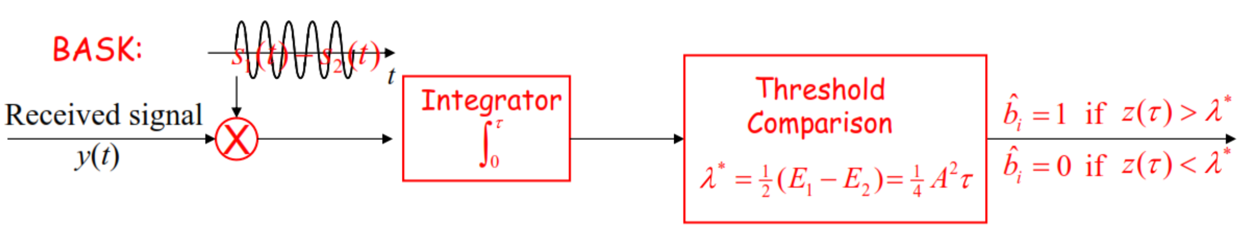
$\cos(2\pi f_c t)$





Binary Amplitude Shift Keying Generation

Optimum Receiver:



Probability of Error

Optimal BER:

$$P_b^* = Q\left(\sqrt{\frac{A^2\tau}{4N_0}}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

Power Spectral Density of Binary ASK

First, we find the statistics of the amplitude variable Z:

- $P(Z=1) = P(Z=0) = 1/2$
- $E(Z) = \mu_Z = 1/2$
- $Var(Z) = \sigma_Z^2 = 1/4$

Then, we find the power spectral density of the baseband binary unipolar non-return to zero signal (On-Off Keying)

$$G_{BOOK} = \frac{1}{\tau}V(f)^2\{\sigma_Z^2 + \frac{{\mu_Z}^2}{\tau}\sum_{m=-\infty}^{\infty}\delta(f-\frac{m}{\tau})\}$$

$$G_{BOOK}(f) = \frac{1}{\tau}\left(A\tau \operatorname{sinc}(f\tau)\right)^2 \cdot \left(\frac{1}{4} + \frac{1}{4\tau}\sum_{m=-\infty}^{\infty}\delta\left(f-\frac{m}{\tau}\right)\right)$$

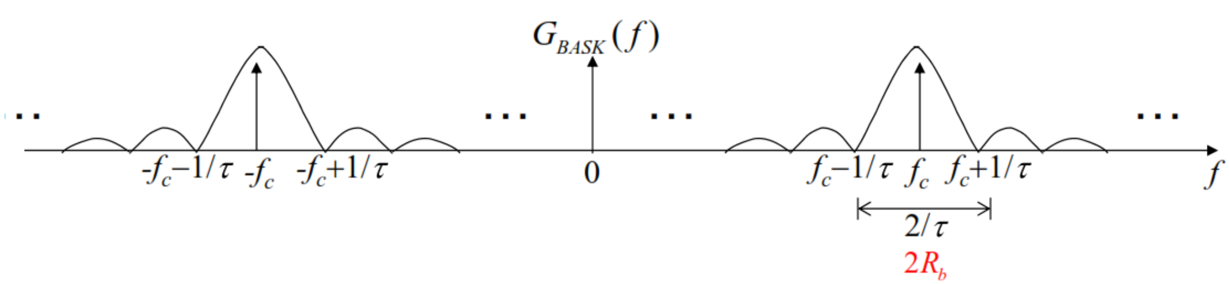
$$G_{BOOK} = \frac{1}{4}A^2\tau(sincf\tau)^2 + \frac{1}{4}A^2\{\delta(f-r_b) + \delta(f+r_b)\}$$

Note that the rest of the terms in the summation on the RHS become zeros since

$$(sincf\tau)^2\delta(f-mr_b) = (sinc\frac{m}{\tau}\tau)^2\delta(f-mr_b) = 0$$

Finally, the power spectral density of the binary ASK is:

$$G_{BASK}(f) = \tfrac{1}{4}[G_{BOOK}(f-f_c) + G_{BOOK}(f+f_c)]$$



The 90% power bandwidth = $\frac{2}{\tau} = 2R_b$ (twice the data rate)

The 95% power bandwidth = $\frac{4}{\tau} = 4R_b$ (four-times the data rate)

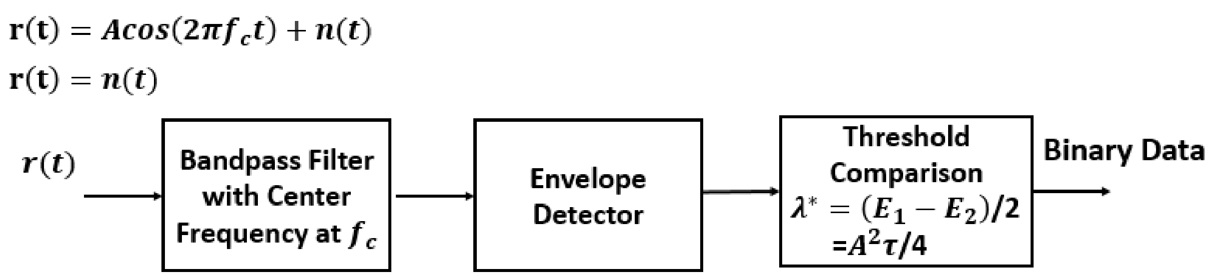
Final Remarks

- **Advantage: Simplicity**
- **Disadvantage: ASK is very susceptible to noise interference – noise usually (only) affects the amplitude. Therefore ASK is the modulation technique most affected by noise**
- **Application: ASK is used to transmit digital data over optical fiber**

Non-Coherent Demodulation for ASK

In this demodulation technique, there is no need for the carrier frequency at the receiver. The basic elements of the receiver are a bandpass filter with center frequency at the carrier, an envelope detector, and a threshold comparator. The receiver is simple, however it is not optimal in terms of the probability of error.

The details are shown in the following block diagram



Non-Coherent Binary ASK Demodulation

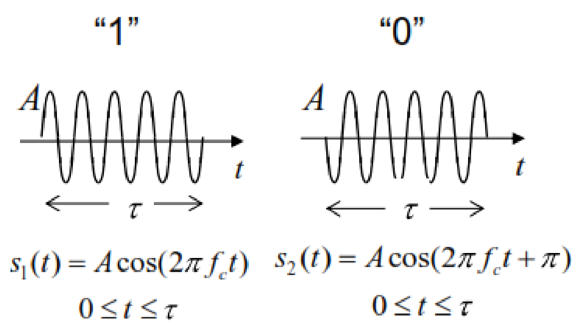
Binary Phase Shift Keying (BPSK)

Signal Representation:

Send: $s_1(t) = Acos(2\pi f_c t)$ if the information bit is "1";

Send: $s_2(t) = Acos(2\pi f_c t + \pi)$

$s_2(t) = -Acos(2\pi f_c t)$ if the information bit is "0";



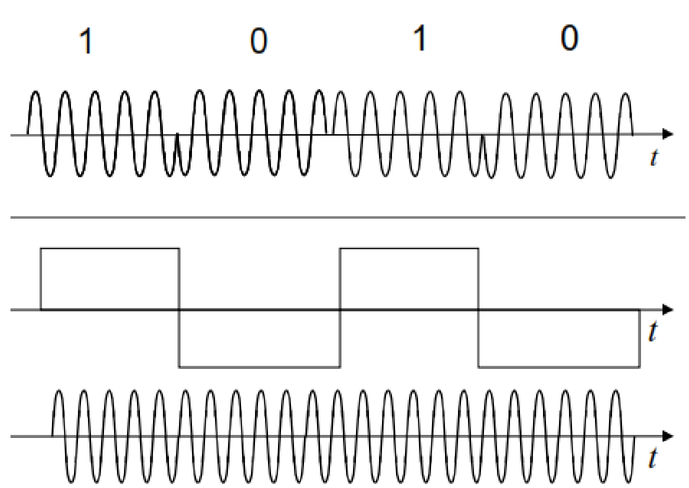
(τ is an integer number of $1/f_c$)

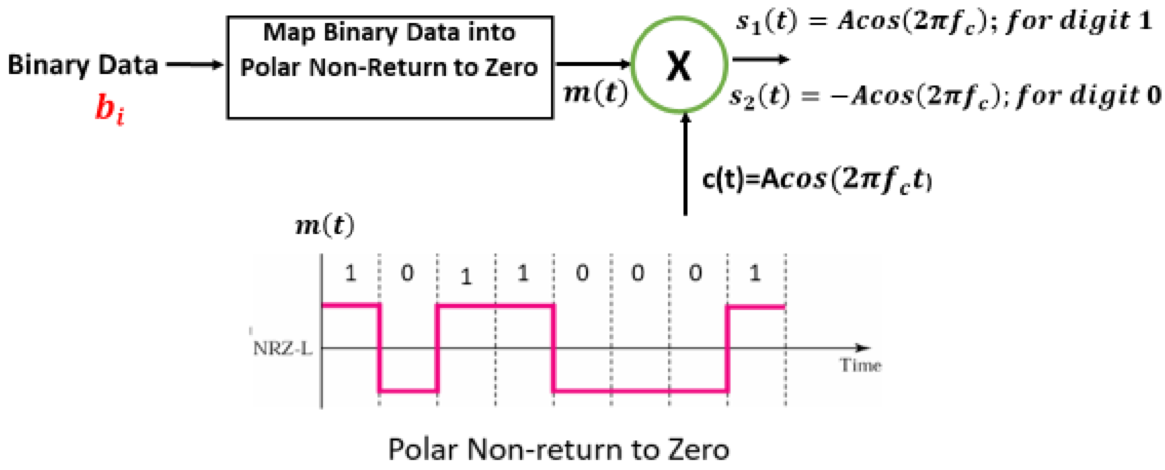
$$s_2(t) = s_1(t+\pi) = -s_1(t)$$

Generating a Binary PSK waveform

The PSK signal is generated by multiplying the polar nonreturn to zero waveform by the sinusoidal carrier.

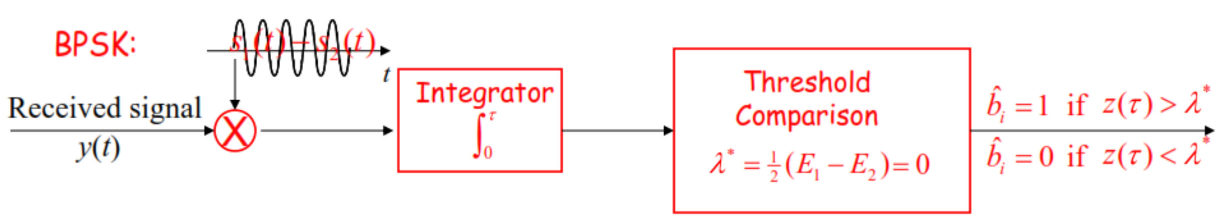
$$s_{BPSK}(t) = s_{BPAM}(t)cos(2\pi f_c t)$$





Binary Phase Shift Keying Generation

The Optimum Receiver



The optimal receiver is also called "coherent receiver" because it must be capable of internally producing a reference signal which is in exact phase and frequency synchronization with the carrier signal$Acos(2\pi f_c t)$.

Probability of Error

Energy of $s_i(t)$: $E_1 = E_2 = \frac{1}{2}A^2\tau$

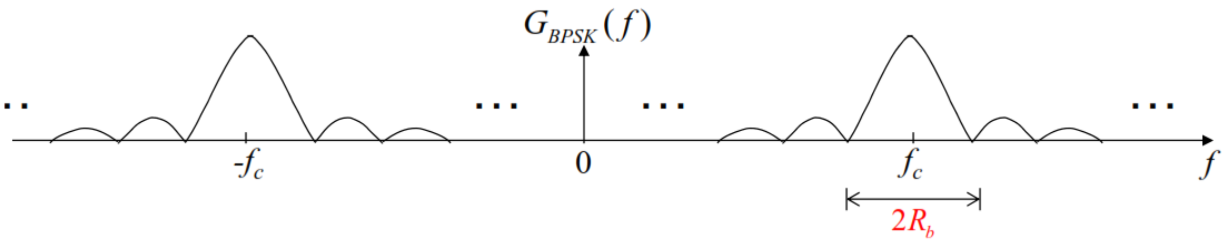
Average Energy per bit: $E_b = \frac{1}{2}(E_1 + E_2) = \frac{1}{2}A^2\tau$

Optimal BER:

$$P_b^* = Q\left(\sqrt{\frac{A^2\tau}{N_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

Power Spectral Density

$$G_{BPSK}(f) = \tfrac{1}{4}[G_{BPAM}(f - f_c) + G_{BPAM}(f + f_c)]$$



Bandwidth

The 90% power bandwidth = $\frac{2}{\tau} = 2R_b$ (twice the data rate); same as that of BASK signal

Final Remarks

- **Advantage: PSK is less susceptible to errors than ASK, while it requires/occupies the same bandwidth as ASK**
- **more efficient use of bandwidth (higher data-rate) are possible, compared to FSK**
- **Disadvantage: more complex signal detection / recovery process, than in ASK and FSK**

Binary Frequency Shift Keying (BFSK)

In binary FSK, the frequency of the carrier signal is varied to represent the binary digits 1 and 0 by two distinct frequencies. The amplitude and frequency remain constant during each bit interval.

Signal Representation (coherent FSK)

Send: $s_1(t) = Acos(2\pi(f_c + \Delta f))t)$ if the information bit is "1";

Send: $s_2(t) = Acos(2\pi(f_c - \Delta f))t)$ if the information bit is "0";

Δf is an offset frequency (from the unmodulated carrier f_c) chosen so that $s_1(t)$ and $s_2(t)$ are orthogonal, i.e.,

$$\int_0^{\tau} s_1(t)s_2(t)dt = 0$$

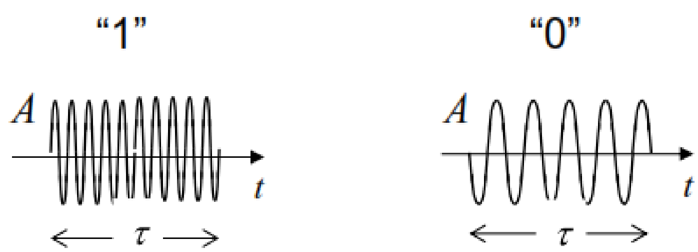
$$\frac{sin(2\pi(2f_c)\tau}{2f_c} + \frac{sin(2\pi(2\Delta f)\tau}{2\Delta f} = 0$$

This condition is satisfied when

$$2f_c = \frac{n}{2\tau} = \frac{nR_b}{2}, n = 1, 2, ...$$

$$2\Delta f = \frac{mR_b}{2}, m = 1, 2, ...$$

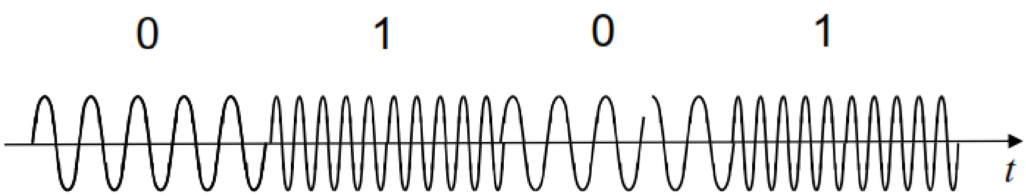
The minimum frequency separation $2\Delta f = R_b/2.$



$s_1(t) = A\cos(2\pi(f_c + \Delta f)t)$ $s_2(t) = A\cos(2\pi(f_c - \Delta f)t)$

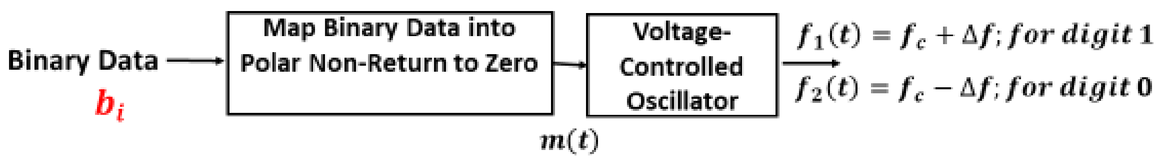
$0 \le t \le \tau$ $0 \le t \le \tau$

(τ is an integer number of $1/(f_c \pm \Delta f)$)

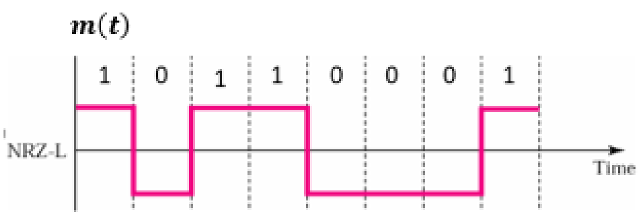


Generating an FSK waveform

Method 1: Using the Voltage Controlled Oscillator (Single Oscillator Method)



$$f_i(t) = f_c + k_f m(t);\ for\ VCO$$



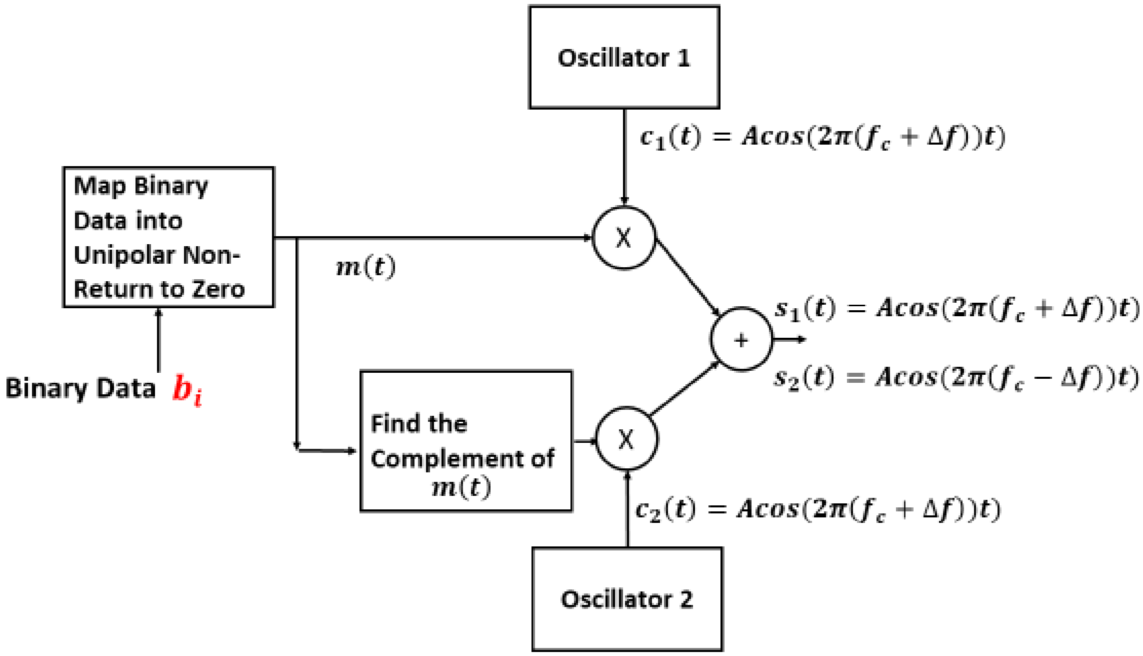
Polar Non-return to Zero

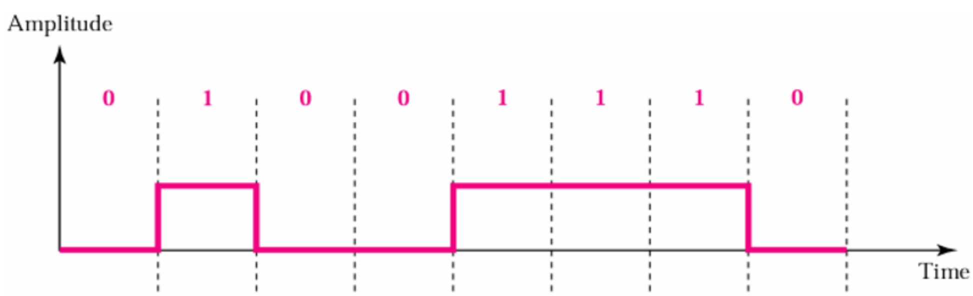
$$s_1(t) = A\cos(2\pi(f_c + \Delta f))t)$$

$$s_2(t) = A\cos(2\pi(f_c - \Delta f))t)$$

ADD THEORY

Method 2: Two-Oscillator Method



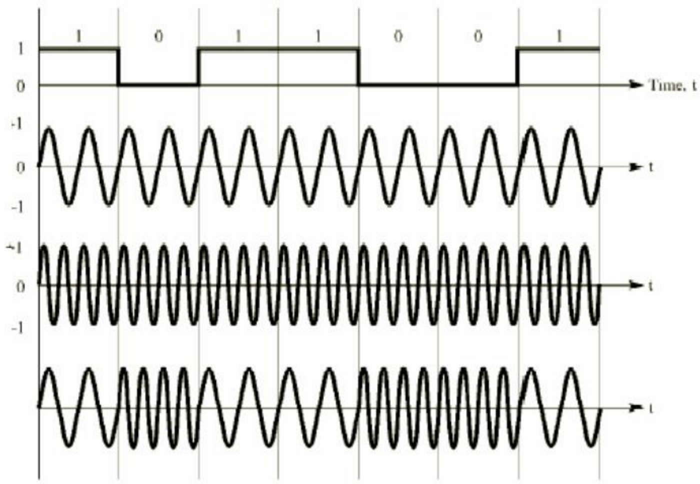


Unipolar Non-Return to Zero Signal

The Binary FSK waveform may be generated as the sum of two binary ASK signals. Let $m(t)$ be the binary unipolar non-return to zero corresponding to the message bits. Then, $\overline{m}(t) = 1 - m(t)$ is the complement to $m(t)$. Therefore, when $m(t) = 1$, $\overline{m}(t) = 0$. The FSK signal is a result of modulating $m(t)$ on the carrier frequencies $Acos(2\pi(f_c + \Delta f))t$ and $\overline{m}(t)$ on $Acos(2\pi(f_c - \Delta f))t$.

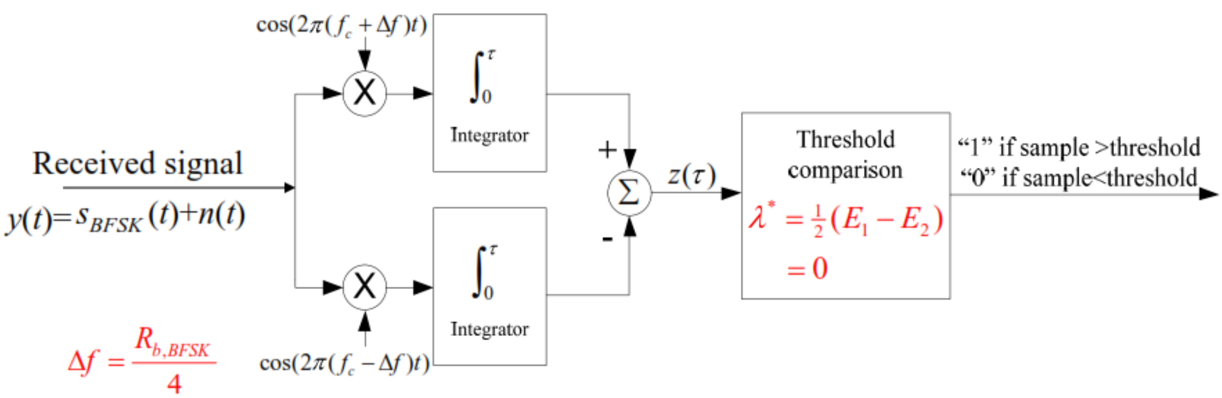
$$s_{BFSK}(t) = ASK\ of\ m(t)\ on\ first\ carrier\ frequency + ASK\ of\ \left(1 - m(t)\right)\ on\ second\ carrier\ frequency$$

$$s_{BFSK}(t) = m(t)Acos\left(2\pi(f_c + \Delta f)\right)t + (1 - m(t))Acos(2\pi(f_c - \Delta f))t$$


1
0
1
1
0
0
1
1
0
Time, t
-1
0
-1
t
-1
0
-1
t
t

Optimum Receiver

The optimum coherent receiver consists of two correlators. The operation of the receiver makes use of the orthogonality condition imposed on the signals $s_1(t)$ and $s_2(t)$. In the absence of noise, if $s_1(t)$ is received, then the ouput of the upper correlator will have a value greater than zero, while the output of the lower correlator is zero. The converse is true when $s_2(t)$ is received. In the presence of noise, the system decides 1 when $z(\tau) > 0$. That is, when the output of the upper correlator is greater than the output of the lower one. Otherwise, it decides 0.



FSK Coherent Receiver

Probability of Error

Energy of $s_i(t)$: $E_1 = E_2 = \frac{1}{2}A^2\tau$

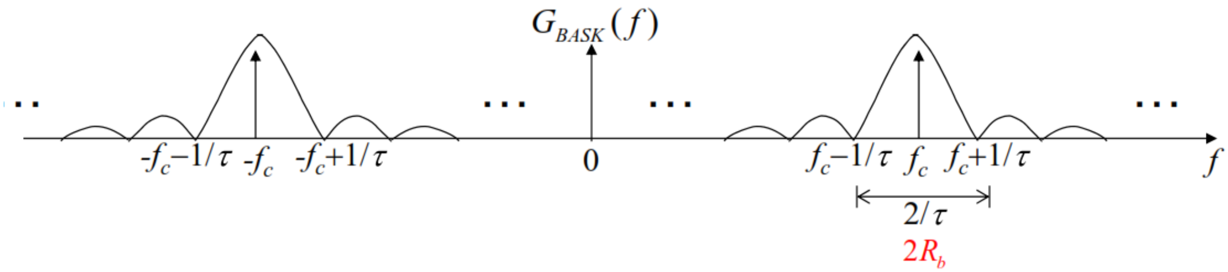
Average Energy per bit: $E_b = \frac{1}{2}(E_1 + E_2) = \frac{1}{2}A^2\tau$

When the signals are orthogonal, i.e., when $\int_0^\tau s_1(t)s_2(t)dt = 0$, the probability of error is given by

$$P_b^* = Q\left(\sqrt{\frac{A^2\tau}{2N_0}}\right) = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

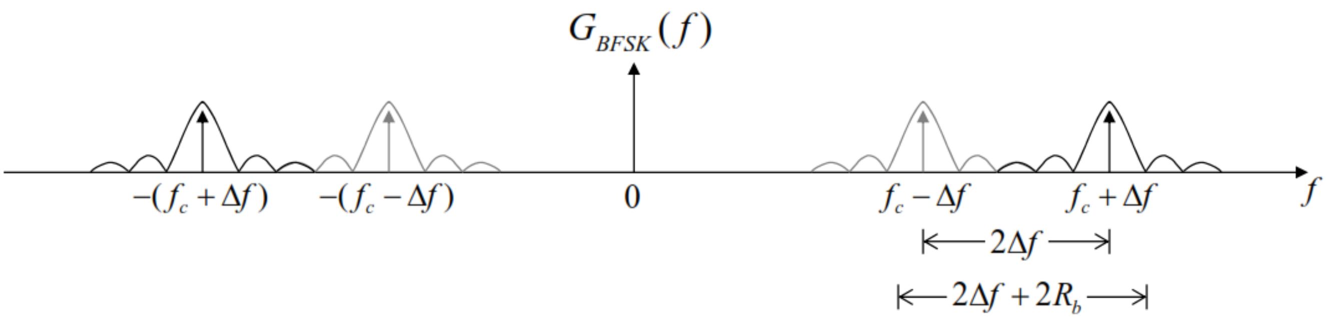
Power Spectral Density

Since the FSK signal is the superposition of two ASK signals on two orthogonal frequencies, the spectrum is also the superposition of that of the ASK signals. We recall that the spectrum of the ASK signal is as shown below



As such, the spectrum of the FSK is as shown below

$$G_{BFSK}(f) = \tfrac{1}{4}[G_{b1,BFSK}(f-(f_c+\Delta f)) + G_{b1,BFSK}(f+(f_c+\Delta f))]$$
$$+\tfrac{1}{4}[G_{b2,BFSK}(f-(f_c-\Delta f)) + G_{b2,BFSK}(f+(f_c-\Delta f))]$$



The required channel bandwidth for 90% in-band power

$$B_{h\_90\%} = 2\Delta f + 2R_b$$

$$\boldsymbol{B.W = (f_1 - f_2) + 2R_b = \frac{R_b}{2} + 2R_b}$$

Final Remarks

- **Advantage: FSK is less susceptible to errors than ASK – receiver looks for specific frequency changes over a number of intervals, so voltage (noise) spikes can be ignored. The FSK signal maintains a constant amplitude.**
- **Disadvantage: FSK spectrum is 2 x ASK spectrum (larger bandwidth)**
- **Application: over voice lines, in high-freq. radio transmission, etc.**

Exercise: Consider the following form of BFSK, in which $\theta_1 \neq \theta_2$

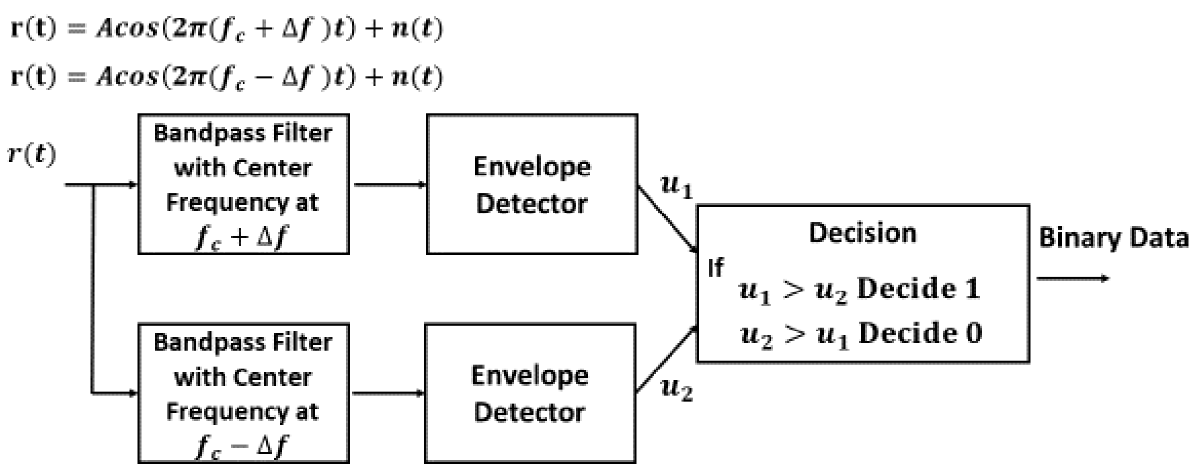
Send: $s_1(t) = A\cos\big(2\pi(f_c + \Delta f)\big)t + \theta_1)$ if the information bit is "1";

Send: $s_2(t) = A\cos(2\pi(f_c - \Delta f))t + \theta_2)$ if the information bit is "0";

Show that the frequency separation that ensures orthogonality is

$$2\Delta f = mR_b, m = 1, 2, ...$$

Non-Coherent Demodulation of Binary FSK



Non-Coherent Binary FSK Demodulation

Quadri-phase Shift Keying (QPSK)

The carrier is transmitted with four possible different carrier phases, allowing each transmitted signal to represent two binary digits.

Signal Representation

$$\text{“1 1”} \quad s_1(t) = A\cos(2\pi f_c t - \pi/4) = +\frac{A}{\sqrt{2}}\cos(2\pi f_c t) + \frac{A}{\sqrt{2}}\sin(2\pi f_c t)$$

$$\text{“1 0”} \quad s_2(t) = A\cos(2\pi f_c t + \pi/4) = +\frac{A}{\sqrt{2}}\cos(2\pi f_c t) - \frac{A}{\sqrt{2}}\sin(2\pi f_c t)$$

$$\text{“0 0”} \quad s_3(t) = A\cos(2\pi f_c t + 3\pi/4) = -\frac{A}{\sqrt{2}}\cos(2\pi f_c t) - \frac{A}{\sqrt{2}}\sin(2\pi f_c t)$$

$$\text{“0 1”} \quad s_4(t) = A\cos(2\pi f_c t + 5\pi/4) = -\frac{A}{\sqrt{2}}\cos(2\pi f_c t) + \frac{A}{\sqrt{2}}\sin(2\pi f_c t)$$

The Modulator:

A QPSK signal can be decomposed into a sum of two PSK signals; an in-phase component and a quadrature component. The serial to parallel converter splits the incoming data sequence into two sequences that consist of the odd (A_k) and even bits (B_k) of the main sequence. The odd bit stream sequence modulates the in-phase carrier, while the even bit stream sequence modulates the quadrature carrier.

$$s_{QPSK}(t) = d_I \frac{A}{\sqrt{2}}\cos(2\pi f_c t) + d_Q \frac{A}{\sqrt{2}}\sin(2\pi f_c t)$$

$$d_I = \begin{cases} 1 & \text{if } b_{2i-1} = 1 \\ -1 & \text{if } b_{2i-1} = 0 \end{cases} \qquad d_Q = \begin{cases} 1 & \text{if } b_{2i} = 1 \\ -1 & \text{if } b_{2i} = 0 \end{cases}$$

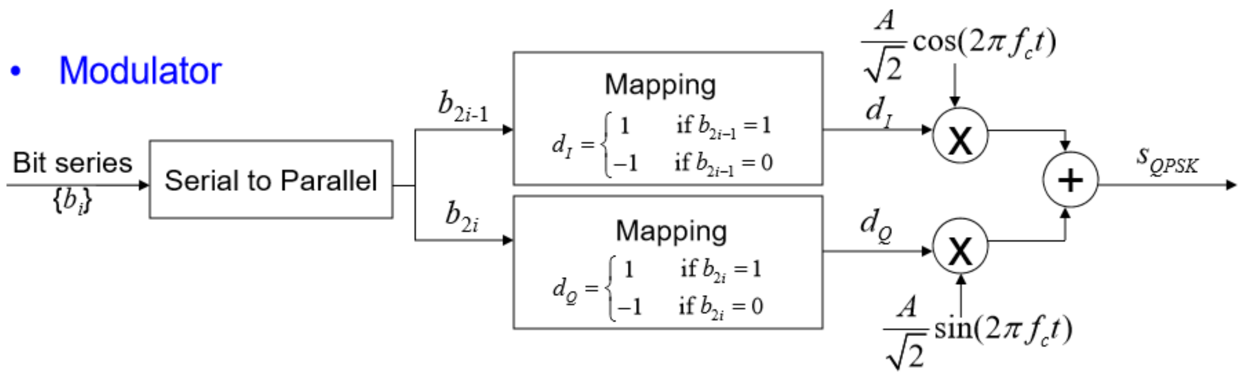
The composite signal $s_{QPSK}(t)$ is sent through the channel, where

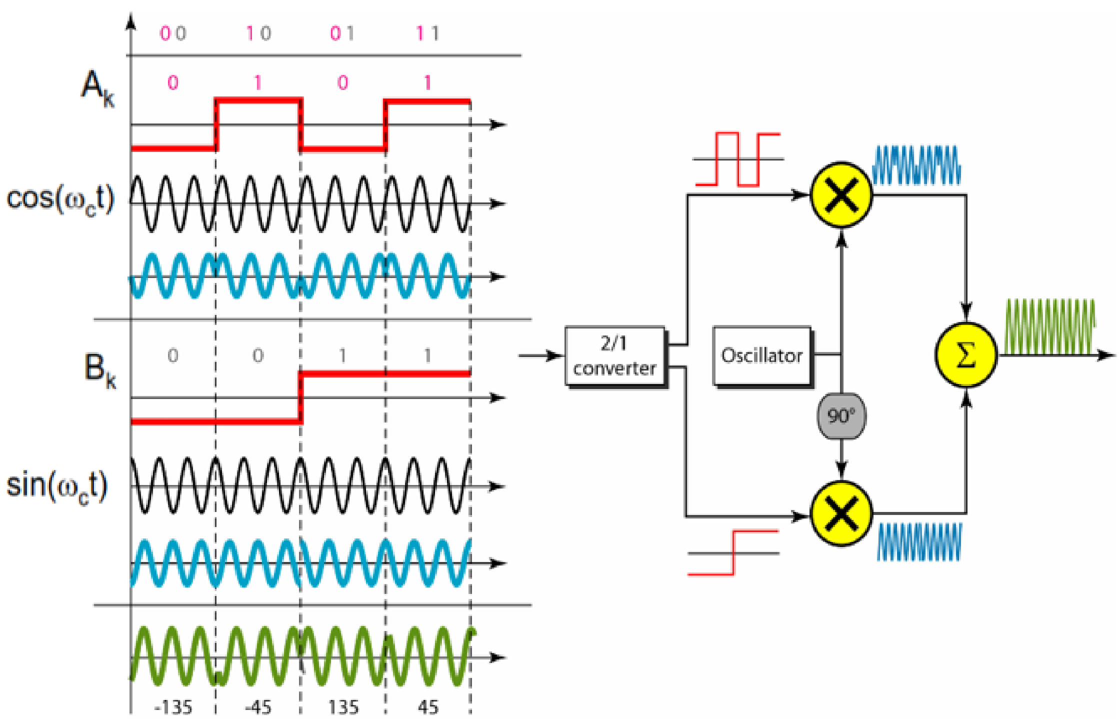
$$s_{QPSK}(t) = A_k cos(2\pi f_c t) + B_k sin(2\pi f_c t)$$

$$s_{QPSK}(t) = Binary\ PSK\ on\ cos(2\pi f_c t) + Binary\ PSK\ on\ sin(2\pi f_c t)$$

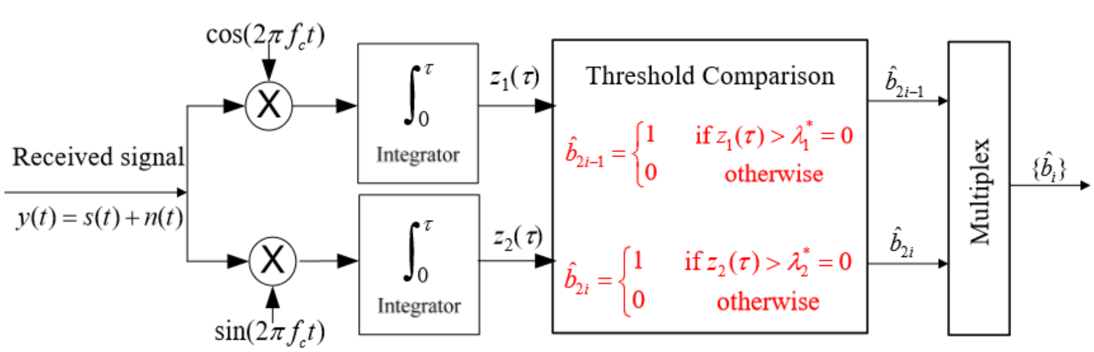
Both of the two BPSK signals occupy the same bandwidth.

- **Modulator**





Optimum Receiver



$$z_1(\tau) = \int_0^\tau [s_{QPSK}(t) + n(t)]cos(2\pi f_c t)dt$$

$$z_1(\tau) = \int_0^\tau [A_k cos(2\pi f_c t) + B_k sin(2\pi f_c t) + n(t)]cos(2\pi f_c t)dt$$

$$z_2(\tau) = \int_0^\tau [A_k cos(2\pi f_c t) + B_k sin(2\pi f_c t) + n(t)]sin(2\pi f_c t)dt$$

$\boldsymbol{z_1(\tau)} = \frac{A_k}{2}\tau + N_1 = \boldsymbol{\frac{Ad_I}{2\sqrt{2}}\tau + N_1}$ signal component proportional to d_i

$\boldsymbol{z_2(\tau)} = \frac{B_k}{2}\tau + N_2; = \boldsymbol{\frac{Ad_Q}{2\sqrt{2}}\tau + N_2}$ signal component proportional to d_q

The odd and even bits can be recovered using the decision rules

$$d_I = \begin{cases} 1, & z_1(\tau) \geq 0 \\ -1, & z_1(\tau) < 0 \end{cases} \Longrightarrow b_{2i-1} = \begin{cases} 1, & d_I = 1 \\ 0, & d_I = -1 \end{cases}; \qquad \text{Odd sequence}$$

$$d_Q = \begin{cases} 1, & z_2(\tau) \geq 0 \\ -1, & z_2(\tau) < 0 \end{cases} \Longrightarrow b_{2i} = \begin{cases} 1, & d_Q = 1 \\ 0, & d_Q = -1 \end{cases}; \qquad \text{Even sequence}$$

Remark: Note that τ, in the formulation above, is the symbol duration, which amounts to two serial data bits. Here, $\boldsymbol{R_s = \frac{1}{\tau} = \frac{R_b}{2}}$**.**

Probability of Error

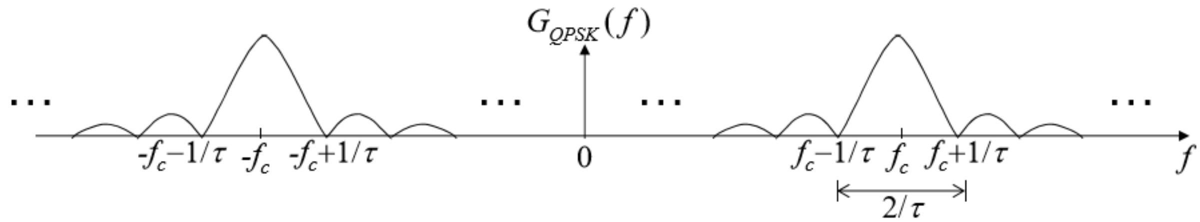
The bit error probability is twice the symbol error probability, and is given as (will also be derived in the next chapter when we consider M-ary PSK).

$$P_b^* = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) = Q\left(\sqrt{\frac{E_s}{N_0}}\right)$$

This is the same as that for binary PSK, provided that both message bits have the same energy. The advantage of QPSK is that it is more bandwidth efficient than BPSK (can transmit twice the data rate)

Power Spectral Density

The power spectral density has the same shape as that for BPSK (the QPSK is the sum of two BPSK signals one modulated on $cos(2\pi f_c t)$ and the other on $sin(2\pi f_c t)$.



Bandwidth

The symbol duration for each BPSK modulator is τ and the symbol rate $\boldsymbol{R_s = \frac{1}{\tau} = \frac{R_b}{2}}$**.**

Bandwidth of QPSK = $\boldsymbol{\frac{2}{\tau} = R_b}$

Therefore, for QPSK, we can transmit **twice** the data that we can transmitted using BPSK. In other words, if a channel bandwidth W is available, then with BPSK, we can transmit $W/2$ bits/sec, while for QPSK, we can transmit W bits/sec.

Spectral Efficiency: Define the spectral efficiency of a digital modulation system as

$$\gamma = \frac{Data\ Rate\ (bit|sec)}{Required\ Channel\ Bandwidth} = \frac{R_b}{B_T}$$

The following table summarizes the spectral efficiency for the bandpass modulation schemes considered in this section.

	Bandwidth Efficiency (90% in-band power)
Binary ASK	0.5
Binary FSK	$0.5 \cdot \frac{1}{1+\Delta f / R_b}$
Binary PSK	0.5
QPSK	1

Final Remarks

- **Advantage: higher data rate than in PSK (2 bits per bit interval), while bandwidth occupancy remains the same**
- **4-PSK can easily be extended to 8-PSK, i.e. n-PSK**
- **However, higher rate PSK schemes are limited by the ability of equipment to distinguish small differences in phase**