# 8 Digital modulation

## 8.1 Model for a digital transmission system

A digital transmission model over an additive White Gaussian noise channel (AWGN) is illustrated in Fig. 37.

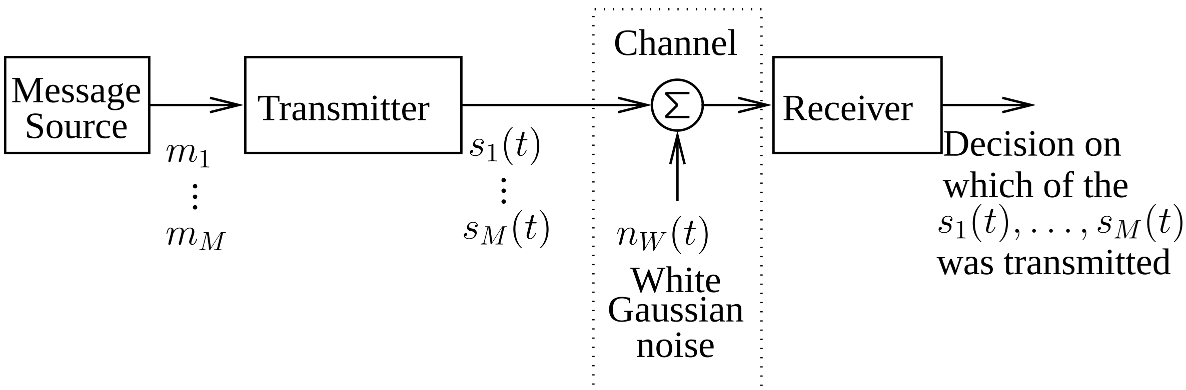


Figure 37: Digital transmission model

Assume that the message source emits one symbol $m$ every $T$ seconds, $m \in \{m_1, \ldots, m_M\}$. The transmitter produces a signal $s_i(t)$ of duration $T$ seconds as the representation of the symbol $m_i$ generated by the message source.

## 8.2 Signal spaces

The Signal Space formalism is a geometric interpretation of continuous time signals as points in an $N$-dimensional space whose axis are orthonormal functions.

### a) The Hilbert space $L^2(a,b)$ and its properties

Let us consider the space formed by all real or complex time functions $x(t)$ that are square integrable on a given time interval $[a,b]$, i.e. $\int_a^b |x(t)|^2 dt < \infty$. This space is called a Hilbert space and it is denoted by $L^2(a,b)$. It is a vector space of infinite dimensions, where the vectors are the functions $x(t)$, and the scalars are the complex numbers if the functions are complex, or the real numbers if the functions are real.

Properties of Hilbert space :

- Closure under addition

$$x(t), y(t) \in L^2(a,b) \quad \Longrightarrow \quad x(t) + y(t) \in L^2(a,b)$$

- Closure under scalar multiplication

$$x(t) \in L^2(a,b) \quad \alpha: \text{ scalar} \quad \Longrightarrow \quad \alpha x(t) \in L^2(a,b)$$

- Inner product:

$$< x(t), y(t) >= \int_a^b x(t) y^*(t) dt \text{ (scalar)} \quad \forall\, x(t), y(t) \in L^2(a,b)$$

norm: $||x(t)|| = (< x(t), x(t) >)^{1/2}$

If $< x(t), y(t) >= 0$, then $x(t)$ is **orthogonal** to $y(t)$ or $x(t)$ and $y(t)$ are orthogonal.

- Span:

The set of functions which can be formed by all the linear combinations of $x_i(t)$, $\sum_{i=1}^{M} \alpha_i x_i(t)$ where $\alpha_i$ are scalars is called the **span** of $x_1(t), \ldots, x_M(t)$ and is denoted span $\{x_1(t), \ldots, x_M(t)\}$. It is a subspace of $L^2(a,b)$.

- Linear independence:

Let $x_1(t), \ldots, x_M(t) \in L^2(a,b)$.

$x_1(t), \ldots, x_M(t)$ are linearly independent if and only if

$$\sum_{i=1}^{M} \alpha_i x_i(t) = 0 \quad \Longrightarrow \quad \alpha_1 = \alpha_2 = \ldots = \alpha_M = 0$$

Otherwise the functions are linearly dependent and at least one function is a linear combination of the others, i.e.

$$x_{i_0}(t) = -\sum_{i=1}^{M} \frac{\alpha_i}{\alpha_{i_0}} x_i(t)$$

- Complete orthogonal set/orthogonal basis for a Hilbert space:

infinite sequence of orthogonal elements in $L^2(a,b)$ $\{u_1(t), u_2(t), \ldots\}$ such that for any $x(t) \in L^2(a,b)$, $x(t)$ can be expressed as

$$x(t) = \sum_{i=1}^{\infty} \alpha_i u_i(t)$$

It can be shown that $\alpha_i = \frac{<x(t), u_i(t)>}{||u_i(t)||^2}$. If furthermore $||u_i(t)|| = 1$, then $\{u_1(t), u_2(t), \ldots\}$ is a complete orthonormal set.

Ex: A basis for $L^2(a,b)$ is $\left\{e^{j\frac{2\pi kt}{b-a}}, k = 0, \pm 1, \pm 2, \ldots\right\}$. A basis for $L^2(-1,1)$ is formed by the Legendre polynomials

$$P_0(t) = 1 \qquad P_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2-1)^n \quad n = 1, 2, \ldots$$

- Projection of $x(t)$ on span $\{u_1(t), u_2(t), \ldots, u_n(t)\}$:

  Let $\{u_1(t), u_2(t), \ldots\}$ be a complete orthonormal set of $L^2(a,b)$. Let $x(t) \in L^2(a,b)$. The projection of $x(t)$ on span$\{u_1(t), u_2(t), \ldots, u_n(t)\}$ is defined as

  $$\hat{x}_n(t) = \sum_{i=1}^{n} < x(t), u_i(t) > u_i(t)$$

  Furthermore

  $$||\hat{x}_n(t)|| \le ||x(t)||$$
  $$\lim_{n\to\infty} ||x(t) - \hat{x}_n(t)|| = 0$$

  Therefore $x(t)$ can be approximated by $\hat{x}_n(t)$, the projection of $x(t)$ on span$\{u_1(t), u_2(t), \ldots, u_n(t)\}$ arbitrarily close.

- Cauchy Schwarz's inequality:

  $$| < x(t), y(t) > |^2 \le ||x(t)||^2 ||y(t)||^2 \quad \forall\, x(t), y(t) \in V \quad \text{(V: vector space)}$$

  with equality when $y(t) = Kx(t)$. Thus applied to $L^2(a,b)$, we obtain

  $$\left|\int_a^b x(t)y^*(t)\right|^2 \le \left(\int_a^b |x(t)|^2 dt\right)\left(\int_a^b |y(t)|^2\, dt\right)$$

**b) Gram-Schmidt Orthogonalization Procedure:**

Let $\{x_1(t), x_2(t), \ldots, x_M(t)\}$ be a linear independent set of functions in $L^2(a,b)$. Let

$$u_1(t) = x_1(t)$$
$$v_1(t) = \frac{u_1(t)}{||u_1(t)||}$$

and

$$u_k(t) = x_k(t) - \sum_{i=1}^{k-1} < x_k(t), v_i(t) > v_i(t) = x_k(t) - \sum_{i=1}^{k-1} \frac{< x_k(t), u_i(t) >}{||u_i(t)||^2} u_i(t)$$

$$v_k(t) = \frac{u_k(t)}{||u_k(t)||}$$

Then

- $\{u_1(t), \ldots, u_M(t)\}$ is an orthogonal set
- $\{v_1(t), \ldots, v_M(t)\}$ is an orthonormal set
- span $\{x_1(t), \ldots, x_k(t)\}$ = span $\{u_1(t), \ldots, u_k(t)\}$ = span $\{v_1(t), \ldots, v_k(t)\}$

Let $\{x_1(t), x_2(t), \ldots, x_M(t)\}$ be a linear **dependent** set of functions in $L^2(a, b)$, then Gram Schmidt procedure yields the orthonormal basis $\{v_1(t), \ldots, v_N(t)\}$ where $N < M$.

### c) Signal space or signal constellation of a modulation

Let $s_1(t), s_2(t), \ldots, s_M(t)$ be the $M$ possible transmitted signals corresponding to a $M$-ary modulation. Let $\phi_1(t), \phi_2(t), \ldots, \phi_N(t)$ ($N \leq M$) be an orthonormal basis such that span $\{\phi_1(t), \phi_2(t), \ldots, \phi_N(t)\}$ = span $\{s_1(t), s_2(t), \ldots, s_M(t)\}$ obtained for example using Gram-Schmidt procedure. The signal space/signal constellation of the modulation is the $N$-dimensional space of axis $\phi_1(t), \ldots, \phi_N(t)$ with the signals $s_1(t), \ldots, s_M(t)$ represented by points with coordinates $(s_{i1}, s_{i2}, \ldots, s_{iN})^T$ in the $N$-dimensional space, where

$$s_{ij} =< s_i(t), \phi_j(t) >$$

Since $s_i(t) = \sum_{i=1}^{N} s_{ij}\phi_j(t)$, $s_i(t)$ is completely characterized by the vector $\boldsymbol{s}_i = (s_{i1}, s_{i2}, \ldots, s_{iN})^T$ and the functions $\{\phi_1(t), \ldots, \phi_N(t)\}$. Note that the set $\{\phi_1(t), \ldots, \phi_N(t)\}$ that can be used to represent $s_1(t), \ldots, s_M(t)$ is not unique.

Example: QPSK modulation

$$s_i(t) = \begin{cases} \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t + (2i-1)\frac{\pi}{4}\right) & 0 \leq t \leq T \quad i = 1, 2, 3, 4 \\ 0 & \text{elsewhere.} \end{cases}$$

Since $s_i(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t) \cos\left((2i-1)\frac{\pi}{4}\right) - \sin(2\pi f_c t) \sin\left((2i-1)\frac{\pi}{4}\right)$, by inspection

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t) \quad 0 \leq t \leq T$$
$$\phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t) \quad 0 \leq t \leq T$$

Assuming $f_c \gg 1$ or $f_c = \frac{n}{T}$, using Gram Schmidt orthogonalization procedure, an equivalent

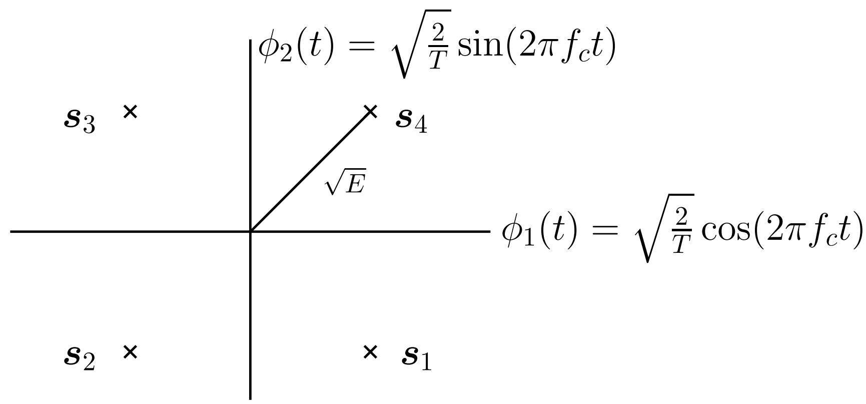


Figure 38: Signal space for QPSK

signal space for QPSK is obtained which is a rotated version of Fig. 38 as illustrated in Fig. 39.

$$||s_1(t)||^2 = \int_0^T \frac{2E}{T}\cos^2\left(2\pi f_c t + \frac{\pi}{4}\right)dt = \int_0^T \frac{E}{T}dt + \int_0^T \frac{E}{T}\cos\left(4\pi f_c t + \frac{\pi}{2}\right)dt = E$$

since the second integral is approximately zero if $f_c \gg 1$ or is equal to zero if $f_c = \frac{n}{T}$.

$$\phi_1^{'}(t) = \frac{u_1(t)}{||u_1(t)||} = \frac{s_1(t)}{||s_1(t)||} = \sqrt{\frac{2}{T}}\cos\left(2\pi f_c t + \frac{\pi}{4}\right)$$

$$\begin{aligned}
< s_2(t), \phi_1^{'}(t) > &= \int_0^T \sqrt{\frac{2E}{T}}\cos\left(2\pi f_c t + \frac{3\pi}{4}\right)\sqrt{\frac{2}{T}}\cos\left(2\pi f_c t + \frac{\pi}{4}\right)dt \\
&= \frac{\sqrt{E}}{T}\left[\int_0^T \cos\left(4\pi f_c t + \pi\right)dt + \int_0^T \cos\left(\frac{\pi}{2}\right)dt\right] \\
&= \frac{\sqrt{E}}{T}\int_0^T \cos\left(4\pi f_c t + \pi\right)dt \\
&= 0 \quad \text{since } f_c \gg 1
\end{aligned}$$

$$u_2(t) = s_2(t) - < s_2(t), \phi_1^{'}(t) > \phi_1^{'}(t) = s_2(t)$$

$$\phi_2^{'}(t) = \frac{u_2(t)}{||u_2(t)||} = \frac{s_2(t)}{||s_2(t)||} = \sqrt{\frac{2}{T}}\cos\left(2\pi f_c t + \frac{3\pi}{4}\right)$$

$$< s_3(t), \phi_1^{'}(t) > = \int_0^T \sqrt{\frac{2E}{T}}\cos\left(2\pi f_c t + \frac{5\pi}{4}\right)\sqrt{\frac{2}{T}}\cos\left(2\pi f_c t + \frac{\pi}{4}\right)dt$$

$$
\begin{aligned}
&= \frac{\sqrt{E}}{T}\left[\int_0^T \cos\left(4\pi f_c t + \frac{3\pi}{2}\right) dt + \int_0^T \cos(\pi)\, dt\right] = -\sqrt{E} \\
< s_3(t), \phi_2^{'}(t) > &= \int_0^T \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t + \frac{5\pi}{4}\right) \sqrt{\frac{2}{T}} \cos\left(2\pi f_c t + \frac{3\pi}{4}\right) dt \\
&= \frac{\sqrt{E}}{T}\left[\int_0^T \cos(4\pi f_c t)\, dt + \int_0^T \cos\left(\frac{\pi}{2}\right) dt\right] = 0
\end{aligned}
$$

$$
\begin{aligned}
u_3(t) &= s_3(t) - < s_3(t), \phi_1^{'}(t) > \phi_1^{'}(t) - < s_3(t), \phi_2^{'}(t) > \phi_2^{'}(t) \\
&= \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t + \frac{5\pi}{4}\right) + \sqrt{E}\sqrt{\frac{2}{T}} \cos\left(2\pi f_c t + \frac{\pi}{4}\right) = 0 \\
\phi_3^{'}(t) &= \frac{u_3(t)}{||u_3(t)||} = 0
\end{aligned}
$$

$$
\begin{aligned}
< s_4(t), \phi_1^{'}(t) > &= \int_0^T \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t + \frac{7\pi}{4}\right) \sqrt{\frac{2}{T}} \cos\left(2\pi f_c t + \frac{\pi}{4}\right) dt \\
&= \frac{\sqrt{E}}{T}\left[\int_0^T \cos(4\pi f_c t)\, dt + \int_0^T \cos\left(\frac{3\pi}{2}\right) dt\right] = 0 \\
< s_4(t), \phi_2^{'}(t) > &= \int_0^T \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t + \frac{7\pi}{4}\right) \sqrt{\frac{2}{T}} \cos\left(2\pi f_c t + \frac{3\pi}{4}\right) dt \\
&= \frac{\sqrt{E}}{T}\left[\int_0^T \cos\left(4\pi f_c t + \frac{5\pi}{2}\right) dt + \int_0^T \cos(\pi)\, dt\right] = -\sqrt{E} \\
< s_4(t), \phi_3^{'}(t) > &=< s_4(t), 0 >= 0
\end{aligned}
$$

$$
\begin{aligned}
u_4(t) &= s_4(t) - < s_4(t), \phi_1^{'}(t) > \phi_1^{'}(t) - < s_4(t), \phi_2^{'}(t) > \phi_2^{'}(t) - < s_4(t), \phi_3^{'}(t) > \phi_3^{'}(t) \\
&= \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t + \frac{7\pi}{4}\right) + \sqrt{E}\sqrt{\frac{2}{T}} \cos\left(2\pi f_c t + \frac{3\pi}{4}\right) = 0 \\
\phi_4^{'}(t) &= 0
\end{aligned}
$$

Note that Fig. 38 is the classical representation of the signal space of QPSK.

**d) Application of signal space in communication**

- The signal space with the basis $\{\phi_i(t)\}_{i=1,\ldots,N}$ completely characterizes the signals $s_i(t)$ and offers a representation of time varying functions (continuous time functions) using vectors (discrete components).

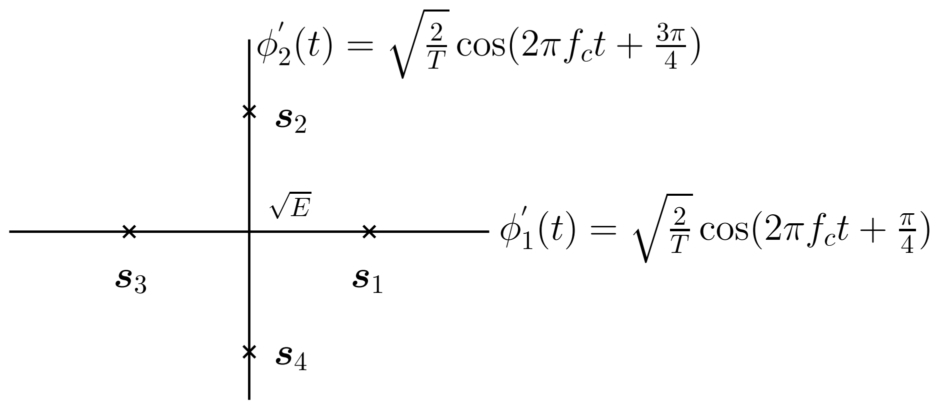


Figure 39: Rotated signal space for QPSK

- The signal space provides a convenient and simplified way to find the energy and Euclidean distance between signals. Note that the Euclidean distance between signals plays a key role in communication over AWGN channels.

$$\begin{aligned}
E_i &\triangleq \int_a^b \left[\sum_{j=1}^N s_{ij}\phi_j(t)\right|^2 dt = \sum_{j=1}^N \sum_{k=1}^N s_{ij}s_{ik}^* \int_a^b \phi_j(t)\phi_k^*(t)dt = \sum_{j=1}^N |s_{ij}|^2 \\
&= ||\boldsymbol{s}_i||^2 \\
< s_i(t), s_k(t) > &\triangleq \int_a^b \left(\sum_{j=1}^N s_{ij}\phi_j(t)\right)\left(\sum_{r=1}^N s_{kr}^*\phi_r^*(t)\right) dt \\
&= \sum_{j=1}^N \sum_{r=1}^N s_{ij}s_{kr}^* \int_a^b \phi_j(t)\phi_r^*(t)dt = \sum_{j=1}^N s_{ij}s_{kj}^* \\
&= \boldsymbol{s}_i^T \boldsymbol{s}_k^* = \boldsymbol{s}_k^\dagger \boldsymbol{s}_i
\end{aligned}$$

Let $\hat{r}_N(t) = \sum_{i=1}^N r_i\phi_i(t)$ be the projection of $r(t)$ on span $\{s_1(t), \ldots, s_M(t)\}$. The squared Euclidean distance between $\hat{r}_N(t)$ and $s_k(t)$ is given by

$$\begin{aligned}
||\hat{r}_N(t) - s_k(t)||^2 &\triangleq \int_a^b |\hat{r}_N(t) - s_k(t)|^2 dt = \int_a^b \left(\hat{r}_N(t) - s_k(t)\right)^* \sum_{i=1}^N (r_i - s_{ki})\,\phi_i(t)dt \\
&= \sum_{i=1}^N (r_i - s_{ki}) \left[\int_a^b \hat{r}_N^*(t)\phi_i(t)dt - \int_a^b s_k^*(t)\phi_i(t)dt\right] \\
&= \sum_{i=1}^N (r_i - s_{ki})\,(r_i^* - s_{ki}^*) = \sum_{i=1}^N |r_i - s_{ki}|^2 \\
&= ||\boldsymbol{r}_N - \boldsymbol{s}_k||^2 \quad \text{with } \boldsymbol{r}_N = [r_1, \ldots, r_N]^T.
\end{aligned}$$

## 8.3 Optimal receiver principles

Let $r(t)$, $0 \le t \le T$ be the received signal. Let $s_1(t), \ldots, s_M(t)$ be the possible transmitted signals. Define the hypothesis

$$H_m : r(t) = f\left(s_m(t), n(t)\right)$$

where $f(\cdot)$ is a function that depends on the type of channels considered and $n(t)$ is a noise term. For example an additive white Gaussian noise (AWGN) channel is characterized by

$$H_m : r(t) = s_m(t) + n_W(t)$$

Based on $r(t)$, $0 \le t \le T$, the receiver has to make a decision on which signal $s_m(t)$ was transmitted. The **optimal** receiver (in the minimum probability sense) is the receiver that minimizes the average probability of error.

Let $P(k|m) = P\left(\text{decide on } s_k(t)|s_m(t) \text{ transmitted}\right)$, be the probability that the receiver decides that $s_k(t)$ was transmitted when in fact $s_m(t)$ was transmitted. The probability of error when $s_m(t)$ is transmitted is given by

$$P_e(m) = \sum_{\substack{k=1 \\ k \neq m}}^{M} P(k|m) = 1 - P_c(m)$$

where $P_c(m)$ is the probability of correct decision when $s_m(t)$ was transmitted. Assume that the probability of transmitting $s_m(t)$ is $P(m)$ (a-priori probability), then the average probability of error is

$$\begin{aligned}
P_e &= \sum_{m=1}^{M} P_e(m)P(m) \quad \text{Bayes's rule} \\
&= \sum_{m=1}^{M} \sum_{\substack{k=1 \\ k \neq m}}^{M} P(k|m)P(m) \\
&= \sum_{m=1}^{M} \left[1 - P_c(m)\right] P(m) = \sum_{m=1}^{M} P(m) - \sum_{m=1}^{M} P_c(m)P(m) = 1 - P_c
\end{aligned}$$

where $P_c$ is the average probability of correct decision. The optimal receiver minimizes $P_e$ or maximizes $P_c$. To find the optimal receiver, we use the signal space representation. First we will consider a finite dimensional approximation to the problem, i.e. we will consider a $K$-dimensional representation of $r(t)$ and $s_m(t)$ and then see what happens when $K$ tends to infinity. Let $\{u_1(t), u_2(t), \ldots\}$ be a complete orthonormal set of $L^2(0,T)$. Let $\hat{r}_K(t)$ and $\hat{s}_{mK}(t)$ be the

$K$-dimensional approximations of $r(t)$ and $s_m(t)$ given by

$$\hat{r}_K(t) = \sum_{i=1}^{K} r_i u_i(t) \qquad\qquad r_i =< r(t), u_i(t) >$$

$$\hat{s}_{mK}(t) = \sum_{i=1}^{K} s_{mi} u_i(t) \qquad\qquad s_{mi} =< s_m(t), u_i(t) >$$

$\hat{r}_K(t)$ and $\hat{s}_{mK}(t)$ are respectively represented by $\boldsymbol{r}_K = (r_1, \ldots, r_K)^T$ and $\boldsymbol{s}_m = (s_{m1}, \ldots, s_{mK})^T$. Let $\mathcal{Z}^K$ denote the $K$-dimensional space of all possible observation vectors (received vectors) called the observation space. $\mathcal{Z}^K \subseteq \mathbb{R}^K$. A decision rule for deciding which $\boldsymbol{s}_m$ was transmitted is a partition of the observation space into $M$ disjoint sets $Z_i, i = 1, \ldots, M$ such that

$$Z_1 \cup Z_2 \ldots \cup Z_M = \mathcal{Z}^K \qquad Z_i \cap Z_m = \emptyset \quad i \neq m$$

The optimum decision rule in the minimum probabilityi of error sense is obtained by finding the partition that minimizes the average probability of error $P_e$, or equivalently that maximizes the average probability of correct decision $P_c$.

The probability of correct decision when $s_m(t)$ was transmitted is given by

$$\begin{aligned} P_c(m) &= P\left(\text{decide on } s_m(t) | s_m(t) \text{ was transmitted}\right) \\ &= P\left[\boldsymbol{r}_K \in Z_m | H_m\right] \\ &= \int_{Z_m} p\left(\boldsymbol{r}_K | H_m\right) d\boldsymbol{r}_K \end{aligned}$$

where $p\left(\boldsymbol{r}_K | H_m\right)$ is the joint probability density function of $\boldsymbol{r}_K$ when hypothesis $H_m$ is satisfied.

$$P_c = \sum_{m=1}^{M} P_c(m) P(m) = \sum_{m=1}^{M} \int_{Z_m} p\left(\boldsymbol{r}_K | H_m\right) P(m) d\boldsymbol{r}_K$$

$P_c$ should be maximized over all possible partitions of the observation space, thus it is maximized if for $m = 1, \ldots, M$, $\int_{Z_m} p\left(\boldsymbol{r}_K | H_m\right) P(m) d\boldsymbol{r}_K$ are maximized, i.e. $Z_m$ should consists of all points of $\mathcal{Z}$ such that

$$p\left(\boldsymbol{r}_K | H_m\right) P(m) > p\left(\boldsymbol{r}_K | H_i\right) P(i) \quad \forall\, i \neq m$$

hence

$$\int_{Z_m} p\left(\boldsymbol{r}_K | H_m\right) P(m) d\boldsymbol{r}_K > \int_{Z_m} p\left(\boldsymbol{r}_K | H_i\right) P(i) d\boldsymbol{r}_K \quad \forall\, i \neq m$$

In this way, each term $p\left(\boldsymbol{r}_K | H_m\right) P(m)$ is integrated over that portion of the observation space where it maximizes its contribution to $P_c$. Thus the optimal decision rule based on $\boldsymbol{f}_K$ is

Decide that $s_{m_o}(t)$ was transmitted if the received signal $r(t)$ is such that

$$P\left(\boldsymbol{r}_K|H_{m_o}\right)P(m_o) = \max_{m=1,2,\ldots,M}\left\{p\left(\boldsymbol{r}_K|H_m\right)P(m)\right\}$$

Equivalently the optimum decision rule can be written as

$$m_o = \arg\max_m \left[p\left(\boldsymbol{r}_K|H_m\right)P(m)\right]$$

*Interpretation:*

Since $p\left(\boldsymbol{r}_K\right)$ is independent of $m$, an equivalent decision rule can be written as

$$\arg\max_m \frac{p\left(\boldsymbol{r}_K|H_m\right)P(m)}{p\left(\boldsymbol{r}_K\right)}$$

From Bayes's rule $\frac{p(\boldsymbol{r}_K|H_m)P(m)}{p(\boldsymbol{r}_K)} = p\left(m|\boldsymbol{r}_K\right)$ is an a posteriori probability. Therefore the Maximum A Posteriori (MAP) rule is

$$\arg\max_m p\left(m|\boldsymbol{r}_K\right)$$

The MAP decision rule maximizes the average probability of correct decision, or equivalently minimizes the average probability of error.

If $P(m)$ are not available or are all equal, the Maximum-Likelihood (ML) decision rule is defined as

$$\arg\max_m p\left(\boldsymbol{r}_K|H_m\right)$$

$p\left(\boldsymbol{r}_K|H_m\right)$ is called the likelihood. The ML decision rule minimizes the average probability of error if all a priori probabilities are equal, i.e. if $P(m) = \frac{1}{M}$.

Remark: Since $p\left(m|\boldsymbol{r}_K\right) \geq 0$ and $\ln(\cdot)$ is monotonic increasing, an equivalent MAP rule is

$$\arg\max_m \ln\left[p\left(m|\boldsymbol{r}_K\right)\right]$$

and an equivalent ML decision rule is

$$\arg\max_m \ln\left[p\left(\boldsymbol{r}_K|H_m\right)\right]$$

$\ln\left[p\left(\boldsymbol{r}_K|H_m\right)\right]$ is called the log-likelihood.

Note that for general random channels, the likelihood or log-likelihood functions may be difficult to calculate. Furthermore, to implement the optimal decision rules, the statistic of the channel needs to be known.

## 8.4 Optimal receiver for additive white Gaussian noise (AWGN) channels

The received signal over an AWGN channel is given by

$$H_m : r(t) = s_m(t) + n_W(t), \quad t \in [0, T], \quad m = 1, 2, \ldots, M$$

where $n_W(t)$ is a zero mean white Gaussian real random process with power spectral density $\frac{N_0}{2}$ Watts/Hz. For sake of simplicity let us assume that $s_1(t), \ldots, s_M(t)$ are real.

Let $\{u_1(t), u_2(t), \ldots\}$ be a complete orthonormal set of $L^2(0, T)$. The set of functions

$$\{s_1(t), s_2(t), \ldots, s_M(t), u_1(t), u_2(t), \ldots\}$$

is complete in $L^2(0, T)$ (in the sense that any signals in $L^2(0, T)$ can be represented as a linear combination of elements of that set) but is not orthonormal. Gram Schmidt procedure performed on $\{s_1(t), s_2(t), \ldots, s_M(t), u_1(t), u_2(t), \ldots\}$ yields a complete orthonormal set $\{v_1(t), v_2(t), \ldots\}$ of $L^2(0, T)$ such that

$$\text{span}\,\{v_1(t), \ldots, v_N(t)\} = \text{span}\,\{s_1(t), \ldots, s_M(t)\} \tag{25}$$

where $N \leq M$.

Let $K \geq N$ and

$$\begin{aligned}
\hat{r}_K(t) &= \sum_{i=1}^{K} r_i v_i(t) & r_i &=< r(t), v_i(t) > \\
\hat{s}_{mK}(t) &= \sum_{i=1}^{K} s_{mi} v_i(t) & s_{mi} &=< s_m(t), v_i(t) > \\
&= \sum_{i=1}^{N} s_{mi} v_i(t) & &\text{since } s_{mi} = 0 \quad \forall\, i > N \\
&= \sum_{i=1}^{\infty} s_{mi} v_i(t) = s_m(t) \\
\hat{n}_K(t) &= \sum_{i=1}^{K} n_i v_i(t) & n_i &=< n_W(t), v_i(t) >
\end{aligned}$$

Proof that $s_{mi} = 0 \,\forall\, i > N$: Let $i > N$, from (25)

$$s_{mi} =< s_m(t), v_i(t) >=< \sum_{k=1}^{N} \alpha_k v_k(t), v_i(t) >= \sum_{k=1}^{N} \alpha_k < v_k(t), v_i(t) >= 0 \quad \text{since } i > N$$

**a) Optimal receiver (MAP receiver)**

The MAP decision rule based on $\boldsymbol{r}_K$ is given by

$$\arg\max_m \; p\left(m|\boldsymbol{r}\right)$$

or equivalently since $p(\boldsymbol{r}_K)$ is independent of $H_m$

$$\arg\max_m \; \left[p\left(\boldsymbol{r}|H_m\right) P(m)\right] \tag{26}$$

where $\boldsymbol{r}_K = [r_1, \ldots, r_K]^T$. Since $r(t) = s_m(t) + n_W(t)$

$$r_i = s_{mi} + n_i \qquad \text{where} \quad n_i =< n_W(t), v_i(t) >= \int_0^T n_W(t)v_i(t)dt$$

Since $n_W(t)$ is a Gaussian random process, $n_1, n_2, \ldots$ are jointly Gaussian random variables with mean

$$E\left[n_i\right] = E\left[\int_0^T n_W(t)v_i(t)dt\right] = \int_0^T E\left[n_W(t)\right]v_i(t)dt = 0 \quad i = 1, 2, \ldots \quad \text{(since } n_W(t) \text{ is zero mean)}$$

and covariance

$$\begin{aligned}
E\left[n_i n_j\right] &= E\left[\int_0^T n_W(t)v_i(t)dt \int_0^T n_W(u)v_j(u)du\right] \\
&= \int_0^T \int_0^T E\left[n_W(t)n_W(u)\right] v_i(t)v_j(u)dtdu \\
&= \int_0^T \int_0^T \frac{N_0}{2}\delta(t-u)v_i(t)v_j(u)dtdu \\
&= \frac{N_0}{2}\int_0^T v_i(t)v_j(t)dt = \begin{cases} \frac{N_0}{2} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}
\end{aligned}$$

Hence the $n_i$'s are zero mean independent Gaussian random variables with variance $\frac{N_0}{2}$. Hence the $r_i$'s ($r_i = s_{mi} + n_i$) are independent Gaussian random variables with mean $s_{mi}$ and variance $\frac{N_0}{2}$. Therefore

$$p\left(\boldsymbol{r}_K|H_m\right) = p\left(\boldsymbol{r}_K|\boldsymbol{r}_K = \boldsymbol{s}_{mK} + \boldsymbol{n}_K\right)$$
$$= \prod_{i=1}^{K}\left[\frac{1}{\sqrt{2\pi\frac{N_0}{2}}}\exp\left\{-\frac{\left(r_i - s_{mi}\right)^2}{2\frac{N_0}{2}}\right\}\right]$$

$$= \left( \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi \frac{N_0}{2}}} \exp\left\{ -\frac{(r_i - s_{mi})^2}{2\frac{N_0}{2}} \right\} \right) \left( \prod_{i=N+1}^{K} \frac{1}{\sqrt{2\pi \frac{N_0}{2}}} \exp\left\{ -\frac{r_i^2}{2\frac{N_0}{2}} \right\} \right) \text{since } \forall\, i > N,\, s_{mi} = 0$$

Since the second product is independent of $m$, an equivalent MAP decision rule is

$$\arg\max_m \left[ p\left(\boldsymbol{r}_N | H_m\right) P(m) \right] \tag{27}$$

Thus $r_1, r_2, \ldots, r_N$ form a sufficient statistic. An equivalent MAP decision rule is

$$\arg\max_m \ln\left[ p\left(\boldsymbol{r}_N | H_m\right) P(m) \right]$$

which means that the receiver decides $s_m(t)$ was transmitted if

$$\begin{aligned}
\ln\left[ p\left(\boldsymbol{r}_N | H_m\right) P(m) \right] &> \ln\left[ p\left(\boldsymbol{r}_N | H_k\right) P(k) \right] \\
-\sum_{i=1}^{N} \frac{(r_i - s_{mi})^2}{N_0} + \ln\bigl(P(m)\bigr) &> -\sum_{i=1}^{N} \frac{(r_i - s_{ki})^2}{N_0} + \ln\bigl(P(k)\bigr) \\
\sum_{i=1}^{N} (r_i - s_{mi})^2 - N_0 \ln\bigl(P(m)\bigr) &< \sum_{i=1}^{N} (r_i - s_{ki})^2 - N_0 \ln\bigl(P(k)\bigr) \\
||\boldsymbol{r}_N - \boldsymbol{s}_m||^2 - N_0 \ln\bigl(P(m)\bigr) &< ||\boldsymbol{r}_N - \boldsymbol{s}_k||^2 - N_0 \ln\bigl(P(k)\bigr) \\
||\hat{r}_N(t) - s_m(t)||^2 - N_0 \ln\bigl(P(m)\bigr) &< ||\hat{r}_N(t) - s_k(t)||^2 - N_0 \ln\bigl(P(k)\bigr)
\end{aligned}$$

Since

$$\sum_{i=1}^{N} (r_i - s_{mi})^2 = ||\boldsymbol{r}_N - \boldsymbol{s}_m||^2 = ||\hat{r}_N(t) - s_m(t)||^2$$

Thus a MAP decision rule over an AWGN channel is

$$\begin{aligned}
&\arg\min_m \left[ ||\boldsymbol{r}_N - \boldsymbol{s}_m||^2 - N_0 \ln\bigl(P(m)\bigr) \right] \\
&\arg\min_m \left[ ||\hat{r}_N(t) - s_m(t)||^2 - N_0 \ln\bigl(P(m)\bigr) \right]
\end{aligned}$$

The receiver implementing the MAP decision rule illustrated in Fig. 40 needs only to consider the sufficient statistic $r_1, r_2, \ldots, r_N$.

The complexity of the implementation of Fig. 40 grows linearly with the dimension of the signal space (with $N$).

To find an equivalent decision rule, let us calculate $||\boldsymbol{r}_N - \boldsymbol{s}_m||^2 = ||\hat{r}_N(t) - s_m(t)||^2$.

$$||\hat{r}_N(t) - s_m(t)||^2 = \int_0^T \left(\hat{r}_N(t) - s_m(t)\right)^2 dt$$

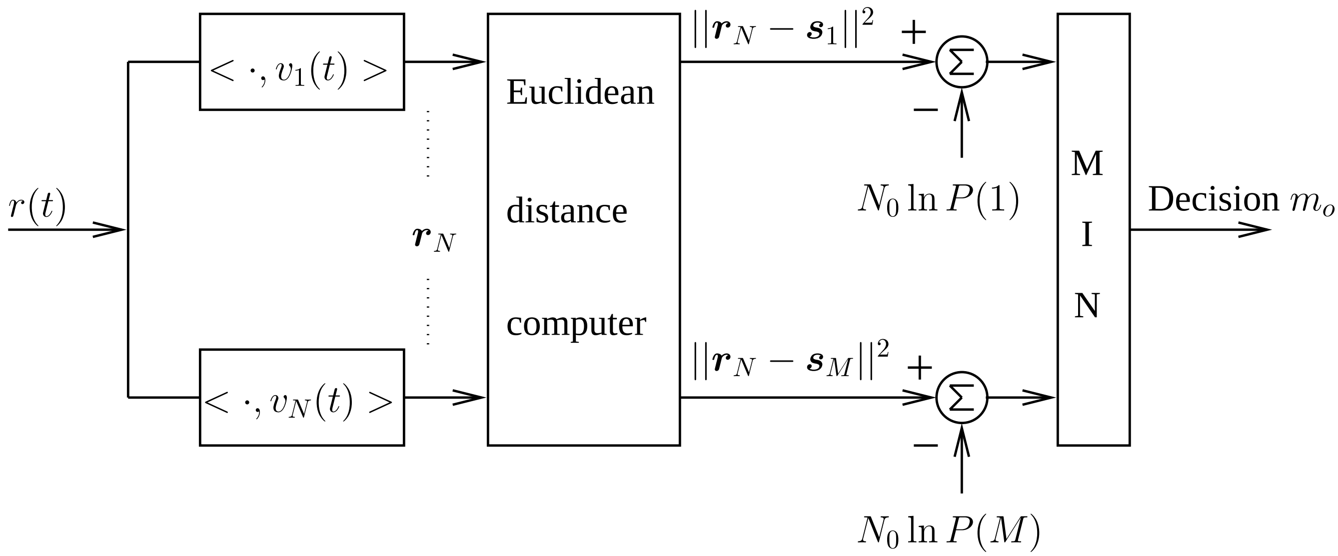


Figure 40: Receiver implementing the MAP decision rule

$$
\begin{aligned}
&= \int_0^T \hat{r}_N^2(t)dt + \int_0^T s_m^2(t)dt - 2\int_0^T \hat{r}_N(t)s_m(t)dt \\
&= \int_0^T \hat{r}_N^2(t)dt + E_m - 2 < \hat{r}_N(t), s_m(t) >
\end{aligned}
$$

Since the first term is independent of $m$ and $< \hat{r}_N(t), s_m(t) >=< r(t), s_m(t) >$ an equivalent MAP decision rule is

$$
\arg\max_m \left[ < r(t), s_m(t) > - \frac{E_m}{2} + \frac{N_0}{2} \ln\big(P(m)\big) \right]
$$

which is illustrated in Fig. 41.

*Proof that* $< \hat{r}_N(t), s_m(t) >=< r(t), s_m(t) >$*:*

Recall that span $\{s_1(t), \ldots, s_M(t)\}$ = span $\{v_1(t), \ldots, v_N(t)\}$

$$
\begin{aligned}
< r(t), s_m(t) > &=< \sum_{i=1}^{\infty} r_i v_i(t), \sum_{k=1}^{N} s_{mk} v_k(t) >= \sum_{i=1}^{\infty}\sum_{k=1}^{N} r_i s_{mk} < v_i(t), v_k(t) > \\
&= \sum_{i=1}^{N}\sum_{k=1}^{N} r_i s_{mk} < v_i(t), v_k(t) > \quad \text{since } < v_i(t), v_k(t) >= 0 \text{ for } i > N \text{ when } k \le N \\
&=< \sum_{i=1}^{N} r_i v_i(t), \sum_{k=1}^{N} s_{mk} v_k(t) >=< \hat{r}_N(t), s_m(t) >
\end{aligned}
$$

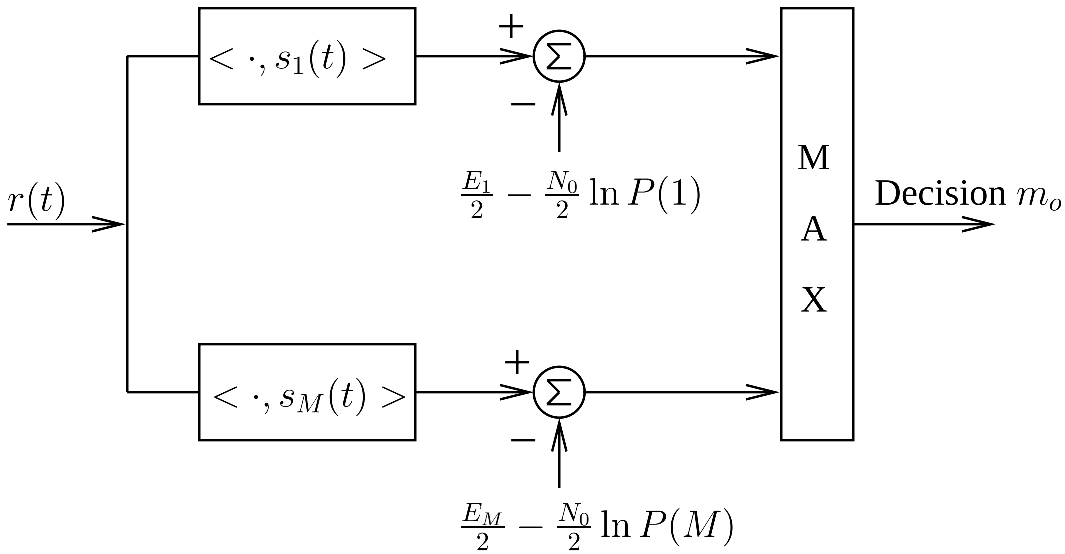


Figure 41: Receiver implementing the MAP decision rule (second implementation)

Both receivers need to implement an inner product. This inner product can be obtained either using a correlator as illustrated in Fig. 42 or using a matched filter $h_i(t)$ sampled at $T$ as illustrated in Fig. 43, where

$$h_i(t) = \begin{cases} v_i(T-t), & t \geq 0 \\ 0, & t < 0. \end{cases}$$

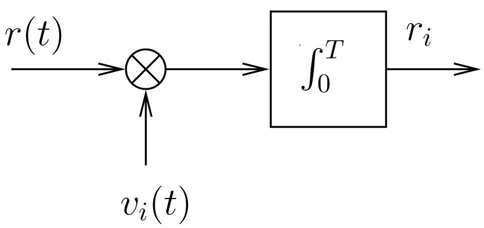


Figure 42: Correlator

The output of the matched filter sampled at $t = T$ is given by

$$\int_{-\infty}^{\infty} r(t) h_i(T-t) dt = \int_0^T r(t) v_i(t) dt = r_i$$

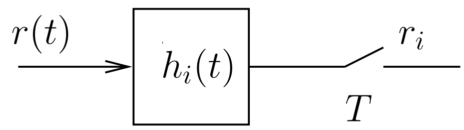


Figure 43: Inner product implemented using a matched filter

**b) ML receiver**

The ML decision rule is

$$\arg\max_m \left[ p\left( \boldsymbol{r}_K | H_m \right) \right] \tag{28}$$

One form of the MAP decision rule is given by (26). Thus the only difference between (28) and (26) is the term $P(m)$. Therefore performing the steps as in a), the ML decision rule over AWGN channels is given by

$$\arg\min_m \left[ ||\boldsymbol{r}_N - \boldsymbol{s}_m||^2 \right]$$
$$\arg\min_m \left[ ||\hat{r}_N(t) - s_m(t)||^2 \right]$$

or equivalently

$$\arg\max_m \left[ < r(t), s_m(t) > - \frac{E_m}{2} \right]$$

Similar to the MAP receiver, the ML receiver needs only to consider the sufficient statistic $r_1, r_2, \ldots, r_N$. Receivers implementing the ML decision rule are illustrated in Fig. 44 and Fig. 45.

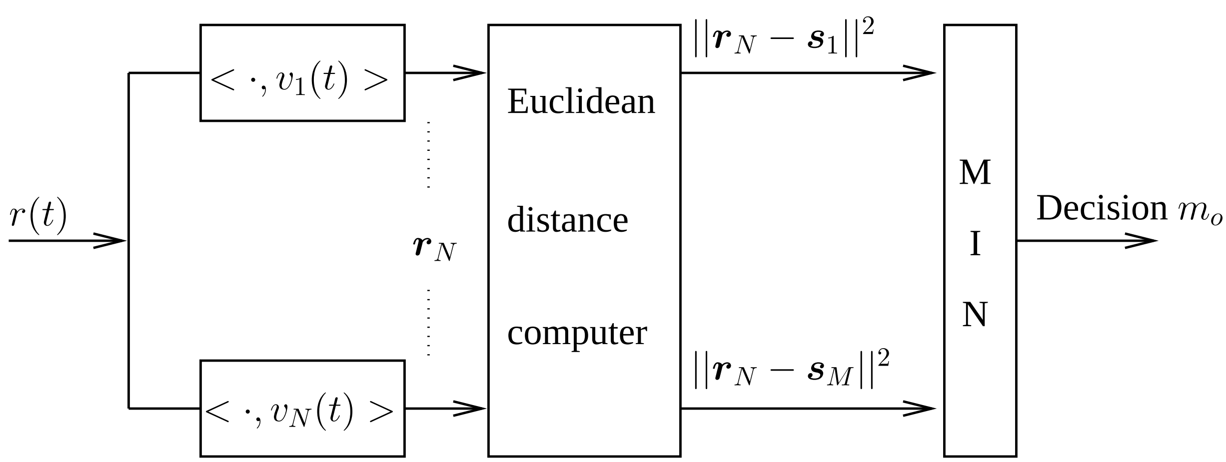


Figure 44: Receiver implementing the ML decision rule

The complexity of the implementation of Fig. 44 also grows linearly with the dimension of the signal space (with $N$).

*Interpretation of the ML receiver:*

Fig. 44 shows that the ML receiver computes the scalar $r_1, r_2, \ldots, r_N$ that forms the sufficient statistic and groups them into a vector $\boldsymbol{r}_N = [r_1, \ldots, r_N]^T$. Computing the Euclidean distances and selecting the minimum one is equivalent to partitioning the signal space span $\{s_1(t), s_2(t), \ldots, s_M(t)\}$ into $M$ disjoint regions $V_1, \ldots, V_M$ called the Voronoi regions such that

$$V_1 \cup V_2 \ldots \cup V_M = \text{span } \{s_1(t), s_2(t), \ldots, s_M(t)\}$$
$$V_i \cap V_k = \emptyset \quad i \neq k$$

where $V_i$ is the set of all the elements of span $\{s_1(t), s_2(t), \ldots, s_M(t)\}$ which are closer in terms of Euclidean distance to $s_i(t)$ than any other $s_k(t)$, $k \neq i$. The vector $\boldsymbol{r}_N$ must belong to one and only one Voronoi region. The receiver decides that $s_{m_o}(t)$ was transmitted if $\boldsymbol{r}_N \in V_{m_o}$. Note that $V_i = Z_i$ if $\forall\, k\ P(k) = \frac{1}{M}$ (equally likely signals).

*Example : QPSK modulation*

$$s_m(t) = a_m \sqrt{\frac{2}{T}} \cos(2\pi f_c t) + b_m \sqrt{\frac{2}{T}} \sin(2\pi f_c t) \quad 0 \leq t \leq T \quad m = 1, 2, \ldots, M$$

where

$$\begin{aligned} a_1 &= b_1 = \sqrt{E} \\ a_2 &= -b_2 = -\sqrt{E} \\ a_3 &= b_3 = -\sqrt{E} \\ a_4 &= -b_4 = \sqrt{E} \end{aligned}$$

The Voronoi regions are formed by the four quadrants of $\mathbb{R}^2$. The decision variables are

$$r_1 =< r(t), \sqrt{\frac{2}{T}} \cos(2\pi f_c t) > \qquad r_2 =< r(t), \sqrt{\frac{2}{T}} \sin(2\pi f_c t) >$$

The decision rule is as follows:

$$\begin{aligned} r_1 \geq 0, r_2 \geq 0 &\Longrightarrow m_o = 1 \\ r_1 < 0, r_2 \geq 0 &\Longrightarrow m_o = 2 \\ r_1 < 0, r_2 < 0 &\Longrightarrow m_o = 3 \\ r_1 \geq 0, r_2 < 0 &\Longrightarrow m_o = 4 \end{aligned}$$

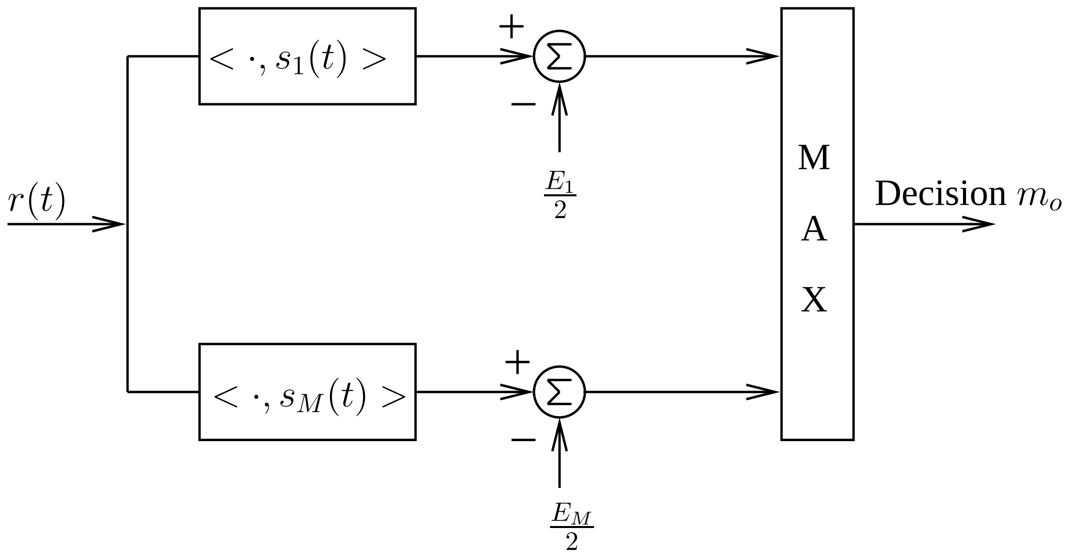


Figure 45: Receiver implementing the ML decision rule (second implementation)

## 8.5 Optimal receiver for binary schemes over an AWGN channel

Assume that the two possible real transmitted signals are $s_1(t), s_2(t) \in L^2(0, T)$ where

$$||s_1(t)||^2 = \int_0^T |s_1(t)|^2 dt = E_1$$
$$||s_2(t)||^2 = \int_0^T |s_2(t)|^2 dt = E_2$$

Let $P(1), P(2)$ be the a-priori probabilities of transmitting $s_1(t)$ and $s_2(t)$. Define the hypotheses $H_1, H_2$ as

$$H_1 : r(t) = s_1(t) + n_W(t) \qquad 0 \le t \le T$$
$$H_2 : r(t) = s_2(t) + n_W(t) \qquad 0 \le t \le T$$

where $n_W(t)$ is a zero mean white Gaussian real random process with power spectral density $\frac{N_0}{2}$ Watts/Hz. Let $\{u_1(t), u_2(t), \ldots\}$ be a complete orthonormal set of $L^2(0, T)$. The set $\{s_1(t), s_2(t), u_1(t), u_2(t), \ldots\}$ is complete but not orthonormal. Gram Schmidt procedure performed on $\{s_1(t), s_2(t), u_1(t), u_2(t), \ldots\}$ yields the complete orthonormal set $\{v_1(t), v_2(t), \ldots\}$, where

$$v_1(t) = \frac{s_1(t)}{||s_1(t)||} = \frac{s_1(t)}{\sqrt{E_1}}$$

$$
\begin{aligned}
v_2(t) &= \frac{s_2(t) - < s_2(t), v_1(t) > v_1(t)}{||s_2(t) - < s_2(t), v_1(t) > v_1(t)||} = \frac{s_2(t) - < s_2(t), \frac{s_1(t)}{\sqrt{E_1}} > \frac{s_1(t)}{\sqrt{E_1}}}{||s_2(t) - < s_2(t), v_1(t) > v_1(t)||} \\
&= \frac{s_2(t) - \sqrt{\frac{E_2}{E_1}}\rho s_1(t)}{\sqrt{\left\| s_2(t) - \sqrt{\frac{E_2}{E_1}}\rho\, s_1(t) \right\|^2}} = \frac{s_2(t) - \sqrt{\frac{E_2}{E_1}}\rho\, s_1(t)}{\sqrt{E_2 + \frac{E_2}{E_1}\rho^2 E_1 - 2\sqrt{\frac{E_2}{E_1}}\rho\sqrt{E_1 E_2}\rho}} \\
&= \frac{1}{\sqrt{1-\rho^2}}\left[\frac{s_2(t)}{\sqrt{E_2}} - \rho\frac{s_1(t)}{\sqrt{E_1}}\right]
\end{aligned}
$$

where $\rho$, the correlation of $s_2(t)$ and $s_1(t)$ is given by

$$
\rho = \frac{< s_2(t), s_1(t) >}{\sqrt{E_1 E_2}} = \frac{1}{\sqrt{E_1 E_2}}\int_0^T s_1(t)s_2(t)dt
$$

Following the methodology of Section 8.4, from (27) the MAP decision rule is given by

$$
\arg\max_{m=1,2} p\left(\boldsymbol{r}_2|H_m\right) P(m)
$$

equivalently written as

$$
p\left(\boldsymbol{r}_2|H_1\right) P(1) \underset{H_2}{\overset{H_1}{\gtrless}} p\left(\boldsymbol{r}_2|H_2\right) P(2)
$$

where $H_i$ indicates the hypothesis corresponding to the decision of the receiver.

$$
\begin{gathered}
\frac{p\left(\boldsymbol{r}_2|H_1\right)}{p\left(\boldsymbol{r}_2|H_2\right)} \underset{H_2}{\overset{H_1}{\gtrless}} \frac{P(2)}{P(1)} = \lambda \quad \text{threshold} \\
\Lambda(\hat{r}_2(t)) \underset{H_2}{\overset{H_1}{\gtrless}} \lambda \\
\Lambda(r(t)) \underset{H_2}{\overset{H_1}{\gtrless}} \lambda \qquad (29)
\end{gathered}
$$

where $\Lambda\left(\hat{r}_2(t)\right) = \Lambda\left(r(t)\right)$ is the likelihood ratio defined as

$$\Lambda\left(\hat{r}_2(t)\right) \triangleq \frac{p\left(\boldsymbol{r}_2|H_1\right)}{p\left(\boldsymbol{r}_2|H_2\right)} = \lim_{K\to\infty} \frac{p\left(\boldsymbol{r}_K|H_1\right)}{p\left(\boldsymbol{r}_K|H_2\right)} \triangleq \Lambda\left(r(t)\right) = \frac{p\left(\boldsymbol{r}|H_1\right)}{p\left(\boldsymbol{r}|H_2\right)}$$

An equivalent decision rule is also obtained by taking the logarithm of (29):

$$\ln\left[\Lambda(r(t))\right] \underset{H_2}{\overset{H_1}{\gtrless}} \ln\lambda$$

Substituting the value of $p\left(\boldsymbol{r}_2|H_1\right)$ and $p\left(\boldsymbol{r}_2|H_2\right)$ and simplifying yields

$$\sum_{i=1}^{2}\left\{(r_i - s_{2i})^2 - (r_i - s_{1i})^2\right\} \underset{H_2}{\overset{H_1}{\gtrless}} N_0 \ln\lambda$$

$$\sum_{i=1}^{2} 2r_i\left(s_{1i} - s_{2i}\right) \underset{H_2}{\overset{H_1}{\gtrless}} N_0\ln\lambda + \sum_{i=1}^{2}\left(s_{1i}^2 - s_{2i}^2\right) \tag{30}$$

$$s_{1i} = \begin{cases} \sqrt{E_1} & i = 1 \\ 0 & \text{else.} \end{cases} \tag{31}$$

$$s_{2i} = \begin{cases} \sqrt{E_2}\rho & i = 1 \\ \sqrt{E_2}\sqrt{1-\rho^2} & i = 2 \\ 0 & \text{else.} \end{cases} \tag{32}$$

Therefore substituting (31-32) into (30) yields

$$r_1\left(\sqrt{E_1} - \sqrt{E_2}\rho\right) - r_2\sqrt{E_2}\sqrt{1-\rho^2} \underset{H_2}{\overset{H_1}{\gtrless}} \left[N_0\ln\lambda + E_1 - E_2\rho^2 - E_2(1-\rho^2)\right]/2 \tag{33}$$

$$< r(t), s_1(t) > -\rho\sqrt{\frac{E_2}{E_1}} < r(t), s_1(t) > - < r(t), s_2(t) > +\rho\sqrt{\frac{E_2}{E_1}} < r(t), s_1(t) >$$

$$\underset{H_2}{\overset{H_1}{\gtrless}} \left[N_0\ln\lambda + E_1 - E_2\right]/2$$

$$< r(t), s_1(t) - s_2(t) > \underset{H_2}{\overset{H_1}{\gtrless}} [N_0 \ln \lambda + E_1 - E_2]/2$$

Let us define $s_\Delta(t) = s_1(t) - s_2(t)$, an equivalent decision rule is

$$\frac{< r(t), s_\Delta(t) >}{||s_\Delta(t)||} \underset{H_2}{\overset{H_1}{\gtrless}} \frac{E_1 - E_2 + N_0 \ln \lambda}{2||s_\Delta(t)||} = \lambda' \tag{34}$$

The optimal receiver structure for binary modulations based on (33) is illustrated in Fig. 46. The

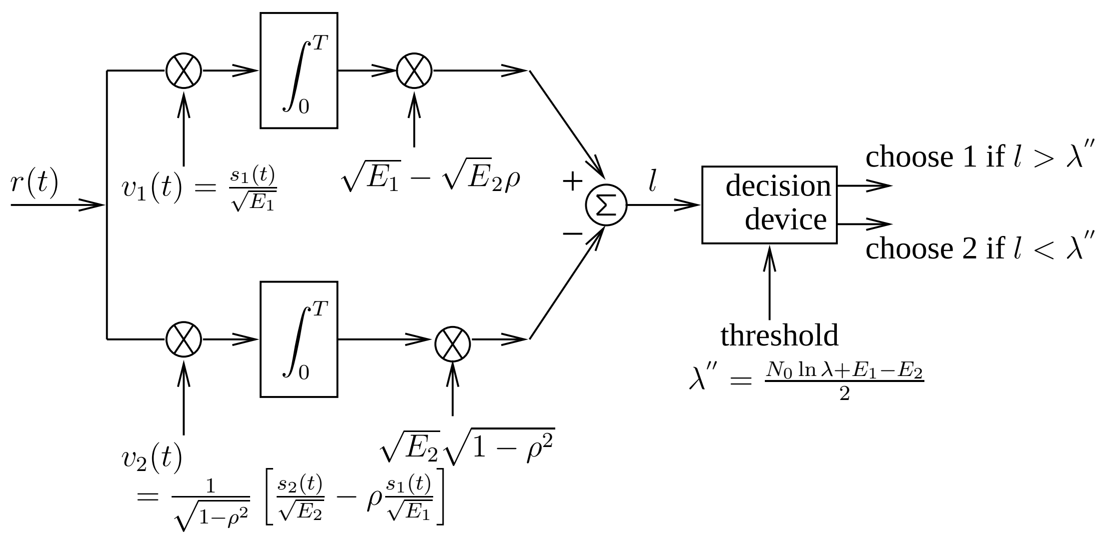


Figure 46: Optimal receiver (MAP receiver) for binary modulations (first implementation)

optimal receiver structure for binary modulation based on (34) is illustrated in Fig. 47 ( correlation receiver) and Fig. 48 (matched filter implementation), where the matched filter $h(t)$ is defined as

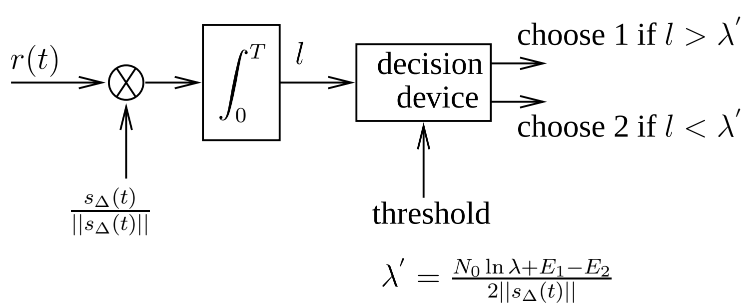


Figure 47: Optimal receiver (MAP receiver) for binary modulation (second implementation)

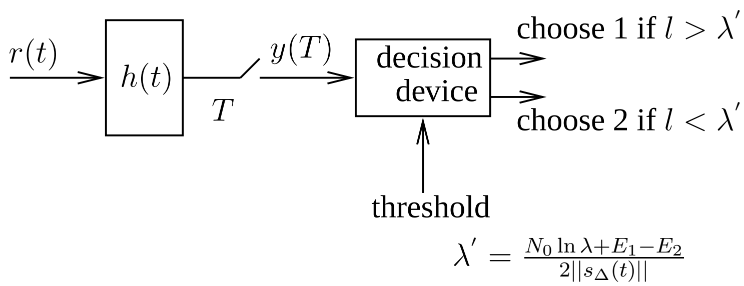


Figure 48: Optimal receiver (MAP receiver) for binary modulation (second implementation)

$$h(t) = \begin{cases} \frac{s_\Delta(t)}{||s_\Delta(t)||} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Remark: In general the optimal receiver structure for $M$-ary transmission ($\{s_1(t), \ldots, s_M(t)\}$ possible transmitted signals) uses $N$ matched filters or correlators ($N \leq M$ such that span $\{v_1(t), \ldots, v_N(t)\} =$ span $\{s_1(t), \ldots, s_M(t)\}$ ). However for binary modulation ($M = 2$), it is possible to use only one correlator or matched filter as shown by the structures of Fig. 47 and Fig. 48.

## 8.6 Probability of symbol error (= probability of bit error) for binary schemes over AWGN channels

Let $P_e$ be the probability of symbol error where the symbols are the $s_i(t)$'s. Since $M = 2$, $P_e = P_{eb}$, where $P_{eb}$ is the probability of bit error. To calculate $P_e$ the decision rule given by (34) will be used. Let $P_e(1)$ be the probability of error when $s_1(t)$ is transmitted. By definition

$$P_e(1) = P\left(r_{s_\Delta} < \lambda' | H_1\right)$$

where

$$r_{s_\Delta} = \frac{< r(t), s_\Delta(t) >}{||s_\Delta(t)||}$$

Under $H_1$ (i.e assuming that $s_1(t)$ was transmitted), $r_{s_\Delta}$ is given by

$$r_{s_\Delta} = \frac{1}{||s_\Delta(t)||} < s_1(t) + n_W(t), s_1(t) - s_2(t) >= \frac{E_1 - \rho\sqrt{E_1 E_2}}{||s_\Delta(t)||} + n_{s_\Delta} = S_1 + n_{s_\Delta}$$

where

$$S_1 = \frac{E_1 - \rho\sqrt{E_1 E_2}}{||s_\Delta(t)||}$$

$$n_{s_\Delta} = \frac{< n_W(t), s_\Delta(t) >}{||s_\Delta(t)||}$$

Since $n_W(t)$ is a zero mean white Gaussian random process, $n_{s_\Delta}$ is a Gaussian random variable with mean

$$E\left[n_{s_\Delta}\right] = E\left[\int_0^T n_W(t)\frac{s_\Delta(t)}{||s_\Delta(t)||}dt\right] = \int_0^T E\left[n_W(t)\right]\frac{s_\Delta(t)}{||s_\Delta(t)||}dt = 0$$

and variance

$$\begin{aligned}
E\left[n_{s_\Delta}^2\right] &= E\left[\int_0^T n_W(t)\frac{s_\Delta(t)}{||s_\Delta(t)||}dt \int_0^T n_W(u)\frac{s_\Delta(u)}{||s_\Delta(u)||}du\right] \\
&= \frac{N_0}{2}\int_0^T\int_0^T \frac{s_\Delta(t)}{||s_\Delta(t)||}\frac{s_\Delta(u)}{||s_\Delta(u)||}\delta(t-u)dt\,du = \frac{N_0}{2}\int_0^T \frac{s_\Delta^2(t)}{||s_\Delta(t)||}dt = \frac{N_0}{2}
\end{aligned}$$

Since $n_{s_\Delta}$ is a zero mean Gaussian random variable with variance $\frac{N_0}{2}$, so is $-n_{s_\Delta}$.

$$\begin{aligned}
P_e(1) &= P\left(S_1 + n_{s_\Delta} < \lambda^{'}\right) = P\left(-n_{s_\Delta} > S_1 - \lambda^{'}\right) \\
&= \int_{S_1-\lambda^{'}}^{\infty} \frac{1}{\sqrt{2\pi\frac{N_0}{2}}}\exp\left\{-\frac{x^2}{2\frac{N_0}{2}}\right\}dx \\
&= \int_{\frac{S_1-\lambda^{'}}{\sqrt{\frac{N_0}{2}}}}^{\infty} \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{u^2}{2}\right\}du \qquad \left(= P\left[\frac{-n_{s_\Delta}}{\sqrt{\frac{N_0}{2}}} > \frac{S_1-\lambda^{'}}{\sqrt{\frac{N_0}{2}}}\right]\right) \\
&\stackrel{\triangle}{=} Q\left[\sqrt{\frac{2}{N_0}}\left(S_1 - \lambda^{'}\right)\right] \\
P_e(1) &= Q\left[\frac{||s_\Delta(t)||}{\sqrt{2N_0}} - \sqrt{\frac{N_0}{2}}\frac{\ln\lambda}{||s_\Delta(t)||}\right]
\end{aligned}$$

since

$$\begin{aligned}
\sqrt{\frac{2}{N_0}}\left(S_1 - \lambda^{'}\right) &= \sqrt{\frac{2}{N_0}}\left\{\frac{E_1 - \rho\sqrt{E_1E_2}}{||s_\Delta(t)||} - \frac{N_0\ln\lambda}{2||s_\Delta(t)||} - \frac{E_1 - E_2}{2||s_\Delta(t)||}\right\} \\
&= \frac{1}{\sqrt{2N_0}||s_\Delta(t)||}\left\{E_1 + E_2 - 2\rho\sqrt{E_1E_2} - N_0\ln\lambda\right\}
\end{aligned}$$

The probability of error when $s_2(t)$ is transmitted is given by

$$P_e(2) = P\left(r_{s_\Delta} > \lambda^{'}|H_2\right) = P\left(S_2 + n_{s_\Delta} > \lambda^{'}\right) = P\left(n_{s_\Delta} > -S_2 + \lambda^{'}\right) = P\left(\frac{n_{s_\Delta}}{\sqrt{\frac{N_0}{2}}} > \frac{-S_2+\lambda^{'}}{\sqrt{\frac{N_0}{2}}}\right)$$

$$= Q\left[\sqrt{\frac{2}{N_0}}\left(-S_2+\lambda^{'}\right)\right] = Q\left[\frac{||s_\Delta(t)||}{\sqrt{2N_0}} + \sqrt{\frac{N_0}{2}}\frac{\ln\lambda}{||s_\Delta(t)||}\right]$$

where

$$S_2 = \frac{\rho\sqrt{E_1E_2} - E_2}{||s_\Delta(t)||}$$

Therefore the probability of (symbol or bit) error is given by

$$\begin{aligned} P_e &= P(1)P_e(1) + P(2)P_e(2) \\ &= P(1)Q\left[\frac{||s_\Delta(t)||}{\sqrt{2N_0}} - \sqrt{\frac{N_0}{2}}\frac{\ln\lambda}{||s_\Delta(t)||}\right] + P(2)Q\left[\frac{||s_\Delta(t)||}{\sqrt{2N_0}} + \sqrt{\frac{N_0}{2}}\frac{\ln\lambda}{||s_\Delta(t)||}\right] \end{aligned}$$

If the signals $s_1(t)$ and $s_2(t)$ are equally likely i.e. $P(1) = P(2) = \frac{1}{2}$ then $\lambda = 1$ and

$$P_e = Q\left[\frac{||s_\Delta(t)||}{\sqrt{2N_0}}\right] \quad \text{(equally likely signals)}$$

$||s_\Delta(t)||$ is the Euclidean distance between the two possible transmitted signals $s_1(t)$ and $s_2(t)$. A better performance is obtained when $||s_\Delta(t)||^2$ is large which corresponds to small $\rho$ for fixed $E_1, E_2$ since

$$||s_\Delta(t)||^2 = E_1 + E_2 - 2\rho\sqrt{E_1E_2}$$

From Cauchy Schwarz's inequality $|\rho|^2 \le \frac{||s_1(t)||^2}{E_1}\frac{||s_2(t)||^2}{E_2}$, thus $-1 \le \rho \le 1$. The minimum value of $\rho$ is $-1$ which corresponds to antipodal signaling, i.e.

$$s_1(t) = -s_2(t) \qquad ||s_\Delta(t)||^2 = \left(\sqrt{E_1} + \sqrt{E_2}\right)^2$$

Assume equal energy signals, i.e. $E_1 = E_2 = E$, then $P_e(1)$ and $P_e(2)$ depend on

$$\begin{aligned} \frac{||s_\Delta(t)||}{\sqrt{2N_0}} &= \sqrt{\frac{E}{N_0}(1-\rho)} \\ &= \begin{cases} \sqrt{2\frac{E}{N_0}} & \text{if } \rho = -1 \\ \sqrt{\frac{E}{N_0}} & \text{if } \rho = 0 \end{cases} \end{aligned}$$

Therefore to have the same $P_e(1)$, $P_e(2)$ orthogonal signaling requires twice as much energy (i.e. 3dB) as antipodal signaling.
*Examples of binary modulation schemes*

- Phase Shift Keying (PSK)

$$s_1(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t) \quad 0 \le t \le T$$
$$s_2(t) = -\sqrt{\frac{2E}{T}} \cos(2\pi f_c t)$$

Assuming $f_c \gg \frac{1}{T}$, the energy of $s_i(t)$ is $||s_i(t)||^2 = E$.

$$s_\Delta(t) = 2s_1(t) \qquad ||s_\Delta(t)|| = 2\sqrt{E}$$

and the probability of error is

$$P_e = P(1)Q\left(\sqrt{\frac{2E}{N_0}} - \frac{1}{2}\sqrt{\frac{N_0}{2E}} \ln \lambda\right) + P(2)Q\left(\sqrt{\frac{2E}{N_0}} + \frac{1}{2}\sqrt{\frac{N_0}{2E}} \ln \lambda\right)$$

which for equally likely signals reduces to

$$P_e = Q\left(\sqrt{\frac{2E}{N_0}}\right)$$

- Frequency Shift Keying (FSK)

$$s_1(t) = \sqrt{\frac{2E}{T}} \cos\left(2\pi \left(f_c - \frac{f_\Delta}{2}\right) t\right) \quad 0 \le t \le T$$
$$s_2(t) = \sqrt{\frac{2E}{T}} \cos\left(2\pi \left(f_c + \frac{f_\Delta}{2}\right) t\right) \quad 0 \le t \le T$$

The cross correlation coefficient of FSK signals is

$$\rho = \frac{1}{\sqrt{E_1 E_2}} \int_0^T s_1(t) s_2(t) dt = \frac{1}{E} \frac{2E}{T} \int_0^T \cos\left(2\pi \left(f_c - \frac{f_\Delta}{2}\right) t\right) \cos\left(2\pi \left(f_c + \frac{f_\Delta}{2}\right) t\right) dt$$
$$= \frac{1}{T}\left[\int_0^T \cos(4\pi f_c t) dt + \int_0^T \cos(2\pi f_\Delta t) dt\right]$$
$$= \frac{\sin(2\pi f_\Delta T)}{2\pi f_\Delta T}$$

Orthogonal FSK corresponds to $f_\Delta = \frac{1}{2T}$ or $f_\Delta = \frac{1}{T}$. The best performance is obtained for

$\rho \approx -0.22$, corresponding to $f_\Delta = \frac{0.715}{T}$. Assume equally likely signals

$$P_e = Q\left(\sqrt{\frac{E}{N_0}(1-\rho)}\right)$$

How many extra or less energy in dB is required to obtain the same probability of error with orthogonal FSK and FSK with $f_\Delta = \frac{0.715}{T}$ [ denoted FSK(0.715)]. Let $E$ be the energy of FSK(0.715) and let $E^{'}$ be the energy of the orthogonal FSK.

$$Q\left(\sqrt{\frac{E}{N_0}\cdot 1.22}\right) = Q\left(\sqrt{\frac{E^{'}}{N_0}}\right)$$

Thus

$$\sqrt{\frac{E}{N_0}\cdot 1.22} = \sqrt{\frac{E^{'}}{N_0}}$$

so

$$\frac{E^{'}}{E} = 1.22 \quad \Longleftrightarrow \quad \left(\frac{E^{'}}{E}\right)_{\text{dB}} = 10\log\left(\frac{E^{'}}{E}\right) = 10\log(1.22) = 0.86dB$$

Hence orthogonal FSK requires 0.86dB extra energy to yield the same probability of error as FSK(0.715).

## 8.7 Probabilities of symbol error for $M$-ary transmission over AWGN channels: union bound on probability of symbol error

Recall that the MAP decision rule is

$$\arg\min_{m=1,\ldots,M}\left[||\boldsymbol{r}_N - \boldsymbol{s}_m||^2 - N_0 \ln P(m)\right]$$

Assume that $s_i(t)$ is transmitted, define the event $A_{ik}$ ($i \neq k$) as

$$||\boldsymbol{r}_N - \boldsymbol{s}_i||^2 - N_0 \ln P(i) \leq ||\boldsymbol{r}_N - \boldsymbol{s}_k||^2 - N_0 \ln P(k)$$

with $s_i(t)$ transmitted.

The probability of correct decision when $s_i(t)$ was transmitted is

$$\begin{aligned}
P_c(i) &= P\left(A_{i1} \cap A_{i2} \ldots \cap A_{ii-1} \cap A_{ii+1} \cap \ldots \cap A_{iM}\right) \\
P_e(i) &= 1 - P_c(i) = P\left(\overline{A_{i1} \cap A_{i2} \ldots \cap A_{ii-1} \cap A_{ii+1} \cap \ldots \cap A_{iM}}\right) \\
&= P\left(\overline{A_{i1}} \cup \overline{A_{i2}} \ldots \cup \overline{A_{ii-1}} \cup \overline{A_{ii+1}} \cup \ldots \overline{A_{iM}}\right) \quad \text{using Morgan law}
\end{aligned}$$

where $\overline{A}$ represents the conjugate event of $A$. From the properties of the probability measure

$$P_e(i) \leq \sum_{\substack{k=1 \\ k \neq i}}^{M} P\left(\overline{A_{ik}}\right)$$

$$\begin{aligned}
P\left(\overline{A_{ik}}\right) &= P\left(||\boldsymbol{r}_N - \boldsymbol{s}_i||^2 - N_0 \ln P(i) > ||\boldsymbol{r}_N - \boldsymbol{s}_k||^2 - N_0 \ln P(k) | s_i(t) \text{ transmitted}\right) \\
&= P\left(r_{s_{\Delta ik}} < \lambda_{ik}^{'} | H_i\right) \qquad \text{similarly to Section 8.6} \\
&= Q\left(\frac{||s_i(t) - s_k(t)||}{\sqrt{2N_0}} - \sqrt{\frac{N_0}{2}} \frac{\ln\left[P(k)/P(i)\right]}{||s_i(t) - s_k(t)||}\right)
\end{aligned}$$

where

$$\begin{aligned}
r_{s_{\Delta ik}} &= \frac{< r(t), s_{\Delta ik}(t) >}{||s_{\Delta ik}(t)||} \\
s_{\Delta ik}(t) &= s_i(t) - s_k(t) \\
\lambda_{ik}^{'} &= \frac{E_i - E_k + N_0 \ln\left[P(k)/P(i)\right]}{2||s_{\Delta ik}(t)||}
\end{aligned}$$

Since $P(e) = \sum_{i=1}^{M} P_e(i) P(i)$, we have

$$P_e = \sum_{i=1}^{M} \sum_{\substack{k=1 \\ k \neq i}}^{M} Q\left(\frac{||s_i(t) - s_k(t)||}{\sqrt{2N_0}} - \sqrt{\frac{N_0}{2}} \frac{\ln\left[P(k)/P(i)\right]}{||s_i(t) - s_k(t)||}\right) P(i)$$

In communication we are interested mostly by the order (magnitude) of $P_e$, whether $P_e$ is $2 \cdot 10^{-5}$ or $3 \cdot 10^{-5}$ is not very important, but whether $P_e$ is $10^{-2}$ or $10^{-5}$ is important. The Q-function decreases very fast when its argument is increased. Therefore at high SNR (small $N_0$), keeping only the largest terms in the sum will still yield a good approximation.

Assume that $s_1(t), s_2(t), \ldots, s_M(t)$ are equally likely then $P(i) = \frac{1}{M}$ and

$$P_e \leq \frac{1}{M} \sum_{i=1}^{M} \sum_{\substack{k=1 \\ k \neq i}}^{M} Q\left(\frac{||s_i(t) - s_k(t)||}{\sqrt{2N_0}}\right)$$

A good approximation is

$$P_e \leq \frac{2n(d_{\min})}{M} Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right)$$

where

$$d_{\min} = \min_{\substack{i=1,\ldots,M \\ i \neq k}} ||s_i(t) - s_k(t)||$$

is the minimum Euclidean distance between the possible transmitted signals $s_i(t)$, and $n(d_{\min})$ is the number of signal pairs with Euclidean distance equal to $d_{\min}$.
*Energies associated with M-ary signaling:*

- Average energy per transmitted signal

  $$E_a = \sum_{i=1}^{M} E_i P(i)$$

  where $E_i$ is the energy of the $i^{th}$ signal and $P(i)$ is the a-priori probability.

- Peak energy

  $$E_p = \max_i E_i$$

- Average energy per transmitted bit:

  Assume that $M = 2^L$ and each transmitted signal is represented by a binary word of $L$ bits, the average energy per transmitted bits is

  $$E_b = \frac{E_a}{L} = \frac{E_a}{\log_2(M)}$$

# References


[1] S. Haykin, *Communication systems*. New York: John Wiley & Sons, 2001.

[2] J. Proakis, *Communication Systems Engineering*. New Jersey: Prentice Hall, 2002.