ELEC-C5310 - Introduction to Estimation, Detection and Learning: Introduction to Detection Theory

Prof. Sergiy Vorobyov

Department of Signal Processing and Acoustics
Aalto University

DETECTION PROBLEM

How the data can be retrieved from the noisy observations?

The process of retrieving data is called *detection*, or *decision making, hypothesis testing, decoding.*

Example: *Decoding* in communication systems is the process of mapping the received signal into one of the possible set of code words or transmitted symbols. Decoder is designed to *minimize average probability of error*.

Bayes Detectors:

Hypothesis 1 (H_1): $Z = N$ (noise alone) $Pr(H_1\text{true}) = p_0$

Hypothesis 2 (H_2): $Z = k + N$ (signal plus noise) $Pr(H_2\text{true}) = 1 - p_0$

BAYES DETECTORS

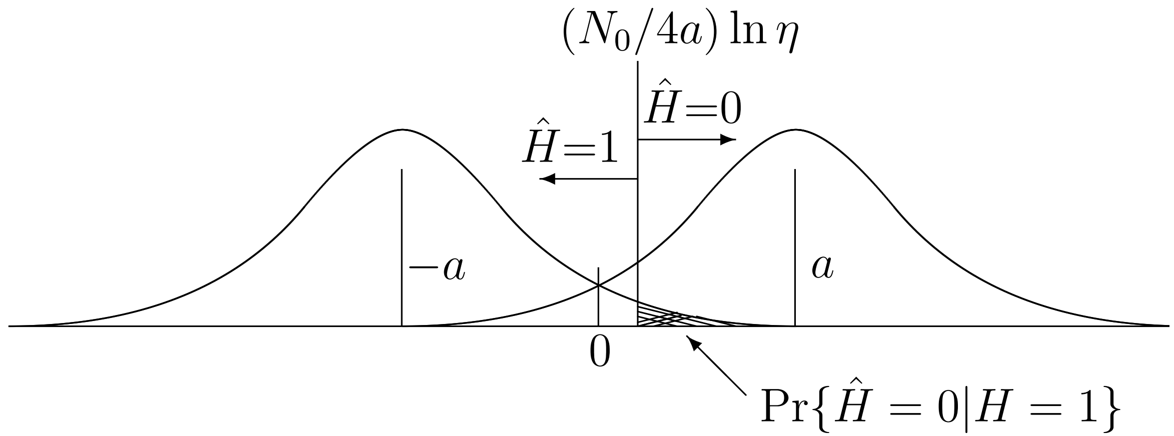
$N \sim \mathcal{N}(0, \sigma_n^2)$ and k is a constant signal.

Under hypothesis H_1 and hypothesis H_2 the pdf's are

$$f_Z(z|H_1) = \frac{e^{-z^2/2\sigma_n^2}}{\sqrt{2\pi\sigma_n^2}}$$

$$f_Z(z|H_2) = \frac{e^{-(z-k)^2/2\sigma_n^2}}{\sqrt{2\pi\sigma_n^2}}$$

Objective: partition the one dimensional observation space Z into two regions R_1 and R_2 such that if Z falls into R_1, we decide hypothesis H_1 is true, while if Z is in R_2, we decide H_2 is true.



Conditional pdf's for two-hypothesis detection problem.

Four types of decisions that we can make and associated costs:

c_{11} – the cost of deciding in favor of H_1 when H_1 is actually true

c_{12} – the cost of deciding in favor of H_1 when H_2 is actually true

c_{21} – the cost of deciding in favor of H_2 when H_1 is actually true

c_{22} – the cost of deciding in favor of H_2 when H_2 is actually true

The conditional average cost of making a decision given that H_1 is true:

$$C(D|H_1) = c_{11} Pr[\text{decide } H_1 | H_1 \text{is true}] + c_{21} Pr[\text{decide } H_2 | H_1 \text{is true}]$$

$$Pr[\text{decide } H_1 | H_1 \text{is true}] = \int_{R_1} f_Z(z|H_1) dz$$

$$Pr[\text{decide } H_2 | H_1 \text{is true}] = \int_{R_2} f_Z(z|H_1) dz$$

Since we are forced to make a decision

$$Pr[\text{decide } H_1 | H_1 \text{is true}] + Pr[\text{decide } H_2 | H_1 \text{is true}] = 1$$

Equivalently,

$$\int_{R_2} f_Z(z|H_1) dz = 1 - \int_{R_1} f_Z(z|H_1) dz$$

The conditional average cost given that H_1 is true:

$$C(D|H_1) = c_{11} \int_{R_1} f_Z(z|H_1)dz + c_{21} \left[1 - \int_{R_1} f_Z(z|H_1)dz \right]$$

The conditional average cost given that H_2 is true:

$$\begin{aligned} C(D|H_2) &= c_{12} Pr[\text{decide } H_1 | H_2 \text{is true}] + c_{22} Pr[\text{decide } H_2 | H_2 \text{is true}] \\ &= c_{12} \int_{R_1} f_Z(z|H_2)dz + c_{22} \int_{R_2} f_Z(z|H_2)dz \\ &= c_{12} \int_{R_1} f_Z(z|H_2)dz + c_{22} \left[1 - \int_{R_1} f_Z(z|H_2)dz \right] \end{aligned}$$

To find the average cost without regard to which hypothesis is actually true, we average $C(D|H_1)$ and $C(D|H_2)$ with respect to the prior probabilities of hypotheses H_1 and H_2, $p_0 = Pr[H_1 \text{ true}]$ and $q_0 = 1 - p_0 = Pr[H_2 \text{ true}]$.

$$C(D) = p_0 C(D|H_1) + q_0 C(D|H_2)$$

The average cost of making a decision:

$$\begin{aligned} C(D) = p_0 &\left(c_{11} \int_{R_1} f_Z(z|H_1) dz + c_{21} \left[1 - \int_{R_1} f_Z(z|H_1) dz \right] \right) \\ + q_0 &\left(c_{12} \int_{R_1} f_Z(z|H_2) dz + c_{22} \left[1 - \int_{R_1} f_Z(z|H_2) dz \right] \right) \end{aligned}$$

- Equivalently,

$$C(D) = [p_0 c_{21} + q_0 c_{22}] + \int_{R_1} \{[q_0(c_{12} - c_{22}) f_Z(z|H_2)] - [p_0(c_{21} - c_{11}) f_Z(z|H_1)]\} \, dz$$

$c_{12} > c_{22}$ and $c_{21} > c_{11}$ because wrong decision should be more costly than right decision.

- Thus, the two bracketed terms within the integral are positive because q_0, p_0, $f_Z(z|H_2)$, and $f_Z(z|H_1)$ are probabilities.
- Hence, all values of z that give a larger value for the second term within the integral than for the first term should be assigned to R_1 because they contribute the negative amount to the integral.

$C(D)$ *is minimized if we follow the rule*

$$q_0(c_{12} - c_{22}) f_Z(z|H_2) \gtrless_{H_1}^{H_2} p_0(c_{21} - c_{11}) f_Z(z|H_1)$$

Equivalently,

$$\Lambda(Z) \gtrless_{H_1}^{H_2} \eta$$

$$\Lambda(Z) \triangleq \frac{f_Z(z|H_2)}{f_Z(z|H_1)}$$

The this ratio of conditional pdf's is called the *likelihood ratio*.

The parameter

$$\eta \triangleq \frac{p_0(c_{21} - c_{11})}{q_0(c_{12} - c_{22})}$$

is called *threshold* of the test.

Example: Let the costs for a Bayes test be $c_{11} = c_{22} = 0$ and $c_{21} = c_{12}$. Consider the pdf's

$$f_Z(z|H_1) = \frac{e^{-z^2/2\sigma_n^2}}{\sqrt{2\pi\sigma_n^2}}, \qquad f_Z(z|H_2) = \frac{e^{-(z-k)^2/2\sigma_n^2}}{\sqrt{2\pi\sigma_n^2}}$$

(i) Find $\Lambda(Z)$.

(ii) Write down the likelihood ration test for $p_0 = q_0 = 0.5$.

(iii) Compare the result of part (ii) with the case $p_0 = \frac{1}{4}$ and $q_0 = \frac{3}{4}$.

To be solved in class.

PERFORMANCE OF BAYES DETECTOR

The conditional probabilities of making wrong decisions:

$$Pr_F = \int_{R_2} f_Z(z|H_1)dz$$

is the *probability of false alarm*, and

$$Pr_M = \int_{R_1} f_Z(z|H_2)dz = 1 - \int_{R_2} f_Z(z|H_2)dz = 1 - Pr_D$$

is the *probability of missed detection*, where Pr_D is the probability of correct detection.

The average cost of making a decision (risk) :

$$C(D) = p_0 c_{21} + q_0 c_{22} + q_0(c_{12} - c_{22})Pr_M - p_0(c_{21} - c_{11})(1 - Pr_F)$$

- We can write these probabilities in terms of the *conditional pdf's of the likelihood ratio $\Lambda(Z)$ given H_1 and H_2*.

$$Pr_M = \int_0^{\eta} f_\Lambda(\lambda|H_2)d\lambda$$

because, given H_2 is true, an erroneous decision is made if $\Lambda(Z) < \eta$.

$$Pr_F = \int_{\eta}^{\infty} f_\Lambda(\lambda|H_1)d\lambda$$

because, given H_1 is true, an error occurs if $\Lambda(Z) > \eta$.

- A plot of $Pr_D = 1 - Pr_M$ versus Pr_F is called the *operating characteristic* of the likelihood ratio test, or applied to communication and radar systems - the *receiver operating characteristic (ROC)*. It provides all the information necessary to evaluate the risk!

Example:

For the conditional pdf's from the previous example, the likelihood ratio test for an arbitrary threshold η is

$$\frac{2kZ - k^2}{2\sigma_n^2} \gtrless_{H_1}^{H_2} \ln\eta \qquad \text{or} \qquad X \gtrless_{H_1}^{H_2} d^{-1}\ln\eta + \frac{1}{2}d$$

$X \triangleq \frac{Z}{\sigma_n}$ is a new random variable, and $d \triangleq \frac{k}{\sigma_n}$ ia a new parameter.

X is obtained from Z by scaling by σ_n as

$$f_X(x|H_1) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}, \qquad f_X(x|H_2) = \frac{e^{-(x-d)^2/2}}{\sqrt{2\pi}}$$

Find Pr_F and Pr_M?

To be solved in class.

THE NEYMAN-PEARSON DETECTOR

- The design of a Bayes detector requires knowledge of the costs and a priory probabilities.
- If these are unavailable, a simple optimization procedure is to *fix Pr_F at some tolerable level α, and maximize Pr_D (or minimize Pr_M) subject to the constraint $Pr_F \leq \alpha$*. It is the *Neyman-Pearson detector*!
- The Neyman-Pearson criterion leads to a likelihood ratio test identical to the aforementioned Bayes test, except that the threshold η is determined by the allowed value of probability of false alarm α.
- This value of η can be obtained from ROC for a given value of Pr_F.
- The slope of a ROC curve at a particular point is equal to the value of the threshold η required to achieve the Pr_D and Pr_F of that point.

MIN PROBABILITY OF ERROR DETECTOR

The Risk if $c_{11} = c_{22} = 0$ and $c_{12} = c_{21} = 1$:

$$\begin{aligned} C(D) &= p_0 \left[1 - \int_{R_1} f_Z(z|H_1) dz \right] + q_0 \int_{R_1} f_Z(z|H_2) dz \\ &= p_0 \int_{R_2} f_Z(z|H_1) dz + q_0 \int_{R_1} f_Z(z|H_2) dz \\ &= p_0 Pr_F + q_0 Pr_M \end{aligned}$$

i.e. zero cost for making right decision, and equal cost for making either type of wrong decision.

It is actually the probability of erroneous decision (probability of error). Thus, the resulting detector is called *minimum probability of error detector*.

MAX A POSTERIORI (MAP) DETECTOR

- Letting $c_{11} = c_{22} = 0$ and $c_{12} = c_{21}$, the Bayes test becomes

$$\frac{f_Z(z|H_2)Pr(H_2)}{f_Z(z)} \gtrless_{H_1}^{H_2} \frac{f_Z(z|H_1)Pr(H_1)}{f_Z(z)}$$

where $Pr(H_1) = p_0$, $Pr(H_2) = q_0$, and

$$f_Z(z) \triangleq f_Z(z|H_1)Pr(H_1) + f_Z(z|H_2)Pr(H_2)$$

- Using Bayes' rule, the test can be simplified as

$$Pr(H_2|Z) \gtrless_{H_1}^{H_2} Pr(H_1|Z)$$

- The probabilities $Pr(H_1|Z)$ and $Pr(H_2|Z)$ are *a posteriori probabilities*, and the detector is the *maximum a posteriori (MAP) detector*.

MATCHED FILTER DETECTION

- The problem of the concern is the detection of a **known deterministic** signal in additive Gaussian noise (the signal and the pdf of noise are known)
- The detector evolving from these assumptions is called **matched filter**
- Applications where the signal is under the designer's control (e.g., coherent communication systems, radar)
- If we want to maximize the probability of detection subject to the constraint on the probability of false alarm: Neumann-Pearson approach
- If we want to minimize the average cost: Bayesian risk approach
- We will use the **minimum-distance criterion**

Problem Formulation

- Transmitter sends only a single symbol, so that the receiver observation

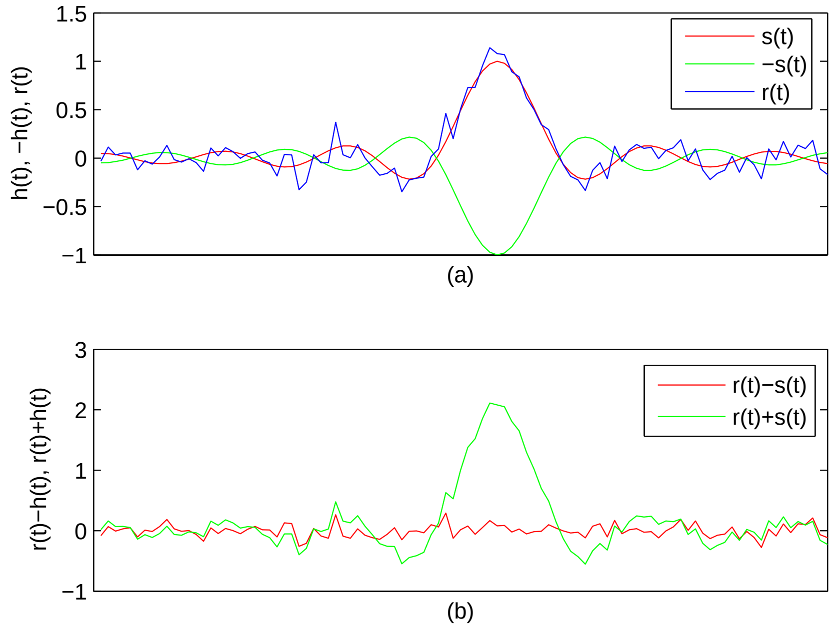
$$r(t) = a \cdot s(t) + n(t)$$

  $a \in \mathcal{A}$ is the transmitted symbol, $s(t)$ is the pulse shape (waveform), $n(t)$ is the noise

- The receiver design problem: infer from $r(t)$ which of the symbols was transmitted
- Minimum distance design strategy: choose the alphabet symbol that best represents the received waveform in a minimum-distance sense

Example: Minimum-Distance Strategy

- Consider the case of binary antipodal signaling with a zero-excess bandwidth pulse and an alphabet $\{\pm 1\}$. Noise variance is 0.1.
- Receiver calculates $\int_{-\infty}^{\infty} |r(t) - s(t)|^2 dt$ and $\int_{-\infty}^{\infty} |r(t) + s(t)|^2 dt$ and compare them



Min-Distance Receiver for Arbitrary Alphabets

- Principle: choose the symbol that best represents the observation in a minimum-distance sense, namely:

$$\hat{a} = \arg\min_{a \in \mathcal{A}} \int_{-\infty}^{\infty} |r(t) - a \cdot s(t)|^2 dt$$

- Minimum-distance terminology is used because signals can be interpreted as vectors in a vector space. Then, the energy in the error between two signals is the squared distance between their corresponding vectors
- Minimum-distance criterion is primarily motivated by noise
- It is "optimal" receiver structure in AWGN
- Must calculate $M = |\mathcal{A}|$ integrals, one for each element of the alphabet

Efficient Implementation

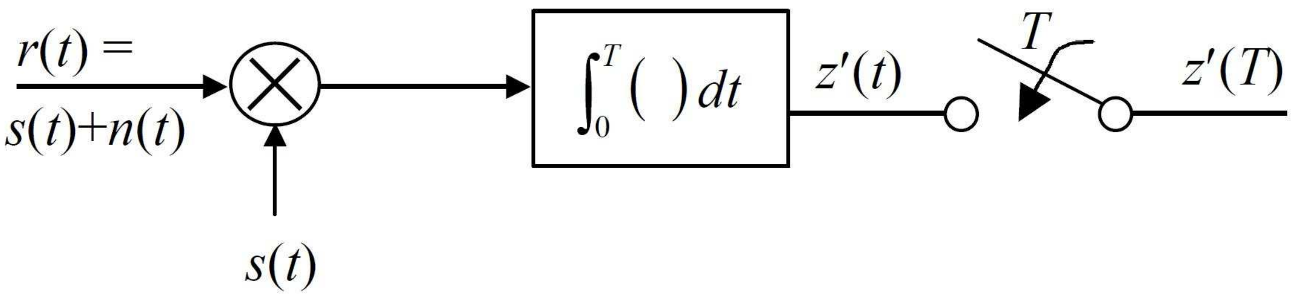
Requires a single integral and uses the cost function:

$$\begin{aligned} J &= \int_{-\infty}^{\infty} |r(t) - a \cdot s(t)|^2 dt \\ &= \underbrace{\int_{-\infty}^{\infty} |r(t)|^2 dt}_{E_r} - 2\mathrm{Re}\left\{ a^* \cdot \int_{-\infty}^{\infty} r(t) s^*(t) dt \right\} \\ &\quad + |a|^2 \underbrace{\int_{-\infty}^{\infty} |s(t)|^2 dt}_{E_s} = E_r - 2\mathrm{Re}\{a^* z\} + |a|^2 E_s \end{aligned}$$

E_r and E_s are the energies of $r(t)$ and $s(t)$, respectively, and $z = \int\limits_{-\infty}^{\infty} r(t) s^*(t) dt$

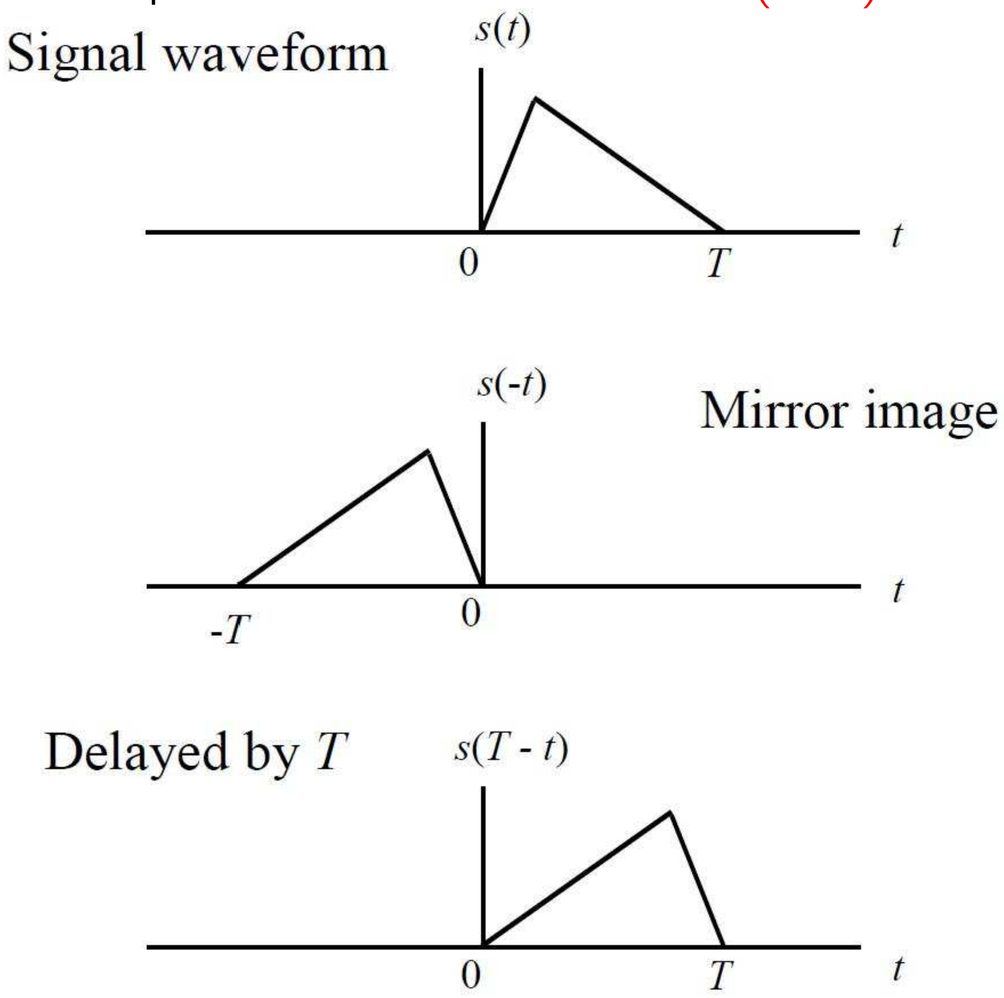
Correlator

- E_r is independent on a, thus, immaterial for minimization
- Only one term depends on the observation waveform $r(t)$, and it does so through the correlation integral z
- z is a **sufficient statistic** for determining the minimum-distance decision. **More details later!**
- Single integral problem: minimize the last two terms of the cost function
- Implementation of the correlation integral z via correlator:



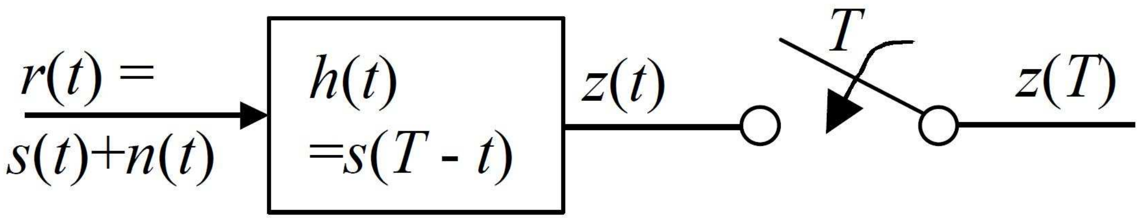
Matched Filter: Idea

Example: Impulse response of matched filter (MF)



Matched Filter: Implementation

- Implementation of the correlation integral z via MF:



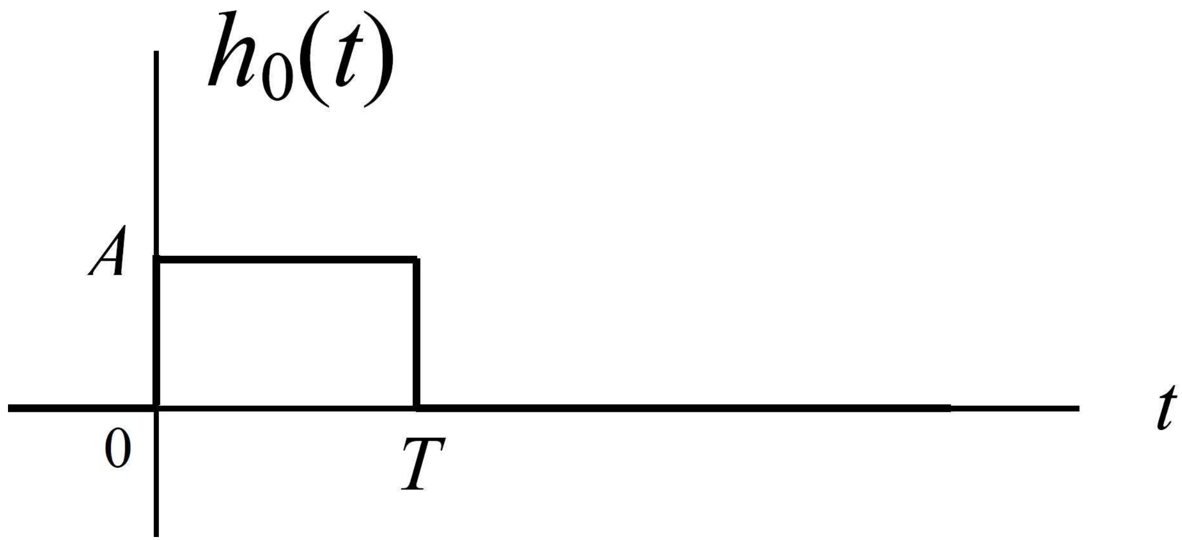
- The correlator and MF are mathematically equivalent, i.e., produce identical outputs
- The MF approach is practically preferred
- The MF is able to compensate for synchronization errors by adjusting the timing of the sampler
- The correlator requires the two inputs $r(t)$ and $s^*(t)$ be synchronized ahead of time

Example: MF

- Consider the pulse signal

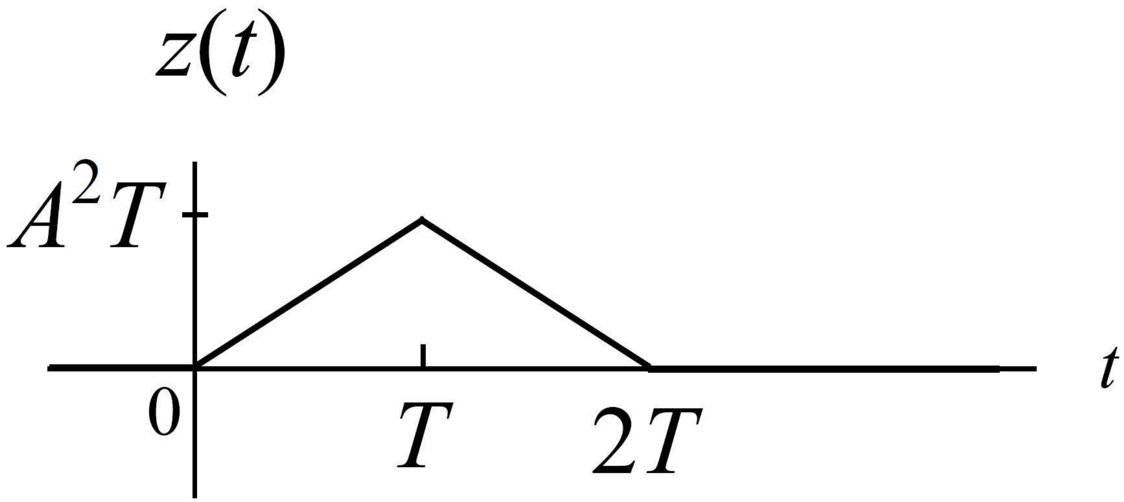
$$s(t) = \begin{cases} a, & 0 \leq t \leq T \\ 0, & \text{otherwise.} \end{cases}$$

- The MF for $s(t)$ is $h_0(t) = s(T - t)$



Example: MF (Continuation)

- The response $z(t)$ of the MF to $s(t)$ is the convolution $z(t) = h_0(t) * s(t)$



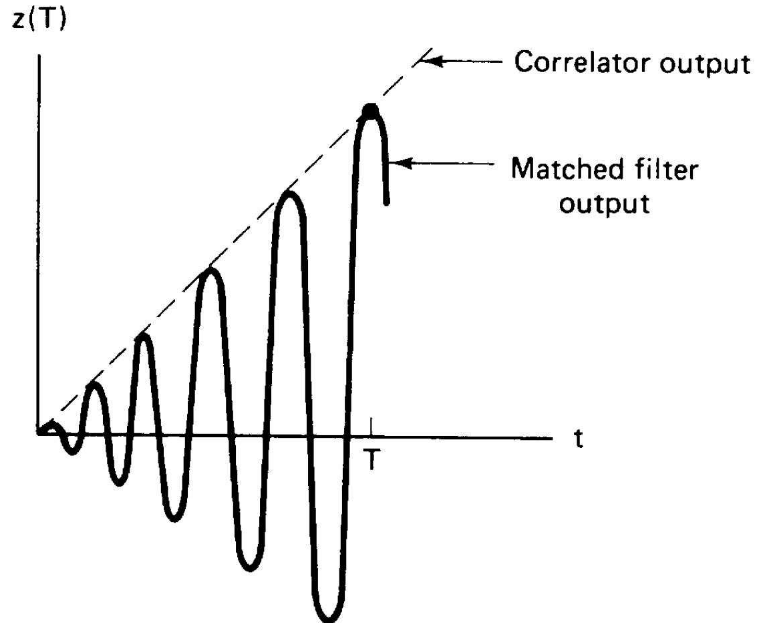
- Note that the peak output signal occurs at $t = T$, which is also the time instant of peak signal-to-noise (SNR) power ratio

Equivalence of the Correlator and MF

- The output of the MF:

$$z(t) = h(t) * r(t) = \int_0^t r(\tau)h(t-\tau)d\tau = \int_0^t r(\tau)s(T-t+\tau)d\tau$$

- When $t = T$, we obtain $z(T) = \int_0^T r(\tau)s(\tau)d\tau$, which is the same as the output of the correlator



Slicer

- Rewrite the cost function by factoring out the term E_s:

$$J = E_s\left(\frac{E_r}{E_s} - 2\mathrm{Re}\{a^*\frac{z}{E_s}\} + |a|^2\right) = E_s\left|\frac{z}{E_s} - a\right|^2 - \frac{|z|^2}{E_s} + E_r$$

- Only the first term depends on a. Then the minimum-distance receiver reduces to:

$$\hat{a} = \arg\min_{a \in \mathcal{A}}\left|\frac{z}{E_s} - a\right|^2$$

- The minimum-distance decision is $a \in \mathcal{A}$ closest to the normalized correlation $y = z/E_s$. The decision can be found by quantizing y to the nearest symbol. The corresponding devise is called a **slicer**

MF maximizes SNR

- Replacing the MF by a more general receiver filter $f(t)$, the sampler output is:

$$z = \int_{-\infty}^{\infty} r(t) f(-t) dt = a \underbrace{\int_{-\infty}^{\infty} s(t) f(-t) dt}_{S} + \underbrace{\int_{-\infty}^{\infty} n(t) f(-t) dt}_{N}$$

- For white nose with PSD N_0, the energy of the noise term is:

$$E\{|N|^2\} = N_0 E_f, \qquad E_f = \int_{-\infty}^{\infty} |f(-t)|^2 dt$$

- SNR is defined as:

$$SNR = \frac{E\{|S|^2\}}{E\{|N|^2\}} = \frac{\left| \int_{-\infty}^{\infty} s(t) f(-t) dt \right|^2}{E_f} \cdot \frac{E_a}{N_0}, \qquad E_a = E\{|a|^2\}$$

MF maximizes SNR (Continuation)

- Cauchy-Schwarz-Bunyakovsky inequality for any two complex integrable functions $s(t)$ and $f(t)$ with energies E_s and E_f, respectively:

$$\left| \int_{-\infty}^{\infty} s(t) f^*(-t) dt \right|^2 \leq E_s E_f$$

with equality if and only if $f(-t) = K s(t)$ for some constant K

- The matched-filter bound on the SNR:

$$SNR \leq E_a E_s / N_0$$

- The equality is reached if and only if $f(t) = K s^*(-t)$. Take $K = 1$.

MF and Inter-Symbol Interference

- When a sequence of pulses is transmitted, using an MF as a receiver filter will generally introduce inter-symbol interference (ISI)
- No ISI, if the received pulse is time-limited to the symbol interval
- Generally, if the overall pulse shape at the output of the MF obeys the Nyquist criterion, then the MF is the optimal receive filter for both the isolated-pulse case and the sequence-of-pulses case, in the sense that is maximizes the SNR

MF and ISI (Continuation)

- The Nyquist criterion in terms of the folder spectrum of the received pulse is:

$$\frac{1}{T}\sum_{m=-\infty}^{\infty}\left|S\left(f-\frac{m}{T}\right)\right|^2=1$$

- $|S(f)|^2$ is the Fourier transform of the overall pulse shape at the output of the MF

THE M-ARY HYPOTHESIS CASE

- *The Bayes decision criterion cam be straightforwardly generalized to $M > 2$ hypotheses.*
- For the M-ary case, M^2 costs and M *a priori* probabilities must be given; M likelihood ratio tests must be carried out in making a decision.
- Consider a special cost assignment used to obtain the MAP detector.
- Then, we have the following MAP decision rule for the M-hypothesis case: *Compute the M posterior probabilities $Pr(H_i|Z)$, $i = 1, 2, \dots, M$, and choose as the correct hypothesis the one corresponding to the largest posterior probability.*
- This decision rule is typically used when M-ary signal detection is considered.

VECTOR OBSERVATIONS

- *If, instead of a single observation Z, we have N observations $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_N)$, all previous results hold* with the exception that N-fold joint pdf's of Z, given H_1 and H_2, are to be used.
- If $Z_1, Z_2, \ldots, Z_N$ are conditionally independent, these joint pdf's are simply the N-fold products of the marginal pdf's.

MAP RECEIVERS FOR DIGITAL SYSTEMS

Example: M-ary communication system:

- *Information source:* One of M possible messages every T seconds; m_i, $i = 1, 2, \ldots, M$.
- *Modulator:* Massage m_i associated with signal $S(t)$, T seconds long; $s_i(t)$, $i = 1, 2, \ldots, M$.
- *Channel:* White Gaussian noise $n(t)$, PSD $= \frac{1}{2}N_0$; $y(t)$.
- *Receiver:* Observes $y(t)$ for T seconds. Guess at transmitted signal every T seconds; *Best guess* $(\min Pr_e)$: $\hat{m}_i(t)$.

For simplicity assume that the messages are produced by the information source with equal *a priori* probability.

The ith signal:

- $$s_i(t) = \sum_{j=1}^{K} a_{ij}\phi_j(t), \qquad i = 1, 2, \ldots, M, \quad K \leq M$$

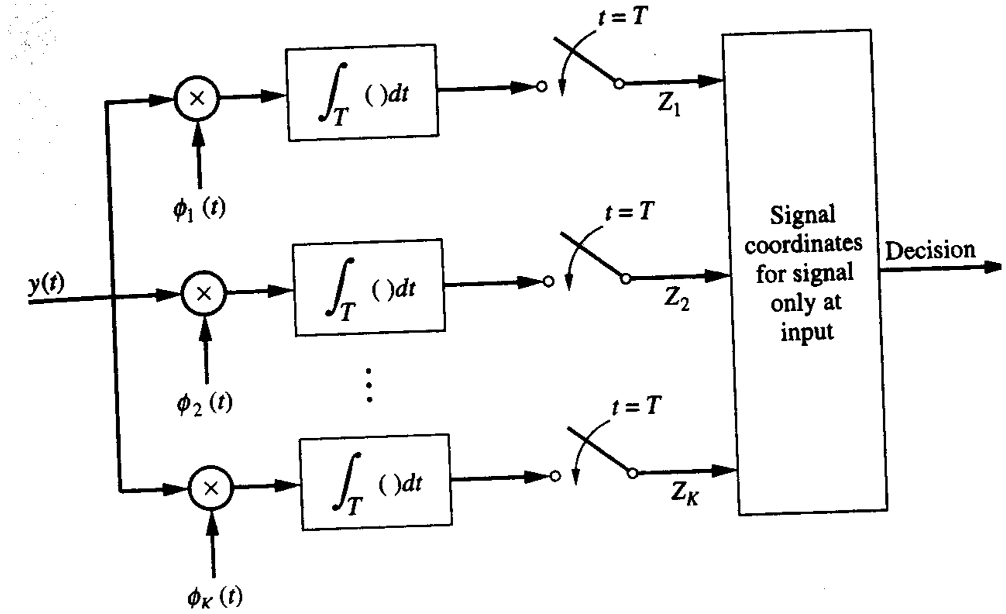
  $\phi_j(t)$'s are orthonormal basis function (chosen according to the *Gram-Schmidt* procedure).

- $$a_{ij} = \int_0^T s_i(t)\phi_j(t)dt = \langle s_i, \phi_j \rangle$$

  are Fourier coefficients for $s_i(t)$.

- Thus, *each possible signal can be represented as a point in K-dimensional signal space with coordinates $(a_{i1}, a_{i2}, \ldots, a_{iK})$, for $i = 1, 2, \ldots, M$.*

Receiver structure for resolving signal into K-dimensional signal space:



The receiver consists of a bank of correlators, and is used to compute the generalized Fourier coefficients for $s_i(t)$. Knowing the coordinates (Fourier coefficients) of $s_i(t)$ is as good as knowing $s_i(t)$ itself.

Detection problem in M-ary communication system

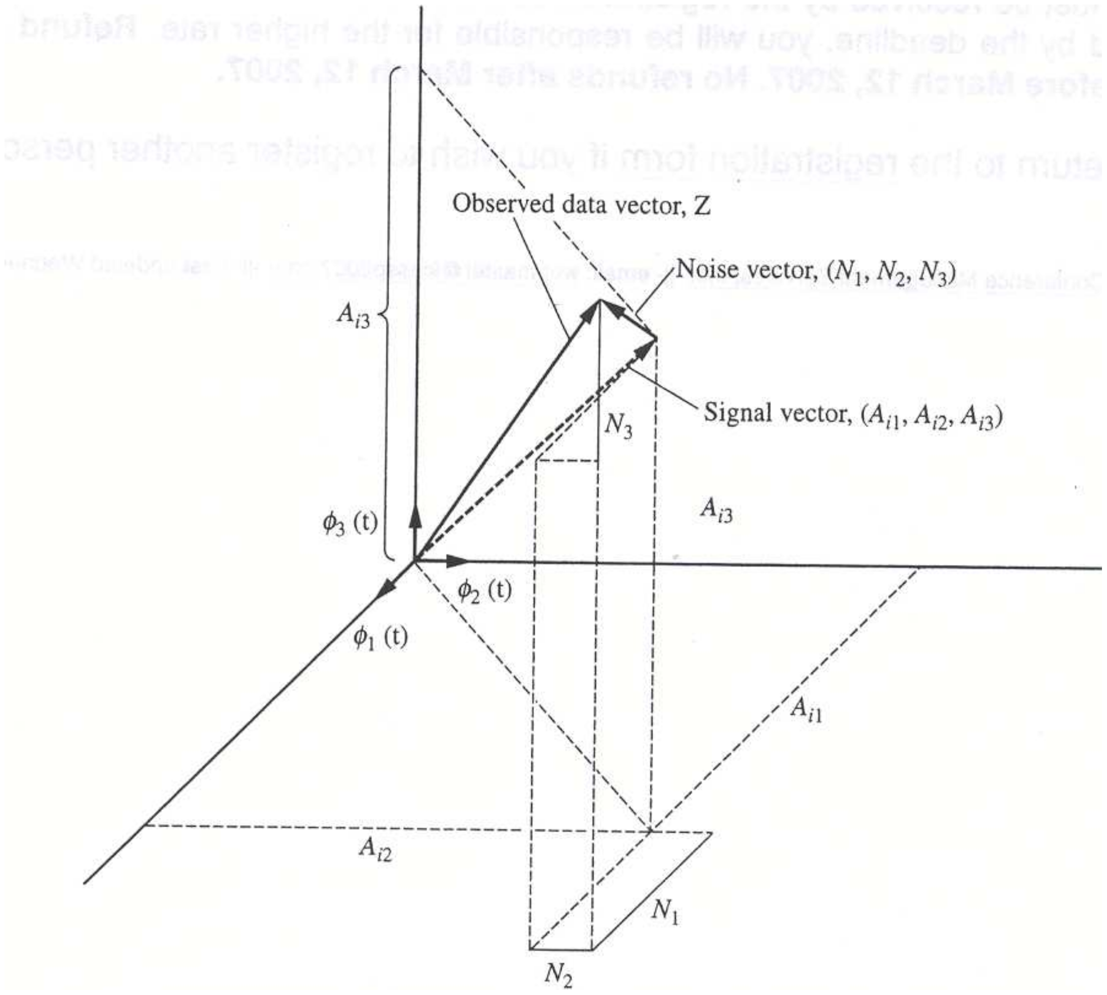
- *The difficulty is that the signal is received in the presence of noise*, i.e., the receiver provides us with noisy coordinates:

$$\boldsymbol{z} = (z_1, z_2, \ldots, z_K) = (a_{i1} + n_1, a_{i2} + n_2, \ldots, a_{iK} + n_K)$$

$$n_j \triangleq \int_0^T n(t)\phi_j dt = \langle n, \phi_j \rangle$$

- $\boldsymbol{z}$ is called *data vector*, and the space of all possible data vectors is called *observation space*.

A three-dimensional observation space:



- Formulation of the detection (decision-making) problem:
  To associate sets of noisy signal points with each possible transmitted signal point in a manner that the average error probability will be minimized, i.e. the observation space must be partitioned into M regions R_i, one associated with each transmitted signal, such that if a received data point falls into region R_l, the decision "$s_l(t)$ transmitted" is made with minimum probability of error.

- Basic principle:

$$\max_{l} Pr(H_l|z_1, z_2, \ldots, z_k), \qquad l = 1, 2, \ldots, M$$

  H_l is the hypothesis "$s_l(t)$ transmitted".

- Assume: $Pr(H_1) = Pr(H_2) = \ldots = Pr(H_M)$.

- Posterior probabilities: using Bayes' rule

$$Pr(H_l|z_1, z_2, \ldots, z_k) = \frac{f_Z(z_1, \ldots, z_K|H_l)Pr(H_l)}{f_Z(z_1, \ldots, z_K)}$$

Since $Pr(H_l)$ and $f_Z(z_1, \ldots, z_K)$ do not depend on l, it is enough to compute $f_Z(z_1, \ldots, z_K|H_l)$ and choose H_l corresponding to the largest.

- The mean of z_j, given hypothesis H_l:

$$E\{z_j|H_l\} = E\{a_{lj} + n_j\} = a_{lj} + \int_0^T E\{n(t)\}\phi_j(t)dt = a_{lj}$$

$$j = 1, 2, \ldots, K$$

- The variance of z_j, given hypothesis H_l:

$$\begin{aligned}
\mathrm{var}\{z_j|H_l\} &= E\{[a_{lj}+n_j]^2\} - a_{lj}^2 = E\{n_j^2\} \\
&= E\left\{\int_0^T n(t)\phi_j(t)dt \int_0^T n(t')\phi_j(t')dt'\right\} \\
&= \int_0^T \int_0^T E\{n(t)n(t')\}\phi_j(t)\phi_j(t')dtdt' \\
&= \int_0^T \int_0^T \frac{N_0}{2}\delta(t-t')\phi_j(t)\phi_j(t')dtdt' \\
&= \int_0^T \frac{N_0}{2}\phi_j^2(t)dt = \frac{1}{2}N_0, \qquad j = 1, 2, \ldots, K
\end{aligned}$$

The orthonormality of the ϕ_j's has been used.

- Similarly, we can find that the covariance of z_j and z_k, for $j \neq k$, is zero.

- *Thus $z_1, z_2, \ldots, z_K$ are uncorrelated Gaussian random variables and, hence, are statistically independent. Thus*

$$f_Z(z_1, \ldots, z_K|H_l) = \prod_{j=1}^{K} \frac{\exp[-(z_j - a_{lj})^2/N_0]}{\sqrt{\pi N_0}}$$

$$= \frac{1}{(\pi N_0)^{K/2}} \exp\left[-\sum_{j=1}^{K}(z_j - a_{lj})^2/N_0\right]$$

$$= \frac{\exp\{-\|z - s_l\|^2/N_0\}}{(\pi N_0)^{K/2}}$$

$$z = \sum_{j=1}^{K} z_j\phi_j(t), \qquad s_l(t) = \sum_{j=1}^{K} a_{lj}\phi_j(t)$$

- Except for the factor independent on l, the expression for $f_Z(z_1, \ldots, z_K|H_l)$ is equivalent to the posteriori probability $Pr(H_l|z_1, \ldots, z_K)$ obtained by applying Bayes' rule.

- *Choosing H_l corresponding to the maximum posterior probability is the same as choosing the signal with coordinates $a_{l1}, a_{l2}, \ldots, a_{jK}$ so as to maximize $f_Z(z_1, \ldots, z_K|H_l)$ or, equivalently, so as to minimize the exponent. But $\|z - s_l\|^2$ is the distance between $z(t)$ and $s_l(t)$!*

- Decision rule:

$$\min_l \|z - s_l\|^2 = \min_l \sum_{j=1}^{K} (z_j - a_{lj})^2, \qquad l = 1, 2, \ldots M$$

Detector for frequency modulated signal used widely in automotive radar

Consider M-ary coherent FSK:

$$s_i(t) = a\cos\{2\pi[f_c + (i-1)\Delta f]t\}, \quad 0 \le t \le T$$

$$\Delta f = \frac{m}{2T}, \qquad m \text{ an integer}, \qquad i = 1, 2, \ldots, M$$

For simplicity assume that $f_c T$ is integer.

(i) Apply Gram-Schmidt orthomormalization to obtain orthonormal basis set. How many orthonormal functions are? How ith signal can be written?

(ii) Denote signal plus noise waveform as $y(t) = s_i(t) + n(t)$. What is the projection of $y(t)$ to the observation space?

(iii) Derive the *decision rule*.

To be solved in class.

SUFFICIENT STATISTICS

- Because of the noise component $n(t)$, $z(t) = \sum_{j=1}^{K} z_j \phi_j(t)$ is not the same as $y(t)$, since an infinite set of basis functions would be required to represent all possible $y(t)$'s.
- *However, only K coordinates, where K is signal space dimension, are required to provide all the information that is relevant to making a decision.*
- $y(t) = \sum_{j=1}^{\infty} y_j \phi_j(t)$ for a complete orthonormal set of basis functions, where the first K of them are chosen using the Gram-Schmidt procedure for the given signal set.

- Given the hypothesis H_l is true, the y_j's are given by

$$y_j = \begin{cases} z_j = a_{lj} + n_j, & j = 1, 2, \ldots K \\ n_j, & j = K+1, K+2, \ldots \end{cases}$$

- The mean and the variance:

$$E\{y_j\} = \begin{cases} a_{lj}, & j = 1, 2, \ldots K \\ 0, & j > K \end{cases} \quad \text{and} \quad \begin{array}{r} \text{var}\{y_j\} = \frac{1}{2}N_0, \quad \text{all } j \\ \text{cov}\{y_j y_k\} = 0, \quad j \neq k \end{array}$$

- The joint pdf of $y_1, y_2, \ldots$, given H_l:

$$f_Y(y_1, \ldots, y_K | H_l) = C \exp\left\{ -\frac{1}{N_0} \left[\sum_{j=1}^{K} (y_j - a_{lj})^2 + \sum_{j=K+1}^{\infty} y_j^2 \right] \right\}$$

$$= C_1 \exp\left\{ -\frac{1}{N_0} \sum_{j=K+1}^{\infty} y_j^2 \right\} f_Z(y_1, \ldots, y_K | H_l)$$

- Since this pdf factors, $y_{K+1}, y_{K+2}, \ldots$ are independent of $y_1, y_2, \ldots, y_K$ and the former provide no information for making a decision. Thus d^2 derived before is a *sufficient statistic*.

Detection of M-ary orthogonal signals

Consider M-ary signaling scheme for which the signal waveforms have equal energies and are orthogonal over signaling interval, i.e.,

$$\int_0^T s_i(t)s_j(t)dt = \begin{cases} E_s & i = j \\ 0 & i \neq j \end{cases} \qquad i = 1, 2, \ldots, M$$

E_s is energy of each signal in $(0, T)$.

Example: We considered before the M-ary FSK:

$$s_i(t) = a\cos\{2\pi[f_c + (i-1)\Delta f]t\}, \quad 0 \leq t \leq T$$

$$\Delta f = \frac{m}{2T}, \qquad m \text{ an integer}, \qquad i = 1, 2, \ldots, M$$

Need $K = M$ orthonormal functions. The receiver has M correlators.

- Decision criterion:

$$\max_{l} z_l = \max_{l} \int_0^T y(t)\phi_l(t)dt$$

i.e., *the signal is chosen that has maximum correlation with the received signal plus noise.*

- Probability of symbol error:

$$\begin{aligned} Pr_E &= \sum_{i=1}^{M} Pr\{E|s_i(t) \text{ sent}\} Pr\{s_i(t) \text{ sent}\} \\ &= \frac{1}{M} \sum_{i=1}^{M} Pr\{E|s_i(t) \text{ sent}\} \end{aligned}$$

where each signal is assumed a priori equal probable.

- We may write

$$Pr\{E|s_i(t)\ \text{sent}\} = 1 - Pr_{ci}$$

Pr_{ci} is the probability of correct decision given that $s_i(t)$ was sent.

- A correct decision results only if

$$z_j = \int_0^T y(t)s_j(t)dt < \int_0^T y(t)s_i(t)dt = z_i$$

for all $i \neq j$.

- Then we can write

$$Pr_{ci} = Pr\{\text{all}\ z_j < z_i,\ j \neq i\}$$

- If $s_i(t)$ is transmitted, then

$$z_i = \int_0^T [\sqrt{E_s}\phi_i(t) + n(t)]\phi_i(t)dt = \sqrt{E_s} + n_i$$

$$n_i = \int_0^T n(t)\phi_i(t)dt$$

- Since $z_j = n_j$, $i \neq j$, given $s_i(t)$ was sent, it follows that

$$Pr_{ci} = Pr\{\text{all } n_j < \sqrt{E_s} + n_i,\ j \neq i\}$$

- Note that n_i is a Gaussian random variable with

$$\mathrm{var}\{n_i\} = E\left\{\left[\int_0^T n(t)\phi_j(t)dt\right]^2\right\} = \frac{N_0}{2}, \quad E\{n_i n_j\} = 0.$$

- For a particular value of n_i, we can write that

$$Pr_{ci} = \prod_{j=1, j \neq i}^{M} Pr\{n_j < \sqrt{E_s} + n_i\} = \left(\int_{-\infty}^{\sqrt{E_s}+n_i} \frac{e^{-n_j^2/N_0}}{\sqrt{\pi N_0}} dn_j \right)^{M-1}$$

because the pdf of n_j is Gaussian zero-mean with variance $\sqrt{N_0/2}$.

- Average over all possible values of n_i gives

$$Pr_{ci} = \int_{-\infty}^{\infty} \frac{e^{-n_i^2/N_0}}{\sqrt{\pi N_0}} \left(\int_{-\infty}^{\sqrt{E_s}+n_i} \frac{e^{-n_j^2/N_0}}{\sqrt{\pi N_0}} dn_j \right)^{M-1} dn_i$$

$$= (\pi N_0)^{-M/2} \int_{-\infty}^{\infty} e^{-y^2} \left(\int_{-\infty}^{\sqrt{E_s/N_0}+y} e^{-x^2} dx \right)^{M-1} dy$$

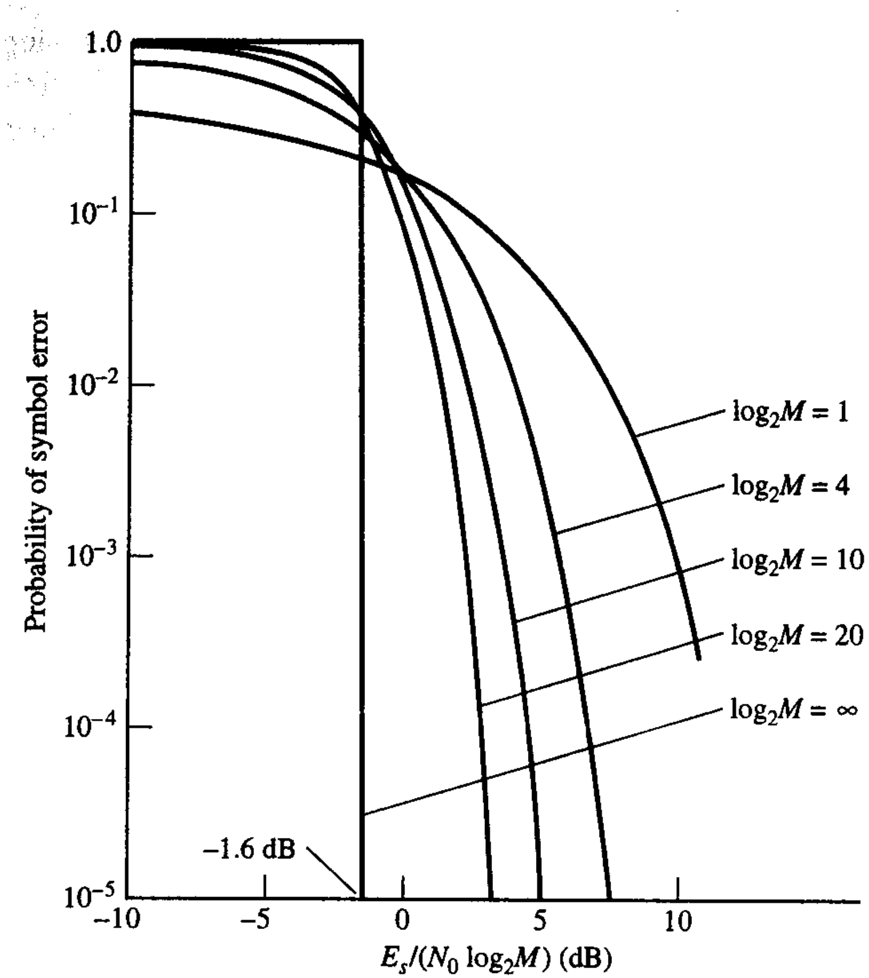
where the substitutions $x = n_j/\sqrt{N_0}$ and $y = n_i/\sqrt{N_0}$ are made.

- Since Pr_{ci} is independent of i

$$Pr_E = 1 - Pr_{ci}$$

- The integral can be computed numerically. See the figure Pr_E versus $E_s/N_0 \log_2 M$ for several values of M.
- *As $M \rightarrow \infty$, error-free transmission can be achieved as long as $E_s/N_0 \log_2 M > \ln 2 = -1.59$ dB. This error-free transmission is achieved at the expense of infinite bandwidth, however, since $M \rightarrow \infty$ means that an infinite number of orthogonal functions are required.*

Probability of symbol error for detection of M-ary orthogonal signals:



End of Introduction to Detection