EE401: Advanced Communication Theory

Professor A. Manikas  
Chair of Communications and Array Processing  
Imperial College London  

Principles of Decision and Detection Theory

Table of Contents

1. Basic Decision Theory
   - Hypothesis Testing in an M-ary Comm System  
   - Terminology  
   - M-ary Decision Criteria  
   - Examples  
     - Example: Gaussian Likelihood Functions (LFs)  
     - Example: Signal + Gaussian Noise  
     - Example: Rectangular and Gaussian LFs

2. Basic Concepts on Optimum Receivers
   - The Concept of "Continuous Sampling"  
   - Optimum M-ary Receivers  
     - Gaussian Multivariable Distribution  
     - Likelihood Functions (Continuous Sampling)

3. Optimum Receivers using Matched Filters
   - Known Signals in White Noise  
   - Signals and Matched Filters in Frequency Domain

4. Optimum Receivers in Non-White Noise
   - Known Signals in Non-White Noise  
   - A Special Case  
   - Output SNR  
   - A Special Case: Output SNR (Signal plus white Noise)  
   - Approximation To Matched Filter Solution

5. Whitening Filter

6. Optimum M-ary Rx based on Signal Constellation
   - Introduction - Orthogonal Signals  
   - The "Approximation Theorem" of an Energy Signal  
   - Comments on the "Approximation" Theorem

7. Gram-Schmidt Orthogonalization
   - M-ary Signals: Energy and Cross-Correlation  
   - Signal Constellation  
   - Distance between two M-ary signals  
   - Examples of Signal Constellation Diagram

8. Examples of Plots of Decision Variables (QPSK-Receiver)
   - Optimum M-ary Receiver using the Signal-Vectors  
   - Encoding M-ary Signals  
   - Important Relationships Between \( p_r \) and \( \rho_{r,e} \)

9. Appendices:
   - Appendix-1: Proof of Equation-26  
   - Appendix-2: Walsh-Hadamard Orthogonal Sets and Signals  
   - Appendix-3: 1595 - Walsh Function of order-64
Basic Decision Theory

- Detection Theory is concerned with determining the existence of a signal in the presence of noise - and can be classified as follows:

![Diagram of Detection Theory]

- **Optimum**
  - **Parametric**: The pdf of input signal \( r(t) \) and some conditional input pdfs (eg pdf, \( r(t) \)) are known. We arrive at a form of the detector \( D \) by utilizing the above known input pdfs.

- **Suboptimal**
  - **Adaptive or Learning**: There is not a complete statistical description of the background noise.
  - **Non-parametric**: Input distribution has a specified shape or form but cannot be classified by a finite No of real parameters.
  - **Distribute Free**: There is not any assumption at all concerning the form of the input distribution.

N.B.:

1. **Adaptive or Learning**:
   - system parameters change as a function of input;
   - difficult to analyze;
   - mathematically complex

2. **Non-Parametric**:
   - fairly constant level of performance because these are based on general assumptions on the input pdf;
   - easier to implement.

3. Consider a parametric detector which has been designed for Gaussian pdf input.
   - if input is actually Gaussian: then parametric’s performance is significantly better;
   - if input ≠ Gaussian but still symmetric: then non-parametric detector performance may be much better than the performance of the parametric detector
Hypothesis Testing in an M-ary Comm System

- A Hypothesis $\triangleq$ a statement of a possible condition
- In an $M$-ary Comm. System we have $M$ hypotheses:

$$r(t) = s_i(t) + n(t)$$

(To determine which signal is present)

$D_i$

where $s_i(t) =$ one of $M$ signal (channel symbols)

- hypotheses:
  
  $H_1$: $s_1(t)$ is present with probability $Pr(H_1)$
  
  $H_2$: $s_2(t)$ is present with probability $Pr(H_2)$
  
  $\ldots$

  $H_M$: $s_M(t)$ is present with probability $Pr(H_M)$

- The statistics of the observed signal

$$r(t) = s_i(t) + n(t)$$

(1)

are affected by the presence of $s_1(t)$, or $s_2(t)$, $\ldots$, or $s_M(t)$

- If $s_i(t)$, $\forall i$ are known

then their distributions are known

and the problem is translated to make a decision about one of the $M$ distributions after having observed $r(t)$

This is called ‘Hypothesis Testing’
Terminology

- **A priori probabilities**: 
  \[ \Pr(H_1), \Pr(H_2), \ldots, \Pr(H_M) \]
  (these are calculated **BEFORE** the experiment is performed)

- **A posterior probabilities**
  \[ \Pr(H_1/r), \Pr(H_2/r), \ldots, \Pr(H_M/r) \]

  That is, if \( r \) = observation variable
  then we have \( M \) Conditional Probabilities
  \[ \Pr(H_i/r), \forall i \in [1, \ldots, M] \]

  known as a **POSTERIOR PROBABILITIES**
  (since these are calculated **AFTER** the experiment is performed).

- **\( \Pr(H_i/r) \ \forall i \): difficult to find.**
  A more natural approach is to find
  \[ \Pr(r/H_i), \forall i \] (2)
  since in general pdf\(_{r/H_i}\), \( \forall i \)
  ▶ are known or
  ▶ can be found

- **Likelihood Functions (LF):**
  \[ \text{pdf}_{r/H_1}(r), \text{pdf}_{r/H_2}(r), \ldots, \text{pdf}_{r/H_M}(r) \] (3)
  i.e. the \( M \) conditional probability density functions pdf\(_{r/H_i}(r)\), \( \forall i \), are known as "Likelihood Functions"

- **Likelihood Ratio (LR):** The ratio
  \[ \frac{\text{pdf}_{r/H_i}(r)}{\text{pdf}_{r/H_j}(r)} \text{ for } i \neq j \] (4)

  is known as "Likelihood Ratio"
**MAP Criterion**

DECISION—choose hypothesis \( H_i \) i.e. \( D_i \):

\[
\text{iff } \Pr(H_i / r) > \Pr(H_j / r), \forall j : j \neq i
\]

\[
\begin{align*}
D_1 : & \quad \text{iff } \Pr(H_1 / r) > \Pr(H_j / r), \forall j \neq 1 \\
D_2 : & \quad \text{iff } \Pr(H_2 / r) > \Pr(H_j / r), \forall j \neq 2 \\
\vdots & \quad \vdots \\
D_M : & \quad \text{iff } \Pr(H_M / r) > \Pr(H_j / r), \forall j \neq M
\end{align*}
\]

*This detection is known as max a posterior probability (MAP) criterion*

---

**Equivalent MAP expressions:**

\[
D_i : \quad \text{iff } \Pr(H_i) \times \text{pdf}_{r/H_i}(r) > \Pr(H_j) \times \text{pdf}_{r/H_j}(r); \forall j : j \neq i
\]

\[
D_i = \max_{j, \forall j} \left( \Pr(H_j) \times \text{pdf}_{r/H_j}(r) \right)
\]

\[\triangleq G_j\]

**Note:**

- The above can be easily proven using the Bayes rule

\[
\Pr(H_i / r) = \frac{\text{pdf}_{r/H_i}(r) \cdot \Pr(H_i)}{\text{pdf}_r(r)}
\]

- \( G_j \) is known as "decision variable"
- In this topic the symbol \( G \) will be used to represent a "decision variable".
• Correlation between two **analogue energy signals** \( r(t) \) and \( s(t) \) of duration \( T_{cs} \):

\[
corr \triangleq \int_0^{T_{cs}} r(t).s(t).dt
\]  

(7)

• Correlation between the **discretised versions** \( r \) and \( s \) of the signals \( r(t) \) and \( s(t) \):

\[
corr \triangleq \underline{r}^T \underline{s}
\]  

(8)

where

\[
\underline{r} = [r_1, r_2, \ldots r_L]^T
\]  

(9)

and

\[
\underline{s} = [s_1, s_2, \ldots s_L]^T
\]  

(10)

---

**M-ary Decision Criteria**

• Consider the sets of parameters \( \mathcal{P}_1 \) and \( \mathcal{P}_2 \) where:

\( \mathcal{P}_1 \) denotes the set of parameters \( \Pr(H_1), \Pr(H_2), \ldots, \Pr(H_M) \)

\( \mathcal{P}_2 \) represents the set of costs/weights \( C_{ij}, \forall i, j \)

associated with the transition probabilities \( \Pr(D_i|H_j) \)

(i.e. one weight for every element of the channel transition matrix \( \mathbb{P} \) - see EE303, Topic on "Comm Channels")

• Estimate/identify the likelihood functions

\[
\text{pdf}_{r|H_1}(r), \ \text{pdf}_{r|H_2}(r), \ldots, \ \text{pdf}_{r|H_M}(r)
\]
- Decision: choose the hypothesis $H_i$ (i.e. $D_i$) with the maximum $G_i(r)$ where $G_i(r)$ depends on the chosen criterion as follows:
  - **Bayes** Criterion
  - **Minimum Probability of Error** $(\min(p_e))$ Criterion
  - **MAP** Criterion
  - **MINIMAX** Criterion
  - Newman-Pearson (N-P) Criterion
  - Maximum Likelihood (ML) Criterion

<table>
<thead>
<tr>
<th>$P_1$</th>
<th>$P_2$</th>
<th>choose Hypothesis with $\max(G_j(r))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>known</td>
<td>known</td>
<td>$G_j(r) \triangleq \text{weight } j \times \Pr(H_j) \times \text{pdf}_{r</td>
</tr>
<tr>
<td>known</td>
<td>unknown</td>
<td>$G_j(r) \triangleq \Pr(H_j) \times \text{pdf}_{r</td>
</tr>
<tr>
<td>unknown</td>
<td>known</td>
<td>$G_j(r) \triangleq \text{weight } j \times \text{pdf}_{r</td>
</tr>
<tr>
<td>unknown</td>
<td>unknown</td>
<td>by solving a constraint optim. problem</td>
</tr>
<tr>
<td>don't care</td>
<td>don't care</td>
<td>$G_j(r) \triangleq \text{pdf}_{r</td>
</tr>
</tbody>
</table>

- **Notes:**
  - **Note-1:**
    if an approximate/initial solution is required then any information about the sets of parameters $P_1$ and/or $P_2$ can be ignored. In this case the Maximum Likelihood (ML) Criterion should be used.
  - **Note-2:**
    \[
    \text{weight } j \triangleq \sum_{i=1}^{M} (C_{ij} - C_{jj})
    \]
    or (since the term for $i = j$ is equal to zero) simply,
    \[
    \text{weight } j = \sum_{i=1}^{M} (C_{ij} - C_{jj})
    \]
  - **Note-3:**
    Sometimes, for convenience, $G_j$ will be used (i.e. $G_j \triangleq G_j(r)$) - i.e. the argument will be ignored.
### Decision Criteria: Mathematical Architecture

**BAYES**
- Evaluate $weight_{1}, Pr(H_i) \cdot pdf_{r/H_i}$
- Evaluate $weight_{2}, Pr(H_j) \cdot pdf_{r/H_j}$
- Evaluate $weight_{u}, Pr(H_u) \cdot pdf_{r/H_u}$
- Compare Decision Variables and choose maximum
- If $G_r^{max}$ then $G_j \rightarrow D_j$

**MAP**
- Evaluate $Pr(H_i) \cdot pdf_{r/H_i}$
- Evaluate $Pr(H_j) \cdot pdf_{r/H_j}$
- Evaluate $Pr(H_u) \cdot pdf_{r/H_u}$
- Compare Decision Variables and choose maximum
- If $G_r^{max}$ then $G_j \rightarrow D_j$

**Minimax**
- Evaluate $pdf_{r/H_i}$
- Evaluate $pdf_{r/H_j}$
- Evaluate $pdf_{r/H_u}$
- Compare Decision Variables and choose maximum
- If $G_r^{max}$ then $G_j \rightarrow D_j$

**ML**
- Evaluate $pdf_{r/H_i}$
- Evaluate $pdf_{r/H_j}$
- Evaluate $pdf_{r/H_u}$
- Compare Decision Variables and choose maximum
- If $G_r^{max}$ then $G_j \rightarrow D_j$

---

### Examples

**Example: Gaussian LFs**

![Gaussian pdfs](image)

- By solving $G_j(r) = G_i(r)$ the decision threshold $r_{th,ji}$ can be estimated
- $\text{area-A} = Pr(D_j | H_i)$ and $\text{area-B} = Pr(D_i | H_j)$
- $p_{e,cs} = \sum_{j=1}^{M} \sum_{i=1, i \neq j}^{M} Pr(D_j, H_i) = \text{symbol error probability/rate (SER)}$
Example: Signal + Gaussian Noise

\[ p_{\text{pdf}_s}(s) = \frac{1}{2} \delta(s - A) + \frac{1}{2} \delta(s + A) \]

\[ p_{\text{pdf}_n}(n) = N(0, \sigma_n) \]

\[ p_{\text{pdf}_r}(r) = \frac{1}{2} N(-A, \sigma_n) + \frac{1}{2} N(A, \sigma_n) \]

 convolution

where

\[ \text{area 1} = \Pr(D_0, H_0) \]
\[ \text{area 2} = \Pr(D_1, H_0) \]

Hypothesis \( H_0 \):

\[ p_{\text{pdf}_{r|H_0}} : N(-A, \sigma_n) \]

Likelihood functions put together

\[ p_{\text{pdf}_{r|H_1}} : N(0, \sigma_n) \]

where

\[ \text{area 3} = \Pr(D_0 | H_1) \]
\[ \text{area 4} = \Pr(D_1 | H_1) \]

Note: \[ \Pr = \frac{\Pr(D_0 | H_1) \cdot \Pr(H_1)}{\Pr(D_0 | H_0) \cdot \Pr(H_0) + \Pr(D_1 | H_0) \cdot \Pr(H_0) + \Pr(D_0 | H_1) \cdot \Pr(H_1) + \Pr(D_1 | H_1) \cdot \Pr(H_1)} \]
Example: Rectangular and Gaussian LFs

- A discrete channel employs two equally probable symbols and the likelihood functions are given by the following expressions:

\[
\text{pdf}_{r|H_0} = \frac{1}{2} \text{rect} \left\{ \frac{r}{2} \right\}
\]

\[
\text{pdf}_{r|H_1} = \mathcal{N} \left( 0, \sigma = \frac{1}{2} \right)
\]

- If the detector employed has been designed in an optimum way, find the decision rule and model the above discrete channel.

Note:

\[
\begin{align*}
\text{area B} &= \text{area C} = T \left\{ \frac{0.483}{1/2} \right\} - T \left\{ \frac{1}{1/2} \right\} = 0.1490; \\
\Pr(D_0/H_1) &= \text{area B + area C} = 0.298 \\
\text{and} \quad \Pr(D_1/H_0) &= \text{area A} = 0.483
\end{align*}
\]
Basic Concepts on Optimum Receivers

- Receiver = Detector + Decision Device

- Consider a $M$-ary communication model in which one of $M$ signals $s_i(t)$, for $i = 1, 2, \ldots, M$, is received in the time interval $(0, T_{cs})$ in the presence of white noise i.e.

$$r(t) = \left\{ \begin{array}{l} s_1(t) \\
\text{or } s_2(t) \\
\ldots \\
\text{or } s_M(t) 
\end{array} \right\} + n(t), \quad 0 \leq t \leq T_{cs}$$  \hspace{1cm} (13)

where $r(t)$ denotes the received (observable) signal.
The Concept of "Continuous Sampling"

Continuous Sampling helps to get a probabilistic description of a continuous signal:

1. assume that $L$ amplitude samples of the received signal

   $$r(t), \quad 0 \leq t \leq T_{cs}$$

   are available, i.e. $r_1, r_2, \ldots, r_L$, (with sampling frequency $\frac{1}{\Delta t}$)

   where $r_k = \begin{cases} s_{1k} \\ or \quad s_{2k} \\ \ldots \\ or \quad s_{Mk} \end{cases} + n_k; \quad k = 1, \ldots, L$

   or, in more compact form,

   $$r = \begin{cases} s_1 \\ or \quad s_2 \\ \ldots \\ or \quad s_M \end{cases} + n$$

   where

   $$r = \begin{bmatrix} r_1 \\ r_2 \\ \cdots \\ r_L \end{bmatrix}^T$$

   $$n = \begin{bmatrix} n_1 \\ n_2 \\ \cdots \\ n_L \end{bmatrix}^T$$

   $$s_i = \begin{bmatrix} s_{i1} \\ s_{i2} \\ \cdots \\ s_{iL} \end{bmatrix}^T$$

   i.e. $r = s_i + n$ (L samples)

2. take the limit as $L \to \infty$, $\Delta t \to 0$, $L \times \Delta t \to T_{cs}$
Optimum M-ary Receivers

- **Objective**: to design a receiver which operates on \( r(t) \) and chooses one of the following \( M \) hypotheses:

\[
\begin{align*}
    &H_1: \quad r(t) = s_1(t) + n(t) \\
    &H_2: \quad r(t) = s_2(t) + n(t) \\
    &\cdots \\
    &H_M: \quad r(t) = s_M(t) + n(t)
\end{align*}
\]

or

\[
\begin{align*}
    &H_1: \quad r = s_1 + n \\
    &H_2: \quad r = s_2 + n \\
    &\cdots \\
    &H_M: \quad r = s_M + n
\end{align*}
\]

- **Optimum Decision Rule**: depends on the likelihood functions \( \text{pdf}_{r|H_i}(r) \)

- **Optimum Receiver - Initial Conceptual Structure**:

\[
\begin{align*}
    r(t) &= s(i) + n(t) \\
    &\Delta t \\
    \text{Compute } M \quad \text{Compare} \\
    \text{likelihood} \quad \text{variables} \\
    \text{functions} \quad \text{and get max} \\
    D_i
\end{align*}
\]

- **Assumptions**:
  - **Noise**: \( \text{PSD}_{n_i}(f) = \frac{N_0}{2} \Rightarrow \text{PSD}_n(f) = \frac{N_0}{2} \text{ rect} \left\{ \frac{f}{2B} \right\} \)
    
    Therefore, the received signal is sampled at intervals
    
    \[
    \Delta t = \frac{1}{2B} \tag{14}
    \]
  - **Sampling**: the samples are uncorrelated and statistically independent

- **Example - ML Receiver with "Cont. Sampling"**: 

\[
\begin{align*}
    &\text{ML} \\
    &r(t) \Delta t \\
    &\text{Evaluate} \quad \text{Evaluate} \quad \text{Evaluate} \\
    &\text{pdf}_{r|H_1} \quad \text{pdf}_{r|H_2} \quad \text{pdf}_{r|H_M} \\
    &\text{Compare} \quad \text{Compare} \\
    &\text{Decision} \quad \text{Decision} \\
    &\text{Variables} \quad \text{Variables} \\
    &\text{and} \quad \text{and} \\
    &\text{choose} \quad \text{choose} \\
    &\text{maximum} \quad \text{maximum} \\
    &\text{If } G_i = \text{max} \quad \text{If } G_i = \text{max} \\
    &\text{then } G_i \rightarrow D_i \quad \text{then } G_i \rightarrow D_i
\end{align*}
\]
Gaussian Multivariable Distribution

- Consider the Gaussian random vector: \( \mathbf{r} = [r_1, r_2, \ldots, r_L]^T \)

Then

\[
\text{pdf}(\mathbf{r}) = \frac{1}{\sqrt{(2\pi)^L \det(\mathbf{R}_{rr})}} \exp \left\{ -\frac{(\mathbf{r} - \mu_r)^T \mathbf{R}_{rr}^{-1} (\mathbf{r} - \mu_r)}{2} \right\} \tag{15}
\]

where

\[
\text{mean} = \mu_r = \mathbb{E}\{\mathbf{r}\} = [\mathbb{E}\{r_1\}, \mathbb{E}\{r_2\}, \ldots, \mathbb{E}\{r_L\}]^T \tag{16}
\]

\[
\text{cov}(\mathbf{r}) = \mathbf{R}_{rr} = \mathbb{E}\left\{ (\mathbf{r} - \mu_r)(\mathbf{r} - \mu_r)^T \right\} \tag{17}
\]

- If \( \mathbf{r} = \mathbf{s} + \mathbf{n} \) then the mean \( \mu_r \) and covariance \( \mathbf{R}_{rr} \) are given as follows:

\[
\mu_r = \mathbf{s} \quad \text{and} \quad \mathbf{R}_{rr} = \mathbf{R}_{nn} \tag{18}
\]

**proof:**

\[
\mathbb{E}\{\mathbf{r}\} = \mu_r \triangleq \mathbb{E}\{\mathbf{s} + \mathbf{n}\} = \mathbb{E}\{\mathbf{s}\} + \mathbb{E}\{\mathbf{n}\} = \mathbb{E}\{\mathbf{s}\} = \mathbf{s} \tag{19}
\]

\[
\text{cov}\{\mathbf{r}\} \triangleq \mathbf{R}_{rr} = \mathbb{E}\left\{ (\mathbf{r} - \mathbb{E}\{\mathbf{r}\})(\mathbf{r} - \mathbb{E}\{\mathbf{r}\})^T \right\}
\]

\[
= \mathbb{E}\left\{ (\mathbf{r} - \mathbf{s})(\mathbf{r} - \mathbf{s})^T \right\}
\]

\[
= \mathbb{E}\{\mathbf{n}\mathbf{n}^T\}
\]

\[
\Rightarrow \text{cov}\{\mathbf{r}\} = \mathbf{R}_{nn} = \sigma_n^2 \mathbf{I}_L = \mathbf{N}_0 \mathbf{B} \mathbf{I}_L \tag{20}
\]

- Note that:

\[
\text{If} \quad \mathbf{R}_{nn} = \sigma_n^2 \mathbf{I}_L \quad \text{then} \quad \det(\mathbf{R}_{nn}) = (\sigma_n^2)^L \tag{21}
\]
Likelihood Functions (Continuous Sampling)

- \( \text{pdf}_{r/H_i}(r) \): Based on the above (for AWGN) and for “Continuous Sampling” we have

\[
\text{pdf}_{r/H_i}(r) = \left( \frac{1}{\sqrt{2\pi\sigma_n^2}} \right)^L \cdot \exp \left\{ -\frac{(r - s_i)^T(r - s_i)}{2\sigma_n^2} \right\} \quad (22)
\]

- However \( \sigma_n^2 = N_0 B \) \( \triangle t = \frac{1}{2B} \) \( \Rightarrow \sigma_n^2 = \frac{N_0}{2 \cdot \triangle t} \)

- Therefore

\[
\text{pdf}_{r/H_i}(r) = \left( \frac{\triangle t}{\pi N_0} \right)^L \cdot \exp \left\{ -\frac{(r - s_i)^T(r - s_i) \cdot \triangle t}{N_0} \right\}
\]

and for “Continuous Sampling”, i.e. \( L \rightarrow \infty, \triangle t \rightarrow 0, L \times \triangle t \rightarrow T_{cs} \) we have

\[
\text{pdf}_{r/H_i}(r(t)) = \text{const} \cdot \exp \left\{ -\frac{1}{N_0} \cdot \int_0^{T_{cs}} (r(t) - s_i(t))^2 dt \right\} \quad (23)
\]

**ML APPROACH** using "Continuous Sampling":

**Rule**: choose the hypothesis \( H_i \) with the maximum likelihood \( \text{pdf}_{r/H_i}(r(t)) \) where

\[
\text{pdf}_{r/H_i}(r(t)) = \text{const} \cdot \exp \left\{ -\frac{1}{N_0} \cdot \int_0^{T_{cs}} (r(t) - s_i(t))^2 dt \right\} \quad (24)
\]

for \( i = 1, \ldots, M \)

**MAP APPROACH** using "Continuous Sampling":

This is a more general approach than the Maximum Likelihood

**Rule**: choose the hypothesis \( H_i \) with the maximum \( \Pr(H_i) \times \text{pdf}_{r/H_i}(r(t)) \) where

\[
\text{pdf}_{r/H_i}(r(t)) = \text{const} \cdot \exp \left\{ -\frac{1}{N_0} \cdot \int_0^{T_{cs}} (r(t) - s_i(t))^2 dt \right\} \quad (25)
\]

for \( i = 1, \ldots, M \)

i.e.

\[
\max \left\{ \Pr(H_i) \times \text{pdf}_{r/H_i}(r(t)) \right\} = \max \left\{ \int_0^{T_{cs}} r(t) s_i^*(t) dt + \frac{N_0}{2} \ln \left( \Pr(H_i) - \frac{1}{2} E_i \right) \right\} \quad (26)
\]
Equation-26 (for proof see Appendix 2) suggests that an $M$-ary receiver will have the following form:

\[
\text{OPTIMUM } M\text{-ary RECEIVER: CORREL. RECEIVER}
\]

\[
\begin{align*}
\text{Compare} & \quad \text{Decision} \\
\text{Variables} & \quad \text{and} \\
\text{choose} & \quad \text{maximum} \\
\text{If } & \quad G_i = \max \\
\text{then } & \quad G_i \rightarrow D_i
\end{align*}
\]

where \( DC_i \equiv \frac{N_0}{2} \ln(\Pr(H_i)) - \frac{1}{2} E_i \) with \( i = 1, 2, \ldots, M \)

Note that if the a priori probabilities are equal, i.e.

\[
\Pr(H_1) = \Pr(H_2) = \cdots = \Pr(H_i)
\]

and the signals have the same energy, i.e.

\[
E_1 = E_2 = \cdots = E_M
\]

then Equation-26 is simplified as follows:

\[
\max \left\{ \Pr(H_i) \times \text{pdf}_{r/H_i}(r(t)) \right\} = \max \left\{ \text{pdf}_{r/H_i}(r(t)) \right\} = \max \left\{ \int_0^{T_{cs}} r(t)s_i^*(t) \cdot dt \right\}
\]

N.B. - ML receiver: similar structure to MAP with \( DC_i = \frac{1}{2} E_i \)
Optimum M-ary Receivers using Matched Filters
Known Signals in White Noise

- Consider the output of one branch of the correlation receiver:

$$r(t) = s_i(t) + n(t)$$

**at point-1** = output = \( \int_0^{T_{cs}} r(t) \cdot s_i(t) \cdot dt \)

- In this section we will try to replace the **correlator** of a correlation receiver with a **linear filter** (known as Matched Filter).

$$r(t) = s_i(t) + n(t)$$

**at point-1** = \( \int_0^t r(u) \cdot h_i(t - u) \cdot dv \)

**at point-2** = output = \( \int_0^{T_{cs}} r(u) \cdot h_i(T_{cs} - u) \cdot du \)

- N.B.:
  compare correlator’s o/p (point-1) with matched filter’s o/p (point-2).
If we choose the impulse response of the linear filter as

$$h_i(t) = s_i(T_{cs} - t) \quad 0 \leq t \leq T_{cs}$$

(29)

then that linear filter, defined by the above equation, is called Matched Filter.

at point-2 $\int_0^{T_{cs}} r(u) \cdot s_i(u) \cdot du$

$\Rightarrow$ at point-2 $\int_0^{T_{cs}} r(t) \cdot s_i(t) \cdot dt$

N.B.: Output Of Correlator = Output Of Matched Filter

only at time $t = T_{cs}$

example:

![Graph showing the example of a matched filter and correlator output](image-url)
CORRELATION RECEIVER:

**OPTIMUM M-ary RECEIVER: CORREL RECEIVER**

\[ r(t) \]

\[ T \]

\[ T_c \]

\[ s_1(t) \]

\[ s_2(t) \]

\[ G_1 \]

\[ G_2 \]

\[ G_M \]

\[ DC_1 \]

\[ DC_2 \]

\[ DC_M \]

Compare Decision Variables and choose maximum

If \( G_i = \max \) then \( G_i \rightarrow D_i \)

or, MATCHED FILTER RECEIVER:

**OPTIMUM M-ary RECEIVER: MATCHED FILTER RECEIVER**

\[ h_a(t) = a(T_c - t) \]

\[ T_c \]

\[ T_a \]

\[ h_a(t) = a(T_c - t) \]

\[ h_a(t) = a(T_c - t) \]

\[ G_1 \]

\[ G_2 \]

\[ G_M \]

\[ DC_1 \]

\[ DC_2 \]

\[ DC_M \]

Compare Decision Variables and choose maximum

If \( G_i = \max \) then \( G_i \rightarrow D_i \)
Signals and Matched Filters in Freq. Domain

- Let \( h(t) = \begin{cases} \ s(T_{cs} - t) & 0 \leq t \leq T_{cs} \\ 0 & \text{elsewhere} \end{cases} \)

- **Time Domain** \( \xrightarrow{\text{FT}} \) **Freq. Domain**

\[
\begin{align*}
    s(t) & \xrightarrow{\text{FT}} S(f) = \int_{0}^{T_{cs}} s(t)e^{-j2\pi ft} dt \\
    h(t) & \xrightarrow{\text{FT}} H(f) = \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft} dt \\
    & = \int_{0}^{T_{cs}} s(T_{cs} - t)e^{-j2\pi ft} dt \\
    & = \int_{0}^{T_{cs}} s(u)e^{-j2\pi ft(T_{cs} - u)} du \\
    & = e^{-j2\pi ftT_{cs}} \int_{0}^{T_{cs}} s(u)e^{j2\pi fu} du \\
    & = e^{-j2\pi ftT_{cs}} \cdot S^*(f)
\end{align*}
\]

therefore

\[
H(f) = e^{-j2\pi ftT_{cs}} \cdot S^*(f)
\] (30)

Known Signals in Non-White Noise

- Now let us assume that the noise \( n(t) \) is described as follows:

\[
n(t) = \begin{cases} \ \text{zero mean} \\ R_{nn}(\tau) = \text{not necessarily white or Gaussian} \end{cases}
\] (31)

- Let

\[
r(t) = \underbrace{s(t)}_{\text{completely known}} + n(t)
\] (32)

- objective: design a linear filter \( h(t) \) such that

\[
\text{SNR}_{\text{out}} = \max_{t=T_{cs}}
\] (33)
\[
\text{output} = \int_0^{T_{cs}} h(\tau) \cdot \sqrt{\frac{\sqrt{\hat{s}(T_{cs})^2}}{\sqrt{\hat{n}(T_{cs})^2}}} \cdot dt \\
= \int_0^{T_{cs}} h(\tau) \cdot s(T_{cs} - \tau) dt + \int_0^{T_{cs}} h(\tau) \cdot n(T_{cs} - \tau) d\tau \\
= \hat{s}(T_{cs}) + \hat{n}(T_{cs}) \\
\Rightarrow \text{SNR}_{out} = \frac{\mathcal{E}\{\hat{s}(T_{cs})^2\}}{\mathcal{E}\{\hat{n}(T_{cs})^2\}} = \frac{\hat{s}(T_{cs})^2}{\mathcal{E}\{\hat{n}(T_{cs})^2\}} 
\]

(34)

\[
\Rightarrow \max (\text{SNR}_{\text{out}}) = \left\{ \min_{\hat{s}(T_{cs}) \text{ constant}} \mathcal{E}\{\hat{n}(T_{cs})^2\} \right\} \quad \text{(35)}
\]

\[
\min_{\hat{s}(T_{cs})} \text{ where } \zeta = \mathcal{E}\{\hat{n}(T_{cs})^2\} - \lambda \cdot \hat{s}(T_{cs}) 
\]

(36)

\[\text{i.e.} \quad \text{optimum impulse response } \overset{\Delta}{=} h_{\text{opt}} = \arg \min_{\hat{s}(T_{cs})} \zeta \quad \text{(37)}\]

where \[\zeta = \mathcal{E}\{\hat{n}(T_{cs})^2\} - \lambda \cdot \hat{s}(T_{cs}) \]

The solution of Equation 37 is the solution of the following

FREDHOLM INTEGRAL EQUATION of 1st KIND:

\[
\int_0^{T_{cs}} h_{\text{opt}}(z) \cdot R_{nn}(\tau - z) \cdot dz - s(T_{cs} - \tau); \quad 0 \leq \tau \leq T_{cs} \quad \text{(38)}
\]

GENERAL EXPRESSION FOR MATCHED FILTERS

Proof of the above result (Equation 38) is not important. What is really important is that the proof does not involve any assumption about the pdf or PSD of the noise (i.e. that the noise is white Gaussian).
A Special Case

- In the case of white noise the above Equation (Equation-38) can be solved easily as follows:

\[
\int_0^{T_{cs}} h_{opt}(z) \cdot \frac{N_0}{2} \cdot \delta(z) \cdot \alpha z = s(T_{cs} - \tau) \quad (39)
\]

\[
\Rightarrow h_{opt}(\tau) \cdot \frac{N_0}{2} = s(T_{cs} - \tau) \Rightarrow
\]

\[
h_{opt} = \frac{2}{N_0} \cdot s(T_{cs} - \tau) \quad (40)
\]

- Be careful - SNR is not influenced by the factor \( \frac{2}{N_0} \)

Output SNR

- Consider next you have got the optimum impulse response \( h_0(t) \) (by using Equation-38). This impulse response provides the maximum output-SNR which can be estimated as follows:

![Matched Filter Diagram](image)

**Matched Filter**

\[ r(t) = s(t) + n(t) \]

\[ Linear \ filter \ h_{opt}(t) \]

\[ T_{cs} \]

- Important Question: \( \text{SNR}_{out} =? \)

we can answer this question as follows:
\[\text{o/p} = \int_0^{T_{cs}} h_{opt}(\tau) \cdot \frac{r(T_{cs} - \tau)}{s(T_{cs} - \tau) + n(T_{cs} - \tau)} \, d\tau\]

\[= \int_0^{T_{cs}} h_{opt}(\tau) \cdot s(T_{cs} - \tau) \, d\tau + \int_0^{T_{cs}} h_{opt}(\tau) \cdot n(T_{cs} - \tau) \, d\tau \tag{41}\]

therefore,

\[SNR_{\text{out}}_{\text{max}} = \frac{\mathcal{E}\{\tilde{s}(T_{cs})^2\}}{\mathcal{E}\{\tilde{n}(T_{cs})^2\}} = \frac{\tilde{s}(T_{cs})^2}{\mathcal{E}\{\tilde{n}(T_{cs})^2\}} = \ldots \ldots \tag{42}\]

\[\ldots \text{(for you)} \ldots\]

\[\text{i.e.} \quad SNR_{\text{out}}_{\text{max}} = \frac{\tilde{s}(T_{cs})^2}{\mathcal{E}\{\tilde{n}(T_{cs})^2\}} = \int_0^{T_{cs}} h_{opt}(z) \cdot s(T_{cs} - z) \, dz \tag{43}\]
A Special Case: Output SNR (Signal plus white Noise)

- For white noise we have seen that: \( h_{opt}(z) = \frac{2}{N_0} \cdot s(T_{cs} - z) \)
- Then Equation-43 becomes:
  \[
  \text{SNR}_{\text{out max}} = \int_0^{T_{cs}} \frac{2}{N_0} \cdot s(T_{cs} - z) \cdot s(T_{cs} - z) \cdot dz \\
  = \frac{2}{N_0} \int_0^{T_{cs}} s^2(T_{cs} - z) \cdot dz = \\
  \implies \text{SNR}_{\text{out max}} = 2 \frac{E}{N_0} \text{ for white noise}
  \] (44)

- Provided that the filter is matched to the signal, it is obvious from Equation-44 that
  \[\text{SNR}_{\text{out max}} \neq \{\text{signal waveform}\} \neq \{\text{signal bandwidth}\} \neq \{\text{peak power}\} \neq \{\text{time duration}\}\]

Approximation To Matched Filter Solution

- We have seen that the matched filter can be found by solving the following integral equation:
  \[
  \int_0^{T_{cs}} h_{opt}(z) \cdot R_{nn}(\tau - z) \cdot dz = s(T_{cs} - \tau); \ 0 \leq \tau \leq T_{cs}
  \]

MATCHED FILTER GENERAL EQUATION
(a Fredholm Integral of the 1st kind.)

- However, it is very difficult to solve the above equation in a general case.

- N.B.: solution=easy when noise=white
- In order to find an approximation to matched filter solution, i.e.
  \[ w_0(z) \simeq h_{opt}(z) \]
  we need to relax the condition \( 0 \leq \tau \leq T_{cs} \) to \( -\infty \leq \tau \leq \infty \)
  \[ 0 \leq \tau \leq T_{cs} \text{ to } -\infty \leq \tau \leq \infty \]

  Then
  \[ \int_{-\infty}^{\infty} w_0(z)R_{nn}(\tau-z) \cdot dz = s(T_0 - \tau) \quad (45) \]
  Maximizes SNR at some time
  However the above integral is a convolution integral

  \[ \therefore W_0(f) \cdot PSD_n(f) = S^*(f) e^{-j2\pi f T_0} \quad (46) \]

  \[ \Rightarrow W_0(f) = \frac{S^*(f) e^{-j2\pi f T_0}}{PSD_n(f)} \quad (47) \]

---

**Whitening Filter**

- \( r(t) = s(t) + n(t) : i/p \rightleftharpoons \text{WHITENING FILTER} \rightarrow \text{MATCHED FILTER} \rightarrow \alpha/p : s(t) + n_i(t) \)
- whitening filter:
  - attenuates regions in frequency domain in which \text{NOISE=LARGE}
  - accentuates regions in frequency domain in which \text{NOISE=LOW}

- combined Whitening Filter-Matched Filter transfer function:

  \[ r(t) \rightarrow \text{WHITENING FILTER} \rightarrow \text{MATCHED FILTER} \rightarrow \text{output} \]

  combined transfer function \( \text{IL}(f) = \frac{S^*(f) e^{-j2\pi f T_0}}{PSD_n(f)} \)
Optimum M-ary Rx based on Signal Constellation

Introduction - Orthogonal Signals

- Two signals $c_i(t)$ and $c_j(t)$ are called orthogonal signals (i.e., $c_i(t) \perp c_j(t)$) in the interval $[a, b]$ iff:

$$\int_a^b c_i(t) \cdot c_j^*(t) \, dt = \begin{cases} E_{c_i} & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$$  \hspace{1cm} (48)

- Note:
  - if $E_{c_i} = 1$ then the signals are called orthonormal
  - $\int_a^b c_i(t) \cdot c_i^*(t) \, dt = E_{c_i} = \text{the energy of } c_i(t) \text{ in the interval } [a, b]$. 

---

The “Approximation Theorem” of an Energy Signal

- **Theorem**: Consider an energy signal $s(t)$. Furthermore consider a set of $D$ orthogonal signals

$$\{c_1(t), c_2(t), \ldots, c_D(t)\}$$

Then

- The signal $s(t)$ can be approximated by a linear combination of the orthogonal signals $c_i(t)$ as follows:

$$\hat{s}(t) = \sum_{i=1}^D w_i^* c_i(t) = w_s^H c(t)$$  \hspace{1cm} (49)

where

$$\begin{align*}
    w_s &= [w_1, w_2, \ldots, w_D]^T \\
    c(t) &= [c_1(t), c_2(t), \ldots, c_D(t)]^T
\end{align*}$$

- $w_s$ is a vector of coefficients or weights.
The best approximation is the one which uses the following coefficients

\[ w_s = \frac{1}{E_c} \odot \int_a^b s(t) \cdot c^*(t) \, dt = \text{optimum weights} \quad (50) \]

where

\[ E_c = [E_{c_1}, E_{c_2}, \ldots, E_{c_E}]^T \quad (51) \]

and \( \odot \) denotes Hadamard multiplication.

The above coefficients are optimum in the sense that they minimize the energy \( E_c \) of the error signal

\[ e(t) = s(t) - \hat{s}(t) \quad (52) \]

Comments on the "Approximation" Theorem

- The above theorem states that any energy signal can be approximated by a linear combination of a set of orthogonal signals and can be equivalent described by the following diagram:
• The energy of the error signal $e(t)$ is minimum when the set of signals
  $\{e(t), c_1(t), c_2(t), ..., c_D(t)\}$ is an orthogonal set
  i.e if
  \[
  \begin{cases}
  e(t) \perp c_i(t), \forall i \\
  \int_a^b (s(t) - w_s^H \zeta(t) - c_i^*(t)) \cdot c_i^*(t) dt = 0
  \end{cases}
  \]
  then $E_e = \text{min}$

• In general
  \[s(t) \approx \tilde{s}(t)\] (54)

• However, if
  \[s(t) = \hat{s}(t)\] (55)
  then $\{c_1(t), c_2(t), ..., c_D(t)\}$ is a complete set of orthogonal signals.

---

**Gram-Schmidt Orthogonalization**

• Suppose that we have a set of energy signals
  \[\{s_1(t), s_2(t), ..., s_M(t)\} = \text{non-orthogonal}\]
  and we wish to construct a set of orthogonal signals
  \[\{c_1(t), c_2(t), ..., c_D(t)\} = \text{orthogonal}\]

• One popular approach to construct this orthogonal set is the **GRAM-SCHMIDT orthogonalisation**, described as follows:
1st orthogonal signal: \( c_1(t) = \frac{s_1(t)}{\sqrt{\text{energy of numerator } E_{1,\text{num}}}} \)

2nd orthogonal signal: \( c_2(t) = \frac{s_2(t) - \alpha_{21} c_1(t)}{\sqrt{\text{energy of numerator } E_{2,\text{num}}}} \)

... or in general, \( k^{th} \) orthogonal signal: \( c_k(t) = \frac{s_k(t) - \sum_{l=1}^{k-1} \alpha_{kl} c_l(t)}{\sqrt{\text{energy of numerator } E_{k,\text{num}}}} \)
M-ary Signals: Energy and Cross-Correlation

- **Binary Com. Systems**: use 2 possible waveforms \( \{s_0(t), s_1(t)\} \) or \( \{s_1(t), s_2(t)\} \)

- **M-ary Com. Systems**: use \( M \) possible waveforms \( \{s_1(t), \ldots, s_M(t)\} \)
  Note: these waveforms are energy signals of duration \( T_{cs} \)

- The \( M \) signals (or channel symbols) are characterized by their energy \( E_i \)
  \[
  E_i = \int_0^{T_{cs}} s_i^2(t) \, dt \quad \forall i
  \]

- Furthermore their similarity (or dissimilarity) is characterized by their cross-correlation
  \[
  \rho_{ij} = \frac{1}{\sqrt{E_i E_j}} \int_0^{T_{cs}} s_i(t) \cdot s_j^*(t) \, dt
  \]

The \( M \) signals may be also expressed, by using the Approx. Theorem described in the previous section, as a linear combination of a set of \( D \) orthogonal signals \( \{c_k(t)\} \)

\[
\begin{align*}
  s_i(t) &= w_{si}^H \mathbf{c}(t); \forall i \\
  \mathbf{c}(t) &= [c_1(t), c_2(t), \ldots, c_D(t)]^T
\end{align*}
\]
Let us define the following \((D \times D)\) matrix

\[
\mathbf{R}_{cc} = \int_{0}^{T_{cs}} \mathbf{c}(t)\mathbf{c}(t)^{H}.dt
\]

\[
= \begin{bmatrix}
\int_{0}^{T_{cs}} c_1(t)c_1^*(t)dt, & \int_{0}^{T_{cs}} c_1(t)c_2^*(t)dt, & \ldots, & \int_{0}^{T_{cs}} c_1(t)c_D^*(t)dt \\
\int_{0}^{T_{cs}} c_2(t)c_1^*(t)dt, & \int_{0}^{T_{cs}} c_2(t)c_2^*(t)dt, & \ldots, & \int_{0}^{T_{cs}} c_2(t)c_D^*(t)dt \\
\vdots & \vdots & \ddots & \vdots \\
\int_{0}^{T_{cs}} c_D(t)c_1^*(t)dt, & \int_{0}^{T_{cs}} c_D(t)c_2^*(t)dt, & \ldots, & \int_{0}^{T_{cs}} c_D(t)c_D^*(t)dt
\end{bmatrix}
\]

i.e.

\[
\mathbf{R}_{cc} = \int_{0}^{T_{cs}} \mathbf{c}(t)\mathbf{c}(t)^{H}.dt = \begin{bmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{bmatrix} = \mathbf{I}_D
\]  \(\text{(60)}\)

Then, by using Equation-56 in conjunction with Equation-58, the energy \(E_i\) of the signal \(s_i(t), \forall i\), can be expressed as follows:

\[
E_i = \int_{0}^{T_{cs}} \underbrace{\mathbf{w}_{s_i}^{H}\mathbf{c}(t)}_{=s_i(t)} \cdot \underbrace{\mathbf{c}(t)^{H}\mathbf{w}_{s_i}}_{=s_i^*(t)} . dt
\]

\[
= \mathbf{w}_{s_i}^{H} \left( \int_{0}^{T_{cs}} \mathbf{c}(t)\mathbf{c}(t)^{H}.dt \right) \mathbf{w}_{s_i}
\]

\[
= \mathbf{w}_{s_i}^{H} \cdot \mathbf{R}_{cc} \cdot \mathbf{w}_{s_i}
\]

\[
= \mathbf{w}_{s_i}^{H} \cdot \mathbf{I}_D \cdot \mathbf{w}_{s_i}
\]

\[
= \mathbf{w}_{s_i}^{H} \mathbf{w}_{s_i}
\]

\[
= \|\mathbf{w}_{s_i}\|^2
\]  \(\text{(61)}\)
Similarly, by using Equation-57 in conjunction with Equation-58 the cross-correlation coef. $\rho_{ij} \forall ij$ becomes

$$
\rho_{ij} = \frac{1}{\sqrt{E_i E_j}} \int_0^{T_c} w_{s_i}^H \cdot c(t) \cdot c(t)^H \cdot w_{s_j} \cdot dt \\
= \frac{1}{\sqrt{E_i E_j}} \cdot w_{s_i}^{T} \cdot \left( \int_0^{T_c} c(t) \cdot c(t)^H \cdot dt \right) \cdot w_{s_j} \\
= \frac{1}{\sqrt{E_i E_j}} w_{s_i}^H \cdot R_{cc} \cdot w_{s_j} \\
= \frac{1}{\sqrt{E_i E_j}} w_{s_i}^H \cdot I_{D} \cdot w_{s_j} \\
= \frac{1}{\sqrt{E_i E_j}} w_{s_i}^H w_{s_j} = \frac{w_{s_i}^H w_{s_j}}{\|w_{s_i}\| \|w_{s_j}\|} 
$$

i.e.

$$
E_i = w_{s_i}^H w_{s_i} = \|w_{s_i}\|^2 \quad (62)
$$

$$
\rho_{ij} = \frac{1}{\sqrt{E_i E_j}} w_{s_i}^H w_{s_j} = \frac{w_{s_i}^H w_{s_j}}{\|w_{s_i}\| \|w_{s_j}\|} \quad (63)
$$
Signal Constellation

- With reference to the figure below

- If the error signal $\varepsilon(t) = 0$
  then the knowledge of the vector $w_i$ is as good as knowing the transmitted signal $s_i(t)$
  or, in the case of $M$ signals ($M$-ary system)
  \{knowledge of $s_i(t)$\} = \{knowledge of $w_{s_i}$\}, $\forall i$

Therefore, we may represent the signal $s_i(t)$
by a point ($w_{s_i}$) in a $D$-dimensional Euclidean space with

$$D \leq M$$

The set of points (vectors) specified by the columns of the matrix

$$W = [w_{s_1}, w_{s_2}, ..., w_{s_M}]$$

is known as “signal constellation”.

---

Prof. A. Manikas (Imperial College)  |  EE.401: Detection Theory  |  v.16  |  65 / 84
Distance between two M-ary signals

- The distance between two signals $s_i(t)$ and $s_j(t)$ is the Euclidean distance between their associate vectors $w_{s_i}$ and $w_{s_j}$

i.e. $d_{ij} = \| w_{s_i} - w_{s_j} \|$

$$= \sqrt{(w_{s_i} - w_{s_j})^H(w_{s_i} - w_{s_j})}$$

$$= \sqrt{w_{s_i}^Hw_{s_i} + w_{s_j}^Hw_{s_j} - 2\text{Re}(w_{s_i}^Hw_{s_j})}$$

$$= \sqrt{E_i + E_j - 2\rho_{ij}\sqrt{E_iE_j}}$$

$$d_{ij}^2 = E_i + E_j - 2\rho_{ij}\sqrt{E_iE_j} \quad (66)$$

It is clear from the above that the Euclidean distance $d_{ij}$ of two signals indicates, like the cross-correlation coefficient, the similarity or dissimilarity of the signals.

An Important Bound involving $d_{ij}$:

- If $p_{e,cs(s_i(t))}$ denotes the prob. of error associated with the channel symbol $s_i(t)$, it can be found that (see S. Haykin p.498)

$$p_{e,cs(s_i(t))} \leq \frac{M}{\sum_{j=1\atop j \neq i}^{M}} T \left\{ \frac{d_{ij}}{\sqrt{2N_0}} \right\} \quad (67)$$

Then, the symbol error probability $p_{e,sc}$ is bounded as follows:

$$p_{e,sc} \leq \frac{1}{M} \sum_{i=1}^{M} \sum_{j=1\atop j \neq i}^{M} T \left\{ \frac{q_{ij}}{\sqrt{2N_0}} \right\} \quad (68)$$

- N.B.: the following expression was used to go from Equ.67 to Equ.68

$$p_{e,sc} = \frac{1}{M} \sum_{i=1}^{M} p_{e,cs(s_i(t))} \quad \text{(see S.Haykin p.498)} \quad (69)$$

- “minimum distance”: $d_{min} \triangleq \min_{i,j} \{ d_{ij} \}$
Examples of Signal Constellation Diagram

Consider an M-ary System having the following signals

\[ \{s_1(t), s_2(t), \ldots, s_M(t)\} \text{ with } 0 \leq t \leq T_{cs} \]

- M-ary ASK
  - channel symbols: \( s_i(t) = A_i \cdot \cos(2\pi F_c t) \) where \( A_i = \text{given}= 2i - 1 - M \) (say)
  - dimensionality of signal space = \( D = 1 \)
  - if \( M = 4 \) then
    \[
    \begin{array}{c|c|c|c|c}
    \sqrt{E_{s_1}} & \sqrt{E_{s_2}} & \sqrt{E_{s_3}} & \sqrt{E_{s_4}} \\
    00 & 01 & 11 & 10 \\
    s_1(t) & s_2(t) & s_3(t) & s_4(t)
    \end{array}
    \]
  - if \( M = 8 \) then
    \[
    \begin{array}{c|c|c|c|c|c|c|c|c}
    \sqrt{E_{s_1}} & \sqrt{E_{s_2}} & \sqrt{E_{s_3}} & \sqrt{E_{s_4}} & \sqrt{E_{s_5}} & \sqrt{E_{s_6}} & \sqrt{E_{s_7}} & \sqrt{E_{s_8}} \\
    000 & 001 & 011 & 010 & 110 & 111 & 101 & 100 \\
    s_1(t) & s_2(t) & s_3(t) & s_4(t) & s_5(t) & s_6(t) & s_7(t) & s_8(t)
    \end{array}
    \]

- M-ary PSK
  - channel symbols: \( s_i(t) = A \cdot \cos\left(\frac{2\pi F_c t}{M} \cdot (i - 1) \cdot \phi\right) \)
    for \( i = 1, 2, \ldots \)
  - dimensionality of signal-space = \( D = 2 \)

<table>
<thead>
<tr>
<th>( M = 4 ) &amp; ( \phi = 0^\circ )</th>
<th>( M = 4 ) &amp; ( \phi = 45^\circ )</th>
<th>( M = 8 ) &amp; ( \phi = 0^\circ )</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
<td><img src="image3" alt="Diagram" /></td>
</tr>
</tbody>
</table>
- M-ary FSK:
  very difficult to be represented using constellation diagram

\[ f_L - f_D = \frac{n}{2T_{c,s}} \]
\[ f_L + f_D = \frac{m}{2T_{c,s}} \]
\[ n, m = \text{integers} \]

Examples of Plots of Decision Variables (QPSK-Receiver)

"good"

"bad"
Optimum M-ary Receiver using the Signal-Vectors

- Let us consider again the general M-ary problem:

\[
\begin{align*}
    r(t) &= s_i(t) + n(t) \\
    \text{Decision rule} &= ? \\
    (\text{To determine which} & \quad D_i \\
    \text{signal is present})
\end{align*}
\]

where \( s_i(t) = \) one of \( M \) signals (channel symbols)

- Hypotheses:

\[
\begin{align*}
    H_1 : & \quad s_1(t) \text{ is present with probability } \Pr(H_1) \\
    H_2 : & \quad s_2(t) \text{ is present with probability } \Pr(H_2) \\
    \quad \vdots \\
    H_M : & \quad s_M(t) \text{ is present with probability } \Pr(H_M)
\end{align*}
\]

we have seen that the MAP criterion, using the "CONTINUOUS SAMPLING" concept, is described by the following rule:

Rule: choose the hypothesis \( H_i \) with the maximum

\[
\Pr(H_i) \times \text{pdf}_{r/H_i}(r(t))
\]

where

\[
\begin{align*}
    \max \left\{ \Pr(H_i) \times \text{pdf}_{r/H_i}(r(t)) \right\} \\
    = \max \left\{ \int_0^{T_{cs}} r(t)s_i^*(t)dt + \frac{N_0}{2} \ln(\Pr(H_i)) - \frac{1}{2}E_i \right\} \quad (70)
\end{align*}
\]

- Remember that this rule can also be described by the following two equivalent receivers:
where

\[ DC_i = \frac{N_0}{2} \ln(Pr(H_i)) - \frac{1}{2} E_i \]  

(71)

Now let us use the Approx. Theorem of energy signals to represent the received signals \( r(t) \) as well as the \( M \)-ary signal \( s_i(t) \), \( \forall i \),

\[ \begin{align*}
    r(t) &= w_r^H c(t) \\
s_i(t) &= w_{s_i}^H c(t) \\
    \Rightarrow s_i^*(t) &= (w_{s_i}^H c(t))^* \\
                    &= c(t)^H w_{s_i} \quad \forall i
\end{align*} \]

In this case the MAP receiver (Equ-70) becomes

\[
\max \left\{ \Pr(H_i) \times \text{pdf}_{r/H_i}(r(t)) \right\} \\
= \max \left\{ \int_0^{T_{cs}} r(t)s_i^*(t)dt + \frac{N_0}{2} \ln(Pr(H_i)) - \frac{1}{2} E_i \right\} \\
= \max \left\{ \int_0^{T_{cs}} w_r^H c(t) \cdot c(t)^H w_{s_i} dt + \frac{N_0}{2} \ln(Pr(H_i)) - \frac{1}{2} E_i \right\} \\
= \max \left\{ w_r^H \left( \int_0^{T_{cs}} c(t) \cdot c(t)^H dt \right) w_{s_i} + \frac{N_0}{2} \ln(Pr(H_i)) - \frac{1}{2} E_i \right\} \\
= \max \left\{ w_r^H w_{s_i} + \frac{N_0}{2} \ln(Pr(H_i)) - \frac{1}{2} E_i \right\}
\]
The above equation may be implemented as follows:

or

### Encoding M-ary Signals

- The performance of $M$-ary systems is evaluated by means of the average probability of symbol error $p_{e,cs}$, which, for $M > 2$, is different than the average probability of bit error (or Bit Error Rate BER), $p_e$.

That is

$$\begin{cases} 
  p_{e,cs} \neq p_e & \text{for } M > 2 \\
  p_{e,cs} = p_e & \text{for } M = 2 
\end{cases} \quad (73)$$

However, because we transmit binary data, the probability of bit error $p_e$ is a more natural parameter for performance evaluation than $p_{e,cs}$.

- Although, these two probabilities are related, i.e. $p_e = f(p_{e,cs})$ their relationship depends on the encoding approach which is employed by the digital modulator for mapping binary digits to $M$-ary signals (channel symbols)

<table>
<thead>
<tr>
<th>Digital Modulator</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$-th word of $\gamma_{cs}$ bits</td>
</tr>
<tr>
<td>$1$st bit</td>
</tr>
</tbody>
</table>

where $\gamma_{cs} = \log_2 M$. 
Important Relationships Between BER and SER

- If the encoder provides a mapping where adjacent symbols differ in one binary digit then it can be proved (see Haykin p500) that
  \[
  \frac{1}{\gamma_{cs}} \cdot p_{e,cs} \leq p_e \leq p_{e,cs}
  \]  
  \[ (74) \]
  
  e.g. \textit{M}-ary PSK (for Gray code, otherwise difficult):
  \[
  BER = p_e = \frac{1}{\gamma_{cs}} \cdot p_{e,cs}
  \]  
  \[ (75) \]
  
  ▶ N.B.: Note that Gray encoder provides a mapping where adjacent symbols differ in one binary digit. This property is very important because the most likely errors (due to channel noise effects) are related with an erroneous selection (wrong decision) of an adjacent symbol.

- If all \textit{M}-ary signals are equally likely, then it can be proved (see Haykin p.500) that
  \[
  p_e = \frac{2^{\gamma_{cs}}}{2^{\gamma_{cs}} - 1} \cdot p_{e,cs}
  \]  
  \[ (76) \]
  
  which, for large $\gamma_{cs}$ is simplified to
  \[
  p_e = \frac{1}{2} \cdot p_{e,cs}
  \]  
  \[ (77) \]
  
  Example:

  \textit{M}-ary FSK:
  \[
  BER = p_e = \frac{2^{\gamma_{cs}}}{2^{\gamma_{cs}} - 1} \cdot p_{e,cs}
  \]  
  \[ (78) \]
Appendix-1: Proof of Equation-26

The above rule can be rewritten as follows:

\[
\max \left\{ \Pr(H_i) \times \text{pdf}_{r/H_i}(r(t)) \right\} \\
= \max \left\{ \ln(\Pr(H_i)) + \ln(\text{const.}) - \frac{1}{N_0} \int_0^{T_{cs}} (r(t) - s_i(t))^2 dt \right\} \\
= \max \left\{ \ln(\Pr(H_i)) - \frac{1}{N_0} \int_0^{T_{cs}} r(t) s_i(t)^2 dt \right\} \\
= \max \left\{ \ln(\Pr(H_i)) - \frac{1}{N_0} \frac{r^2(t)}{dt} - \frac{1}{N_0} \int_0^{T_{cs}} s_i^2(t) dt + \frac{2}{N_0} \int_0^{T_{cs}} r(t) s_i^*(t) dt \right\} \\
= \max \left\{ \ln(\Pr(H_i)) - \frac{1}{N_0} E_r - \frac{1}{N_0} E_i + \frac{2}{N_0} \int_0^{T_{cs}} r(t) s_i^*(t) dt \right\} \\
= \max \left\{ \frac{N_0}{2} \ln(\Pr(H_i)) - \frac{1}{2} E_i + \int_0^{T_{cs}} r(t) s_i^*(t) dt \right\} \\
= \max \left\{ \int_0^{T_{cs}} r(t) s_i^*(t) dt + \frac{N_0}{2} \ln(\Pr(H_i)) - \frac{1}{2} E_i \right\}
\]

Appendix-2: Walsh-Hadamard Orthogonal Sets and Signals

- \( \mathbf{H}_1 = [1] \)
- \( \mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \) (2 \( \times \) 2) matrix
- \( \mathbf{H}_{64} = \mathbf{H}_2 \otimes \mathbf{H}_2 \otimes \mathbf{H}_2 \otimes \mathbf{H}_2 \otimes \mathbf{H}_2 \otimes \mathbf{H}_2 \) \( 6 \)-times

where \( \otimes \) denotes the Kronecker product of two matrices
e.g. for two matrices \( \mathbf{A} \) and \( \mathbf{B} \)

if \( \mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \) then \( \mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} A_{11} \mathbf{B} & A_{12} \mathbf{B} \\ A_{21} \mathbf{B} & A_{22} \mathbf{B} \end{bmatrix} \)

- Properties:
  - \( \mathbf{H}_{64}^T \mathbf{H}_{64} = 64 \mathbf{I}_{64} \)
  - Let \( \mathbf{h}_i \) denote the \( i \)-th column of \( \mathbf{H}_{64} \)
i.e. \( \mathbf{H}_{64} = \begin{bmatrix} \mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{64} \end{bmatrix} \)
The following example illustrates an application of “Walsh Hadamard” in a mobile communications where (each user in the same cell and same frequency channel is assigned one column of Walsh matrix.

Transmitter

\[ ... m_1, m_2, m_3, ..., m_k, ... \]

2nd column

\[ h_v \]

\[ \rightarrow \]

\[ ... m_1 h_v^T, m_2 h_v^T, m_3 h_v^T, ..., m_k h_v^T, ... \]

output binary sequence

Receiver-A

Walsh

2nd column

\[ h_v/64 \]

\[ \rightarrow \]

\[ ... m_1, m_2, m_3, ..., m_k, ... \]

output binary sequence

Receiver-B

Walsh

3rd column

\[ h_v/64 \]

\[ \rightarrow \]

\[ ... 0, 0, 0, ..., 0, ... \]

output binary sequence

Appendix-3: IS95 - Walsh Function of order-64

Walsh functions of order-64 as defined in IS-95 Section 5.2.1. The Walsh functions, \( m_1, \ldots, m_k \), all have \( 0 \) parity, \( \ell = 1 \), and are normalized to \( h_v' = h_v/\sqrt{64} \).

\[ m_{1} \quad m_{2} \quad m_{3} \quad m_{4} \quad m_{5} \quad m_{6} \quad m_{7} \quad m_{8} \]

\[ m_{9} \quad m_{10} \quad m_{11} \quad m_{12} \quad m_{13} \quad m_{14} \quad m_{15} \quad m_{16} \]

\[ m_{17} \quad m_{18} \quad m_{19} \quad m_{20} \quad m_{21} \quad m_{22} \quad m_{23} \quad m_{24} \]

\[ m_{25} \quad m_{26} \quad m_{27} \quad m_{28} \quad m_{29} \quad m_{30} \quad m_{31} \quad m_{32} \]

\[ m_{33} \quad m_{34} \quad m_{35} \quad m_{36} \quad m_{37} \quad m_{38} \quad m_{39} \quad m_{40} \]

\[ m_{41} \quad m_{42} \quad m_{43} \quad m_{44} \quad m_{45} \quad m_{46} \quad m_{47} \quad m_{48} \]

\[ m_{49} \quad m_{50} \quad m_{51} \quad m_{52} \quad m_{53} \quad m_{54} \quad m_{55} \quad m_{56} \]

\[ m_{57} \quad m_{58} \quad m_{59} \quad m_{60} \quad m_{61} \quad m_{62} \quad m_{63} \quad m_{64} \]