

Update/Key Lecture Materials

- **Have Considered**

- Geometrical representation of signals.
- Signal space, signals as vectors.
- Signal space dimension, independent vectors.
- Basis Vectors/functions, Orthogonal/orthonormal signals.
- Systematic determination of an orthogonal basis set (Gram-Schmidt Orthogonalization or GSO process).

- » **We will now consider**

- Optimum signal detection.
- We will determine the optimum receiver (in the sense of minimizing P_e) for general M-ary signaling in the presence of AWGN.



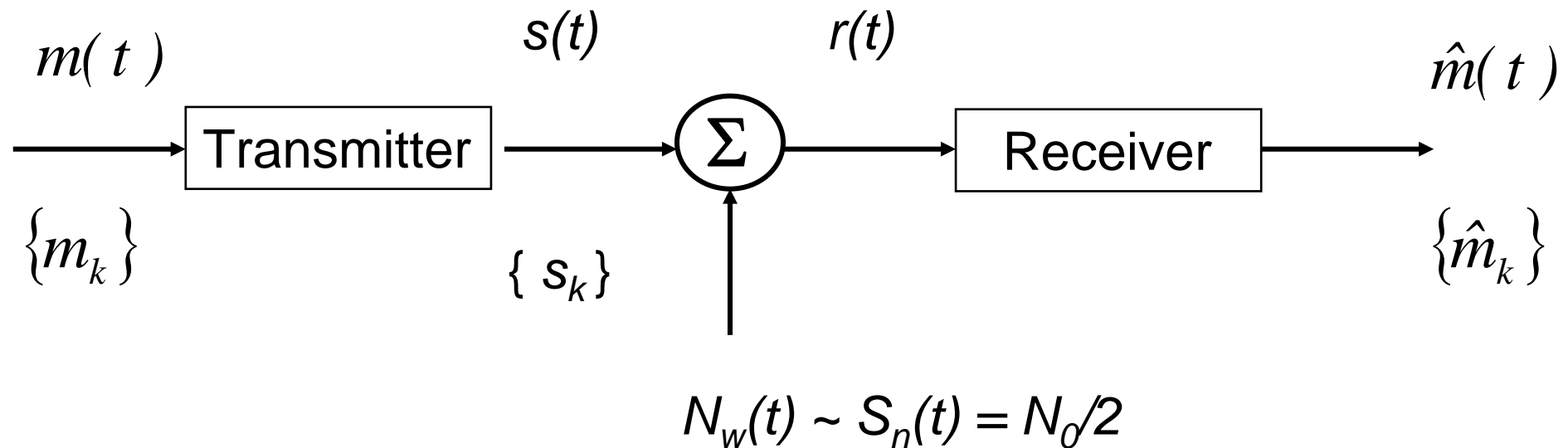
Optimal Signal Detection

Key Assumptions

- Transmitter transmits a sequence of symbols or messages from a set of M symbols m_1, m_2, \dots, m_M .
- The symbols are represented by finite energy waveforms $S_1(t), S_2(t), \dots, S_M(t)$.
- The noise is zero mean AWGN.
- The signal space is assumed to be of dimension N .
- $\phi_k(t)$ for $k=1, \dots, N$ will denote an orthonormal basis.

Optimum Receivers for AWGN Channels

Consider the following communication model:



Signal Detection

$$s_k(t) = \sum_{j=1}^N s_{kj} \varphi_j(t) \quad \text{where} \quad s_{kj} = \int_{-\infty}^{\infty} s_k(t) \varphi_j(t) dt$$

Next note that $n_w(t)$ can be written as: $n_w(t) = n(t) + n_0(t)$

where

$n(t)$ = Projection of $n_w(t)$ on N -dim space

$n_0(t) = n_w(t) - n(t) \quad \longrightarrow \quad N_0(t) \perp \text{Space}$

Signal Detection

Thus $n(t) = \sum_{j=1}^N n_j \varphi_j(t)$ with $n_j = \int_{-\infty}^{\infty} n(t) \varphi_j(t) dt$

Next, we have

$$\begin{aligned} r(t) &= s_k(t) + n_w(t) \\ &= s_k(t) + n(t) + n_o(t) \\ &= Z(t) + n_o(t) \end{aligned}$$

$Z(t)$ = Projection of $r(t)$ on N -dim signal space.

Signal Detection

$$Z(t) = \sum_{j=1}^N (s_{kj} + n_j) \varphi_j(t) \quad \text{Or} \quad \underline{Z} = (Z_1, Z_2, \dots, Z_N)$$

$$= (\underline{S} + \underline{n})$$

$$= (s_{k1} + n_1, s_{k2} + n_2, \dots, s_{kN} + n_N)$$

Next, note that

$$n_k = \int_{-\infty}^{\infty} n_w(t) \varphi_k(t) dt$$



$$\overline{n_j n_k} = E \left\{ \int_{-\infty}^{\infty} n_w(\alpha) \varphi_j(\alpha) d\alpha \int_{-\infty}^{\infty} n_w(\beta) \varphi_k(\beta) d\beta \right\}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \overline{n_w(\alpha) n_w(\beta)} \varphi_j(\alpha) \varphi_k(\beta) d\alpha d\beta$$

Signal Detection

$$\begin{aligned}\overline{n_j n_k} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{n_w}(\alpha - \beta) \varphi_j(\alpha) \varphi_k(\beta) d\alpha d\beta \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{N_0}{2} \delta(\alpha - \beta) \varphi_j(\alpha) \varphi_k(\beta) d\alpha d\beta \\ &= \frac{N_0}{2} \int_{-\infty}^{\infty} \varphi_j(\alpha) \varphi_k(\alpha) d\alpha \\ &= \begin{cases} \frac{N_0}{2} & j = k \\ 0 & j \neq k \end{cases}\end{aligned}$$

Signal Detection

Note that n_j and n_k are **uncorrelated Gaussian** random variables.

➡ They are **independent** with **zero mean** and **variance $N_0/2$** .

$$f_n(\underline{n}) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{n_j^2}{2\sigma^2}}$$

$$= \prod_{j=1}^N \frac{1}{\sqrt{2\pi(\frac{N_0}{2})}} e^{-\frac{n_j^2}{2(\frac{N_0}{2})}}$$

$$= \frac{1}{(\pi N_0)^{N/2}} e^{-\left\{ \frac{n_1^2 + n_2^2 + \dots + n_N^2}{N_0} \right\}} = (\pi N_0)^{-\frac{N}{2}} e^{-\left\{ \frac{\|\underline{n}\|^2}{N_0} \right\}}$$

M-ary Signaling Optimal Detection

We wish to design a signal detector such that the probability of making an error is minimized (or correct decision is maximized) given the received signal z

Introduce the *a posteriori* probabilities defined as

$$P(m_k | z) \quad \text{P(signal } m_k \text{ was transmitted given } z \text{ received)}$$

Known as *a posteriori* since the decision is made after (or given) the observation z

Different from *a priori* where some information about the decision is known in advance of the observation

M-ary Signaling Optimal Detection

A good decision criteria is to decide in favor of the signal corresponding to the maximum of the *a posteriori* probabilities $P(m_k | z)$ given z is received


This decision criterion is called the **maximum *a posteriori* probability (MAP)** criterion

The optimum receiver would then be

$$\begin{aligned} & \text{Select } \hat{m} = m_k \text{ if} \\ & P(m_k / \underline{Z} = z) > P(m_j / \underline{Z} = z) \quad \forall j \neq k \end{aligned}$$

M-ary Signaling Optimal Detection

Now by Baye's Rule

$$P(m_k / \underline{Z} = z) = \frac{f(z / m_k) P(m_k)}{f(z)}$$


Not relevant in maximization since $f(z)$ is same for all m_k

Requires knowledge of **a priori** probabilities and conditional pdf's

M-ary Signaling Optimal Detection

Optimum receiver sets $\hat{m} = m_k \iff$

$$f(z / m_k)P(m_k) > f(z / m_i)P(m_i) \quad \forall i \neq k$$

Practical implementation

But, $f(\underline{z} / m_k) \equiv f_n(\underline{z} - \underline{s}_k)$


$$\begin{aligned} &= \prod_{i=1}^N \frac{\exp[-(z_i - s_{ki})^2 / N_0]}{\sqrt{\pi N_0}} = \frac{\exp[-\sum_{i=1}^N (z_i - s_{ki})^2 / N_0]}{(\pi N_0)^{N/2}} \\ &= \frac{\exp[-\|\underline{z} - \underline{s}_k\|^2 / N_0]}{(\pi N_0)^{N/2}} \end{aligned}$$

M-ary Signaling Optimal Detection

Recall that

$$\underline{z} \sim Z(t) = \sum_{i=1}^N z_i \varphi_i(t)$$
$$\underline{s}_k \sim s_k(t) = \sum_{i=1}^N s_{ki} \varphi_i(t)$$

By dropping $(\pi N_0)^{N/2}$ and taking $\log[f(\underline{z}/\underline{s}_k)P(m_k)]$, we conclude that

| |
|--|
| Maximizing $f(\underline{z}/\underline{s}_k)P(m_k)$ |
|  |
| Minimizing $\ \underline{z} - \underline{s}_k\ ^2 - N_0 \log[P(m_k)]$ |

M-ary Signaling Optimal Detection

i.e. Optimum receiver sets $\hat{m} = m_k$



If for all $i \neq k$

$$\sum_{j=1}^N (z_j - s_{kj})^2 - N_0 \log[P(m_k)] < \sum_{j=1}^N (z_j - s_{ij})^2 - N_0 \log[P(m_i)]$$

M-ary Signaling Optimal Detection

For **equally likely signals**, $P(m_k)=P(m_i)$, \forall i and k ,

Therefore we can just consider the quantities

$$f(z / m_k)$$

The conditional pdf $f(z/m_k)$ or any function of it is known as the likelihood function.

The decision criteria based on the maximum of $f(z/m_k)$ over the M signals is called the **maximum-likelihood (ML) criterion**

M-ary Signaling Optimal Detection

➡ Decision Rule becomes

$$\text{Minimize} \quad \left\| \underline{z} - \underline{s}_k \right\|^2 = \sum_{i=1}^N (z_i - s_{ki})^2$$

i.e. **Optimum receiver** chooses $\hat{m} = m_k$ iff Received vector \underline{z} is closer to \underline{s}_k in terms of **Euclidean distance** than to any other s_i for $i \neq k$.


M-ary Signaling Optimal Detection

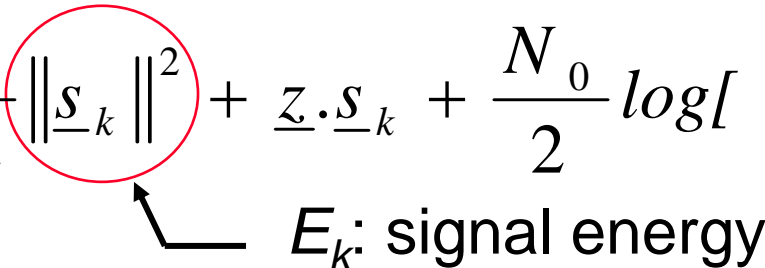
From this expression we can develop a receiver structure using following derivation:

Maximize

$$- \|\underline{z} - \underline{s}_k\|^2 + N_0 \log[P(m_k)]$$

$$= 2 \frac{ \left\{ -\|\underline{z}\|^2 - \|\underline{s}_k\|^2 + 2 \underline{z} \cdot \underline{s}_k + N_0 \log[P(m_k)] \right\} }{2}$$


$$- \frac{1}{2} \|\underline{z}\|^2 - \frac{1}{2} \|\underline{s}_k\|^2 + \underline{z} \cdot \underline{s}_k + \frac{N_0}{2} \log[P(m_k)]$$

 E_k : signal energy

M-ary Signaling Optimal Detection

Let $\alpha_k = \frac{1}{2} [N_0 \log[P(m_k)] - E_k]$

Decision function becomes

$$\alpha_k + \underline{z} \cdot \underline{s}_k - \frac{1}{2} \|\underline{z}\|^2$$

Common to all M decision

\therefore can be omitted

\therefore New decision function:

$$b_k = a_k + \underline{z} \cdot \underline{s}_k$$

M-ary Signaling Optimal Detection

Now, if $z(t)$ is applied at the input of a linear system characterized by $h(t)$.

Then at $t = T$, the output is
$$\int_{-\infty}^{\infty} Z(\lambda) h(T - \lambda) d\lambda$$

Let $h(t) = s_k(T - \lambda)$

i.e., $h(t)$ is matched to $s_k(t)$. Then, $h(T - \lambda) = s_k(\lambda)$

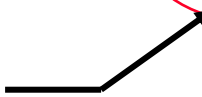
$$\text{blue arrow } output = \int_{-\infty}^{\infty} z(\lambda) s_k(\lambda) d\lambda = \underline{z} \cdot \underline{s}_k$$

i.e. $\underline{z} \cdot \underline{s}_k$ is the output at $t = T$ of a filter matched to $s_k(t)$ when $Z(t)$ is the input.

M-ary Signaling Optimal Detection

Now, recall that

$$\begin{aligned} r(t) &= s_k(t) + n_w(t) \\ &= s_k(t) + n(t) + n_0(t) \\ &= Z(t) + n_0(t) \end{aligned}$$

Not relevant 

Because

$$\int_{-\infty}^{\infty} n_0(t) s_k(t) dt = 0 \quad (n_0(t) \perp \text{signal space})$$

M-ary Signaling Optimal Detection

So

$$\begin{aligned}\underline{z} \cdot \underline{s}_k &= \int_{-\infty}^{\infty} z(t) s_k(t) dt + \int_{-\infty}^{\infty} n_0(t) s_k(t) dt \\ &= \int_{-\infty}^{\infty} [z(t) + n_0(t)] s_k(t) dt \\ &= \int_{-\infty}^{\infty} r(t) s_k(t) dt\end{aligned}$$

≡ 0 (Orthogonal)



Thus $r(t)$ can be used in place of $Z(t)$ at the input.

Optimum M -ary Receiver

Apply the incoming signal $r(t)$ to a parallel bank of matched filters (or correlation receivers) and sample the output of filters at $t = T_0$.

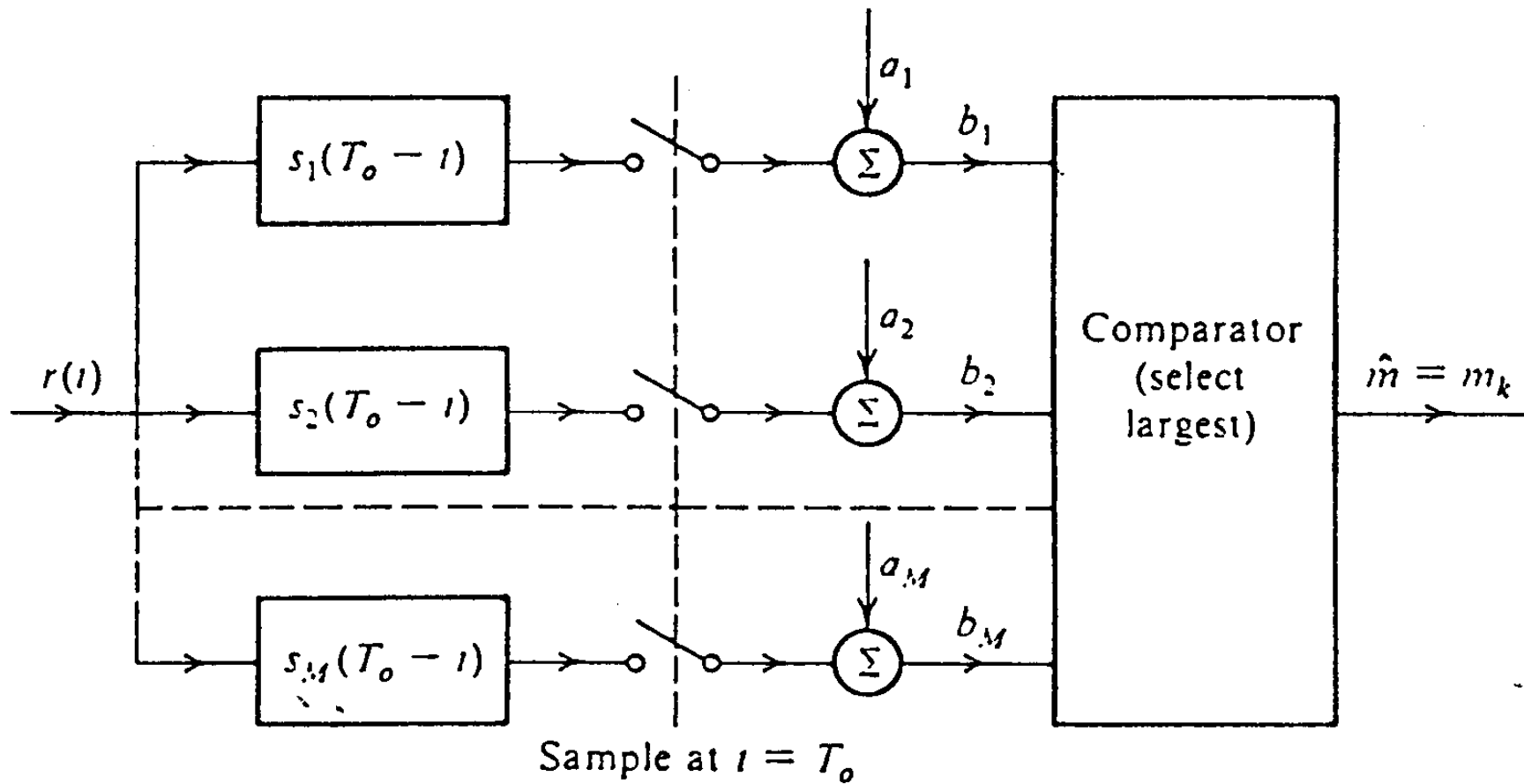
To this a constant a_k is added to the k -th filter output sample.

The resulting outputs are compared and the decision is made in favor of the signal for which the output is the largest.

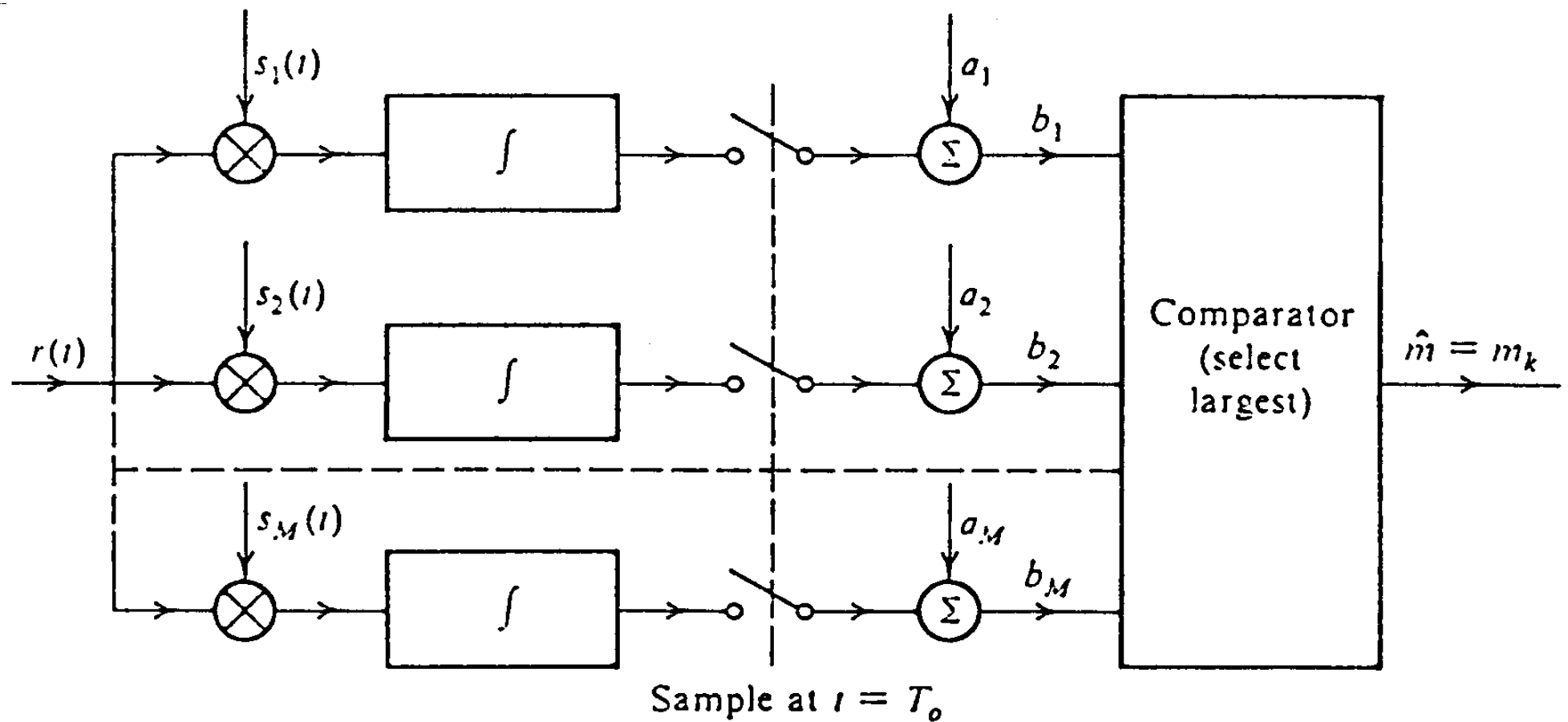
In brief, have shown that matched filters form the Optimum receiver that minimize the symbol error probability P_e (assuming AWGN).

Also note that the system is linear although it was not constrained to be so.

Optimum M-ary Receiver: Matched-filter detector



Optimum M-ary Receiver: Correlation detector



Example

In an additive white Gaussian noise channel with a noise power-spectral density of $N_0/2$, two equiprobable messages are transmitted by

$$s_1(t) = \begin{cases} \frac{At}{T} & 0 \leq t \leq T \\ 0 & \text{otherwise} \end{cases}$$

$$s_2(t) = \begin{cases} A - \frac{At}{T} & 0 \leq t \leq T \\ 0 & \text{otherwise} \end{cases}$$

1. Determine the structure of the optimal receiver.

Optimum M-ary Receiver

To develop an alternative receiver structure, recall that

$$\underline{z} \cdot \underline{s}_k = \sum_{j=1}^N r_j s_{kj} \quad \text{with} \quad r_j = \int_{-\infty}^{\infty} r(t) \varphi_j(t) dt = z_j$$

where s_{kj} are known and satisfy

$$s_k(t) = \sum_{j=1}^N s_{kj} \varphi_j(t)$$
$$s_{kj} = \int_{-\infty}^{\infty} s_k(t) \varphi_j(t) dt$$

Optimum M-ary Receiver

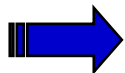


M correlation detectors (or matched filters) can be replaced by N filters matched to $\varphi_1(t)$, $\varphi_2(t)$, ..., $\varphi_N(t)$,

Clearly, both receivers perform identically.
Choice depends on circumstances.

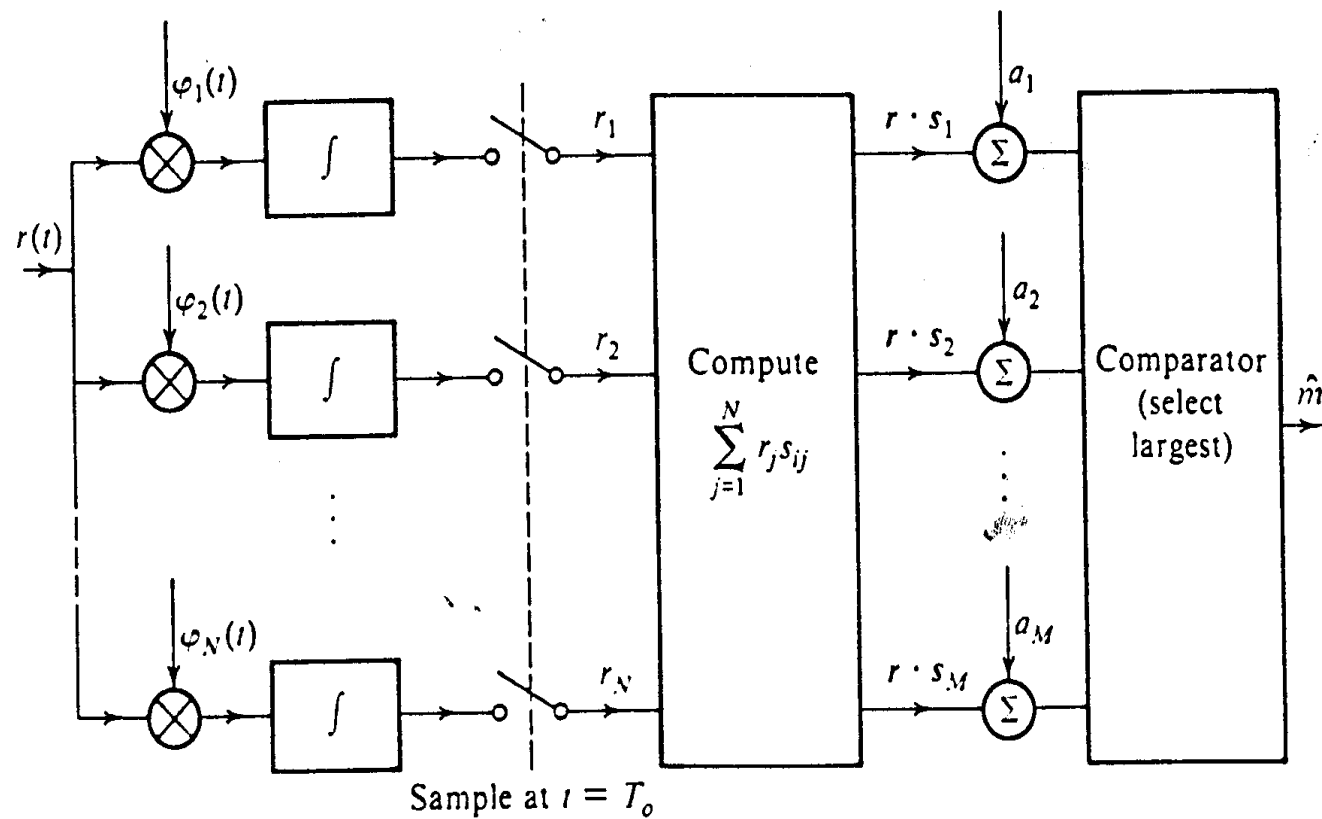
Example:

If $N \ll M$, and $\{\varphi_j(t)\}$ are easier to generate than $\{s_j(t)\}$

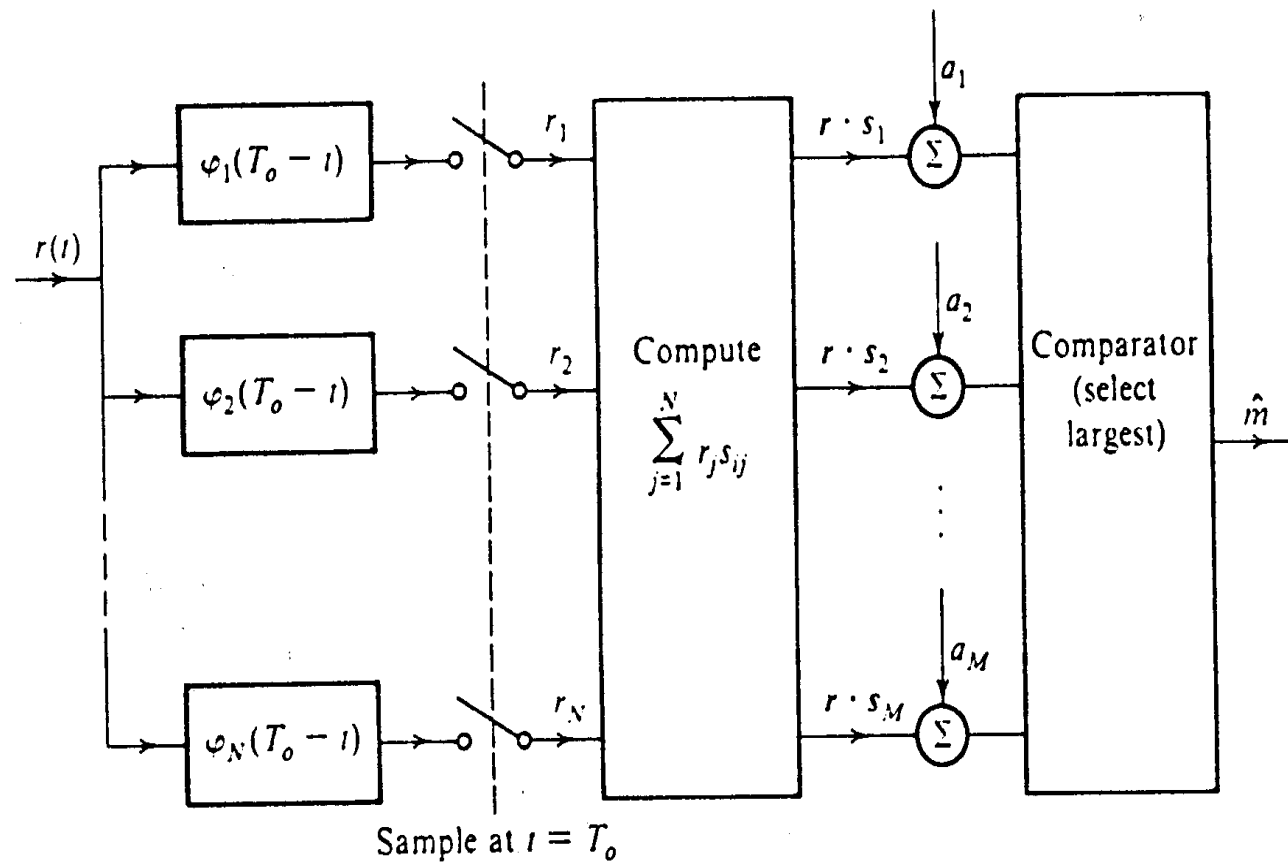


Choice is obvious.

Another form of Optimum M-ary Receiver: Correlation detector



Another form of Optimum M-ary Receiver: Matched-filter detector



Graphical Interpretation - Decision Regions

Recall that optimum receiver sets $\hat{m} = m_k$ iff

$P(m_k | z)$ or $\|z - \underline{s}_k\|^2 - N_0 \log[P(m_k)]$ is minimized.

An alternative view of the receiver operation is:

Divide the Signal space into M disjoint decision regions R_1, R_2, \dots, R_M such that

|| If $z \in R_k$  Decide m_k was transmitted

The optimal receiver thus sets the decision regions so that P_e is minimized.

Decision Regions – Equiprobable Signals

For simplicity, if one assumes $P(m_k) = 1/M \forall k$.

Optimum receiver sets $\hat{m} = m_k$ if $\|\underline{z} - \underline{s}_k\|^2$ is minimized.

Geometrically, this means

Take projection of $r(t)$ in signal space (i.e., \underline{z}). Then, decision is made in favor of signal that is the closest to \underline{z} .

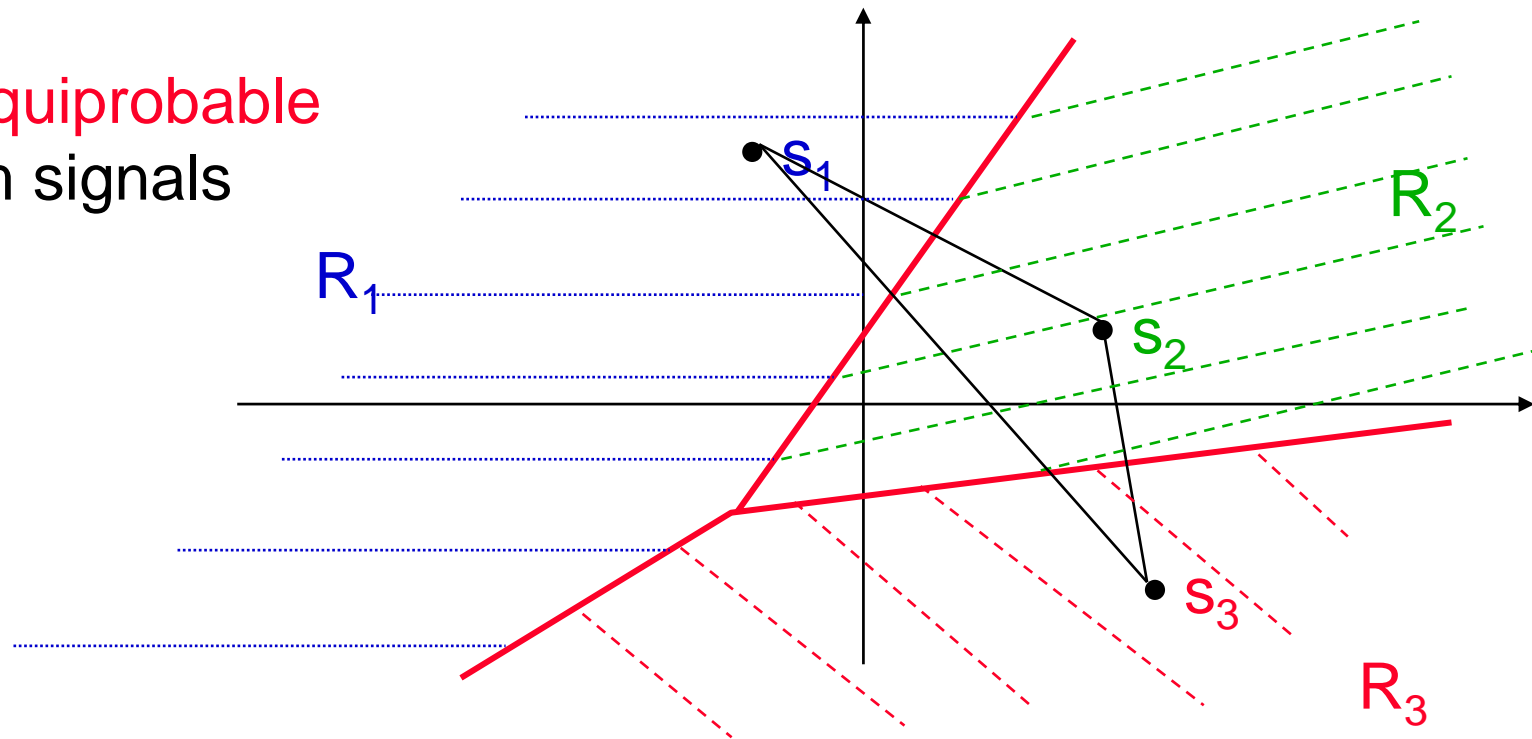
Equivalently,

The decision region R_k of \underline{s}_k consists of all points in the signal space which are closest to \underline{s}_k in Euclidean distance.

Determining the optimum decision regions

In general, boundaries of decision regions are **perpendicular bisectors** of the lines joining the original transmitted signals. In particular, the **pairwise** decision regions of two signal points are the **half-spaces** containing the signal points divided by the bisector.

Three **equiprobable**
2-dim signals



Example

Three equally probable messages m_1 , m_2 , and m_3 are to be transmitted over an AWGN channel with noise power-spectral density $N_0 / 2$. The messages are

$$s_1(t) = \begin{cases} 1 & 0 \leq t \leq T \\ 0 & \text{otherwise} \end{cases}$$
$$s_2(t) = -s_3(t) = \begin{cases} 1 & 0 \leq t \leq \frac{T}{2} \\ -1 & \frac{T}{2} \leq t \leq T \\ 0 & \text{otherwise} \end{cases}$$

1. What is the dimensionality of the signal space ?
2. Find an appropriate basis for the signal space (Hint: You can find the basis without using the Gram-Schmidt procedure).
3. Draw the signal constellation for this problem.
4. Sketch the optimal decision regions R_1 , R_2 , and R_3 .

Decision Regions – Non-equibable Signals

Example:

Consider binary data transmission over AWGN channel with $S_n(f)=N_0/2$ with

$$s_1(t) = -s_2(t) = \sqrt{E}\varphi(t)$$

Assume $P(m_1) \neq P(m_2)$.

Determine optimum receiver and P_e .

Decision Regions

Decision is m_1 if

$$\left\| \underline{z} - \underline{s}_1 \right\|^2 - N_0 \log[P(m_1)] < \left\| \underline{z} - \underline{s}_2 \right\|^2 - N_0 \log[P(m_2)]$$

Otherwise, decision is m_2

Let

$$\left. \begin{aligned} d_1 &= \left\| \underline{z} - \underline{s}_1 \right\| \\ d_2 &= \left\| \underline{z} - \underline{s}_2 \right\| \end{aligned} \right\} \Rightarrow$$

Decision Regions

Decision is m_1 if

$$d_1^2 - d_2^2 < N_0 \underbrace{\log \left[\frac{P(m_1)}{P(m_2)} \right]}_c$$

Decide m_1 if $d_1^2 - d_2^2 < c$ (R_1)

else decide m_2 (R_2)

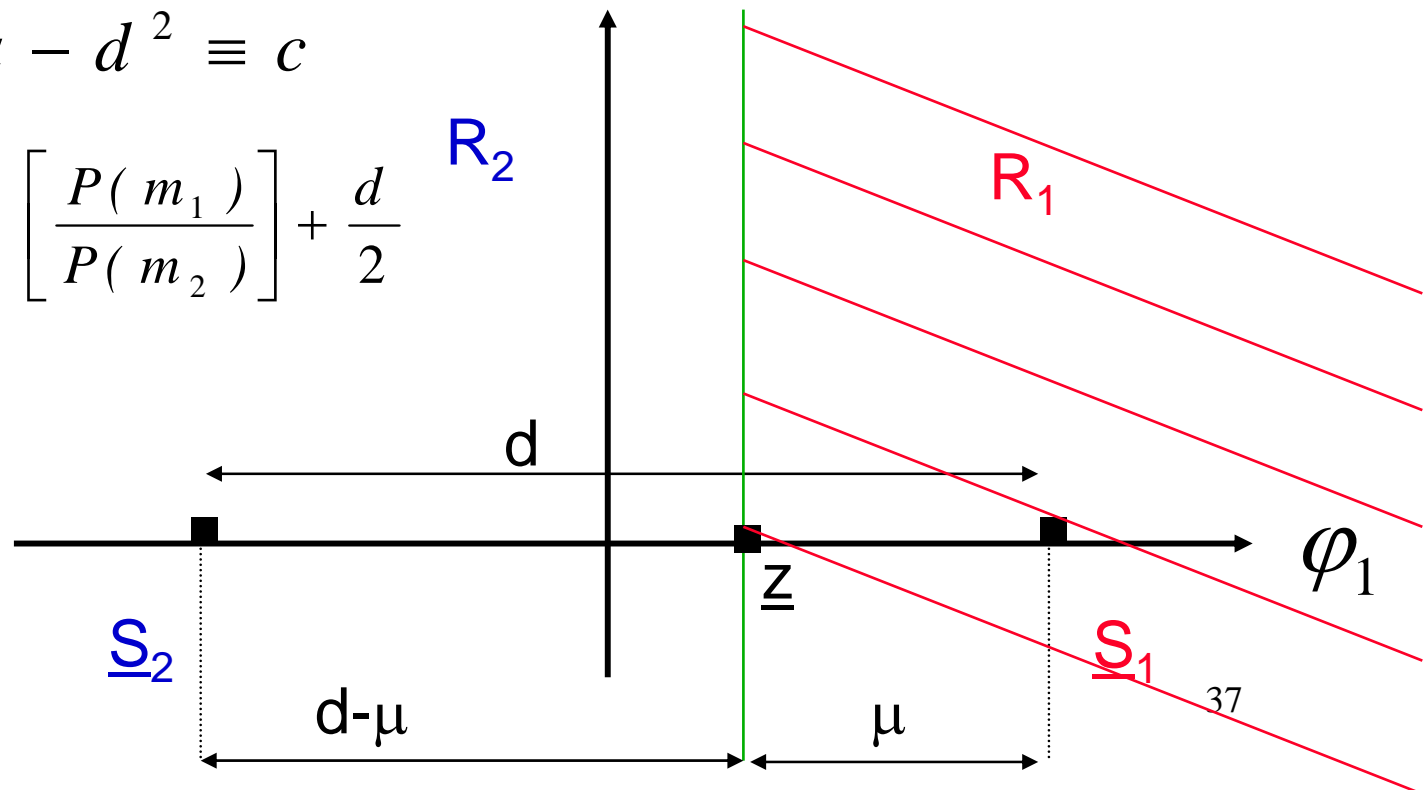
Decision Regions

Now consider our example with z on decision boundary

$$d = d_1 + d_2 \qquad d_1^2 = \mu^2 \qquad d_2^2 = (d - \mu)^2$$

$$d_1^2 - d_2^2 = 2d\mu - d^2 \equiv c$$

$$\mu = \frac{c + d^2}{2d} = \frac{N_0}{2d} \log \left[\frac{P(m_1)}{P(m_2)} \right] + \frac{d}{2} \quad \text{R}_2$$

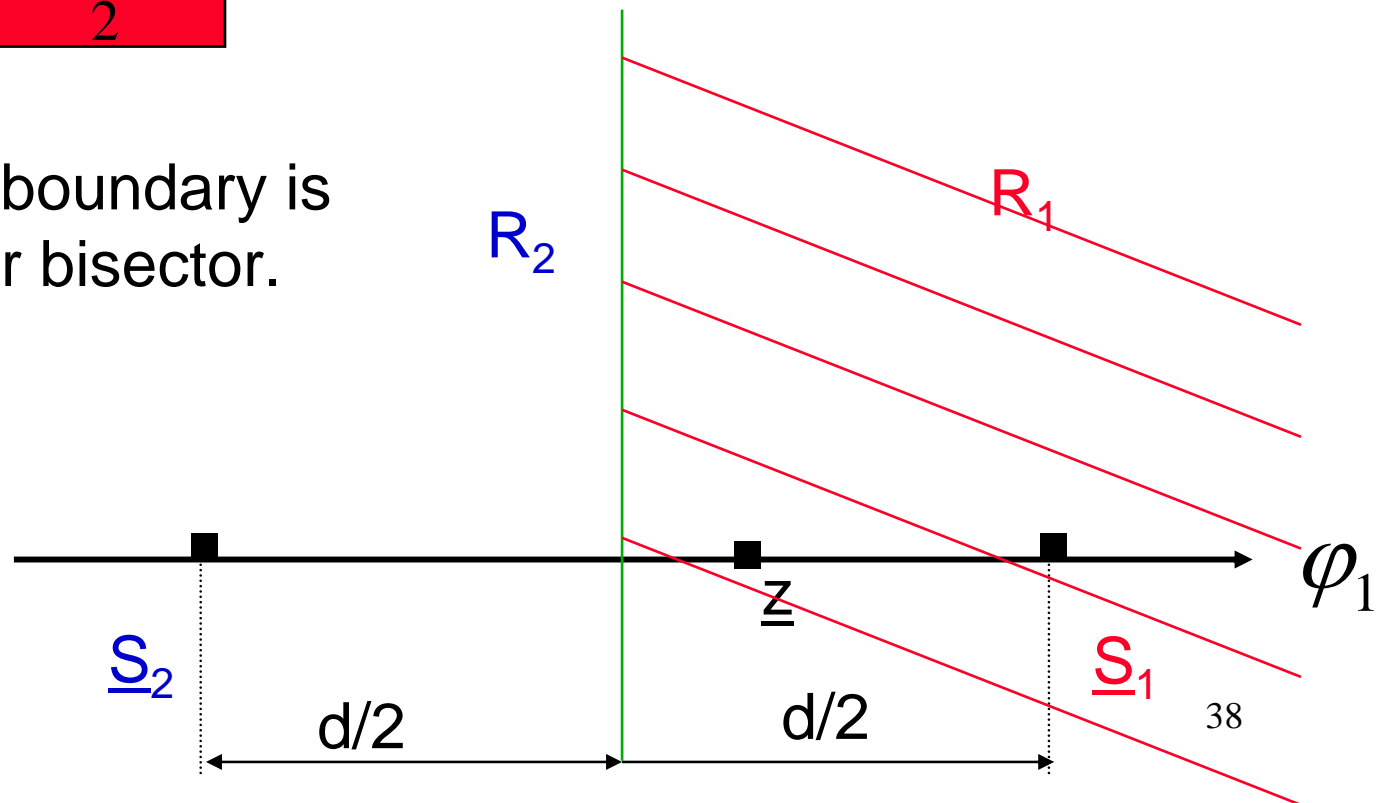


Decision Regions

For $P(m_1)=P(m_2)$

$$\mu = \sqrt{E} \equiv \frac{d}{2}$$

Hence, decision boundary is the perpendicular bisector.



Decision Regions

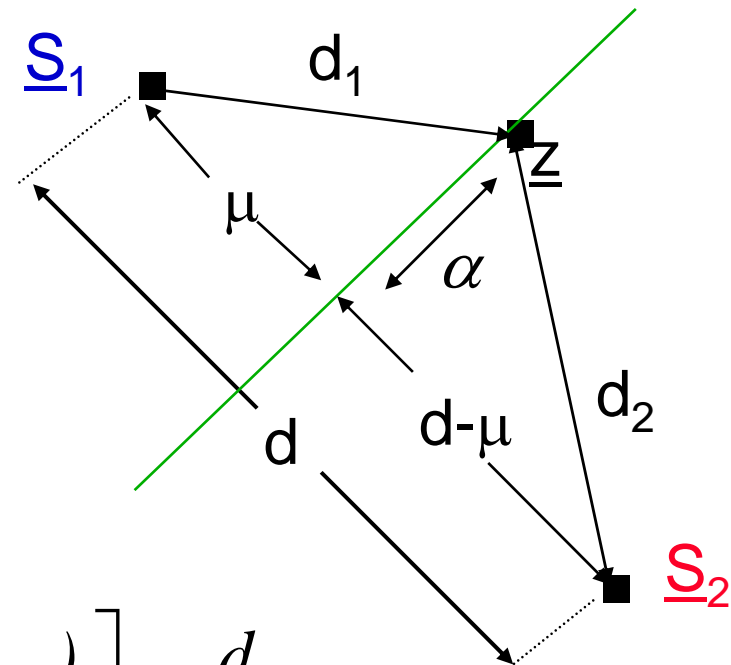
For 2-dimensional situation

$$d_1^2 = \alpha^2 + \mu^2$$

$$d_2^2 = \alpha^2 + (d - \mu)^2$$

$$\Rightarrow d_1^2 - d_2^2 = 2d\mu - d^2 \equiv c$$

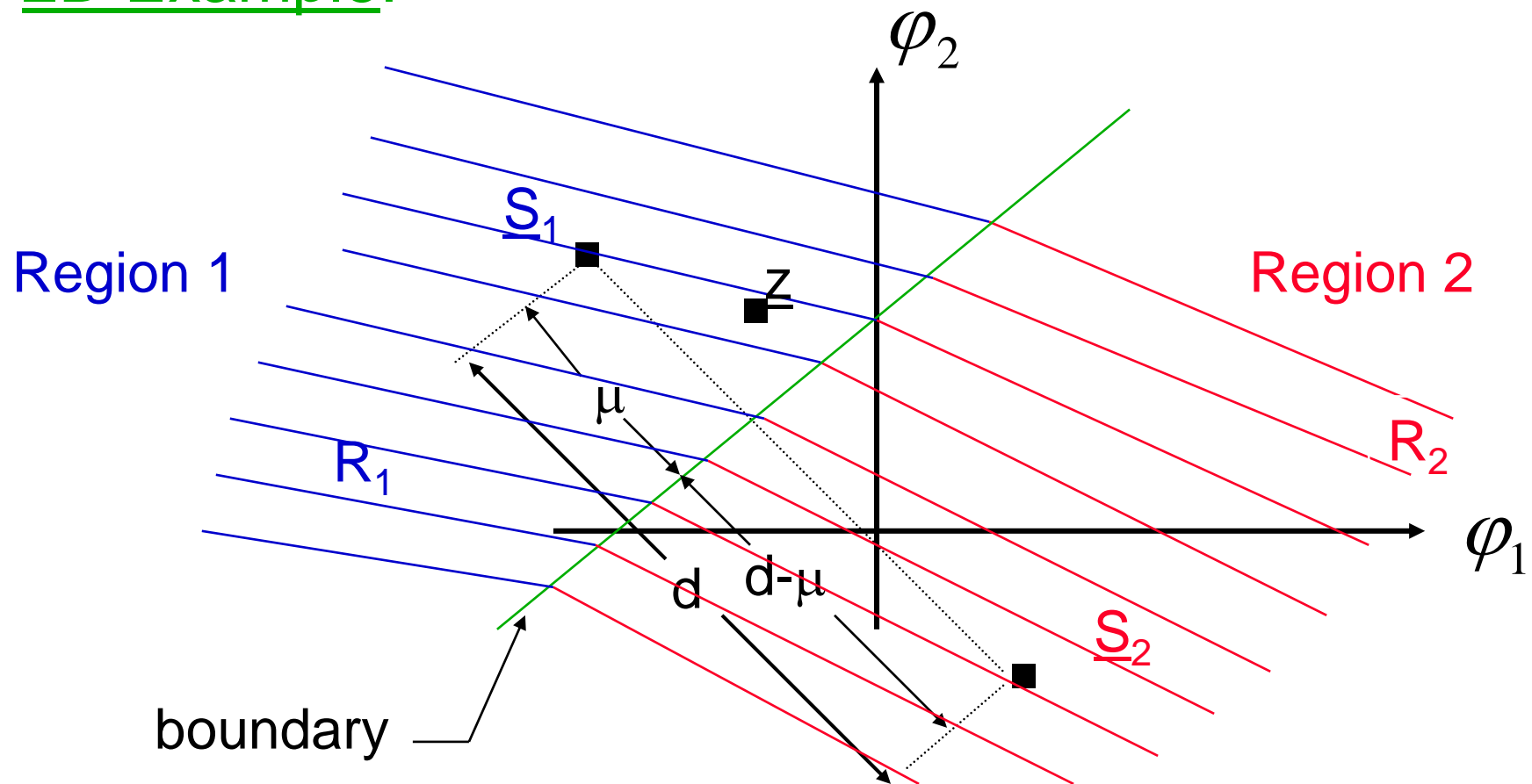
$$\Rightarrow \mu = \frac{c + d^2}{2d} = \frac{N_0}{2d} \log \left[\frac{P(m_1)}{P(m_2)} \right] + \frac{d}{2}$$



We note that similar lines can be used to derive decision regions for $M > 2$.

Decision Regions

2D Example:



Probability of Error

We can also deduce the probability of error expressions from the graphical interpretation of decision regions

$$P(C / m_k) = P(z \in R_k)$$

and
$$P(C) = \sum_{k=1}^M P(C / m_k) P(m_k)$$



$$P_e = 1 - P(C)$$



Remember Q-function

Remember that the Q-function is a standard form for expressing error probabilities without closed form expression

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

$$x \geq 0$$

Noise has probability distribution

$$f_n(\underline{n}) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{n_j^2}{2\sigma^2}}$$

Probability of Error

By inspection, decide m_1 if $z = s_1 + n \in R_1$

$$\xrightarrow{\text{blue arrow}} Z > d' \xrightarrow{\text{blue arrow}} n > -\mu$$

$$\begin{aligned} P(C / s = s_1) &= P(n > -\mu) \\ &= 1 - Q\left[\frac{\mu}{\sigma}\right] = 1 - Q\left[\frac{\mu}{\sqrt{N_0/2}}\right] \end{aligned}$$

Likewise

$$P(C / s = s_2) = 1 - Q\left[\frac{d - \mu}{\sqrt{N_0/2}}\right]$$

Decision Regions

$$\begin{aligned}
 P(C) &= P(m_1) \left\{ 1 - Q \left[\frac{\mu}{\sqrt{N_0/2}} \right] \right\} + P(m_2) \left\{ 1 - Q \left[\frac{d - \mu}{\sqrt{N_0/2}} \right] \right\} \\
 &= 1 - P(m_1) Q \left[\frac{\mu}{\sqrt{N_0/2}} \right] - P(m_2) Q \left[\frac{d - \mu}{\sqrt{N_0/2}} \right] \\
 \Rightarrow \quad P_e &= P(m_1) Q \left[\frac{\mu}{\sqrt{N_0/2}} \right] + P(m_2) Q \left[\frac{d - \mu}{\sqrt{N_0/2}} \right]
 \end{aligned}$$


where

$$d = 2\sqrt{E} \quad \text{and} \quad \mu = \frac{N_0}{4\sqrt{E}} \log \left[\frac{P(m_1)}{P(m_2)} \right] + \sqrt{\frac{E}{44}}$$

Decision Regions

Note that when $P(m_1) = P(m_2)$,

$$\mu = \sqrt{E} \equiv \frac{d}{2} \quad \leftarrow$$


$$\begin{aligned} P_e &= Q \left[\frac{d/2}{\sqrt{N_0/2}} \right] = Q \left[\sqrt{\frac{d^2}{2N_0}} \right] = Q \left[\sqrt{\frac{2E}{N_0}} \right] \\ &= Q \left[\sqrt{\frac{E_1 + E_2 - 2\rho_{12}\sqrt{E_1E_2}}{2N_0}} \right] = Q \left[\sqrt{\frac{2E}{N_0}} \right] \end{aligned}$$

Decision Regions

- This example demonstrate an interesting fact
- When optimum receiver is used, P_e does not depend upon the specific waveform used.
- P_e depends only on their geometrical representation in signal space.
- In particular, P_e depends on signal waveforms only through their energies (distance).

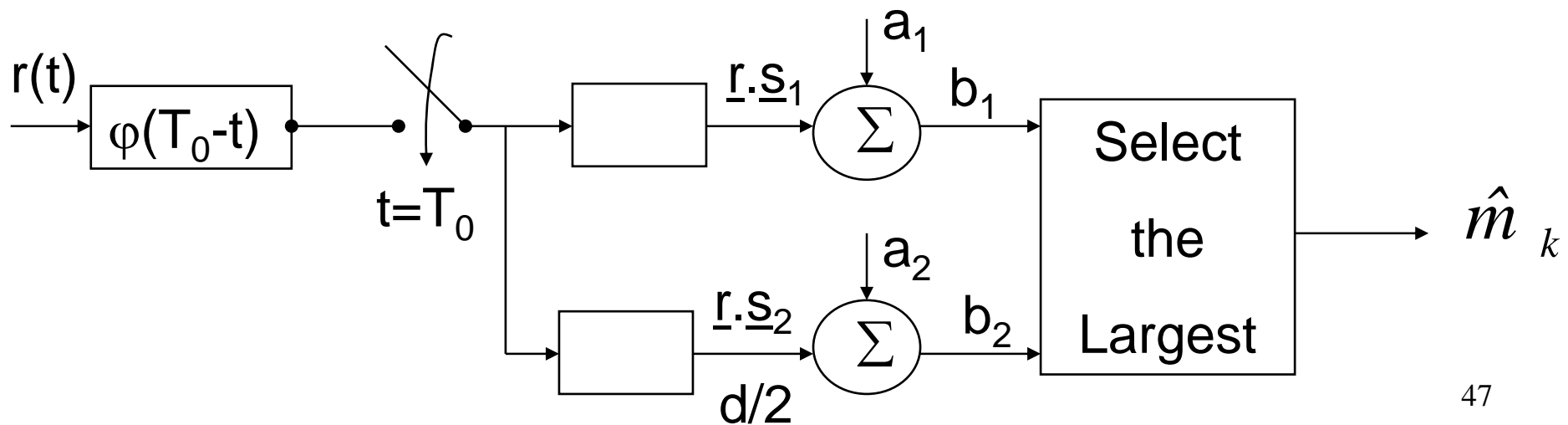
Binary Modulation Revisited

Next note that $N = 1$ and $M = 2$

$$s_{11} = \sqrt{E} \quad s_{21} = -\sqrt{E}$$

$$\underline{r} \cdot \underline{s}_1 = r_1 s_{11} = \sqrt{E} r \quad (r \equiv r_1)$$

$$\underline{r} \cdot \underline{s}_2 = r_2 s_{21} = -\sqrt{E} r$$



Binary Modulation Revisited

Decision criterion, Select m_1 if

$$\sqrt{E} r + a_1 > \sqrt{E} r + a_2$$

$$\text{or} \quad r > \frac{a_2 - a_1}{2\sqrt{E}} \quad \text{select} \quad m_1$$

$$\text{Select } m_1 \text{ if } r > \underbrace{\frac{N_0}{4\sqrt{E}} \log \left[\frac{P(m_2)}{P(m_1)} \right]}_{\substack{\text{Threshold} = 0 \\ \text{if } P(m_1) = P(m_2)}}$$

Binary Modulation Revisited

Note that receiver output $r = z$

 If $r > d'$, $z \in R_1$  Select m_1 .

But

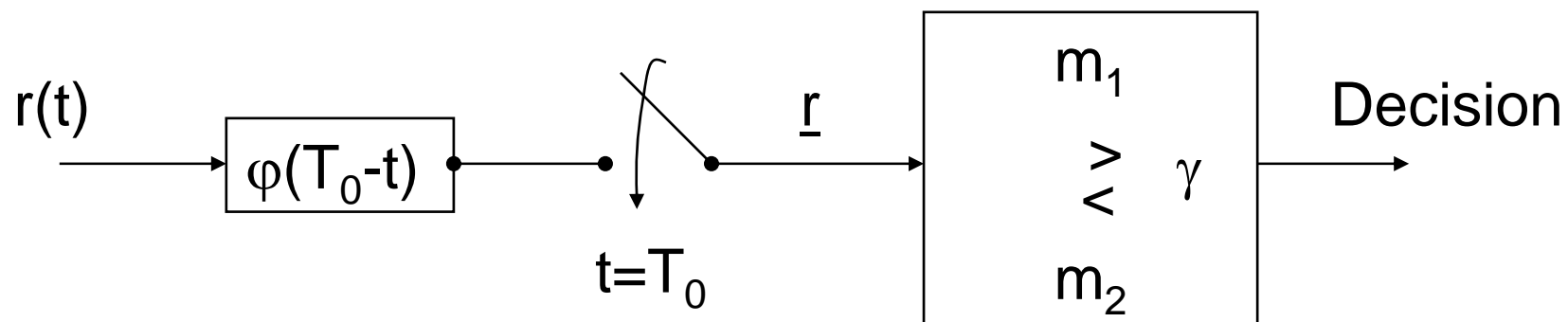
$$d' = \sqrt{E} - \mu = \frac{N_0}{4\sqrt{E}} \log \left[\frac{P(m_2)}{P(m_1)} \right]$$

i.e., result can be obtained almost **by inspection**

Binary Modulation Revisited

Note: This example demonstrates that the optimum binary receivers we derived in Elec 214 are special cases of the optimum M-ary receiver

$$\gamma = \frac{N_0}{4\sqrt{E}} \log \left[\frac{P(m_2)}{P(m_1)} \right]$$



Example

Three equally probable messages m_1 , m_2 , and m_3 are to be transmitted over an AWGN channel with noise power-spectral density $N_0 / 2$. The messages are

$$s_1(t) = \begin{cases} 1 & 0 \leq t \leq T \\ 0 & \text{otherwise} \end{cases}$$
$$s_2(t) = -s_3(t) = \begin{cases} 1 & 0 \leq t \leq \frac{T}{2} \\ -1 & \frac{T}{2} \leq t \leq T \\ 0 & \text{otherwise} \end{cases}$$

1. What is the dimensionality of the signal space ?
2. Find an appropriate basis for the signal space (Hint: You can find the basis without using the Gram-Schmidt procedure).
3. Draw the signal constellation for this problem.
4. Sketch the optimal decision regions R_1 , R_2 , and R_3 .
5. Which of the three messages is more vulnerable to errors and why ? In other words, which of $p(\text{Error} | m_i \text{ transmitted})$, $i = 1, 2, 3$ is larger ?

Probability of Error

As a point of interest we wish to prove that ML is the optimum receiver in the sense that it minimizes the probability of error (so far we have only been considering pdf's)

Error occurs when a decision is made in favor of another when the signal falls outside the decision region R_k

$$\longrightarrow P(e | m_k) = \int_{R_k^c} f(z | m_k) dz$$

where the superscript c refers to complement set
Average overall P(e) is thus

$$P(e) = \sum_{k=1}^M \frac{1}{M} P(e | m_k) = \sum_{k=1}^M \frac{1}{M} \int_{R_k^c} f(z | m_k) dz = \sum_{k=1}^M \frac{1}{M} \left[1 - \int_{R_k} f(z | m_k) dz \right]$$

We note that P(e) is minimized by selecting the largest $f(z|m_k)$ for all k and z