Chapter 5  Signal - Space Analysis

5.1 Introduction

Figure 5.1 Block diagram of a generic digital communication system.

The message source emits one symbol $m_i$

$m_i \in \{m_1, m_2, \ldots, m_M\}$

The probability the symbol $m_i$ is emitted is

$$P_i = P(m_i) = \frac{1}{M} \text{ for } i = 1, 2, \ldots, M \quad (5.1)$$

The transmitter codes $m_i$ into $s_i(t)$

The energy of $s_i(t)$ is

$$E_i = \int_0^T s_i^2(t) dt \ , \ i = 1, 2, \ldots, M \quad (5.2)$$
Assuming that the channel is linear and the channel noise, $w(t)$, is AWGN
\[ x(t) = s_i(t) + w(t) \text{, for } 0 \leq t \leq T \text{ and } i = 1, 2, \ldots, M \]  
(5.3)

![Diagram](Image)

**Figure 5.2** Additive white Gaussian noise (AWGN) model of a channel.

At the receiver the average prob. of symbol error

\[ P_e = \sum_{i=1}^{M} p_i P(\hat{m} \neq m_i|m_i) \]  
(5.4)

$m_i$ : the transmitted symbol

$\hat{m}$ : the estimate
Geometric Representation of Signals

Any set of $M$ signals $\{s_i(t)\}$ can be represented by linear combinations of $N$ orthonormal basis functions, where $N \leq M$

$$s_i(t) = \sum_{j=1}^{N} s_{ij} \phi_j(t) \quad i = 1, 2, \ldots, M \quad (5.5)$$

$$s_{ij} = \int_{0}^{T} s_i(t) \phi_j(t) \, dt \quad i = 1, 2, \ldots, M \quad j = 1, 2, \ldots, N \quad (5.6)$$

$$\int_{0}^{T} \phi_i(t) \phi_j(t) \, dt = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (5.7)$$

\(\{s_{ij}\}_{j=1}^{N}\) is an $N$-dimensional vector
Figure 5.3 (a) Synthesizer for generating the signal $s_i(t)$. (b) Analyzer for generating the set of signal vectors $\{s_i\}$. 
Figure 5.4 Illustrating the geometric representation of signals for the case when $N = 2$ and $M = 3$. 

\[ s_i = \begin{bmatrix} s_{i1} \\ s_{i2} \\ \vdots \\ s_{iN} \end{bmatrix} \quad i = 1,2,\ldots,M \]  

$\phi_1(t), \phi_2(t), \ldots, \phi_N(t)$ form the signal space.
The length (norm) is
\[ \|s_i\|^2 = s_i^T s_i \]
\[ = \sum_{j=1}^{N} s_{ij}^2 \quad i = 1, 2, \ldots, M \quad (5.9) \]

\[ E_i = \int_0^T s_i^2(t)dt \]
\[ = \int_0^T \left[ \sum_{j=1}^{N} s_{ij} \phi_j(t) \right] \left[ \sum_{k=1}^{N} s_{jk} \phi_k(t) \right] dt \quad (5.10) \]

\[ E_i = \sum_{j=1}^{N} \sum_{k=1}^{N} s_{ij} s_{ik} \int_0^T \phi_j(t) \phi_k(t)dt \quad (5.11) \]

Recall \( \int_0^T \phi_i(t) \phi_j(t) dt = \delta_{ij} \)

\[ E_i = \sum_{j=1}^{N} s_{ij}^2 = \|s_i\|^2 \quad (5.12) \]

\[ \int_0^T s_i(t)s_k(t)dt = \int_0^T \left[ \sum_{l=1}^{N} s_{il} \phi_l(t) \right] \left[ \sum_{j=1}^{N} s_{kj} \phi_j(t) \right] dt \]
\[ = \sum_{l=1}^{N} \sum_{j=1}^{N} s_{il} s_{kj} \int_0^T \phi_l(t) \phi_j(t)dt \]
\[ = \sum_{l=1}^{N} s_{il} s_{kl} = s_i^T s_k \quad (5.13) \]

invariant to the choice of \( \{\phi_i(t)\} \)
The Euclidean distance $d_{ik}$ is defined as

$$d_{ik}^2 = \left\| s_i - s_k \right\|^2$$

$$= \sum_{j=1}^{N} (s_{ij} - s_{kj})^2$$

$$= \int_{0}^{T} (s_i(t) - s_k(t))^2 \, dt$$

The angle $\theta_{ik}$ between $S_i$ and $S_k$ is defined as

$$\cos \theta_{ik} = \frac{s_i^T s_k}{\left\| s_i \right\| \left\| s_k \right\|} \quad (5.15)$$
Example 5.1 Schwarz Inequality

Let \( s_1(t) = s_{11}\phi_1(t) + s_{12}\phi_2(t) \)
\( s_2(t) = s_{21}\phi_1(t) + s_{22}\phi_2(t) \)

\[
\int_{-\infty}^{\infty} \phi_i(t)\phi_j(t)dt = \delta_{ij}
\]

\[
s_1 = \begin{bmatrix} s_{11} \\ s_{12} \end{bmatrix}, \quad s_2 = \begin{bmatrix} s_{21} \\ s_{22} \end{bmatrix}
\]

\[-1 \leq \cos \Theta = \frac{s_1^T s_2}{\|s_1\| \|s_2\|} = \frac{\int_{-\infty}^{\infty} s_1(t)s_2(t)dt}{(\int_{-\infty}^{\infty} s_1^2(t)dt)^{1/2}(\int_{-\infty}^{\infty} s_2^2(t)dt)^{1/2}} \leq 1\]

\[
\Rightarrow (\int_{-\infty}^{\infty} s_1(t)s_2(t)dt)^2 \leq (\int_{-\infty}^{\infty} s_1^2(t)dt)(\int_{-\infty}^{\infty} s_2^2(t)dt)
\]

For complex value signals

\[
\left| \int_{-\infty}^{\infty} s_1(t)s_2^*(t)dt \right|^2 \leq \int_{-\infty}^{\infty} |s_1(t)|^2 dt \int_{-\infty}^{\infty} |s_2(t)|^2 dt
\]
Figure 5.5 Vector representations of signals $s_1(t)$ and $s_2(t)$, providing the background picture for providing the Schwarz inequality.
Gram-Schmidt Orthogonalization Procedure (GSOP)

(gsop) generates \( \{ \phi_i(t) \} \) from \( \{ s_j(t) \} \)

Define \( \phi_1(t) = \frac{s_1(t)}{\sqrt{E_1}} \) \hspace{1cm} (5.19)

\[ s_1(t) = \sqrt{E_1} \phi_1(t) = s_{11} \phi_1(t) \] \hspace{1cm} (5.20)

where \( s_{11} = \sqrt{E_1} \)

define \( s_{21} = \int_0^r s_2(t)\phi_1(t)\,dt \) \hspace{1cm} (5.21)

let \( g_2(t) = s_2(t) - s_{21}(t)\phi_1(t) \) \hspace{1cm} (5.22)

\[ \phi_2(t) = \frac{g_2(t)}{\sqrt{\int_0^r g_2^2(t)\,dt}} \] \hspace{1cm} (5.23)

\[ \int_0^r \phi_2^2(t)\,dt = 1 \]
\[
\int \phi_1(t) \phi_2(t) dt = 0
\]

In general, we define

\[
g_i = s_i(t) - \sum_{j=1}^{i-1} s_{ij} \phi_j(t)
\] (5.25)

\[
s_{ij} = \int s_i(t) \phi_j(t) dt , j = 1,2,\ldots,i-1
\] (5.26)

We may now define \( \phi_i(t) \) as

\[
\phi_i(t) = \frac{g_i(t)}{\sqrt{\int g_i^2(t) dt}} , i = 1,2,\ldots,N
\] (5.27)

where \( \{s_i(t)\} , i = 1,2,\ldots,M \) are linear indp.

then \( M = N \) otherwise \( M > N \)

Important properties of \( \{\phi_i(t)\} \)

1. The form of basis function is not specified

2. \( s_i(t) = \sum_{i=1}^{N} s_{ij} \phi_j(t) \) is an exact expression
Example 5.2 2B1Q Code

The 2B1Q code was described in Chapter 4 as the North American line code for digital subscriber lines. It represents a quaternary PAM signal as shown in the Gray-encoded alphabet of Table 5.1. The four possible signals, \( s_1(t), s_2(t), s_3(t), \) and \( s_4(t) \), are amplitude-scaled versions of a Nyquist pulse. Each signal represents a dabit. We wish to find the vector representation of the 2B1Q code.

This example is simple enough for us to solve it by inspection. Let \( \phi_1(t) \) denote the Nyquist pulse, normalized to have unit energy. The \( \phi_1(t) \) so defined is the only basis function for the vector representation of the 2B1Q code. Accordingly, the signal-space representation of this code is as shown in Figure 5.6. It consists of four signal vectors \( s_1, s_2, s_3, \) and \( s_4 \), which are located on the \( \phi_1 \)-axis in a symmetric manner about the origin. In this example, we thus have \( M = 4 \) and \( N = 1 \).

We may generalize the result depicted in Figure 5.6 for the 2B1Q code as follows. The signal-space diagram of an \( M \)-ary pulse-amplitude modulated signal, in general, is one-dimensional with \( M \) signal points uniformly positioned on the only axis of the diagram.

![Figure 5.6 Signal-space representation of the 2B1Q code.](image)
Figure 4.20
Output of a quaternary system. (a) Waveform. (b) Representation of the 4 possible dibits, based on Gray encoding.
Figure 5.1
Block diagram of a generic digital communication system.
5.3 Conversion of the continuous AWGN channel into a vector channel

\[ s_i(t) \]

\[ + \]

\[ x(t) \]

\[ \sum \]

White Gaussian noise \( w(t) \)

Figure 5.2 Additive white Gaussian noise (AWGN) model of a channel.

Figure 5.3 (a) Synthesizer for generating the signal \( s_i(t) \). (b) Analyzer for generating the set of signal vectors \( \{s_i\} \).
\[ x(t) = s_i(t) + w(t) \quad , \quad i = 1, 2, ..., M , \quad 0 \leq t \leq T \quad (5.28) \]
Note that \( w(t) \) is AWGN and has infinite dimensions.

The output of correlator \( j \), \( x_j \), is

\[ x_j = \int_0^T x(t)\phi_j(t)dt \]

\[ = s_{ij} + w_j \quad , \quad j = 1, 2, ..., N \quad (5.29) \]

where

\[ s_{ij} = \int_0^T s_i(t)\phi_j(t)dt \quad (5.30) \]

\[ w_j = \int_0^T w(t)\phi_j(t)dt \quad (5.31) \]

consider that

\[ x'(t) = x(t) - \sum_{j=1}^{N} x_j\phi_j(t) \quad (5.32) \]
\[ x'(t) = s_i(t) + w(t) - \sum_{j=1}^{N} (s_{ij} + w_j) \phi_j(t) \]
\[ = w(t) - \sum_{j=1}^{N} w_j \phi_j(t) \]
\[ = w'(t) \quad \text{(5.33)} \]

The sample function \( x'(t) \) depends on \( w(t) \) only.

From (5.32)
\[ x(t) = \sum_{j=1}^{N} x_j \phi_j(t) + x'(t) \]
\[ = \sum_{j=1}^{N} x_j \phi_j(t) + w'(t) \]

remainder, which is not in the signal space

and will not affect the receiver performance
Let $X_j$ denote the random variable of the output of correlator $j$.

The mean of $X_j$ for signal $s_i(t)$ is (from 5.29)

$$
\mu_{x_j} = E[X_j] = E[s_{ij}] + E[W_j] = s_{ij}.
$$
\[
\sigma_{xj}^2 = E[(X_j - s_{ij})^2] \\
= E[W_j^2]
\]

According to (5.31)

\[
W_j = \int_0^T W(t)\phi_j(t)dt
\]

\[
\sigma_{xj}^2 = E\left[\int_0^T W(t)\phi_j(t)dt\int_0^T W(u)\phi_j(u)du\right]
\]

\[
= \int_0^T \int_0^T \phi_j(t)\phi_j(u)E[W(t)W(u)]dtdu
\]

\[
= \int_0^T \int_0^T \phi_j(t)\phi_j(u)R_w(t,u)dtdu
\]

\[
= \frac{N_0}{2} \int_0^T \phi_j^2(t)dt = \frac{N_0}{2}
\]

Because \(R_w(t,u) = \frac{N_0}{2} \delta(t-u)\) (5.39)
\[
\text{cov} \left[ X_j X_k \right] = E \left[ (X_j - \mu_{x_j})(X_k - \mu_{x_k}) \right] \\
= E \left[ (X_j - S_{ij})(X_k - S_{ik}) \right] \\
= E \left[ W_j W_k \right] \\
= \int_0^T \int_0^T \phi_j(t) \phi_k(u) R_w(t,u) \, dt \, du \\
= \frac{N_0}{2} \int_0^T \int_0^T \phi_j(t) \phi_k(u) \delta(t-u) \, dt \, du \\
= \frac{N_0}{2} \int_0^T \phi_j(t) \phi_k(t) \, dt \\
= 0 \quad , \quad j \neq k
\]

The \( X_j \) are mutually uncorrelated.

Since the \( X_j \) are Gaussian, they are statistically independent. i.e. the sampled correlator outputs are independent Gaussian random variables.
If the random variables $X$ and $Y$ are jointly normal, their joint probability function is

$$f(x, y) = \frac{1}{2\pi\sigma_x \sigma_y \sqrt{1-r^2}} \exp \left\{ -\frac{1}{2(1-r^2)} \left[ \frac{(x-\mu_x)^2}{\sigma_x^2} - 2r \frac{(x-\mu_x)(y-\mu_y)}{\sigma_x \sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right] \right\}$$

where $\mu_x$ and $\mu_y$ are the means of $X$ and $Y$, and $\sigma_x^2$ and $\sigma_y^2$ are their variances.

$r = \text{correlation coefficient}$

$$r = \frac{E[XY] - \mu_x \mu_y}{\sigma_x \sigma_y}$$

If $X$ and $Y$ are uncorrelated and normal

$$r = 0 \quad E[XY] = \mu_x \mu_y$$

$$\Rightarrow f(x, y) = f(x)f(y)$$

$X$ and $Y$ are independent

(If $E[XY] = 0$, $X$ and $Y$ are called orthogonal)
Define the vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

(5.43)

$x$ is called the observation vector.

The elements are indep. Gaussian RVs with means $s_{ij}$ and $\frac{N_0}{2}$ variances (for sample $s_i(t)$).

The conditional pdf given that $S_i(t)$ or $(m_i)$ was transmitted is

$$f_x(x|m_i) = \prod_{j=1}^{N} f_{x_j}(x_j|m_i) , \ i = 1,2,...,M$$

(5.44)

Any channel that satisfies (5.44) is called a memoryless channel.
\[ f_{x_j}(x_j|m_i) = \frac{1}{\sqrt{\pi N_0}} \exp \left[ -\frac{1}{N_0} (x_j - s_{ij})^2 \right], \quad j = 1, 2, \ldots, N \]
\[ f_x(x|m_i) = (\pi N_0)^{-\frac{N}{2}} \exp \left[ -\frac{1}{N_0} \sum_{j=1}^{N} (x_j - s_{ij})^2 \right], \quad i = 1, 2, \ldots, M \]

Recall
\[
x(t) = \sum_{j=1}^{N} x_j \phi_j(t) + w'(t)
\]

\[ w'(t) \text{ is a zero - mean Gaussian process and indep. of } \{X_j\} \] (prob. 5.10)

\[
E[X_jw'(t_k)] = E[(S_{ij} + W_j)w'(t_k)]
= E[W_jw'(t_k)] = w(t_k) - \sum_{j=1}^{N} W_i \phi_i(t_k)
= 0
\]

⇒ Theorem of irrelevance: Only the projections of the noise onto the basis functions affects the detections, the remainder is irrelevant

⇒ The AWGN channel is equivalent to an N-dim. vector channel
\[
x = s_i + w, \quad i = 1, 2, \ldots, M
\]
5.4 Likelihood Functions

Given the observation vector $\mathbf{x}$, we have to estimate the transmitted symbol $m_i$.

Denote the likelihood function by $L(m_i)$

$$L(m_i) = f_x(x|m_i), \quad i = 1,2,\ldots,M \quad (5.49)$$

$$= (\pi N_0)^{-N/2} \exp \left[ -\frac{1}{N_0} \sum_{j=1}^{N} (x_j - S_{ij})^2 \right] \quad i = 1,2,\ldots,M$$

For convenience, we define the log-likelihood function

$$l(m_i) = \log L(m_i), \quad i = 1,2,\ldots,M \quad (5.50)$$

1. A pdf is always nonnegative, so $L(m_i)$ is nonnegative.

2. Log function is a monotonical function.

⇒ $l(m_i)$ bears a one-to-one relationship to $L(m_i)$.

From (5.46) and (5.50), we have

$$l(m_i) = -\frac{1}{N_0} \sum_{j=1}^{N} (x_j - s_{ij})^2 \quad i = 1,2,\ldots,M \quad (5.51)$$

where we ignore the constant $-\left(\frac{N}{2}\right) \log(\pi N_0)$. 

24
5.5 Coherent Detection of Signals in Noise: Maximum Likelihood Decoding.

The set of transmitted signals is called a signal constellation.

Figure 6.46 Signal constellation for (a) $M$-ary PSK and (b) corresponding $M$-ary QAM, for $M = 16$. 
Figure 5.7 Illustrating the effect of noise perturbation, depicted in (a), on the location of the received signal point, depicted in (b).
Detection problem:

Given \( x \), perform a mapping from \( x \) to an estimate \( \hat{m}_i \) of \( m_i \), in a way that would minimize the probability of error.

The prob.of error denoted by \( P_e(m_i|x) \) is

\[
P_e(m_i|x) = P(m_i \text{ not sent} | x) = 1 - P(m_i \text{ sent} | x)
\]

(5.52)

The optimum decision rule is

set \( \hat{m} = m_i \) if

\[
P(m_i \text{ sent} | x) \geq P(m_k \text{ sent} | x) \quad \text{for all } k \neq i
\]

(5.53)

which is also called the maximum a posteriori probability (MAP) rule.
In terms to the a priori prob. of \( \{m_i\} \), using Bayes' rule, we may restate the MAP rule as

\[
\hat{m} = m_i \text{ if } \frac{P_k f_x(x|m_k)}{f_x(x)} \text{ is maximum for } k = i \quad (5.54)
\]

\{(5.53) \Rightarrow P(m_k \text{ sent} | X) \text{ is maximum for } i = k \}

where \( p_k \) is the a priori prob. of \( m_k \)

Note that

1. \( f_x(x) \) is indep. of \( \{m_i\} \)

2. If \( \{m_i\} \) are equally likely, \( p_k = p_i = p \)

3. \( f_x(x|m_k) \) bears one-to-one relationship to \( l(m_k) \)

Then we can restate the decision rule as

\[
\hat{m} = m_i \text{ if } l(m_k) \text{ is maximum for } k = i \quad (5.55)
\]
(5.55) is called as the maximum likelihood rule. The maximum likelihood decoder differs from the maximum a posteriori decoder (it assumes equally likely message symbols).

Let $Z$ denote the $N$-dim space (observation space). We may partition $Z$ into $M$-decision regions denoted by $Z_1, Z_2, \ldots, Z_M$.

Observation vector $x$ lies in $Z_i$ if

$$l(m_k) \text{ is max. for } k = i$$

Recall (5.51) $l(m_k) = -\frac{1}{N_0} \sum_{j=1}^{N} (x_j - s_{kj})^2, \ i = 1, 2, \ldots, M$ (5.52)

minimum this term to maximize $l(m_k)$ by the choice $i = k$
x lies in $Z_i$ if
\[
\sum_{j=1}^{N} (x_j - s_{kj})^2 = \left\| x - s_k \right\|^2 \quad \text{is min. for } k = i \quad (5.57)
\]
\[\Rightarrow x \in Z_i \, , \, \text{if } \left\| x - s_k \right\| \text{ is min. for } k = i \quad (5.59)\]
\[\Rightarrow \text{to choose the message point closest to the received signal point.}\]
\[
\sum_{j=1}^{N} (x_j - s_{kj})^2 = \sum_{j=1}^{N} x_j^2 - 2 \sum_{j=1}^{N} x_j s_{kj} + \sum_{j=1}^{N} s_{kj}^2 \quad (5.60)
\]

indep. of $k$  \quad energy of $s_k(t) = E_k$
Equivalently we have

\[ x \in Z_i \text{ if } \sum_{j=1}^{N} x_j s_{kj} - \frac{1}{2} E_k \text{ is max. for } k = i \]

**Figure 5.8** Illustrating the partitioning of the observation space into decision regions for the case when \( N = 2 \) and \( M = 4 \); it is assumed that the \( M \) transmitted symbols are equally likely.
5.6 Correlation Receiver \[ \text{when} \{s_i(t)\} \text{ are equally likely} \]

The correlation receiver consists of two parts:

* detector (correlator)
* decoder
Figure 5.9 (a) Detector or demodulator. (b) Signal transmission decoder.
Equivalence of correlation and matched filter (receivers)

Recall the output of the matched filter

$$y_j(t) = \int_{-\infty}^{\infty} x(\tau)h_j(t-\tau)d\tau$$  \hspace{1cm} (5.63)

let $$h_j(t) = \phi_j(T-t)$$  \hspace{1cm} (5.64)

$$y_j(t) = \int_{-\infty}^{\infty} x(\tau)\phi_j(T-t+\tau)d\tau$$  \hspace{1cm} (5.65)

sample at $$t = T$$, and $$\phi_j(t) = 0$$, $$t < 0$$, or $$t > T$$

$$y_j(T) = \int_{-\infty}^{\infty} x(\tau)\phi_j(\tau)d\tau$$

$$= \int_{0}^{T} x(\tau)\phi_j(\tau)d\tau$$  \hspace{1cm} (5.66)

correlator
Figure 5.10 Detector part of matched filter receiver; the signal transmission decoder is as shown in Fig. 5.9b.
5.7 Prob. of Error

The average prob. of symbol error

\[ P_e = \sum_{i=1}^{M} p_i P(x \text{ does not lie in } Z_i \mid m_i \text{ sent}) \]

\[ = \frac{1}{M} \sum_{i=1}^{M} P(x \text{ does not lie in } Z_i \mid m_i \text{ sent}) \] (5.67)

\[ = 1 - \frac{1}{M} \sum_{i=1}^{M} P(x \text{ lies in } Z_i \mid m_i \text{ sent}) \] (5.68)
Invariance of $P_e$ to Rotation and Translation

Changes of coordinates does not affect $P_e$

1. $P_e$ depends on $\|x - s_k\|$

2. Gaussian noise is spherically symmetric

The effect of a rotation is equivalent to

$$s_{i,\text{rotate}} = Qs_i \quad i = 1,2,...,M \quad (5.70)$$

where $QQ^T = I \quad (5.69)$

The noise vector becomes

$$w_{\text{rotate}} = Qw \quad (5.71)$$
1. $w_{\text{rotate}}$ is Gaussian

2. $E[w_{\text{rotate}}] = E[Qw]$
   \[ = 0 \quad \text{(5.72)} \]

3. Since $E[ww^T] = \frac{N_0}{2} I$ \quad \text{(5.73)}

   \[
   E[w_{\text{rotate}} w_{\text{rotate}}^T] = E[QW(QW)^T]
   = QE[ww^T]Q^T
   = \frac{N_0}{2} QQ^T
   = \frac{N_0}{2} I \quad \text{(5.74)}
   \]

$w_{\text{rotate}}$ are Gaussian with zero mean and variance $\frac{N_0}{2}$

The observation vector

\[
x_{\text{rotate}} = Qs_i + w \quad i = 1, 2, \ldots, M \quad \text{(5.75)}
\]
\[ \| x_{\text{rotate}} - s_{i,\text{rotate}} \| = w = \| x - s_i \| \quad \text{for all } i \] (5.76)

Principle of rotational invariance

If a signal constellation is rotated by an orthonormal transformation, that is,

\[ s_{i,\text{rotate}} = Qs_i, \quad i = 1,2,\ldots,M \]

where \( Q \) is an orthonormal matrix, then the probability of symbol error \( P_e \) incurred in maximum likelihood signal detection over an AWGN channel is completely unchanged.

Figure 5.11 A pair of signal constellations for illustrating the principle of rotational invariance.
For Translation

\[ s_{i, \text{translate}} = s_i - a, \ i = 1,2,\ldots,M \quad (5.77) \]

\[ x_{\text{translate}} = x - a \quad (5.78) \]

\[ \left\| x_{\text{translate}} - s_{i, \text{translate}} \right\| = \left\| x - a - s_i + a \right\| = \left\| x - s_i \right\| \quad \text{for all} \ i \quad (5.79) \]

The principle of translational invariance:

If a signal constellation is translated by a constant vector amount, then the probability of symbol error \( P_e \) incurred in maximum likelihood signal detection over an AWGN channel is completely unchanged.
Minimum Energy Signals

Figure 5.12 A pair of signal constellations for illustrating the principle of translational invariance.

consider \( \{ m_i \} \) represented by \( \{ s_i \} \)

The average energy of \( \{ s_i \} \) translated by a

\[
\xi_{\text{translate}} = \sum_{i=1}^{M} \| s_i - a \|^2 p_i \quad (5.80)
\]

\( p_i : \) the prob. of \( m_i \)

\[
\| s_i - a \|^2 = \| s_i \|^2 - 2a^T s_i + \| a \|^2
\]
\[ \xi_{\text{translate}} = \sum_{i=1}^{M} \| s_i \|^2 p_i - 2 \sum_{i=1}^{M} a^T s_i p_i + \| a \|^2 \sum_{i=1}^{M} p_i \]

\[ = \xi - 2 a^T E[s] + \| a \|^2 \]  

The energy of original constellation

\[ E[s] = \sum_{i=1}^{M} s_i p_i \]  

\[ \frac{\partial \xi_{\text{translate}}}{\partial a} = 0 \implies a_{\text{min}} = E[s] \]

\[ \xi_{\text{translate, min}} = \xi - \| a_{\text{min}} \|^2 \]

The minimum energy translate:

Given a signal costellation \( \{ s_i \}_{i=1}^{M} \), the corresponding signal constellation with minimum average energy is obtained by subtracting from each signal vector \( s_i \) in the given constellation an amount equal to the constant vector \( E[s] \), where \( E[s] \) is defined by Equation (5.82)
Union Bound on $P_e$

Recall $p_e = 1 - \frac{1}{M} \int_{Z_i} f_x(x \mid m_i) \, dx$  \hspace{1cm} (5.68)

$$f_x(x \mid m_i) = (\pi N_0)^{-N/2} \exp \left[ -\frac{1}{N_0} \sum_{j=1}^{N} (x_j - s_{ij})^2 \right], \quad i = 1, 2, \ldots, M \hspace{1cm} (5.46)$$

The integral may be difficult to obtain.

Let $A_{ik} = \{ x \text{ is closer to } s_k \text{ than to } s_i \}$

The conditional prob.

$$p_e(m_i) \leq \sum_{k=1}^{M} P(A_{ik}), \quad i = 1, 2, \ldots, M \hspace{1cm} (5.85)$$

$$p_e(m_i) \leq \sum_{k=1}^{M} P_2(s_i, s_k), \quad i = 1, 2, \ldots, M \hspace{1cm} (5.86)$$

where $P_2(s_i, s_k)$ is the pairwise error prob. if only $s_i$ and $s_k$ are used and the receiver mistaking $s_i$ for $s_k$. 
Figure 5.13 Illustrating the union bound. (a) Constellation of four message points. (b) Three constellations with a common message point and one other message point retained from the original constellation.
\[ P_2(s_i, s_k) = P(\text{x is closer to } s_k \text{ than } s_i \mid s_i \text{ is sent}) \]

\[ = \int_{d_{ik}/2}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{v^2}{N_0}\right) dv \]

\[ d_{ik} = \|s_i - s_k\| \]

\[ P_2(s_i, s_k) = \frac{1}{2} \text{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right) \quad (5.89) \]

\[ P_e(m_i) \leq \frac{1}{2} \sum_{k \neq i}^{M} \text{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right), i = 1, 2, \ldots, M \quad (5.90) \]

\[ P_e = \sum_{i=1}^{M} p_i P_e(m_i) \leq \frac{1}{2} \sum_{i=1}^{M} \sum_{k=1}^{M} p_i \text{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right) \quad (5.91) \]
1. If the constellation is circularly symmetric, $P_e(m_i)$ is the same for all $i$ (e.g. M-ary PSK)

$$P_e \leq \frac{1}{2} \sum_{k=1}^{M} \text{erfc} \left( \frac{d_{ik}}{2\sqrt{N_0}} \right), \quad \text{for all } i \quad (5.92)$$

2. Define $d_{\min} = \text{minimum } d_{ik}$

$$\text{erfc} \left( \frac{d_{ik}}{2\sqrt{N_0}} \right) \leq \text{erfc} \left( \frac{d_{\min}}{2\sqrt{N_0}} \right)$$

$$P_e \leq \frac{(M-1)}{2} \text{erfc} \left( \frac{d_{\min}}{2\sqrt{N_0}} \right)$$

$$\leq \frac{(M-1)}{2\sqrt{\pi}} \exp \left( -\frac{d_{\min}^2}{4N_0} \right) \quad (5.97)$$

\[ \therefore \text{erfc}(u) \leq \frac{\exp(-u^2)}{\sqrt{\pi u}} \leq \frac{\exp(-u^2)}{\sqrt{\pi}} \quad \text{for large } u \]
BER v.s. $P_e$

**Case 1. (Gray code)**

\[
P_e = P \left( \bigcup_{i=1}^{\log_2 M} \{ \text{ith bit is in error} \} \right) 
\leq \sum_{i=1}^{\log_2 M} P(\text{ith bit is in error})
= \log_2 M \cdot \text{BER}
\]

\[
P_e \geq \text{BER}
\]

\[
\frac{P_e}{\log_2 M} \leq \text{BER} \leq P_e
\]

**Case 2. $M = 2^k$**

\[
\text{BER} = \left(\frac{2^{k-1}}{2^k - 1}\right) P_e = \left(\frac{M/2}{M-1}\right) P_e
\]

For large $M$, $\text{BER} \rightarrow \frac{P_e}{2}$
Case 2

Let \( M = 2^k \)

If all symbol errors are equally likely,

\[
\text{Symbol error} = \frac{P_e}{M - 1} = \frac{P_e}{2^k - 1}
\]

where \( P_e \) is the average symbol error.

The probability that the \( i \)th bit is in error is \( \frac{1}{2} \).

So if the \( i \)th bit is in error then it results \( 2^{k-1} \) symbol errors,

\[
\text{BER} = 2^{k-1} \text{ symbol error}
\]

\[
= \frac{2^{k-1}}{2^k - 1} P_e
\]

\[
= \frac{M/2}{M-1} P_e
\]

For large \( M \), \( \text{BER} \approx \frac{1}{2} P_e \)

The bit errors are not independent

\[
P (\text{ith and } i\text{th bits are in error}) = \frac{2^{k-2}}{2^k - 1} P_e \neq (\text{BER})^2
\]