Chapter 5 Signal-Space Analysis

Lecturer: David Shiung

Date: 2009/06/09

Signal-space Analysis (1/1)

- Geometric representation of signals with finite energy
- Maximum likelihood procedure for the detection of a signal in AWGN channel
- Derivation of the correlation receiver that is equivalent to the matched filter receiver
- Probability of symbol error and the union bound for its approximate calculation

Introduction (1/5)

Block diagram of a generic digital communication system

- A message source emits one symbol every T seconds, with the symbols belonging to an alphabet of M symbols denoted by $m_1, m_2, \ldots, m_M$.

Introduction (2/5)

- Assume that the M symbols of the alphabet are *equally likely*.

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

- $s_i(t)$ is a real-valued energy signal (i.e., a signal with finite energy)

$$E_i = \int_0^T s_i^2(t)\, dt, \qquad i = 1, 2, \ldots, M$$

Introduction (3/5)

- The channel is assumed to have two characteristics:

1. The channel is linear with a bandwidth that is wide enough to accommodate the transmission of signal si(t) with negligible or no distortion.
2. The channel noise, *w(t)*, is the sample function of a zero-mean white Gaussian noise process.

- received signal is

$$x(t) = s_i(t) + w(t), \quad \begin{cases} 0 \le t \le T \\ i = 1, 2, \ldots, M \end{cases}$$

- Making a best estimate of the transmitted signal $s_i(t)$ or, equivalently, the symbol m_i

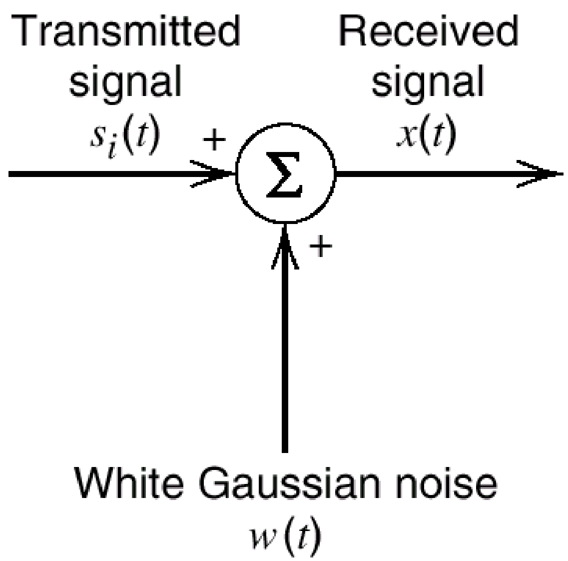
Introduction (4/5)

- The requirement is therefore to design the receiver so as to minimize the average probability of symbol error, defined as

$$P_e = \sum_{i=1}^{M} p_i \, P(\hat{m} \neq m_i | m_i)$$

- m_i is the transmitted symbol, $\hat{m}$ is the estimate produced by the receiver, and $P(\hat{m} \neq m_i | m_i)$ is the conditional error probability given that the *i*th symbol was sent
- The resulting receiver is said to be *optimum in the minimum probability of error* sense.

Introduction (5/5)



Additive white Guassian noise (AWGN) model of a channel

Geometric Representation of Signals (1/7)

- The essence of geometric representation of signals is to represent any set of M energy signals $\{s_i(t)\}$ as linear combinations of N orthonormal basis functions, where $N \leq M$. That is, given a set of real-valued energy signals $s_1(t)$, $s_2(t)$,…, $s_M(t)$, each with duration T, we write

$$s_i(t) = \sum_{j=1}^{N} s_{ij}\phi_j(t), \qquad \begin{cases} 0 \leq t \leq T \\ i = 1, 2, \ldots, M \end{cases}$$

$$s_{ij} = \int_0^T s_i(t)\phi_j(t)\,dt, \qquad \begin{cases} i = 1, 2, \ldots, M \\ j = 1, 2, \ldots, N \end{cases}$$

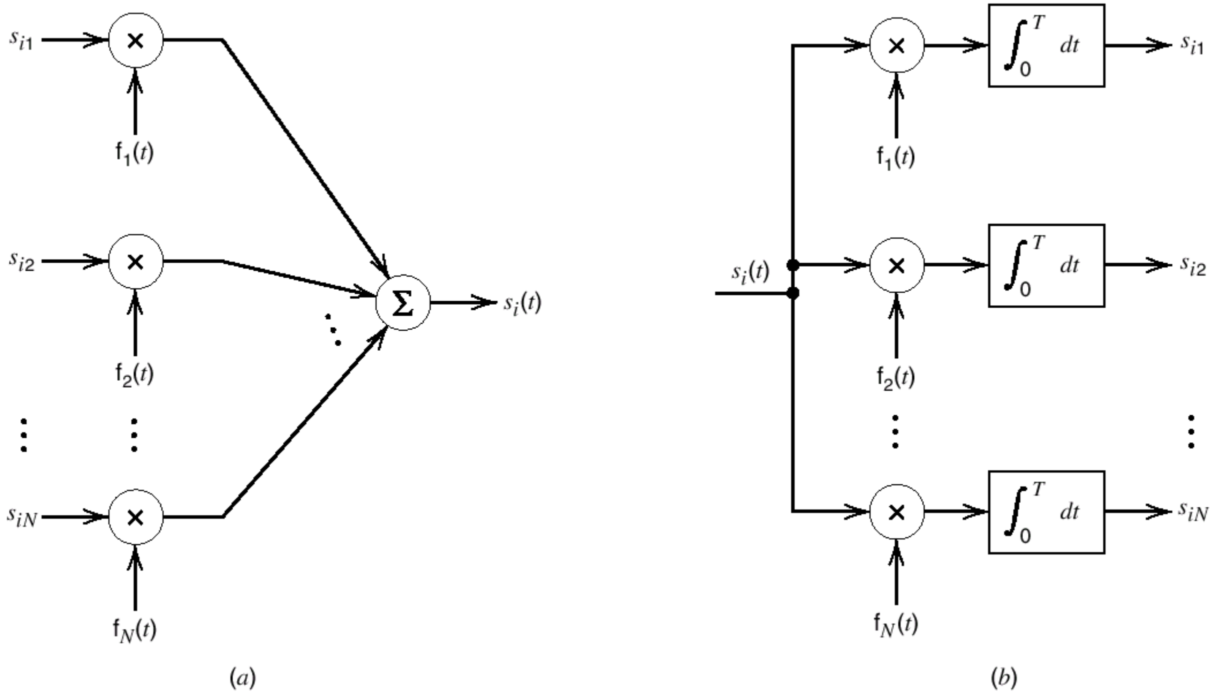
Geometric Representation of Signals (2/7)

- The real-valued basis functions $\phi_1(t), \phi_2(t), \ldots, \phi_N(t)$ are *orthonormal* if

$$\int_0^T \phi_i(t)\phi_j(t)\, dt = \delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

- δ_{ij} is the Kronecker delta. δ_{ij} implies both *normalized* and *orthogonal*.
- The set of coefficients $\{s_{ij}\}_{j=1}^{N}$ may be viewed as an *N-dimensional vector* denoted by $\mathbf{s}_i$

Geometric Representation of Signals (3/7)



(*a*) Synthesizer for generating the signal $s_i(t)$. (*b*) Analyzer for generating the set of signal vectors $\{s_i\}$.

Geometric Representation of Signals (4/7)

$$\mathbf{s}_i = \begin{bmatrix} s_{i1} \\ s_{i2} \\ \vdots \\ s_{iN} \end{bmatrix}, \qquad i = 1, 2, \ldots, M$$

- The vector $\mathbf{s}_i$ is called a *signal vector*.
- Extend two- and three-dimensional Euclidean spaces to an *N*-dimensional Euclidean space with *N* mutually perpendicular axes $\phi_1, \phi_2, \ldots, \phi_N$.
- The squared-length of $\mathbf{s}_i$ is defined as the inner product of $\mathbf{s}_i$

$$\begin{aligned} \| \mathbf{s}_i \|^2 &= \mathbf{s}_i^T \mathbf{s}_i \\ &= \sum_{j=1}^{N} s_{ij}^2, \qquad i = 1, 2, \ldots, M \end{aligned}$$

Geometric Representation of Signals (5/7)

- The energy of a signal $s_i(t)$ of duration T seconds is

$$E_i = \int_0^T s_i^2(t)\, dt$$

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

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

$$E_i = \sum_{j=1}^{N} s_{ij}^2$$

$$= \|\mathbf{s}_i\|^2$$

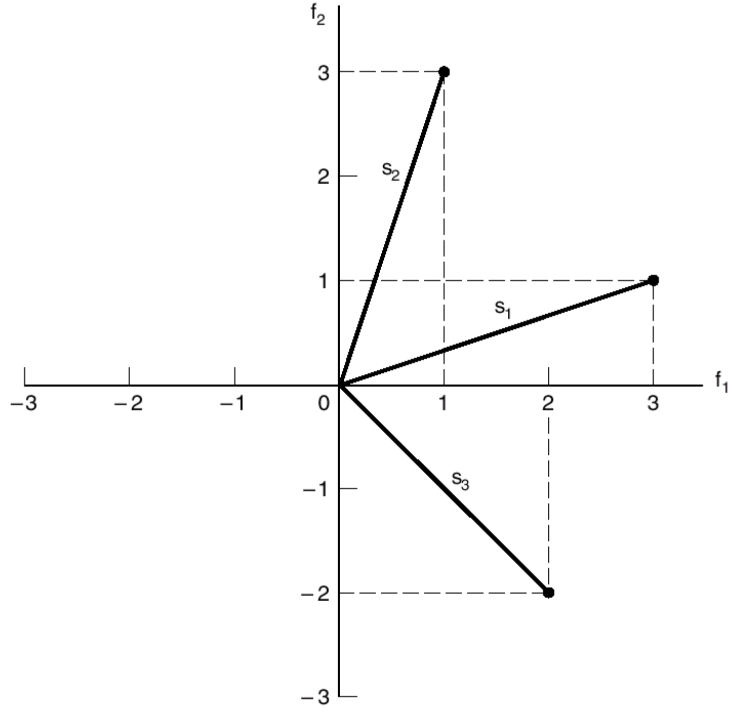
$$\int_0^T s_i(t)s_k(t)\, dt = \mathbf{s}_i^T\mathbf{s}_k$$

Geometric Representation of Signals (6/7)

- The inner product of the signals $s_i(t)$ and $s_k(t)$ over the interval [0,T] is equal to the inner product of their respective vector representations $\mathbf{s}_i$ and $\mathbf{s}_k$.
- Inner product of $s_i(t)$ and $s_k(t)$ is *invariant* to the choice of basis function $\{\phi_j(t)\}_{j=1}^{N}$ in that it only depends on the components of the signals $s_i(t)$ and $s_k(t)$ projected onto each of the basis functions.
- $\|\mathbf{s}_i - \mathbf{s}_k\|$ is the Euclidean distance between $\mathbf{s}_i$ and $\mathbf{s}_k$

$$\|\mathbf{s}_i - \mathbf{s}_k\|^2 = \sum_{j=1}^{N} (s_{ij} - s_{kj})^2$$
$$= \int_0^T (s_i(t) - s_k(t))^2 dt$$

Geometric Representation of Signals (7/7)



- Illustrating the geometric representation of signals for the case when $N = 2$ and $M = 3$.

Schwarz Inequality (1/2)

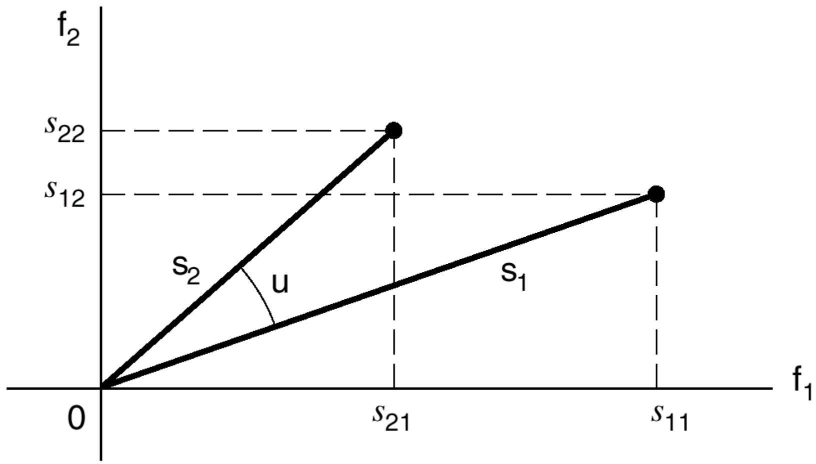
- Consider two energy signals $s_1(t)$ and $s_2(t)$. The Schwarz inequality states that

$$\left(\int_{-\infty}^{\infty} s_1(t)s_2(t)dt\right)^2 \le \left(\int_{-\infty}^{\infty} s_1^2(t)dt\right)\left(\int_{-\infty}^{\infty} s_2^2(t)dt\right)$$

- The equality holds if and only if $s_2(t)=cs_1(t)$, where c is any constant.
- θ is angle subtended between the vector $\mathbf{s}_1$ and $\mathbf{s}_2$,

$$\cos\theta = \frac{\mathbf{s}_1^T\mathbf{s}_2}{\|\mathbf{s}_1\|\,\|\mathbf{s}_2\|} = \frac{\int_{-\infty}^{\infty} s_1(t)s_2(t)dt}{\left(\int_{-\infty}^{\infty} s_1^2(t)dt\right)^{1/2}\left(\int_{-\infty}^{\infty} s_2^2(t)dt\right)^{1/2}}$$

Schwarz Inequality (2/2)



- Vector representations of signals $s_1(t)$ and $s_2(t)$, providing the background picture for proving the Schwarz inequality.
- For complex-valued signals:

$$\left| \int_{-\infty}^{\infty} s_1(t) s_2^*(t) dt \right|^2 \leq \left(\int_{-\infty}^{\infty} |s_1(t)|^2 dt \right) \left(\int_{-\infty}^{\infty} |s_2(t)|^2 dt \right)$$

Gram-Schmidt Orthogonalization Procedure (1/3)

- We need a complete orthonormal set of basis functions.
- Suppose we have a set of *M* energy signals denoted by $s_1(t), s_2(t), \ldots, s_M(t)$. The first basis function is

$$\phi_1(t) = \frac{s_1(t)}{\sqrt{E_1}}$$

- The second basis function is

$$s_{21} = \int_0^T s_2(t)\phi_1(t)dt$$

$$g_2(t) = s_2(t) - s_{21}\phi_1(t)$$

$$\phi_2(t) = \frac{g_2(t)}{\sqrt{\int_0^T g_2^2(t)dt}}$$

Gram-Schmidt Orthogonalization Procedure (2/3)

- For general case:

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

$$s_{ij} = \int_0^T s_i(t)\phi_j(t)dt, \qquad j = 1, 2, \ldots, i-1$$

$$\phi_i(t) = \frac{g_i(t)}{\sqrt{\int_0^T g_i^2(t)dt}}, \qquad i = 1, 2, \ldots, N$$

Gram-Schmidt Orthogonalization Procedure (3/3)

- The dimension N is less than or equal to the number of given signals, M, depending on one of two possibilities:
  - The signals $s_1(t), s_2(t), \ldots, s_M(t)$ form a linearly independent set, in which case $N=M$.
  - The signals $s_1(t), s_2(t), \ldots, s_M(t)$ are not linearly independent, in which case $N<M$, and the intermediate function $g_i(t)$ is zero for $i>N$.
- We do not restrict the Gram-Schmidt orthogonalization procedure to be in terms of sinusoidal functions or sinc functions of time.

Conversion of the Continuous AWGN Channel into a Vector Channel (1/1)

- From Fig. 5.2 and Fig. 5.3(b), we find that the output of correlator *j*, say, is the sample value of a random variable x_j, as shown by

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

$$= s_{ij} + w_j, \qquad j = 1, 2, \ldots, N$$

$$s_{ij} = \int_0^T s_i(t)\phi_j(t)dt$$

$$w_j = \int_0^T w(t)\phi_j(t)dt$$

- We may express the received signal as

$$x(t) = \sum_{j=1}^{N} x_j\phi_j(t) + x'(t)$$

Statistical Characterization of the Correlator Outputs (1/5)

- Let *X(t)* denote the random process, a sample function of which is represented by the received signal *x(t)*.
- Let W_j denote the random variable represented by the sample value w_j produced by the *j*th correlator in response to the white Gaussian noise component *w(t)*. The random variable W_j has zero mean, because the noise process *W(t)* represented by *w(t)* in the AWGN model of Figure 5.2 has zero mean by definition.

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

Statistical Characterization of the Correlator Outputs (2/5)

- We have

$$\sigma^2_{X_j} = E\left[\int_0^T W(t)\phi_j(t)dt \int_0^T W(u)\phi_j(u)du\right]$$
$$= E\left[\int_0^T \int_0^T \phi_j(t)\phi_j(u)W(t)W(u)dtdu\right]$$

- Interchanging the order of integration and expectation

$$\sigma^2_{X_j} = \int_0^T \int_0^T \phi_j(t)\phi_j(u)E[W(t)W(u)]dtdu$$
$$= \int_0^T \int_0^T \phi_j(t)\phi_j(u)R_W(t, u)dtdu$$

- $R_W(t,u)$ is the autocorrelation function of the noise process $W(t)$.

Statistical Characterization of the Correlator Outputs (3/5)

- Since the power spectral density of $W(t)$ is $N_0/2$, the Fourier transform of $N_0/2 \cdot \delta(t)$ is $N_0/2$. We have

$$R_W(t, u) = \frac{N_0}{2}\,\delta(t-u)$$

$$\sigma_{X_j}^2 = \frac{N_0}{2}\int_0^T\int_0^T \phi_j(t)\phi_j(u)\,\delta(t-u)\,dt\,du$$

$$= \frac{N_0}{2}\int_0^T \phi_j^2(t)\,dt$$

$$\sigma_{X_j}^2 = \frac{N_0}{2} \qquad \text{for all } j$$

- All the correlator outputs denoted by X_j with $j=1,2,\ldots,N$, have a variance equal to the power spectral density $N_0/2$ of the noise process $W(t)$.

Statistical Characterization of the Correlator Outputs (4/5)

- The correlator outputs X_j are mutually uncorrelated

$$
\begin{aligned}
\operatorname{cov}[X_j X_k] &= E[(X_j - \mu_{X_j})(X_k - \mu_{X_k})] \\
&= E[(X_j - s_{ij})(X_k - s_{ik})] \\
&= E[W_j W_k] \\
&= E\left[\int_0^T W(t)\phi_j(t)dt \int_0^T W(u)\phi_k(u)du\right] \\
&= \int_0^T \int_0^T \phi_j(t)\phi_k(u)R_W(t, u)dtdu \\
&= \frac{N_0}{2}\int_0^T \int_0^T \phi_j(t)\phi_k(u)\ \delta(t - u)dtdu \\
&= \frac{N_0}{2}\int_0^T \phi_j(t)\phi_k(t)dt \\
&= 0, \qquad j \neq k
\end{aligned}
$$

Statistical Characterization of the Correlator Outputs (5/5)

- Theorem of irrelevance:

  As signal detection in additive white Gaussian noise is concerned, only the projections of the noise onto the basis functions of the signal set $\{s_1(t), s_2(t), \ldots, s_M(t)\}$ affects the sufficient statistics (statistical inference 統計推論) of the detection problem; the remainder of the noise is irrelevant.
- Sufficient statistic: In statistics, a sufficient statistic for a statistical model is relevant to statisical inference within the content of the model. (http://www.scholarpedia.org/article/Sufficient_statistic)

Likelihood Functions (1/2)

- Define

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$$

whose elements are independent Gaussian random variables with mean values equal to s_{ij} and variances equal to $N_0/2$. N denotes number of basis functions. $j=1,\ldots,N$

- Given m_i is transmitted, **x** and x_j are sample values of the random vector **X** and random variable X_j .

$$f_{\mathbf{X}}(\mathbf{x}|m_i) = \prod_{j=1}^{N} f_{X_j}(x_j|m_i), \qquad i = 1, 2, \ldots, M$$

Likelihood Functions (2/2)

$$f_{X_j}(x_j|m_i) = \frac{1}{\sqrt{\pi N_0}} \exp\left[-\frac{1}{N_0}(x_j - s_{ij})^2\right], \qquad \begin{matrix} j = 1, 2, \ldots, N \\ i = 1, 2, \ldots, M \end{matrix}$$

$$f_{\mathbf{X}}(\mathbf{x}|m_i) = (\pi N_0)^{-N/2} \exp\left[-\frac{1}{N_0}\sum_{j=1}^{N}(x_j - s_{ij})^2\right], \qquad i = 1, 2, \ldots, M$$

- Likelihood function:

$$L(m_i) = f_{\mathbf{X}}(\mathbf{x}|m_i), \qquad i = 1, 2, \ldots, M$$

- Log-likelihood function:

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

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

Coherent Detection of Signals in Noise: Maximum Likelihood Decoding (1/8)

- Suppose that in each time slot of duration T seconds, one of the M possible signals $s_1(t)$, $s_2(t),\ldots,s_M(t)$ is transmitted with equal probability, $1/M$. For geometric signals representation, the signal $s_i(t)$, $i=1,2,\ldots,M$, is applied to a bank of correlators, with a common input and supplied with an appropriate set of N orthonormal basis functions.
- The set of message points corresponding to the set of transmitted signals $\{s_i(t)\}_{i=1}^{M}$ is called signal constellation.

Coherent Detection of Signals in Noise: Maximum Likelihood Decoding (2/8)

- The vector **x** differs from the signal vector s_i by the noise **w** whose orientation is completely random. The noise vector **w** is completely characterized by the noise *w(t)*.
- Signal detection problem:

  Given the observation vector **x**, perform a mapping from **x** to an estimate $\hat{m}$ of the transmitted symbol, m_i, in a way that would minimize the probability of error in the decision-making process.
- The probability of error in this decision is

$$
\begin{aligned}
P_e(m_i|\mathbf{x}) &= P(m_i \text{ not sent}|\mathbf{x}) \\
&= 1 - P(m_i \text{ sent}|\mathbf{x})
\end{aligned}
$$

Coherent Detection of Signals in Noise: Maximum Likelihood Decoding (3/8)

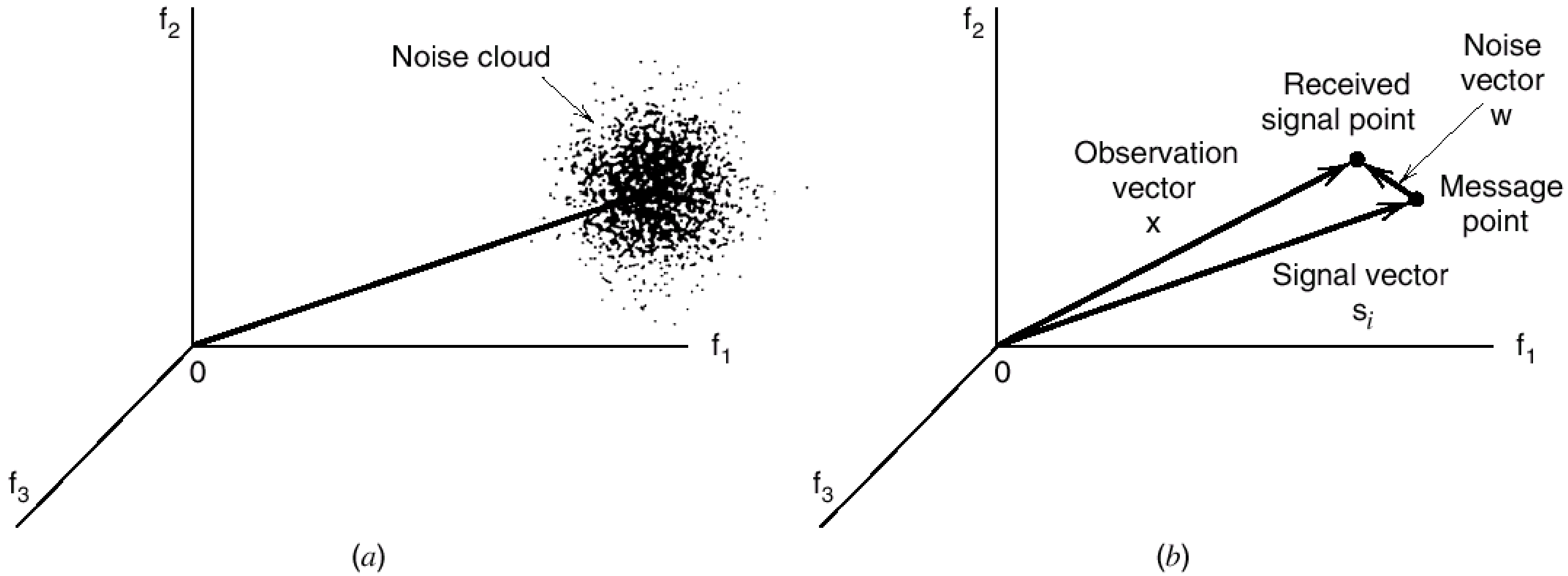
- The decision-making criterion is to minimize the probability of error in mapping each given observation vector **x** into a decision. The optimum decision rule is

$$\text{Set } \hat{m} = m_i \text{ if}$$
$$P(m_i \text{ sent} \mid \mathbf{x}) \geq P(m_k \text{ sent} \mid \mathbf{x}) \qquad \text{for all } k \neq i$$

- k=1,2,…,M. The decision rule is referred to as the maximum *a posteriori* probability (MAP) rule.
- An equivalent MAP rule statement:

$$\text{Set } \hat{m} = m_i \text{ if}$$
$$\frac{p_k f_{\mathbf{X}}(\mathbf{x} \mid m_k)}{f_{\mathbf{X}}(\mathbf{x})} \text{ is maximum for } k = i$$

Coherent Detection of Signals in Noise: Maximum Likelihood Decoding (4/8)



- Illustrating the effect of noise perturbation, depicted in (a), on the location of the received point, depicted in (b).

Coherent Detection of Signals in Noise: Maximum Likelihood Decoding (5/8)

- If the following three conditions are satisfied, MAP rule is equivalent to maximum likelihood (ML) rule.

1. $f_{\mathbf{X}}(\mathbf{x})$ is independent of the transmitted symbol
2. The a priori probability $p_k = p_i$ when all the source symbols are transmitted with equal probability
3. $f_{\mathbf{X}}(\mathbf{x}/m_k)$ bears a one-to-one relationship to the log-likelihood function $l(m_k)$

- Maximum likelihood rule:

$$\text{Set } \hat{m} = m_i \text{ if}$$
$$l(m_k) \text{ is maximum for } k = i$$

- Note: $l(m_i)$=log $f_{\mathbf{X}}(\mathbf{x}/m_i)$

Coherent Detection of Signals in Noise: Maximum Likelihood Decoding (6/8)

- The maximum likelihood decoder differs from the maximum *a posteriori* decoder in that it assumes equally likely message symbols.
- The total observation space Z is correspondingly partitioned into M-decision regions, denoted by Z_1, Z_2,…,Z_M. Maximum likelihood rule can be rewritten as

Observation vector $\mathbf{x}$ lies in region Z_i if

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

, or

Observation vector $\mathbf{x}$ lies in region Z_i if

$$\sum_{j=1}^{N} (x_j - s_{kj})^2 \text{ is minimum for } k = i$$

Coherent Detection of Signals in Noise: Maximum Likelihood Decoding (7/8)

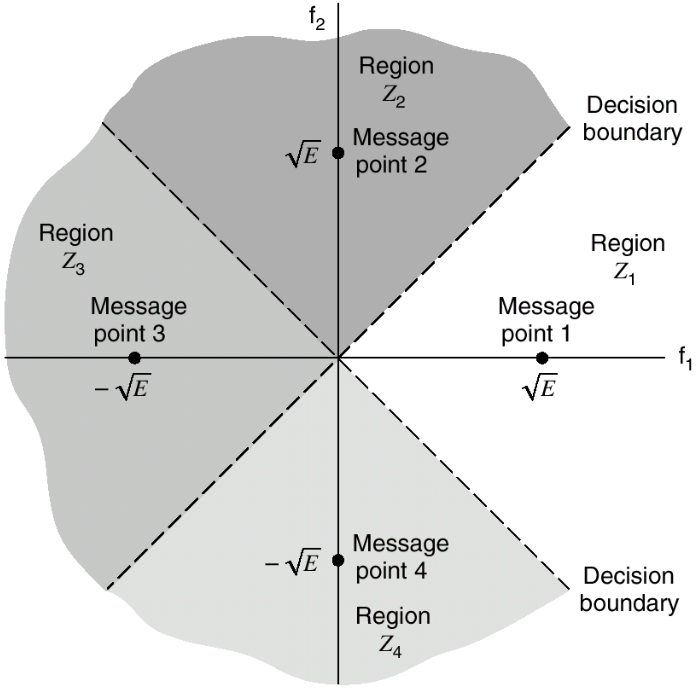
- Or, equivalently,

  Observation vector $\mathbf{x}$ lies in region Z_i if
  the Euclidean distance $\| \mathbf{x} - \mathbf{s}_k \|$ is minimum for $k = i$

- The maximum likelihood decision rule is simply to choose the message point closest to the received signal point.

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

Coherent Detection of Signals in Noise: Maximum Likelihood Decoding (8/8)

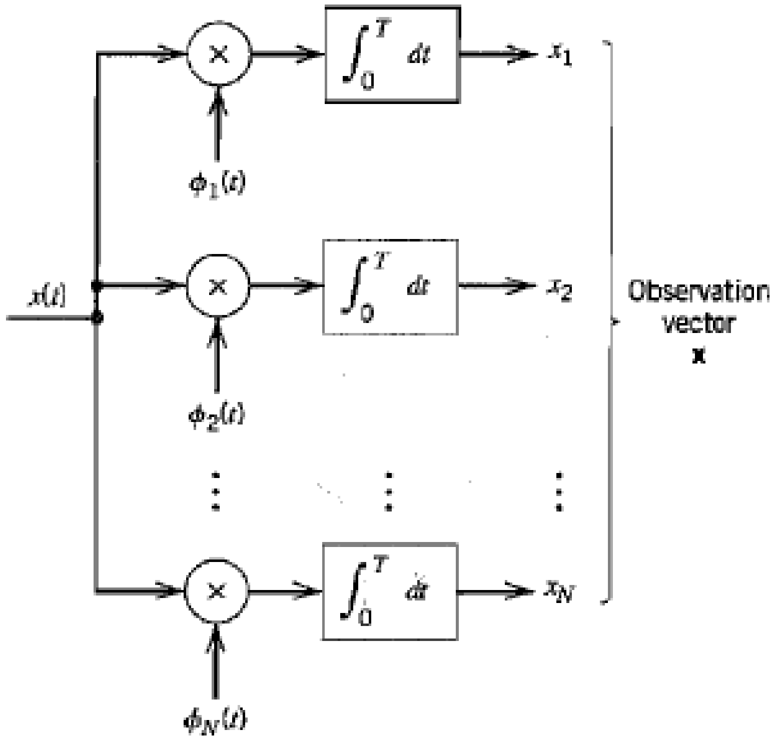


- Illustrating the partitioning of the observation space into decision regions for the case when N=2 and M=4; it is assumed that the M transmitted symbols are equally likely.

Correlation Receiver (1/4)

- A kind of optimum receiver
- The correlation receiver consists of two parts:
  the detector part and the signal transmission decoder part
- The detector part consists of a bank of M product-integrators or correlators.
- The signal transmission decoder is implemented in the form of a maximum-likelihood decoder that operates on the observation vector **x** to produce an estimate, $\hat{m}$, of the transmitted symbol m_i, in a way that would minimize the average probability of symbol error.

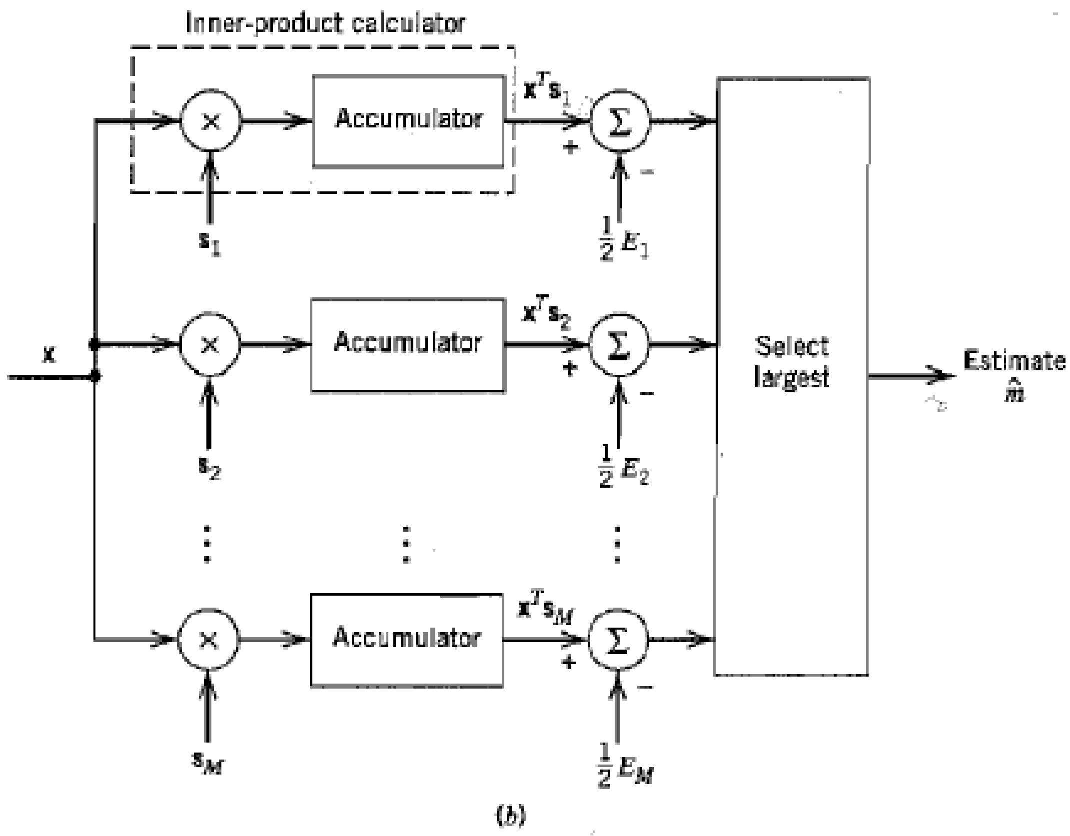
Correlation Receiver (2/4)



(a)

- Fig. 5.9 (a) Detector or demodulator.

Correlation Receiver (3/4)



- Fig. 5.9 (b) Signal transmission decoder.

Correlation Receiver (4/4)

- The optimum receiver of Fig. 5.9 is commonly referred to as a correlation receiver.

Equivalence of Correlation and Matched Filter Receivers (1/3)

- To demonstrate the equivalence of a correlator and a matched filter, consider a linear time-invariant filter with impulse response $h_j(t)$.
- Note that the impulse response $h_j(t)$ of a linear time-invariant filter matched to an input signal $\Phi_j(t)$ is a time-reversed and delayed version of the input $\Phi_j(t)$.

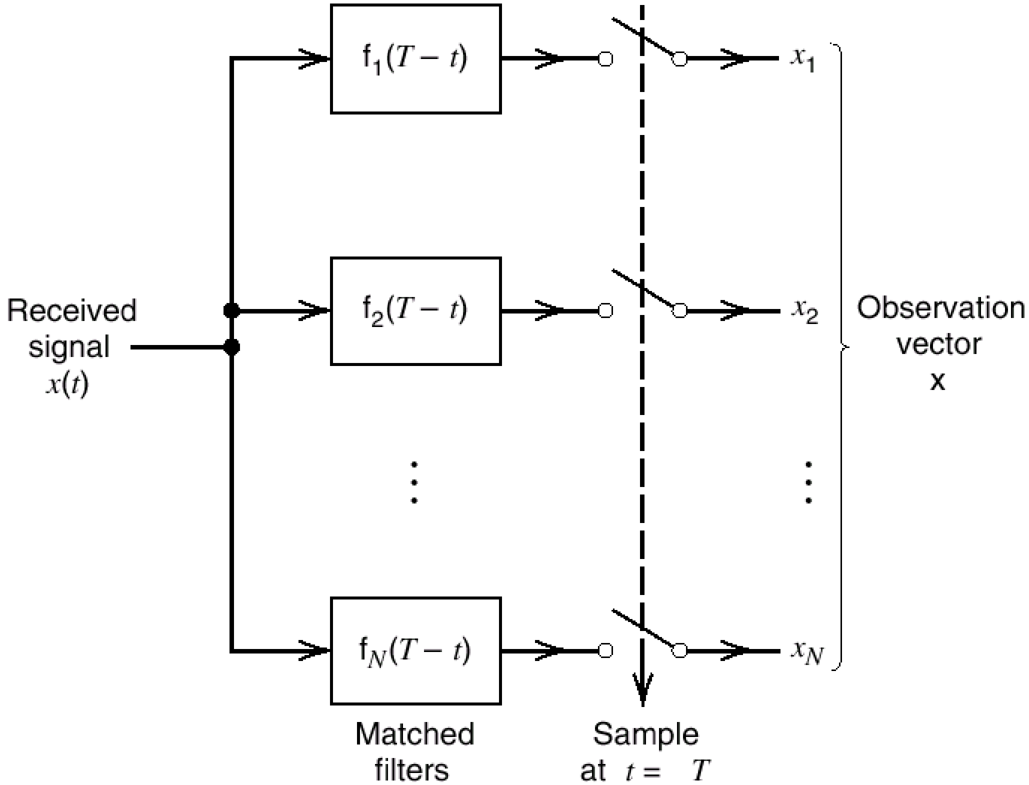
$$y_j(t) = \int_{-\infty}^{\infty} x(\tau)h_j(t - \tau)d\tau$$

$$h_j(t) = \phi_j(T - t)$$

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

$$y_j(T) = \int_{-\infty}^{\infty} x(\tau)\phi_j(\tau)d\tau \qquad y_j(T) = \int_{0}^{T} x(\tau)\phi_j(\tau)d\tau$$

Equivalence of Correlation and Matched Filter Receivers (2/3)



- Fig. 5.10 Detector part of matched filter receiver; the signal transmission decoder is as shown in Fig. 5.9b

Equivalence of Correlation and Matched Filter Receivers (3/3)

- The output of each correlator in Fig. 5.9a is equivalent to the output of a corresponding matched filter in Fig. 5.10 only when that output is sampled at time $t=T$.

Probability of Error (1/1)

$$
\begin{aligned}
P_e &= \sum_{i=1}^{M} p_i \, P(\mathbf{x} \text{ does not lie in } Z_i \,|\, m_i \text{ sent}) \\
&= \frac{1}{M} \sum_{i=1}^{M} P(\mathbf{x} \text{ does not lie in } Z_i \,|\, m_i \text{ sent}) \\
&= 1 - \frac{1}{M} \sum_{i=1}^{M} P(\mathbf{x} \text{ lies in } Z_i \,|\, m_i \text{ sent})
\end{aligned}
$$

$$
P_e = 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{Z_i} f_{\mathbf{X}}(\mathbf{x}\,|\,m_i) \, d\mathbf{x}
$$

Invariance of the Probability of Error to Rotation and Translation (1/5)

- Change in the orientation of the signal constellation with respect to both the coordinate axes and origin of the signal space do not affect the probability of symbol error P_e.
- In maximum likelihood detection, the probability of symbol error P_e depends solely on the relative Euclidean distances between the message points in the constellation.
- The additive white Gaussian noise is spherically symmetric in all directions in the signal space.

Invariance of the Probability of Error to Rotation and Translation (2/5)

- The effect of a rotation applied to all the message points in a constellation is equivalent to multiplying the N-dimensional signal vector $\mathbf{s_i}$ by an *N*-by-*N* orthonormal matrix denoted by **Q** for all *i*. The matrix **Q** satisfies the condition

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$$

  where $\mathbf{I}$ is the identity matrix.

- **w**: *N*-by-1 noise vector

$$\mathbf{w}_{\text{rotate}} = \mathbf{Q}\mathbf{w}$$

Invariance of the Probability of Error to Rotation and Translation (3/5)

- We have

$$
\begin{aligned}
E[\mathbf{w}_{\text{rotate}}] &= E[\mathbf{Q}\mathbf{w}] \\
&= \mathbf{Q}E[\mathbf{w}] \\
&= \mathbf{0} \\
E[\mathbf{w}\mathbf{w}^T] &= \frac{N_0}{2}\mathbf{I} \\
E[\mathbf{w}_{\text{rotate}}\mathbf{w}_{\text{rotate}}^T] &= E[\mathbf{Q}\mathbf{w}(\mathbf{Q}\mathbf{w})^T] \\
&= E[\mathbf{Q}\mathbf{w}\mathbf{w}^T\mathbf{Q}^T] \\
&= \mathbf{Q}E[\mathbf{w}\mathbf{w}^T]\mathbf{Q}^T \\
&= \frac{N_0}{2}\mathbf{Q}\mathbf{Q}^T \\
&= \frac{N_0}{2}\mathbf{I}
\end{aligned}
$$

Invariance of the Probability of Error to Rotation and Translation (4/5)

- We have

$$\| \mathbf{x}_{\text{rotate}} - \mathbf{s}_{i,\text{rotate}} \| = \| \mathbf{x} - \mathbf{s}_i \| \qquad \text{for all } i$$

- Principle of rotational invariance:

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

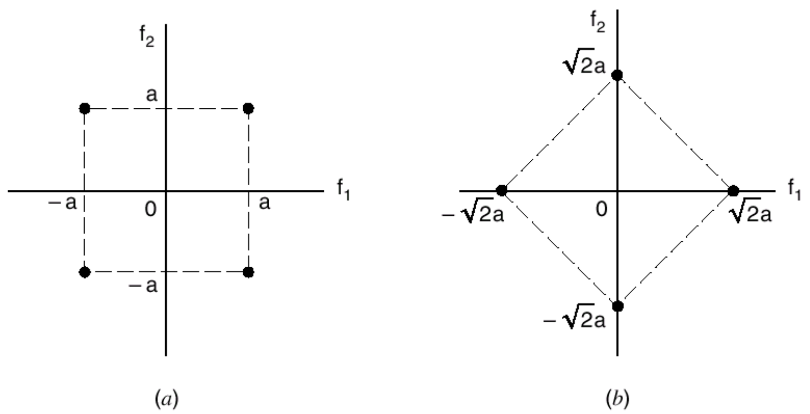
$$\mathbf{s}_{i,\text{rotate}} = \mathbf{Q}\mathbf{s}_i, \qquad i = 1, 2, \ldots, M$$

  where Q is an orthonormal matrix, when the probability of symbol error Pe incurred in maximum likelihood signal detection over an AWGN channel is completely unchanged.

Invariance of the Probability of Error to Rotation and Translation (5/5)

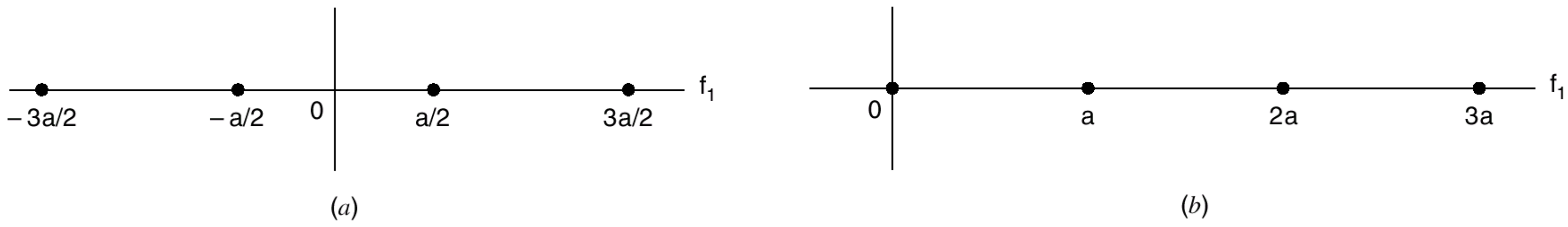
- Principle of translational invariance:

  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.



- Fig. 5.11 A pair of signal constellations for illustrating the principle of rotational invariance.

Minimum Energy Signals (1/2)



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

$$\mathcal{E}_{\text{translate}} = \sum_{i=1}^{M} \| \mathbf{s}_i - \mathbf{a} \|^2 p_i$$

$$\| \mathbf{s}_i - \mathbf{a} \|^2 = \| \mathbf{s}_i \|^2 - 2\mathbf{a}^T \mathbf{s}_i + \| \mathbf{a} \|^2$$

$$\mathcal{E}_{\text{translate}} = \sum_{i=1}^{M} \| \mathbf{s}_i \|^2 p_i - 2 \sum_{i=1}^{M} \mathbf{a}^T \mathbf{s}_i p_i + \| \mathbf{a} \|^2 \sum_{i=1}^{M} p_i$$

$$= \mathcal{E} - 2\mathbf{a}^T E[\mathbf{s}] + \| \mathbf{a} \|^2$$

Minimum Energy Signals (2/2)

$$E[\mathbf{s}] = \sum_{i=1}^{M} \mathbf{s}_i p_i$$

$$\mathbf{a}_{\min} = E[\mathbf{s}]$$

$$\mathcal{E}_{\text{translate,min}} = \mathcal{E} - \| \mathbf{a}_{\min} \|^2$$

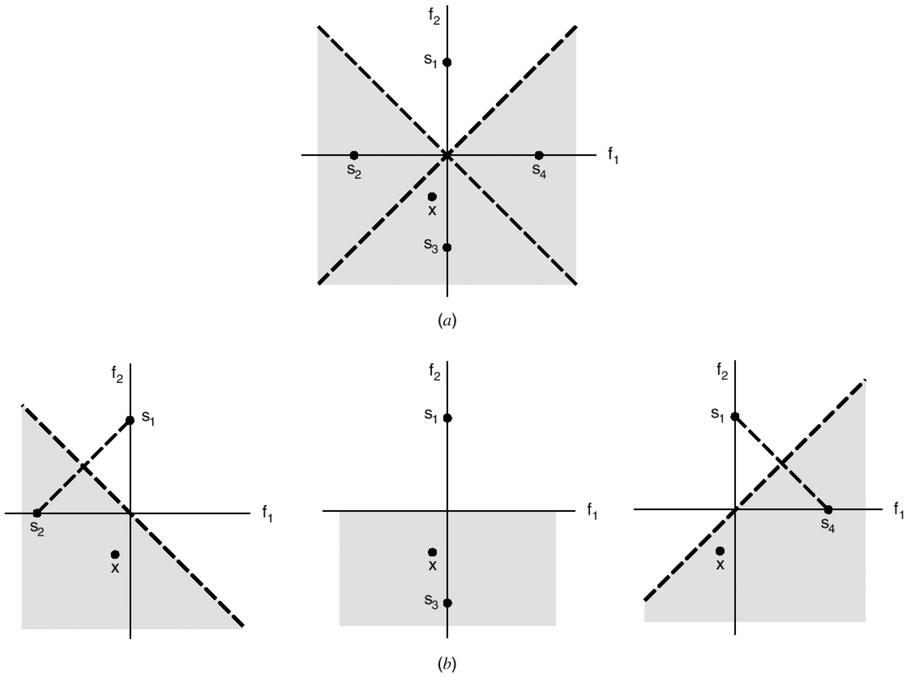
- Procedure to find the minimum energy translate:
  Given a signal constellation $\{\mathbf{s}_i\}_{i=1}^{M}$, the corresponding signal constellation with minimum average energy is obtained by subtracting from each signal vector $\mathbf{s}_i$, in the given constellation an amount equal to the constant vector E[**s**].

Union Bound on the Probability of Error (1/4)

- Let A_{ik}, with (i,k)=1,2,…,M, denote the event that the observation vector **x** is closer to the signal vector $\mathbf{s}_k$ than to $\mathbf{s}_i$, when the symbol m_i (vector $\mathbf{s}_i$) is sent. The conditional probability of symbol error when symbol mi is sent, $P_e(m_i)$, is equal to the probability of the union of events, A_{i1}, A_{i2},…,$A_{i,i-1}$,$A_{i,i+1}$,…,$A_{i,M}$. From probability theory we know that the probability of a finite union of events is overbound by the sum

$$P_e(m_i) \le \sum_{\substack{k=1 \\ k \ne i}}^{M} P(A_{ik}), \qquad i = 1, 2, \ldots, M$$

Union Bound on the Probability of Error (2/4)



- 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.

Union Bound on the Probability of Error (3/4)

- White Gaussian noise is identically distributed along any set of orthogonal axes.
- The corresponding decision boundary is represented by the bisector that is perpendicular to the line joining the points $\mathbf{s}_i$ and $\mathbf{s}_k$.

$$P_2(\mathbf{s}_i, \mathbf{s}_k) = P(\mathbf{x} \text{ is closer to } \mathbf{s}_k \text{ than } \mathbf{s}_i, \text{ when } \mathbf{s}_i \text{ is sent})$$

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

$$d_{ik} = \| \mathbf{s}_i - \mathbf{s}_k \|$$

$$\operatorname{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^{\infty} \exp(-z^2)\, dz$$

$$P_2(\mathbf{s}_i, \mathbf{s}_k) = \frac{1}{2} \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right)$$

Union Bound on the Probability of Error (4/4)

$$P_e(m_i) \le \sum_{\substack{k=1 \\ k \ne i}}^{M} P_2(\mathbf{s}_i, \mathbf{s}_k), \qquad i = 1, 2, \ldots, M$$

$$P_e(m_i) \le \frac{1}{2} \sum_{\substack{k=1 \\ k \ne i}}^{M} \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right), \qquad i = 1, 2, \ldots, M$$

$$P_e = \sum_{i=1}^{M} p_i P_e(m_i)$$

$$\le \frac{1}{2} \sum_{i=1}^{M} \sum_{\substack{k=1 \\ k \ne i}}^{M} p_i \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right)$$

- p_i is the probability of transmitting symbol m_i

Bit Versus Symbol Error Probability (1/1)

- For Gray code with symbol error, the most probable number of bit errors is one.

$$P_e = P\left(\bigcup_{i=1}^{\log_2 M} \{i\text{th bit is in error}\}\right)$$

$$\leq \sum_{i=1}^{\log_2 M} P(i\text{th bit is in error})$$

$$= \log_2 M \cdot (\text{BER})$$

$$P_e \geq P(i\text{th bit is in error}) = \text{BER}$$

$$\frac{P_e}{\log_2 M} \leq \text{BER} \leq P_e$$

Exercise

5.3, 5.5, 5.9, 5.10, 5.11, 5.12, 5.13, 5.17, 5.19