Unit 4, DC, 4th Sem

This Unit has 3 topics only:

1. Maximum Likehood Sequence Detection (Notes Attached)

2. Equalization — Already Covered in Unit 2

3. Synchronization & Carrier Recovery — Already Covered in Unit 3

Note: These 03 Topics are very important. Always Come in the examination.

5.5 Coherent Detection of Signals in Noise: Maximum Likelihood Decoding

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 signal 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 resulting correlator outputs define the *signal vector* $\mathbf{s}_i$. Since knowledge of the signal vector $\mathbf{s}_i$ is as good as knowing the transmitted signal $s_i(t)$ itself, and vice versa, we may represent $s_i(t)$ by a point in a Euclidean space of dimension $N \leq M$. We refer to this point as the *trans-*

Simon Haykins

mitted signal point or *message point*. The set of message points corresponding to the set of transmitted signals $\{s_i(t)\}_{i=1}^{M}$ is called a *signal constellation*.

However, the representation of the received signal $x(t)$ is complicated by the presence of additive noise $w(t)$. We note that when the received signal $x(t)$ is applied to the bank of N correlators, the correlator outputs define the observation vector $\mathbf{x}$. From Equation (5.48), the vector $\mathbf{x}$ differs from the signal vector $\mathbf{s}_i$ by the *noise vector* $\mathbf{w}$ whose orientation is completely random. The noise vector $\mathbf{w}$ is completely characterized by the noise $w(t)$; the converse of this statement, however, is not true. The noise vector $\mathbf{w}$ represents that portion of the noise $w(t)$ that will interfere with the detection process; the remaining portion of this noise, denoted by $w'(t)$, is tuned out by the bank of correlators.

Now, based on the observation vector $\mathbf{x}$, we may represent the received signal $x(t)$ by a point in the same Euclidean space used to represent the transmitted signal. We refer to this second point as the *received signal point*. The received signal point wanders about the message point in a completely random fashion, in the sense that it may lie anywhere inside a Gaussian-distributed "cloud" centered on the message point. This is illustrated in Figure 5.7*a* for the case of a three-dimensional signal space. For a particular realization of the noise vector $\mathbf{w}$ (i.e., a particular point inside the random cloud of Figure 5.7*a*), the relationship between the observation vector $\mathbf{x}$ and the signal vector $\mathbf{s}_i$ is as illustrated in Figure 5.7*b*.

We are now ready to state the signal detection problem:

> Given the observation vector $\mathbf{x}$, perform a mapping from $\mathbf{x}$ 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.

Suppose that, given the observation vector $\mathbf{x}$, we make the decision $\hat{m} = m_i$. The probability of error in this decision, which we denote by $P_e(m_i|\mathbf{x})$, is simply

$$P_e(m_i|\,\mathbf{x}) = P(m_i \text{ not sent}\,|\,\mathbf{x}) = 1 - P(m_i \text{ sent}\,|\,\mathbf{x}) \tag{5.52}$$

The decision-making criterion is to minimize the probability of error in mapping each given observation vector $\mathbf{x}$ into a decision. On the basis of Equation (5.52), we may therefore state the *optimum decision rule*:

$$\text{Set } \hat{m} = m_i \text{ if} \qquad P(m_i \text{ sent}\,|\,\mathbf{x}) \geq P(m_k \text{ sent}\,|\,\mathbf{x}) \quad \text{for all } k \neq i \tag{5.53}$$

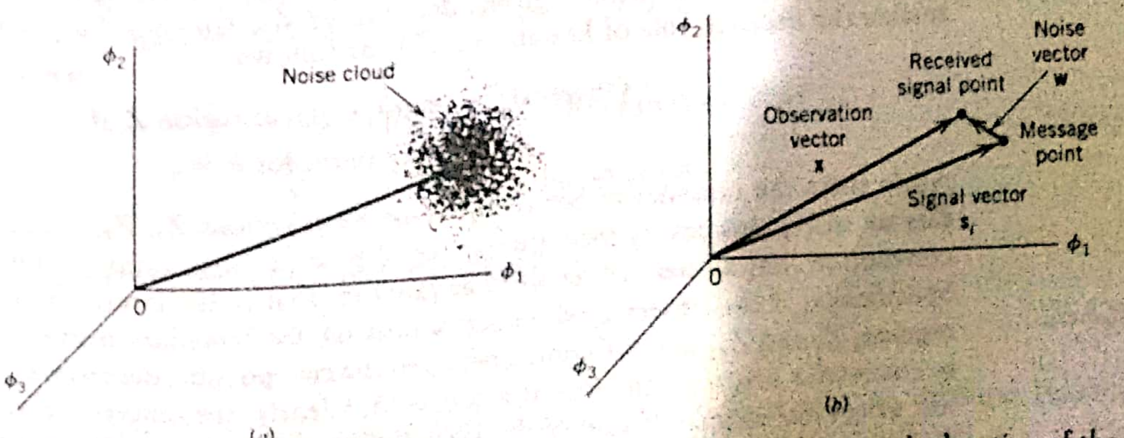

FIGURE 5.7 Illustrating the effect of noise perturbation, depicted in (*a*), on the location of the received signal point, depicted in (*b*).

where $k = 1, 2, \ldots, M$. This decision rule is referred to as the *maximum a posteriori probability (MAP) rule.*

The condition of Equation (5.53) may be expressed more explicitly in terms of the *a priori* probabilities of the transmitted signals and in terms of the likelihood functions. Using Bayes' rule in Equation (5.53), and for the moment ignoring possible ties in the decision-making process, we may restate the MAP rule as follows:

$$\text{Set } \hat{m} = m_i \text{ if}$$
$$\frac{p_k f_{\mathbf{X}}(\mathbf{x}|m_k)}{f_{\mathbf{X}}(\mathbf{x})} \text{ is maximum for } k = i \tag{5.54}$$

where p_k is the *a priori* probability of transmitting symbol m_k, $f_{\mathbf{X}}(\mathbf{x}|m_k)$ is the conditional probability density function of the random observation vector $\mathbf{X}$ given the transmission of symbol m_k, and $f_{\mathbf{X}}(\mathbf{x})$ is the unconditional probability density function of $\mathbf{X}$. In Equation (5.54) we may note the following:

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

Accordingly, we may restate the decision rule of Equation (5.54) in terms of $l(m_k)$ simply as follows:

$$\text{Set } \hat{m} = m_i \text{ if}$$
$$l(m_k) \text{ is maximum for } k = i \tag{5.55}$$

This decision rule is referred to as the *maximum likelihood rule*, and the device for its implementation is correspondingly referred to as the *maximum likelihood decoder*. According to Equation (5.55), a maximum likelihood decoder computes the log-likelihood functions as metrics for all the M possible message symbols, compares them, and then decides in favor of the maximum. Thus the maximum likelihood decoder differs from the maximum *a posteriori* decoder in that it assumes equally likely message symbols.

It is useful to have a graphical interpretation of the maximum likelihood decision rule. Let Z denote the N-dimensional space of all possible observation vectors $\mathbf{x}$. We refer to this space as the *observation space*. Because we have assumed that the decision rule must say $\hat{m} = m_i$, where $i = 1, 2, \ldots, M$, the total observation space Z is correspondingly partitioned into *M-decision regions*, denoted by $Z_1, Z_2, \ldots, Z_M$. Accordingly, we may restate the decision rule of Equation (5.55) as follows:

$$\text{Observation vector } \mathbf{x} \text{ lies in region } Z_i \text{ if}$$
$$l(m_k) \text{ is maximum for } k = i \tag{5.56}$$

Aside from the boundaries between the decision regions $Z_1, Z_2, \ldots, Z_M$, it is clear that this set of regions covers the entire space of possible observation vectors $\mathbf{x}$. We adopt the convention that all ties are resolved at random; that is, the receiver simply makes a guess. Specifically, if the observation vector $\mathbf{x}$ falls on the boundary between any two decision regions, Z_i and Z_k, say, the choice between the two possible decisions $\hat{m} = m_i$ and $\hat{m} = m_k$ is resolved *a priori* by the flip of a fair coin. Clearly, the outcome of such an event does not affect the ultimate value of the probability of error since, on this boundary, the condition of Equation (5.53) is satisfied with the equality sign.

The maximum likelihood decision rule of Equation (5.55) or its geometric counterpart described in Equation (5.56) is of a generic kind, with the channel noise $w(t)$ being additive as the only restriction imposed on it. We next specialize this rule for the case when $w(t)$ is both white and Gaussian.

From the log-likelihood function defined in Equation (5.51) for an AWGN channel we note that $l(m_k)$ attains its maximum value when the summation term

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

is minimized by the choice $k = i$. Accordingly, we may formulate the maximum likelihood decision rule for an AWGN channel as

$$\begin{gathered}\text{Observation vector } \mathbf{x} \text{ lies in region } Z_i \text{ if} \\ \sum_{j=1}^{N} (x_j - s_{kj})^2 \text{ is minimum for } k = i\end{gathered} \tag{5.57}$$

Next, we note from our earlier discussion that (see Equation (5.14) for comparison)

$$\sum_{j=1}^{N} (x_j - s_{kj})^2 = \| \mathbf{x} - \mathbf{s}_k \|^2 \tag{5.58}$$

where $\| \mathbf{x} - \mathbf{s}_k \|$ is the Euclidean distance between the received signal point and message point, represented by the vectors $\mathbf{x}$ and $\mathbf{s}_k$, respectively. Accordingly, we may restate the decision rule of Equation (5.57) as follows:

$$\begin{gathered}\text{Observation vector } \mathbf{x} \text{ lies in region } Z_i \text{ if} \\ \text{the Euclidean distance } \| \mathbf{x} - \mathbf{s}_k \| \text{ is minimum for } k = i\end{gathered} \tag{5.59}$$

Equation (5.59) states that *the maximum likelihood decision rule is simply to choose the message point closest to the received signal point*, which is intuitively satisfying.

In practice, the need for squarers in the decision rule of Equation (5.59) is avoided by recognizing that

$$\sum_{j=1}^{N} (x_j - s_{kj})^2 = \sum_{j=1}^{N} x_j^2 - 2 \sum_{j=1}^{N} x_j s_{kj} + \sum_{j=1}^{N} s_{kj}^2 \tag{5.60}$$

The first summation term of this expansion is independent of the index k and may therefore be ignored. The second summation term is the inner product of the observation vector $\mathbf{x}$ and signal vector $\mathbf{s}_k$. The third summation term is the energy of the transmitted signal $s_k(t)$. Accordingly, we may formulate a decision rule equivalent to that of Equation (5.59) as follows:

$$\begin{gathered}\text{Observation vector } \mathbf{x} \text{ lies in region } Z_i \text{ if} \\ \sum_{j=1}^{N} x_j s_{kj} - \frac{1}{2} E_k \text{ is maximum for } k = i\end{gathered} \tag{5.61}$$

where E_k is the energy of the transmitted signal $s_k(t)$:

$$E_k = \sum_{j=1}^{N} s_{kj}^2 \tag{5.62}$$

From Equation (5.61) we deduce that, for an AWGN channel, the decision regions are regions of the N-dimensional observation space Z, bounded by linear [$(N-1)$-dimensional hyperplane] boundaries. Figure 5.8 shows the example of decision regions for

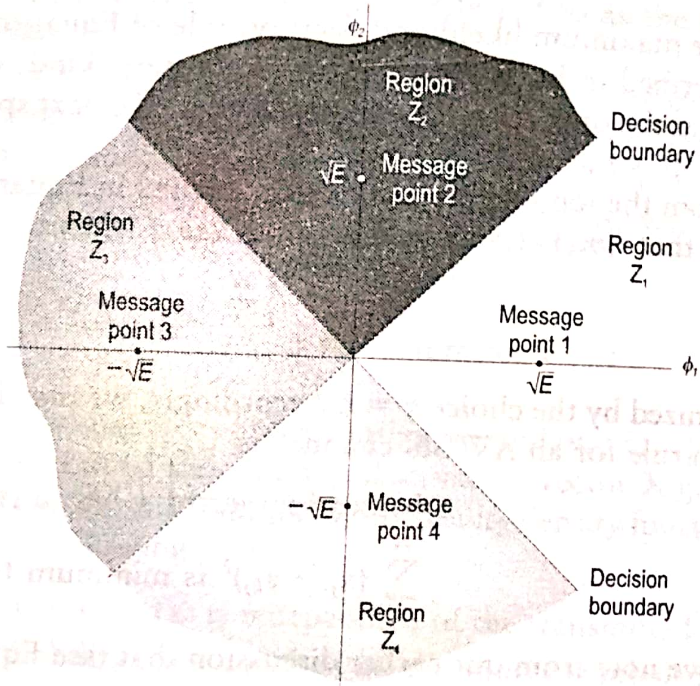

FIGURE 5.8 Illustrating the partitioning of the observation space into decision regions for the case when $N = 2$ and $M = 4$; it is assumed that the M transmitted symbols are equally likely.

$M = 4$ signals and $N = 2$ dimensions, assuming that the signals are transmitted with equal energy, E, and equal probability.