5

Discussion: Detection of Gaussian Signals

In Chapters 2 through 4 we studied the problem of detecting Gaussian signals in the presence of Gaussian noise. In this chapter we first discuss some related topics. We then summarize the major results of our detection theory discussion.

5.1 RELATED TOPICS

5.1.1 M-ary Detection: Gaussian Signals in Noise

All of our discussion in Chapters 2 through 4 dealt with the binary detection problem. In this section we discuss briefly the M-hypothesis problem. The general Gaussian M-ary problem is

\[ r(t) = r_i(t), \quad T_b \leq t \leq T_f; H_i, \quad i = 1, \ldots, M, \] (1)

where

\[ E[r_i(t) \mid H_i] = m_i(t) \] (2)

and

\[ E[(r_i(t) - m_i(t))(r_i(u) - m_i(u)) \mid H_i] = K_{H_i}(t, u). \] (3)

Most of the ideas from the binary case carry over to the M-ary case with suitable modifications. As an illustration we consider a special case of the general problem.

The problem of interest is described by the following model. The received waveforms on the M hypotheses are

\[ r(t) = s_i(t) + n(t), \quad T_b \leq t \leq T_f; H_i, \quad i = 1, 2, \ldots, M. \] (4)

The additive noise \( n(t) \) is a sample function from a zero-mean Gaussian process with covariance function \( K_n(t, u) \). A white noise term is not
necessarily present. The signal processes are sample functions from Gaussian processes and are statistically independent of the noise process. The signal processes are characterized by

\[ E[s_i(t)] = m_i(t), \quad T_b \leq t \leq T_f, \quad i = 1, \ldots, M \]  

(5)

and

\[ E\{[s_i(t) - m_i(t)][s_i(u) - m_i(u)]\} = K_{s_i}(t, u), \quad T_b \leq t, u \leq T_f, \quad i = 1, \ldots, M. \]  

(6)

The a-priori probability of the \( i \)th hypothesis is \( P_i \) and the criterion is minimum \( \text{Pr}(\varepsilon) \). We assume that each pair of hypotheses would lead to a nonsingular binary test. The derivation of the optimum receiver is similar to the derivation for the binary case, and so we shall simply state the results. The reader can consult [1]-[3] or Problem 5.1.1 for details of the derivation.

To perform the likelihood ratio test we compute a set of \( M \) sufficient statistics, which we denote by \( l_i, i = 1, \ldots, M \). The first component of the \( i \)th sufficient statistic is

\[ l_{R_i} = \int_{T_b}^{T_f} r(t)h_i(t, u)r(u) \, dt \, du \]  

(7)

where \( h_i(t, u) \) is specified by the integral equation

\[ \int_{T_b}^{T_f} K_n(t, x)h_i(x, y)[K_n(y, u) + K_{s_i}(y, u)] \, dx \, dy = K_{s_i}(t, u), \quad T_b \leq t, u \leq T_f. \]  

(8)

The second component of the \( i \)th sufficient statistic is

\[ l_{D_i} = \int_{T_b}^{T_f} r(t)[g_i(t) - \int_{T_b}^{T_f} h_i(t, u)m_i(u) \, du] \, dt, \]  

(9)

where \( g_i(t) \) is specified by the integral equation

\[ \int_{T_b}^{T_f} K_n(t, u)g_i(u) \, du = m_i(t), \quad T_b \leq t \leq T_f. \]  

(10)

The bias component of the \( i \)th sufficient statistic is

\[ l_{B_i} = -\frac{1}{2} \sum_{k=1}^{\infty} \ln (1 + \lambda_{ik}^*), \]  

(11)

where the \( \lambda_{ik}^* \) are the eigenvalues of the kernel,

\[ K_{s_k}^*(t, u) = \int_{T_b}^{T_f} K_n^{-\frac{1}{4}}(t, x)K_n(x, y)K_n^{-\frac{1}{4}}(u, y) \, dx \, dy. \]  

(12)
The complete $i$th sufficient statistic is

$$I_i = l_{R_i} + l_{D_i} + l_{B_i}, \quad i = 1, \ldots, M. \tag{13}$$

The test consists of computing

$$l_i + \ln P_i, \quad i = 1, \ldots, M \tag{14}$$

and choosing the largest.

A special case of (4) that occurs frequently is

$$K_n(t, u) = \frac{N_0}{2} \delta(t - u), \tag{15}$$

$$m_i(t) = 0, \quad i = 1, \ldots, M. \tag{16}$$

Then $h_i(t, u)$ satisfies the equation

$$\frac{N_0}{2} h_i(t, u) + \int_{T_b}^{T_f} h_i(t, y) K_{s_i}(y, u) \, dy = K_{s_i}(t, u), \quad T_b \leq t, u \leq T_f, \quad i = 1, \ldots, M. \tag{17}$$

All of the canonical realizations in Chapter 2 are valid for this case. The bias term is

$$l_{B_i} = - \frac{1}{N_0} \int_{T_b}^{T_f} \xi_{P_i} \left( t \mid s_i(\cdot), \frac{N_0}{2} \right) \, dt, \tag{18}$$

where $\xi_{P_i}(t \mid s_i(\cdot),)$ is defined as in (2.137).

The performance calculation for the general $M$-ary case is difficult. We would anticipate this because, even in the known signal case, exact $M$-ary performance calculations are usually not feasible.

An important problem in which we can get accurate bounds is that of digital communication over a Rayleigh channel using $M$-orthogonal signals. The binary version of this problem was discussed in Examples 3 and 4 of Chapter 4 (see pages 111–117). We now indicate the results for the $M$-ary problem.

The transmitted signal on the $i$th hypothesis is

$$s_i(t) = \sqrt{\frac{2E_i}{T}} \sin \omega_i t, \quad T_b \leq t \leq T_f; H_i \tag{19}$$

the signal passes over a fluctuating Rayleigh channel. The received waveform on the $i$th hypothesis is

$$r(t) = s_i(t) + w(t), \quad T_b \leq t \leq T_f; H_i \tag{20}$$

The $i$th signal $s_i(t)$ is a sample function of a bandpass process centered at $\omega_i$, whose spectrum is symmetric around $\omega_i$.† The signal processes

† The symmetric assumption is included to keep the notation simple. After we introduce complex notation in Chapter 9, we can handle the asymmetric case easily.
are essentially disjoint in frequency. The additive noise is a sample function from a zero-mean white Gaussian random process with spectral height $N_0/2$. The low-pass spectra of the signal processes are identical. We denote them by $S_s(\omega)$. The total received power in the signal is

$$P_r = 2\sigma_y^2 \int_{-\infty}^{\infty} S_s(\omega) \frac{d\omega}{2\pi}. \quad (21)$$

Kennedy [4] and Viterbi [5] have studied the performance for this case. Our discussion follows the latter's. Starting from a general result in [6], one can show that

$$\Pr (\epsilon) \leq \exp [\rho TR] \frac{\left[D_{\varphi} \left(\frac{2}{N_0}\right)\right]^\rho}{\left[D_{\varphi} \left(\frac{2\rho}{N_0(1 + \rho)}\right)\right]^{1+\rho}}, \quad 0 \leq \rho \leq 1, \quad (22)$$

where $D_{\varphi}(z)$ is the Fredholm determinant of the low-pass process, $T$ is the length of interval, and

$$R = \ln M \quad (23)$$

is the transmission rate in nats per second. The parameter $\rho$ is used to optimize the bound. When the observation time is long, we can use (1-3.182) to obtain

$$\ln D_{\varphi}(z) = T \int_{-\infty}^{\infty} \ln (1 + zS_s(\omega)) \frac{d\omega}{2\pi}. \quad (24)$$

We now define

$$E_0(\rho) \triangleq (1 + \rho) \int_{-\infty}^{\infty} \ln \left[1 + \frac{\rho}{1 + \rho} \frac{2S_s(\omega)}{N_0}\right] \frac{d\omega}{2\pi}$$

$$- \rho \int_{-\infty}^{\infty} \ln \left[1 + \frac{2S_s(\omega)}{N_0}\right] \frac{d\omega}{2\pi} \quad (25)$$

and

$$E(R) = \max_{0 \leq \rho \leq 1} [E_0(\rho) - \rho R]. \quad (26)$$

Comparing (25) and (4.21), we observe that

$$- \frac{T}{2} \left. \frac{E_0(\rho)}{1 + \rho}\right|_{\rho = (1-s)/z} = \mu_{s,s}(s) \quad (27)$$

and, in particular,

$$-TF_0(1) = 4\mu_{s,s}(1) = 2\mu_{1^*} = \mu_{1^*,1}(1) = \mu_{BS,BP,\infty}(1/2). \quad (28)$$
Using (24)–(26), (22) reduces to

\[ \Pr (\varepsilon) \leq e^{-T E(R)} \quad (29) \]

The final step is to perform the maximization indicated in (26). The result is obtained as follows:

1. If

\[ \hat{E}_0(\rho) \triangleq \frac{\partial E_0(\rho)}{\partial \rho} = R \quad (30) \]

has a solution for \( 0 \leq \rho \leq 1 \), we denote it as \( \rho_m \). Then

\[ E(R) = E_0(\rho_m) - \rho_m \hat{E}_0(\rho_m), \quad (31) \]

and

\[ \hat{E}_0(1) \leq R \leq C_{0R} = \hat{E}_0(0). \quad (32) \]

2. If (30) does not have a solution in the allowable range of \( \rho \), the maximum is at \( \rho = 1 \), and

\[ E(R) = E_0(1) - R, \quad (34) \]

\[ 0 < R < \hat{E}_0(1). \quad (35) \]

The results in (31) and (34) provide the exponents in the \( \Pr (\varepsilon) \) expression. In the problems, we include a number of examples to illustrate the application of these results.

This concludes our brief discussion of the \( M \)-ary problem. For a large class of processes we can find the optimum receiver, but, except for orthogonal signal processes, the performance evaluation is usually difficult.

5.1.2 Suboptimum Receivers

We have been able to find the optimum receiver to implement the likelihood ratio test for a large class of Gaussian signal processes. Frequently, the filters in the receiver are time-varying and may be difficult to implement. This motivates the search for suboptimum receivers, which are simpler to implement than the optimum receiver but perform almost as well as the optimum receiver. To illustrate this idea we consider a simple example.

Example. Consider the simple binary detection example discussed on page 104. The received waveforms on the two hypotheses are

\[ r(t) = s(t) + w(t), \quad T_i \leq t \leq T_f: H_1, \]

\[ r(t) = w(t), \quad T_i \leq t \leq T_f: H_0, \quad (36) \]
where \( w(t) \) is a white Gaussian process with spectral height \( N_0/2 \) and \( s(t) \) is a Gaussian process with spectrum

\[
S_s(\omega) = \frac{2kP}{\omega^2 + k^2}.
\]  

(37)

We saw that the optimum receiver could be implemented as a cascade of a time-varying filter, a square-law device, and an integrator. The difficulty arises in implementing the time-varying filter.

A receiver that is simpler to implement is shown in Fig. 5.1. The structure is the same as the optimum receiver, but the linear filter is time-invariant,

\[
h_{\text{sub}}(\tau) = e^{-\beta \tau} u_{-1}(\tau), \quad -\infty < \tau < \infty.
\]  

(38)

We choose \( \beta \) to optimize the performance. From our results in Section 4.1 (page 104) we know that if

\[
\beta = k \left( 1 + \frac{4P}{kN_0} \right)^{-\frac{1}{2}},
\]  

(39)

then the suboptimum receiver will be essentially optimum for long observation times. For arbitrary observation times, some other choice of \( \beta \) might give better performance. Thus, the problem of interest is to choose \( \beta \) to maximize the performance.

With this example as motivation, we consider the general question of suboptimum receivers. The choice of the structure for the suboptimum receiver is strongly dependent on the particular problem. Usually one takes the structure of the optimum receiver as a starting point, tries various modifications, and analyzes the resulting performance. In this section we discuss the performance of suboptimum receivers.

To motivate our development, we first recall the performance results for the optimum receiver. The optimum receiver computes \( I \), the logarithm of the likelihood ratio, and compares it with a threshold. The error probabilities are

\[
P_F = \Pr (\epsilon | H_0) = \int_{-\infty}^{\infty} p_{\| H_0} (L | H_0) \, dL
\]  

(40)

and

\[
P_M = \Pr (\epsilon | H_1) = \int_{-\infty}^{\gamma} p_{\| H_1} (L | H_0) \, dL.
\]  

(41)

All of our performance discussion in the Gaussian signal problem has been based on \( \mu(s) \), which is defined as

\[
\mu(s) = \ln M_{\| H_0}(s) = \ln \int_{-\infty}^{\infty} e^{sL} p_{\| H_0} (L | H_0) \, dL,
\]  

(42)
the logarithm of the moment-generating function of \( l \), given that \( H_0 \) is true. Since \( l \) is the logarithm of the likelihood ratio, we can also write \( \mu(s) \) in terms of \( M_{l|H_1}(s) \),

\[
\mu(s) = \ln M_{l|H_1}(s - 1) \tag{43}
\]

(see pages I-118–I-119). Thus we can express both \( P_M \) and \( P_F \) in terms of \( \mu(s) \).

A suboptimum receiver computes a test statistic \( l_x \) and compares it with a threshold \( \gamma_x \) in order to make a decision. The statistic \( l_x \) is not equivalent to \( l \) and generally is used because it is easier to compute. For suboptimum receivers, the probability densities of \( l_x \) on \( H_1 \) and \( H_0 \) are not uniquely related, and so we can no longer express \( P_M \) and \( P_F \) in terms of a single function. This forces us to introduce two functions analogous to \( \mu(s) \) and makes the performance calculations more involved.

To analyze the suboptimum receiver, we go through a development parallel to that in Sections I-2.7 and III-2.2. Because the derivation is straightforward, we merely state the results. We define

\[
\mu_0(s) \triangleq \ln M_{l_x|H_0}(s) = \int_{-\infty}^{\infty} e^{s l_{x}} P_{l_x|H_0}(l | H_0) dL 
\]

and

\[
\mu_1(s) \triangleq \ln M_{l_x|H_1}(s) = \int_{-\infty}^{\infty} e^{s l_{x}} P_{l_x|H_1}(l | H_1) dL. 
\]

The Chernoff bounds are

\[
\Pr (\epsilon | H_0) \leq \exp \left[ \mu_0(s_0) - s_0 \gamma \right], \quad s_0 > 0, \tag{46}
\]

\[
\Pr (\epsilon | H_1) \leq \exp \left[ \mu_1(s_1) - s_1 \gamma \right], \quad s_1 < 0, \tag{47}
\]

where

\[
\hat{\mu}_0(s_0) = \gamma, \quad s_0 > 0 \tag{48}
\]

and

\[
\hat{\mu}_1(s_1) = \gamma, \quad s_1 < 0. \tag{49}
\]

The equations (48) and (49) will have a unique solution if

\[
E[l_x | H_0] \leq \gamma \leq E[l_x | H_1]. \tag{50}
\]

The first-order asymptotic approximations are

\[
\Pr (\epsilon | H_0) \simeq \text{erfc}_* (s_0 \sqrt{\hat{\mu}_0(s_0)}) \exp \left[ \mu_0(s_0) - s_0 \hat{\mu}_0(s_0) + \frac{s_0^2}{2} \hat{\mu}_0^2(s_0) \right] \tag{51}
\]

and

\[
\Pr (\epsilon | H_1) \simeq \text{erfc}_* (-s_1 \sqrt{\hat{\mu}_1(s_1)}) \exp \left[ \mu_1(s_1) - s_1 \hat{\mu}_1(s_1) + \frac{s_1^2}{2} \hat{\mu}_1^2(s_1) \right]. \tag{52}
\]
where \( s_0 \) and \( s_1 \) satisfy (48) and (49). Equations (51) and (52) are analogous to (179) and (188). Results similar to (181) and (189) follow easily.

The results in (44)–(52) are applicable to an arbitrary detection problem. To apply them to the general Gaussian problem, we must be able to evaluate \( \mu_0(s) \) and \( \mu_1(s) \) efficiently. The best technique for evaluating \( \mu_0(s) \) and \( \mu_1(s) \) will depend on the structure of the suboptimum receiver. We demonstrate the technique for the general filter-squarer-integrator (FSI) receiver shown in Fig. 5.2. The filter may be time-varying. For this structure the techniques that we developed in Section 2.2 (pages 35–44) are still valid. We illustrate the procedure by finding an expression for \( \mu_1(s) \).

**Calculation of \( \mu_1(s) \) for an FSI Receiver.** To find \( \mu_1(s) \), we expand \( y(t) \), the input to the squarer under \( H_1 \), in a Karhunen-Loève expansion. Thus

\[
y(t) = \sum_{i=1}^{\infty} y_i \phi_{1i}(t), \quad T_i \leq t \leq T_f; H_1,
\]

where the \( \phi_{1i}(t) \) are the eigenfunctions of \( y(t) \) on \( H_1 \). The corresponding eigenvalues are \( \lambda_{1i} \). We assume that the eigenvalues are ordered in magnitude so that \( \lambda_{11} \) is the largest. From (45),

\[
\begin{align*}
\mu_1(s) &= \ln \left\{ \mathbb{E}[e^{\mu_1 s} \mid H_1] \right\} \\
&= \ln \left\{ \mathbb{E}\left[ \exp \left( s \sum_{i=1}^{\infty} y_i^2 \right) \mid H_1 \right] \right\} \\
&= -s \sum_{i=1}^{\infty} \ln (1 - 2s\lambda_{1i}), \quad s < \frac{1}{2\lambda_{11}}. \quad (54)
\end{align*}
\]

The expectation is a special case of Problem I-4.4.2. The sum can be written as a Fredholm determinant,\(^\dagger\)

\[
\begin{align*}
\mu_1(s) &= -\frac{1}{2} \ln D_{\phi\mid H_1}(-2s), \quad s < \frac{1}{2\lambda_{11}}. \quad (55)
\end{align*}
\]

A similar result follows for \( \mu_0(s) \),

\[
\begin{align*}
\mu_0(s) &= -\frac{1}{2} \ln D_{\phi\mid H_0}(-2s), \quad s < \frac{1}{2\lambda_{01}}. \quad (56)
\end{align*}
\]

We now have \( \mu_0(s) \) and \( \mu_1(s) \) expressed in terms of Fredholm determinants. The final step is to evaluate these functions. Three cases in which we can evaluate \( D_{\phi\mid H_i}() \) are the following:

1. Stationary processes, long observation time.
2. Separable kernels.

\(^\dagger\) This result is due to Kac and Siegert [8] (e.g., [9, Chapter 9]).
The procedure for the first two cases is clear. In the third case, we can use the algorithms in section 2.2.1 or section 2.2.3 to evaluate $\mu_0(s)$ and $\mu_1(s)$. The important point to remember is that the state equation that we use to compute $\mu_1(s)$ corresponds to the system that produces $y_1(t)$ when driven by white noise. Similarly, the state equation that we use to compute $\mu_0(s)$ corresponds to the system that produces $y_0(t)$ when driven by white noise.

In this section we have developed the performance expressions needed to analyze suboptimum receivers. Because the results are straightforward modifications of our earlier results, our discussion was brief. The analysis based on these results is important in the implementation of practical receiver configurations. A number of interesting examples are developed in the problems. In Chapter 11, we encounter suboptimum receivers again and discuss them in more detail.

### 5.1.3 Adaptive Receivers

A complete discussion of adaptive receivers would take us too far afield. On the other hand, several simple observations are useful.

All of our discussion of communication systems assumed that we made a decision on each baud. This decision was independent of those made on previous bauds. If the channel process is correlated over several bauds, one should be able to exploit this correlation in order to reduce the probability of error. Since the optimum "single-baud" receiver is an estimator-correlator, a logical approach is to perform a continuous channel estimation and use this to adjust the receiver filters and gains. An easy way to perform the channel estimation is through the use of decision-directed feedback. Here we assume that all past decisions are correct in order to perform the channel estimation. As long as most of the decisions are correct, this reduces the channel estimation problem to that of a "known" signal into an unknown channel. Decision feedback schemes for simple channels have been studied by Proakis and Drouilhet [10]. More complicated systems have been studied by Glaser [11], Jakowitz, Shuey, and White [12], Scudder [13, 14], Boyd [15], and Austin [16]. Another procedure to exploit the correlation of the channel process would be to devote part of the available energy to send a known signal to measure the channel.

There has been a great deal of work done on adaptive systems. In almost all cases, the receivers are so complicated and difficult to analyze that one cannot make many useful general statements. We do feel the reader should recognize that many of these systems are logical extrapolations from the general Gaussian problem we have studied. References that deal with various types of adaptive systems include [17]–[30].
5.1.4 Non-Gaussian Processes

All of our results have dealt with Gaussian random processes. When the processes involved are non-Gaussian, the problems are appreciably more difficult. We shall divide our comments on these problems into four categories:

1. Processes derived from Gaussian processes.
2. Structured non-Gaussian processes.
3. Unspecified non-Gaussian processes.
4. Analysis of fixed receivers.

We shall explain the descriptions in the course of our discussion.

Processes Derived from Gaussian Processes. We have emphasized cases in which the received waveform is conditionally Gaussian. A related class of problems comprises those in which \( r(t) \) is a sample function of a process that can be derived from a Gaussian process. A common case is one in which either the mean-value function or the covariance function contains a random parameter set. In this case, we might have \( m(t, \theta_m) \) and \( K_r(t, u; \theta_k) \). If the probability densities of \( \theta_m \) and \( \theta_k \) are known, the parameters are integrated out in an obvious manner (conceptually, at least). Whether we can actually carry out the integration depends on how the parameters enter into the expression.

If either \( \theta_m \) or \( \theta_k \) is a nonrandom variable, we can check to see if a uniformly most powerful test exists. If it does not, a generalized likelihood ratio test may be appropriate.

Structural Non-Gaussian Processes. The key to the simplicity in Gaussian problems is that we can completely characterize the process by its mean-value function and covariance function. We would expect that whenever the processes involved could be completely characterized in a reasonably simple manner, one could find the optimum receiver. An important example of such a processes is the Poisson process. References [31]–[35] discuss this problem. A second important example is Markov processes (e.g., [2-21]–[2-24]).

Unspecified Non-Gaussian Processes. In this case we would like to make some general statements about the optimum receiver without restricting the process to have a particular structure. One result of this type is available in the LEC case that we studied in Section 4.3. Middleton [36], [37], derives the LEC receiver without requiring that the signal process be Gaussian. (See [39] for a different series expansion approach.) A
second important result concerning unspecified Gaussian processes is
given in [38]. Here, Kailath extends the realizable estimator-correlator
receiver to include non-Gaussian processes.

Analysis of Fixed Receivers. In this case, we consider a fixed receiver
structure and analyze its performance in the presence of non-Gaussian
signals and noise. Suitable examples of this type of analysis are contained
in [40], [41].

These four topics illustrate some of the issues involved in the study of
non-Gaussian processes. The selection was intended to be representative,
not exhaustive.

5.1.5 Vector Gaussian Processes

We have not discussed the case in which the received signal is a vector
random process. The formal extension of our results to this case is
straightforward. In fact, all of the necessary equations have been developed
in the problem sections of Chapters 2-4. The important issues in the
vector case are the solution of the equations specifying the optimum
receiver and its performance and the interpretation of the results in the
context of particular physical situations. In the subsequent volume [42],
we shall study the vector problem in the context of array processing in
sonar and seismic systems. At that time, we shall discuss the above issues
in detail.

5.2 SUMMARY OF DETECTION THEORY

In Chapters 2 through 5 we have studied the detection of Gaussian
signals in Gaussian noise in detail. The motivation of this detailed study
is to provide an adequate background for actually solving problems we
encounter when modeling physical situations.

In Chapter 2 we considered the simple binary problem. The first step
was to develop the likelihood ratio test. We saw that the likelihood ratio
contained three components. The first was obtained by a nonlinear
operation on the received waveform and arose because of the randomness
in the signal. The second was obtained by a linear operation on the
received waveform and was due to the deterministic part of the received
signal. This component was familiar from our earlier work. The third
component was the bias term, which had to be evaluated in order to
conduct a Bayes test.
We next turned our attention to the problem of realizing the nonlinear operation needed to generate $I_R$. Four canonical realizations were developed:

1. The estimator-correlator receiver.
2. The filter-correlator receiver.
3. The filter-squarer receiver.
4. The optimum realizable filter receiver.

The last realization was particularly appealing when the process had finite state-variable representation. In this case we could use all of the effective state-variable procedures that we developed in Section I-6.3 actually to find the receiver.

A more difficult issue was the performance of the optimum receiver. As we might expect from our earlier work, an exact performance calculation is not feasible in many cases. By building on our earlier work on bounds and approximate expressions in Section I-2.7, we developed performance results for this problem. The key to the results was the $\mu(s)$ function defined in (2.148). We were able to express this in terms of both a realizable filtering error and the logarithm of the Fredholm determinant. We have effective computational procedures to evaluate each of these functions.

We next turned to the general binary problem in Chapter 3, where the received waveform could contain a nonwhite component on each hypothesis. The procedures were similar to the simple binary case. A key result was (3.33), whose solution was the kernel of the nonlinear part of the receiver. The modifications of the various canonical realizations were straightforward, and the performance bounds were extended. A new issue that we encountered was that of singularity. We first derived simple upper and lower bounds on the probability of error in terms of $\mu(\frac{1}{2})$. We then showed that a necessary and sufficient condition for a nonsingular test was that $\mu(\frac{1}{2})$ be finite. This condition was then expressed in terms of a square-integrability requirement on a kernel. As before, singularity was never an issue when the same white noise component was assumed to be present on both hypotheses.

In Chapter 4 we considered three special cases that led to particularly simple solutions. In Section 4.1 we looked at the stationary-process-long-time-interval case. This assumption enabled us to neglect homogeneous solutions in the integral equation specifying the kernel and allowed us to solve this equation using Fourier transform techniques. Several practical examples were considered. The separable kernel case was studied in Section 4.2. We saw that this was a suitable model for pulsed radars.
with slowly fluctuating targets, ionospheric communications over resolvable multipath channels, and frequency-diversity systems. The solution for this case was straightforward. Finally, in Section 4.3, we studied the low-energy-coherence case, which occurs frequently in passive sonar and radar astronomy problems. The energy in the signal process is spread over a large number of coordinates so that each eigenvalue is small when compared to the white noise level. This smallness enabled us to obtain a series solution to the integral equation. In this particular case we found that the output signal-to-noise ratio \( d^2 \) is an accurate performance measure. In addition to these three special cases, we had previously developed a complete solution for the case in which the processes have a finite state representation. A large portion of the physical situations that we encounter can be satisfactorily approximated by one of these cases.

In Section 5.1 we extended our results to the M-ary problem. The optimum receiver is a straightforward extension of our earlier results, but the performance calculation for the general problem is difficult. A reasonably simple bound for the case of M-orthogonal processes was presented. In Section 5.2 we derived performance expressions for suboptimum receivers.

Our discussion of the detection problem has been lengthy and, in several instances, quite detailed. The purpose is to give the reader a thorough understanding of the techniques involved in solving actual problems. In addition to the references we have cited earlier, the reader may we wish to consult [43]–[48] for further reading in this area. In the next two chapters we consider the parameter estimation problem that was described in Chapter 1.

### 5.3 PROBLEMS

#### P.5.1 Related Topics

**M-ary Detection**

**Problem 5.1.1.** Consider the model described in (4)–(6). Assume that \( n(t) \) contains a white noise component with spectral height \( N_0/2 \). Assume that \( m_i(t) = 0, \ i = 1, \ldots, M. \)

1. Derive (7)–(8) and (11)–(14).
2. Draw a block diagram of the optimum receiver.

**Problem 5.1.2.** Generalize the model in Problem 5.1.1 to include nonzero means and unequal costs. Derive the optimum Bayes receiver.
5.3 Problems

Problem 5.1.3. Consider the model in (15)-(17).

Assume that

\[ K_{t_i}(t, u) = iK_s(t - u), \quad i = 1, \ldots, M \]

and that the SPLOT condition is valid. The hypotheses are equally likely.

1. Draw a block diagram of the optimum receiver.
2. Consider the case in which \( M = 3 \). Derive a bound on the Pr (\( \epsilon \)).

Problem 5.1.4. Consider the communication system using \( M \)-orthogonal signals that is described in (19)-(21). On pages 1-263-1-264, we derived a bound on the Pr (\( \epsilon \)) in an \( M \)-ary system in terms of the Pr (\( \epsilon \)) in a binary system.

1. Extend this technique to the current problem of interest.
2. Compare the bound in part 1 with the bound given by (22)-(35). For what values of \( R \) is the bound in part 1 useful?

Problem 5.1.5. Consider the problem of detecting one of \( M \)-orthogonal bandpass processes in the presence of white noise. Assume that each process has \( N \) eigenvalues.

1. We can immediately reduce the problem to one with \( MN \) dimensions. Denote this resulting vector as \( \mathbf{R} \). Compute \( P_{r|H_m}(\mathbf{R} | H_m) \).
2. In [6], Gallager derived the basic formula for a bound on the error probability,

\[
\Pr(\epsilon | H_m) \leq \int \cdots \int d\mathbf{R} \left[ P_{r|H_m}(\mathbf{R} | H_m) \right]^{1/(1+\rho)} \times \left[ \sum_{k \neq m} P_{r|H_k}(\mathbf{R} | H_k) \right]^{\rho}, \quad \rho \geq 0. \quad (P.1)
\]

Use (P.1) to derive (22). \((\text{Hint: Use the fact that } E[x^p] \leq (E[x])^p, 0 \leq \rho \leq 1.\)

Problem 5.1.6.

1. Verify the results in (29)-(35).
2. One can show that \( C_{or} \) is the capacity of the channel for this type of communication system (i.e., we require \( M \)-orthogonal signals and use rectangular signal envelopes). Assume

\[ S_{t_i}(\omega) = \frac{2k}{\omega^2 + k^2} \quad (P.1) \]

and

\[ P_r \triangleq 2\sigma_b^2. \quad (P.2) \]

Plot \( C_{or} \) as a function of

\[ \Lambda_{\eta} = \pi P_r \frac{k}{kN_0}. \quad (P.3) \]

3. Repeat for

\[ S_{q_i}(\omega) = \begin{cases} \frac{\pi}{k}, & |\omega| \leq k, \\ 0, & |\omega| > k. \end{cases} \quad (P.4) \]
Problem 5.1.7. The error exponent, $E(R)$, is defined by (31) and (34).

1. Plot $E(R)$ as a function of $N_0 R/P_r$ for the bandlimited message spectrum in (P.4) of Problem 5.1.6.
2. Plot $E(R)$ as a function of $N_0 R/P_r$ for the one-pole message spectrum in (P.1) of Problem 5.1.6.

Problem 5.1.8. Assume that we want to signal at low rates so that

$$E(R) \approx E(0).$$

1. Consider the one-pole message spectrum in (P.1) of Problem 5.1.6. Plot $E(0)/(P_r/N_0)$ as a function of $\Lambda_B$. What value of $\Lambda_B$ maximizes $E(0)/(P_r/N_0)$?
2. Repeat part 1 for the ideal bandlimited message spectrum in (P.4) of Problem 5.1.6.
3. Compare the results in parts 1 and 2 with those in (4.60) and (4.69).

Problem 5.1.9. Assume that we want to signal at the rate

$$R = \frac{1}{10} C_\infty = \frac{1}{10} \frac{P_r}{N_0}.$$ 

We want to maximize $E(R)/(P_r/N_0)$ by choosing $\Lambda_B$.

1. Carry out this maximization for the one-pole spectrum.
2. Carry out this maximization for the ideal bandlimited spectrum.
3. Compare your results with those in Problem 5.1.8.

Problem 5.1.10. Define

$$E^*_n(R) = \max_{\Lambda_B} \left[ \frac{E(R)}{P_r/N_0} \right].$$

1. Find $E^*_n(R)$ as $R$ varies from 0 to $C_\infty$ for the ideal bandlimited spectrum.
2. Repeat part 1 for the one-pole spectrum.

Problem 5.1.11 [6]. Assume that each signal process has a non-zero mean. Specifically,

$$E\{s_i(t)\sqrt{2} \cos (\omega_i t)\}_{L_p} = m(t),$$
$$E\{s_i(t)\sqrt{2} \sin (\omega_i t)\}_{L_p} = 0.$$

Show that the effect of the non-zero mean is to add a term to $E_0(\rho)$ in (25), which is

$$E_{0m}(\rho) = \frac{\rho}{2N_0 (1 + \rho)} \int_{-\infty}^{\infty} \left[ |M(j\omega)|^2 \left( 1 + \left( \frac{\rho}{1 + \rho} \right)^2 \frac{2S_{s_i}(\omega)}{N_0} \right)^{-1} \right] \frac{d\omega}{2\pi}.$$

Problem 5.1.12 [7]. Consider the special case of Problem 5.1.11 in which

$$S_{s_i}(\omega) = 0.$$

1. Prove

$$\frac{E(R)}{C_\infty} = \begin{cases} \frac{1}{2} - \frac{R}{C_\infty}, & 0 \leq \frac{R}{C_\infty} \leq \frac{1}{2}, \\ \left(1 - \frac{R}{\sqrt{C_\infty}}\right)^2, & \frac{1}{2} \leq \frac{R}{C_\infty} < 1. \end{cases}$$
where
\[ C_\infty = \frac{P_r}{N_0} - \frac{1}{N_0} \int_{-\infty}^{\infty} m^2(t) \, dt. \]

2. Discuss the significance of this result.

**Suboptimum Receivers**

**Problem 5.1.13.** Consider the definitions of \( \mu_0(s) \) and \( \mu_1(s) \) given in (44) and (45).

1. Derive the Chernoff bounds in (46)–(49).
2. Derive the approximate error expressions in (51) and (52).

**Problem 5.1.14.** Consider the simple binary detection problem described on page 151 and the filter-squarer-integrator receiver in Fig. 5.1. The filter is time-invariant with transfer function
\[ H(j\omega) = \frac{1}{j\omega + \beta}. \]
The message spectrum is given in (37).

1. Write the state equations that are needed to evaluate \( \mu_1(s) \) and \( \mu_0(s) \).
2. Assume that the long-time-interval approximation is valid. Find \( \mu_{1\omega}(s) \) and \( \mu_{0\omega}(s) \). Verify that the value of \( \beta \) in (39) is optimum.

**Problem 5.1.15.**

1. Repeat part 1 of Problem 5.1.14 for the case in which \( s(t) \) is a Wiener process,
\[ s(0) = 0, \quad t \geq 0, \]
\[ E[s^2(t)] = \sigma^2 t, \quad t \geq 0. \]
2. Find the optimum value of \( \beta \) for long observation times.

**Problem 5.1.16.** Consider the binary symmetric communication problem whose model was given in Section 3.4.3. The quantities \( r_{C_1}(t) \), \( r_{S_1}(t) \), \( r_{C_0}(t) \), and \( r_{S_0}(t) \) were defined in Fig. 3.9. We operate on each of these waveforms as shown in Fig. 3.10. Instead of the optimum filter \( h_{C_1}(t, u) \), we use some arbitrary filter \( h_{mu}(x) \) in each path. Denote the output of the top branch as \( I_1 \) and the output of the bottom branch as \( I_0 \). Define
\[ I_e = I_1 - I_0. \]
The bias terms are both zero and
\[ \ln \eta \Delta \gamma_e = 0. \]
Define
\[ \mu_{1j}(s) = \ln E[e^{sI_j} | H_j], \quad j = 0, 1, \]
and
\[ \mu_{0j}(s) = \ln E[e^{sI_j} | H_j], \quad j = 0, 1. \]

1. Prove
\[ \mu_{BS,1}(s) = \mu_{11}(s) + \mu_{01}(-s), \]
\[ \mu_{BS,0}(s) = \mu_{00}(-s) + \mu_{10}(s) = \mu_{BS,1}(-s). \]

2. Prove
\[ \Pr (\epsilon) < \frac{1}{2} \exp (\mu_{BS,1}(s_m)), \]
where
\[ \mu_{BS,1}(s)|_{s=s_m} = 0. \]
3. Prove
\[ \Pr (c) \approx \frac{1}{2(2\pi \mu_{BS,1}(s_m))^{1/2}s_m(1 - s_m)} \exp \mu_{BS,1}(s_m). \]

4. Express \( \mu_{BS,1}(s) \) in terms of Fredholm determinants.

**Problem 5.1.17.** Consider the binary communication system described in Problem 5.1.16. Assume that \( s_1(t) \) is a sample function of a stationary process whose low-pass equivalent spectrum is \( S_{s_1}(\omega) \) and \( h(t, \tau) \) is a time-invariant filter with a rational transfer function. Assume that the SPLOT condition is valid.

1. Find an expression for \( \mu_{BS,1,\infty}(s) \) in terms of \( S_{s_1}(\omega) \), \( H(j\omega) \), and \( N_0 \).
2. Verify that \( \mu_{BS,1,\infty}(s) \) reduces to \( \mu_{OPT,\infty}(s) \) when \( H(j\omega) \) is chosen optimally.
3. Plot \( \mu_{BS,1,\infty}(s) \) for the case in which
\[ S_{s_1}(\omega) = \frac{kP_r}{\omega^2 + k^2} \]
and
\[ H(j\omega) = \frac{1}{j\omega + \beta}. \]

Find \( s_m \).

**Problem 5.1.18 (continuation).** Consider the binary communication system discussed in Problems 5.1.16 and 5.1.17. We are interested in the case discussed in part 3 of Problem 5.1.17.

One of the problems in designing the optimum receiver is that \( P_r \) may be unknown or may vary slowly. Assume that we think that
\[ P_r = P_{rn} \]
and design the optimum receiver.

1. Evaluate \( \mu_{BS,1,\infty}(s_m) \) and \( \mu_{OPT,\infty}(\frac{1}{2}) \) for this receiver when
\[ \Lambda_1 \Delta \frac{2P_{rn}}{kN_0} = 100 \]
2. Now assume that
\[ 0.1P_{rn} \leq P_r \leq 10P_{rn}. \]

Plot \( \mu_{BS,1,\infty}(s_m) \). The receiver design is fixed.

3. Assume that the receiver is redesigned for each \( P_r \). Compare \( \mu_{OPT,\infty}(\frac{1}{2}) \) with \( \mu_{BS,1,\infty}(s_m) \).

**Problem 5.1.19.** The LEC receiver was derived in Section 4.3 and was shown in Fig. 4.21. This receiver is sometimes used when the LEC condition is not satisfied.

1. Derive an approximate expression for the performance of this receiver.
2. Assume that \( s(t) \) has a finite-dimensional state representation. Find a state equation for \( \mu_1(s) \) and \( \mu_0(s) \).
3. Assume that the SPLOT condition is valid. Find a simple expression for \( \mu_1(s) \) and \( \mu_0(s) \).
REFERENCES


166 References


