Optimum Reception of M-ary Gaussian Signals in Gaussian Noise

By T. T. KADOTA

(Manuscript received May 28, 1965)

The problem of optium reception of M-ary Gaussian signals in Gaussian noise is to specify, in terms of the observable waveform, a scheme for deciding among M alternative mean and covariance functions with minimum error probability. Although much literature on the problem exists, a mathematically rigorous solution has yet to appear. By formulating the problem as optimum discrimination of M Gaussian measures in function space induced by the mean and covariance functions, this paper presents such a solution.

Let $m_k(t)$ and $r_k(s,t)$, $k = 1, \cdots, M$, be the alternative mean and covariance functions of the Gaussian signal, and let $m_0(t)$ and $r_0(s,t)$ be the mean and covariance functions of the Gaussian noise. If, for each $k = 1, \cdots, M$, the integral equations,

$$\int r_0(s,t)g_k(s)\, ds = m_k(t)$$

and

$$\iint r_0(s,u)h_k(u,v)[r_0(v,t) + r_k(v,t)]\, du\, dv = r_k(s,t),$$

admit a square-integrable solution $g_k(t)$ and a symmetric, square-integrable solution $h_k(s,t)$, then the following decision scheme is optimum: given an observable waveform $x(t)$,

choose $m_k(t)$ and $r_k(s,t)$ if $I_k(x)$ is the largest among all $I_j(x)$, $j = 1, \cdots, M$,

where I_k is defined by

$$I_k(x) = \frac{1}{2} \iint x(s)h_k(s,t)x(t)\, ds\, dt + \int x(t)f_k(t)\, dt + c_k$$

in which

$$f_k(t) = g_k(t) - \int h_k(s,t)[m_0(s) + m_k(s)]\, ds$$

and c_k *is a constant determined by the mean and covariance functions* $m_0(t)$, $m_k(t)$, $r_0(s,t)$ *and* $r_k(s,t)$ *as well as the a priori probability associated with* $m_k(t)$ *and* $r_k(s,t)$.

The first section introduces and defines the problem and the second presents the solution with pertinent discussions while a precise mathematical treatment is left to the appendix.

I. INTRODUCTION

Before we formulate the general problem of optimum reception of M-ary signals in noise, let us review a simplified version of the classical problem: optimum reception of M-ary *sure* signals in Gaussian noise. Suppose there are M sure signals $m_k(t)$, $k = 1, \cdots, M$, with *a priori* probabilities α_k, $0 < \alpha_k < 1$ and $\sum_{k=1}^{M} \alpha_k = 1$, for transmission. The received waveform $x(t)$ consists of one of these M signals and an additive Gaussian noise $n(t)$, i.e.

$$x(t) = m_k(t) + n(t).$$

In order to simplify the problem, we "represent" the signals and noise by certain finite sequences $m_{k1}, \cdots, m_{kn}$, $k = 1, \cdots, M$, and $n_1, \cdots, n_n$ respectively.* Then the representing sequence $x_1, \cdots, x_n$ of the received waveform is given by

$$x_i = m_{ki} + n_i, \qquad i = 1, \cdots, n. \tag{1}$$

It is assumed that the signal sequences are linearly independent vectors in an n-dimensional space R_n while the elements, $n_1, \cdots, n_n$ of the noise sequence are statistically independent, identically distributed Gaussian variables with mean zero and variance one. The task of the receiver is to observe the received sequence $x_1, \cdots, x_n$ and to decide which one of M signal sequences must have been transmitted. For each possible erroneous decision, there is an associated probability, and the average of all these probabilities weighted by the *a priori* probabilities is the so-called (average) error probability. Then, the problem of optimum reception in this simplified form is to specify in terms of the observable sequence $x_1, \cdots, x_n$ a scheme for choosing the value of the index k such that the error probability is minimum over all possible decision schemes.

---

* These sequences may be regarded as the sample values of the waveforms or the Fourier coefficients of certain orthonormal expansions of the waveforms.

Now, from the assumption concerning the noise sequence, the joint probability density function of $n_1, \cdots, n_n$ is

$$p_0(\nu_1, \cdots, \nu_n) = (2\pi)^{-(n/2)} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \nu_i^2\right). \quad (2)$$

Hence, through the use of the relation (1), the joint density function of $x_1, \cdots, x_n$ becomes

$$p_k(\nu_1, \cdots, \nu_n) = (2\pi)^{-(n/2)} \exp\left[-\frac{1}{2}\sum_{i=1}^{n} (\nu_i - m_{ki})^2\right]. \quad (3)$$

Note that since the observable sequence is a point in R_n, a decision scheme in terms of it is equivalent to a division of R_n into M non-overlapping regions $\hat{\Lambda}_1, \cdots, \hat{\Lambda}_M$ so that $k = j$ is chosen if $(x_1, \cdots, x_n)$ belongs to $\hat{\Lambda}_j$. Then, the error probability associated with such a division $(\hat{\Lambda}_1, \cdots, \hat{\Lambda}_M)$ is given by

$$P_e = 1 - \sum_{k=1}^{M} \alpha_k \int_{\hat{\Lambda}_k} p_k(\nu_1, \cdots, \nu_n)\, d\nu_1 \cdots d\nu_n,$$

which can be rewritten as

$$P_e = 1 - \int_{R_n} \left[\sum_{k=1}^{M} \chi_{\hat{\Lambda}_k}(\nu_1, \cdots, \nu_n)\alpha_k p_k(\nu_1, \cdots, \nu_n)\right] d\nu_1 \cdots d\nu_n, \quad (4)$$

where $\chi_{\hat{\Lambda}_k}$ is the indicator function of the set $\hat{\Lambda}_k$, $k = 1, \cdots, M$. Clearly, P_e is minimum over all divisions if the integrand in the above is maximum at every point $(\nu_1, \cdots, \nu_n)$ in R_n. This can be achieved by arranging a division $(\hat{\Lambda}_1, \cdots, \hat{\Lambda}_M)$ so that, in the region $\hat{\Lambda}_k$, the $\alpha_k p_k$ with the same index is the largest among all $\alpha_1 p_1, \cdots, \alpha_M p_M$. Namely, the region with index k consists of a set of points at which $\alpha_k p_k$ is the largest. If more than one $\alpha_k p_k$ is the largest at the same point, that point is to be assigned to the region whose index is the smallest of all those k's so that all the regions remain disjoint. That is, a division $(\hat{S}_1, \cdots, \hat{S}_M)$ of R_n specified by

$$\hat{S}_k = \left\{ (\nu_1, \cdots, \nu_n): \begin{array}{l} \alpha_k p_k(\nu_1, \cdots, \nu_n) > \alpha_j p_j(\nu_1, \cdots, \nu_n), \quad j < k, \\ \alpha_k p_k(\nu_1, \cdots, \nu_n) \geqq \alpha_j p_j(\nu_1, \cdots, \nu_n), \quad j > k \end{array} \right\}, \quad (5)$$

is the division which minimizes P_e. Or the optimum decision scheme is the following:

Given an observable sequence $(x_1, \cdots, x_n)$, choose the minimum value of k for which $\alpha_k p_k(x_1, \cdots, x_n)$ is maximum as a function of k.*

Now by expanding the exponent in (3),

$$\alpha_k p_k(x_1, \cdots, x_n) = \alpha_k \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} x_i m_{ki} - \frac{1}{2}\sum_{i=1}^{n} m_{ki}^2\right).$$

Since the first sum in the bracket is common to all $\alpha_k p_k$, if we put

$$\hat{I}_k = \sum_{i=1}^{n} x_i m_{ki} - \frac{1}{2}\sum_{i=1}^{n} m_{ki}^2 + \log \alpha_k, \tag{6}$$

then the optimum decision scheme is equivalent to choosing the minimum value of k for which $\hat{I}_k$ is maximum.

In the special case of "equi-probable and equi-energetic" transmitted signal sequences, i.e.

$$\alpha_i = \cdots = \alpha_M \qquad \text{and} \qquad \sum_{i=1}^{n} m_{1i}^2 = \cdots = \sum_{i=1}^{n} m_{Mi}^2,$$

$\hat{I}_k$ can be effectively replaced by its first term, i.e. $\sum_{i=1}^{n} x_i m_{ki}$. In other words, the optimum decision scheme in this case consists in performing correlation of the observable sequence with M signal sequences and choosing the smallest of the k values corresponding to the largest correlation sums.†

In the general problem of optimum reception of M-ary Gaussian signals in Gaussian noise, the observable waveform (received waveform) $x(t)$ is expressed by

$$x(t) = y_k(t) + n(t)$$

where $y_k(t)$ is one of M possible Gaussian signals which are characterized by mean and covariance functions just as $n(t)$ is. We assume that each Gaussian signal is statistically independent of the noise and it cannot be detected "perfectly" in the presence of this noise.‡ Again, the task of the receiver is to observe the waveform $x(t)$ for a finite time, say $0 \leqq t \leqq 1$, and to decide which one of M Gaussian signals must have been received. Then, by defining the error probability as before, the problem of optimum reception becomes that of specifying, in terms of the

---

* That is, suppose for a given $(x_1, \cdots, x_n)$, $\alpha_k p_k(x_1, \cdots, x_n)$ as a function of k assumes its maximum at $k = k_1, \cdots, k_j$ where $k_1 < \cdots < k_j$. Then, we choose $k = k_1$.

† For classical references, see Refs. 1 and 2.

‡ This is the assumption of "non-singular detection". A necessary and sufficient condition of non-singular detection is given by (13) in Appendix.

observable waveform $x(t)$, a scheme for choosing the index k such that the error probability is minimum over all possible decision schemes.

One mathematical idealization of the above problem is the following: Let $\{y_t, 0 \leqq t \leqq 1\}$ be a Gaussian process whose mean and covariance functions are one of M possible pairs of $m_k(t)$, $0 \leqq t \leqq 1$, and $r_k(s,t)$, $0 \leqq s, t \leqq 1$, $k = 1, \cdots, M$, where $m_k(t)$ and $r_k(s,t)$ are assumed to be continuous.* Similarly, let $\{n_t, 0 \leqq t \leqq 1\}$ be a Gaussian process whose mean and covariance functions are $m_0(t)$, $0 \leqq t \leqq 1$, and $r_0(s,t)$, $0 \leqq s,t \leqq 1$, where $m_0(t)$ is assumed to be continuous while $r_0(s,t)$ is positive-definite as well as continuous. It is further assumed that $\{y_t, 0 \leqq t \leqq 1\}$ and $\{n_t, 0 \leqq t \leqq 1\}$ are mutually independent for every $k = 1, \cdots, M$. Now define a new process $\{x_t, 0 \leqq t \leqq 1\}$ by $x_t = y_t + n_t$. Then, from the mutual independence assumption, the mean and covariance functions of $\{x_t, 0 \leqq t \leqq 1\}$ are one of M possible pairs of $m_0(t) + m_k(t)$ and $r_0(s,t) + r_k(s,t)$, $k = 1, \cdots, M$. Let P_k, $k = 1, \cdots, M$, be the Gaussian (probability) measure corresponding to the pair $m_0(t) + m_k(t)$ and $r_0(s,t) + r_k(s,t)$, and let P_0 be the one corresponding to $m_0(t)$ and $r_0(s,t)$. It is assumed that $m_0(t)$, $m_k(t)$, $r_0(s,t)$ and $r_k(s,t)$ are such that the two measures P_0 and P_k are equivalent, i.e. $P_0 \equiv P_k$, $k = 1, \cdots, M$.† Denote by H_k and α_k, $k = 1, \cdots, M$, the hypothesis and *a priori* probability that $m_k(t)$ and $r_k(s,t)$ are the pertinent mean and covariance functions of $\{y_t, 0 \leqq t \leqq 1\}$. Let $x(t)$ be the sample function of $\{x_t, 0 \leqq t \leqq 1\}$. Then, specification of the decision scheme amounts to dividing the space Ω of all sample functions $x(t)$ into M disjoint sets, $\Lambda_1, \cdots, \Lambda_M$, so that, if $x(t) \in \Lambda_k$, then H_k is to be chosen. Moreover, the error probability associated with such a division (or a decision scheme) is given by

$$P_e = 1 - \sum_{k=1}^{M} a_k P_k(\Lambda_k). \tag{7}$$

Thus, the problem of optimum reception is to specify in terms of $x(t)$ such a division of Ω that its associated error probability is minimum over all possible divisions.

Unlike the previous simple case, the sample (the observable) in this general case is the sample *function* $x(t)$ of the Gaussian process $\{x_t, 0 \leqq t \leqq 1\}$ instead of the sample *sequence* $x_1, \cdots, x_n$ of the finite sequence of Gaussian variables. Thus, the sample space which is to be divided into M non-overlapping regions, is the function space Ω instead of the n-dimensional sequence space R_n. Hence, we no longer have at our

* Note: $r_k(s,t) = E_k\{(x_s - m_k(s))(x_t - m_k(t))\}$, $k = 1, \cdots, M$.
† This corresponds to the assumption of non-singular detection.

disposal the joint density functions $p_k(\nu_1, \cdots, \nu_n)$, $k = 1, \cdots, M$, through which the optimum decision scheme is constructed. Nevertheless, there exists a certain generalization to this basic approach. Note from (2) that $p_0(\nu_1, \cdots, \nu_n) > 0$, $-\infty < \nu_i < \infty$, $i = 1, \cdots, n$. Hence, (4) can be rewritten as

$$P_e = 1 - \int_{R_n} \left[\sum_{k=1}^{M} \chi_{\hat{\Lambda}_k}(\nu_1, \cdots, \nu_n) \alpha_k \frac{p_k(\nu_1, \cdots, \nu_n)}{p_0(\nu_1, \cdots, \nu_n)} \right] p_0(\nu_1, \cdots, \nu_n) \times d\nu_1 \cdots d\nu_n .$$

Then, the optimum division of R_n is specified in terms of $\alpha_k[p_k(x_1, \cdots, x_n)/p_0(x_1, \cdots, x_n)]$ instead of $\alpha_k p_k(x_1, \cdots, x_n)$, though the two are obviously equivalent. Now, in the general case where the sample space is Ω instead of R_n, the likelihood ratio $p_k(x_1, \cdots, x_n)/p_0(x_1, \cdots, x_n)$ is replaced by its generalized version dP_k/dP_0, the Radon-Nikodym derivative (of P_k with respect to P_0), which is a function of $x(t)$, and $p_0(\nu_1, \cdots, \nu_n)d\nu_1 \cdots d\nu_n$ is replaced by dP_0. Thus, the error probability in the general case can be expressed as

$$P_e = 1 - \int_{\Omega} \left[\sum_{k=1}^{M} \chi_{\Lambda_k}(x) \alpha_k \frac{dP_k}{dP_0}(x) \right] dP_0(x),$$

where $(\Lambda_1, \cdots, \Lambda_M)$ is a nonoverlapping division of Ω. Then an optimum division of Ω, which is analogous to (5), can be specified by $\alpha_\kappa(dP_k/dP_0)$, $k = 1, \cdots, M$, and dP_k/dP_0 can in turn be expressed in terms of certain functionals of $x(t)$.

II. SUMMARY OF MAIN RESULTS AND DISCUSSIONS

The foundation for solution of the general problem stated in the preceding section consists of the two following facts:[3,4]

(*i*) If two Gaussian measures P_0 and P_k are equivalent for each $k = 1, \cdots, M$, then there exists random variables dP_k/dP_0 so that the optimum division $(S_1, \cdots, S_M)$ of the sample space Ω can be specified by

$$S_k = \left\{ x(t) \colon \alpha_k \frac{dP_k}{dP_0}(x) > \alpha_j \frac{dP_j}{dP_0}(x), j < k; \right. \\ \left. \alpha_k \frac{dP_k}{dP_0}(x) \geqq \alpha_j \frac{dP_j}{dP_0}(x), j > k \right\}. \tag{8}$$

(*ii*) If the integral equations

$$\int_0^1 r_0(s,t) g_k(s)\, ds = m_k(t), \qquad 0 \leqq t \leqq 1, \tag{9}$$

and

$$\int_0^1 \int_0^1 r_0(s,u) h_k(u,v)[r_0(v,t) + r_k(v,t)]\, du\, dv = r_k(s,t), \qquad 0 \leqq s,t \leqq 1, \tag{10}$$

have a square-integrable solution $g_k(t)$ and a symmetric, square-integrable solution $h_k(s,t)$ respectively,* then

$$\frac{dP_k}{dP_0}(x) = \beta_k^{\frac{1}{2}} \exp\left[\frac{1}{2}\int_0^1\int_0^1 [x(s) - m_0(s) - m_k(s)]h_k(s,t) \cdot [x(t) - m_0(t) - m_k(t)]\, ds\, dt + \int_0^1 [x(t) - m_0(t) - \tfrac{1}{2}m_k(t)]g_k(t)\, dt\right], \tag{11}$$

for almost all sample functions under all hypotheses H_k, $k = 1, \cdots, M$, where $\beta_k^{-1} = \prod_{i=1}^{\infty} \lambda_i^{(k)}$ and $\lambda_i^{(k)} > 0$, $i = 1,2, \cdots$, are the eigenvalues of the operators $R_0^{-\frac{1}{2}}(R_0 + R_k)R_0^{-\frac{1}{2}}$ and R_0 and R_k are integral operators whose kernels are $r_0(s,t)$ and $r_k(s,t)$ respectively.

Then, upon combination of (*i*) and (*ii*), the optimum decision scheme can be specified as follows:

Choose the minimum value of k for which I_k is maximum where

$$I_k = \frac{1}{2}\int_0^1\int_0^1 x(s)h_k(s,t)x(t)\, ds\, dt + \int_0^1 x(t)f_k(t)\, dt + \frac{1}{2}\int_0^1\int_0^1 [m_0(s) + m_k(s)]\, h_k(s,t)\, [m_0(t) + m_k(t)]\, ds\, dt - \int_0^1 [m_0(t) + \tfrac{1}{2}\, m_k(t)]\, g_k(t)\, dt + \log \alpha_k \beta_k^{\frac{1}{2}}, \tag{12}$$

in which $f_k(t)$, $k = 1, \cdots, M$, are defined by

$$f_k(t) = g_k(t) - \int_0^1 h_k(s,t)[m_0(s) + m_k(s)]\, ds, \qquad 0 \leqq t \leqq 1.$$

Needless to say, the condition for the above decision scheme to be optimum is the existence of such $g_k(t)$ and $h_k(s,t)$, $k = 1, \cdots, M$, as described in (*ii*). It should be remarked that the existence of $g_k(t)$

* If such solutions exist, they are necessarily unique. It should be remarked that square-integrability of $h(s,t)$ is in the sense of

$$\int_0^1 \int_0^1 h^2(s,t) ds\, dt < \infty.$$

and $h_k(s,t)$ implies that P_0 and P_k are equivalent for each k. That is, if these $g_k(t)$ and $h_k(s,t)$ exist, then none of the Gaussian signals $y_k(t)$ can be detected perfectly in the presence of noise $n(t)$.

Physical interpretation of the above optimum decision scheme is straightforward, at least in principle. Suppose given $m_0(t)$, $m_k(t)$, $r_0(s,t)$ and $r_k(s,t)$, $k = 1, \cdots, M$, are such that the integral equations (9) and (10) admit a square-integrable solution $g_k(t)$ and a symmetric, square-integrable solution $h_k(s,t)$, then the optimum decision scheme consists in performing the single and the double integrals involving the received waveform $x(t)$ as specified by (12), and adding to these integrals the predetermined constants, the last three terms of (12), and finally choosing the minimum value of k for which the sum of these integrals and the constants is maximum.

It is instructive to consider the following two special cases:

Case 1 $\quad m_0(t) \equiv 0, \qquad r_k(s,t) \equiv 0, \qquad k = 1, \cdots, M.$

This is the case of "M-ary sure signals in noise", a simplified version of which has already been discussed in the introduction. Here the integral equation (10) always has a symmetric, square-integrable solution for each $k = 1, \cdots, M$, namely, the trivial solution:

$$h_k(s,t) \equiv 0.$$

Furthermore, $\lambda_i^{(k)} = 1$, $i = 1,2, \cdots$; $k = 1, \cdots, M$. Thus,

$$\beta_k = 1, \qquad k = 1, \cdots, M.$$

Hence, I_k of (12) is reduced to

$$I_k' = \int_0^1 x(t)g_k(t) - \frac{1}{2}\int_0^1 m_k(t)g_k(t)\,dt + \log \alpha_k ,$$

provided the square-integrable solutions $g_k(t)$, $k = 1, \cdots, M$, exist for the integral equations (9).* Note that I_k' is the function-space counterpart to $\hat{I}_k$ of (6) in the sequence-space case. With additional conditions that

$$\alpha_1 = \cdots = \alpha_M \quad \text{and} \quad \int_0^1 m_1(t)g_1(t)\,dt = \cdots = \int_0^1 m_M(t)g_M(t)\,dt,\dagger$$

the optimum decision scheme is reduced to choosing the minimum value

* The form of I_k' agrees with the result formally obtained in Ref. 5.

† For example, choose $m_k(t) = \sqrt{\sigma_k}\,\psi_k(t)$, $k = 1, \cdots, M$, where $\sigma_1 \geqq \sigma_2 \geqq \cdots$, and $\psi_1(t)$, $\psi_2(t)$, $\cdots$ are the eigenvalues and the orthonormalized eigenfunctions of R_0.

of k for which the "correlation integral"

$$\int_0^1 x(t)g_k(t)\,dt$$

is maximum.

Case 2 $\qquad m_0(t) \equiv 0, \qquad m_k(t) \equiv 0, \qquad k = 1, \cdots, M.$

This is the case commonly termed as "M-ary Gaussian signals in noise". Here, the integral equation (9) always has a square-integrable solution, namely, the trivial solution:

$$g_k(t) \equiv 0, \qquad k = 1, \cdots, M.$$

Hence, I_k is reduced to

$$I_k'' = \frac{1}{2}\int_0^1\int_0^1 x(s)h_k(s,t)x(t)\,ds\,dt + \log \alpha_k\beta_k^{\frac{1}{2}}.$$

Thus, the optimum decision scheme consists in choosing the minimum value of k for which I_k'' is maximum, provided that the symmetric, square-integrable solutions $h_k(s,t)$, $k = 1, \cdots, M$, of (10) exist.

(Remark)

It is interesting to note that formal substitution of $r_0(s,t) = \delta(s - t)$ into (10) yields the result which is consistent with those obtained previously by Price.[6]

APPENDIX

Mathematical Supplement

In the preceding, mathematical precision has been somewhat compromised for intuitive appeal. It is the purpose of this appendix to clarify the content of the preceding sections by supplying a brief mathematical summary with pertinent remarks.

Let Ω be the space of all real-valued functions on [0,1] and let $\tilde{P}_0$ and $\tilde{P}_k$, $k = 1, \cdots, M$, be the Gaussian measures induced by m_0 and r_0 and by $m_0 + m_k$ and $r_0 + r_k$ respectively on a σ-field generated by the class of all intervals in Ω, where m_0 and m_k, $k = 1, \cdots, M$, are real-valued, continuous functions on [0,1], while r_0 and r_k, $k = 1, \cdots, M$, are real-valued, symmetric, positive-definite, continuous functions on $[0,1] \times [0,1]$.* Then, without loss of generality, there exists a real,

* r_k, $k = 1, \cdots, M$, can be only nonnegative-definite.

separable (with respect to $\tilde{P}_0$) and measurable process $\{x_t, 0 \leqq t \leqq 1\}$.*
Let $\mathcal{B}$ be the minimal σ-field with respect to which every x_t, $t \in [0,1]$, is measurable, and P_0 and P_k, $k = 1, \cdots, M$, be the restrictions on $\mathcal{B}$ of $\tilde{P}_0$ and $\tilde{P}_k$ respectively. We assume that $P_0 \equiv P_k$, $k = 1, \cdots, M$, which immediately implies that $P_i \equiv P_j$; $i,j = 1, \cdots, M$, and $\{x_t, 0 \leqq t \leqq 1\}$ is separable with respect to P_k, $k = 1, \cdots, M$, also. For a necessary and sufficient condition for $P_0 \equiv P_k$, we cite the following:†

$$\lim_{n\to\infty} \operatorname{tr} [R_0^{(n)} (R_0^{(n)} + R_k^{(n)})^{-1} - 2I + (R_0^{(n)} + R_k^{(n)}) (R_0^{(n)})^{-1}] < \infty,$$

$$\lim_{n\to\infty} \operatorname{tr} (R_0^{(n)})^{-1} M_k^{(n)} < \infty, \tag{13}$$

where $R_0^{(n)}$, $R_k^{(n)}$ and $M_k^{(n)}$, $k = 1, \cdots, M$, are $n \times n$-matrices defined by

$$\begin{aligned} (R_0^{(n)})_{ij} &= r_0(t_i, t_j), \\ (R_k^{(n)})_{ij} &= r_k(t_i, t_j), \qquad\qquad i,j = 1, \cdots, n, \\ (M_k^{(n)})_{ij} &= m_k(t_i) m_k(t_j), \end{aligned}$$

and t_i, $i = 1, \cdots, n$, are a finite subset of any sequence dense in [0,1]. Now the two fundamental facts for solution are:

(*i*) Let dP_k/dP_0, $k = 1, \cdots, M$, be the Radon-Nikodym derivatives of P_k with respect to P_0, and let $(S_1, \cdots, S_M)$ be the partition of Ω defined by

$$S_k = \left\{\omega\colon \alpha_k \frac{dP_k}{dP_0}(\omega) > \alpha_j \frac{dP_j}{dP_0}(\omega), j < k\right\} \cap \left\{\omega\colon \alpha_k \frac{dP_k}{dP_0}(\omega) \geqq \alpha_j \frac{dP_j}{dP_0}(\omega), j > k\right\}.$$

Then, for any partition $(\Lambda_1, \cdots, \Lambda_M)$ of Ω,

$$\sum_{k=1}^{M} \alpha_k P_k(S_k) \geqq \sum_{k=1}^{M} \alpha_k P_k(\Lambda_k).$$

(*ii*) For each $k = 1, \cdots, M$, if the integral equations (9) and (10) admit a square-integrable solution g_k and a solution h_k satisfying

$$h_k(s,t) = h_k(t,s), \qquad \int_0^1 \int_0^1 h_k^2(s,t)\, ds\, dt < \infty,$$

* For a detailed justification, see Ref. 4.
† See Ref. 4.

then

(a) such solutions g_k and h_k are unique,

(b) $P_0 \equiv P_k$, $k = 1, \cdots, M$,

(c) there exist eigenvalues $\lambda_i^{(k)} > 0$, $i = 1, 2, \cdots$, of

$$R_0^{-\frac{1}{2}}(R_0 + R_k)R_0^{-\frac{1}{2}}$$

such that $\prod_{i=1}^{\infty} \lambda_i^{(k)}$ converges to β_k^{-1}, $0 < \beta_k < \infty$, and

$$\frac{dP_k}{dP_0} = \beta_k^{\frac{1}{2}} \exp\left[\frac{1}{2}\int_0^1\int_0^1 [x_s - m_0(s) - m_k(s)]h_k(s,t)[x_t - m_0(t) - m_k(t)]\,ds\,dt + \int_0^1 [x_t - m_0(t) - \tfrac{1}{2}m_k(t)]g_k(t)\,dt\right], \qquad \text{a.e. } (P_0).$$

Hence, specification of S_k, $k = 1, \cdots, M$, is obtained by combining (*i*) and (*ii*), and the hypothesis in (*ii*) is the condition that such a partition $(S_1, \cdots, S_M)$ exists and minimizes (7).

REFERENCES

1. Woodward, P. M., *Probability and Information Theory with Applications to Radar*, Pergamon Press, London, 1953.
2. Kotel'nikov, V. A., *The Theory of Optimum Noise Immunity*, McGraw-Hill Book Co., New York, 1959.
3. Kadota, T. T., Generalized Maximum Likelihood Test and Minimum Error Probability, (to appear in IEEE Trans. on Information Theory, Jan., 1966).
4. Kadota, T. T., Optimum Reception of Binary Sure and Gaussian Signals, B.S.T.J., *44*, Oct., 1965, pp. 1621–1658.
5. Thomas, J. B., Wolf, J. K., On the Statistical Detection Problem for Multiple Signals, IRE Trans., *IT-8*, pp. 274–280.
6. Price, R., Optimum Detection of Stochastic Signals in Noise with Applications to Scatter-Multipath Communications, IRE Trans., *IT-2*, pp. 125–135.