Optimum Reception of Binary Gaussian Signals

By T. T. KADOTA

(Manuscript received May 4, 1964)

The problem of optimum reception of binary Gaussian signals is to specify, in terms of the received waveform, a scheme for deciding between two alternative covariance functions with minimum error probability. Although a considerable literature already exists on the problem, an optimum decision scheme has yet to appear which is both mathematically rigorous and convenient for physical application. In the context of a general treatment of the problem, this article presents such a solution. The optimum decision scheme obtained consists in comparing, with a predetermined threshold k, a quadratic form (of function space) in the received waveform $x(t)$, namely,

$$\text{choose} \quad r_0(s,t) \quad \textit{if} \quad \iint x(s)h(s,t)x(t)\, ds\, dt < k,$$

$$\text{choose} \quad r_1(s,t) \quad \textit{if} \quad \iint x(s)h(s,t)x(t)\, ds\, dt \geqq k,$$

where $r_0(s,t)$ and $r_1(s,t)$ are the covariance functions while $h(s,t)$ is given as a solution of the integral equation,

$$\iint r_0(s,u)h(u,v)r_1(v,t)\, du\, dv = r_1(s,t) - r_0(s,t).$$

This may be regarded as a generalization of the "correlation detection" in the case of binary sure signals in noise.

Section I defines the problem, reviews the literature, and, together with certain pertinent remarks, summarizes principal results. A detailed mathematical treatment follows in Section II and the Appendices.

I. INTRODUCTION AND SUMMARY

1.1 *Definition and Nature of Problem*

The problem of optimum reception of binary Gaussian signals arises as a mathematical idealization of a common communication problem.

Consider a radio communication link containing a random medium. The transmitter sends one of two possible signals with known frequency rates (a priori probabilities), and the receiver decides which one of the two has been transmitted. Even if the transmitted signals are deterministic, the observable waveforms at the receiver appear to be random owing to effects of the random medium and the ever-present thermal noise at the receiver. The task of the so-called optimum (or ideal) receiver is to decide, upon observation of the received waveform for a finite time, which one of the two signals has been transmitted in such a way as to minimize the so-called probability of error. Thus, the problem of optimum reception amounts to specifying in terms of the received waveform such an optimum decision scheme for given a priori probabilities.

It is assumed that the values of the received waveforms at arbitrary instants of time during the observation interval, say $0 \leqq t \leqq 1$, are jointly Gaussian distributed with means zero and a covariance matrix which is determined by either one of two known covariance functions, depending upon which one of the two signals is transmitted. Then, the above problem may be stated as one of testing simple hypotheses as follows: Suppose there are two ensembles of real functions of time t, $0 \leqq t \leqq 1$, which are statistically characterized as being Gaussian distributed with identically vanishing mean functions and two distinct covariance functions. A sample (function) $x(t)$ is drawn either from the first ensemble with probability α (the null hypothesis: H_0) or from the second with probability $1 - \alpha$ (the alternative hypothesis: H_1). Determine a "critical region" Λ_α (a subset of a space of real functions Ω) for rejecting H_0 (or accepting H_1) if $x(t)$ belongs to Λ_α and accepting H_0 if $x(t)$ does not, in such a way that the associated error probability,

$$P_e(\Lambda_\alpha) = \alpha P_0(\Lambda_\alpha) + (1 - \alpha)P_1(\Omega - \Lambda_\alpha), \tag{1}$$

is no greater than $P_e(\Lambda)$ for an arbitrary $\Lambda \subset \Omega$; where P_0 and P_1 are two Gaussian (probability) measures defined on (measurable) subsets of Ω by the two zero mean functions and two covariance functions. Thus, the problem of optimum reception amounts to dividing the function space into two parts in such a way that the weighted probabilities on them specified by (1) are minimum among all possible divisions.

There are two features worth noting in this formulation. One is the lack of uniqueness of the optimum division as a consequence of adopting the minimum error probability as the optimality criterion. Namely, it is immaterial whether a certain set N (of functions) with both probabilities zero, i.e., $P_0(N) = 0 = P_1(N)$, should be a part of Λ_α or $\Omega -$

Λ_α, since it does not contribute to the error probability P_e. Thus, those sets upon which P_0 and P_1 vanish can effectively be ignored. The other feature is a stipulation that the division be specified in terms of the general sample (function), namely, the general element ω of the function space Ω, so that each sample (a received waveform) can be classified as a member of Λ_α or $\Omega - \Lambda_\alpha$. From the probability theoretical point of view, these features dictate specification of the division (or the decision scheme) to be made in terms of the "almost all sample functions" (or "almost surely," "with probability one," etc.) proposition. While this offers flexibility in one sense, it presents a restriction in another. For example, anticipating the forthcoming results, if the division of Ω is made by means of a certain ω function on Ω, this function can be arbitrary or even undefined on the sets of ω upon which P_0 and P_1 vanish. Yet, if the function is defined as a certain limit (or, obtained by a limit operation), then the sense of convergence must be at least "for almost all sample functions," but not "in quadratic mean (in the mean)," "in probability," and "in distribution," which are in general weaker.

The problem of optimum reception of binary Gaussian signals may be regarded as a generalization of an almost classical problem in communication theory, namely, optimum detection of binary *sure* signals in Gaussian noise. It is well known that such detection consists in comparing, with a preassigned threshold, the correlation integral of the received waveform and a certain function determined by the two signals and noise characteristics. More precisely, let $\{x_t, 0 \leqq t \leqq 1\}$ be a Gaussian process whose covariance function is $r(s,t)$, $0 \leqq s,t \leqq 1$, continuous and positive-definite, and whose mean function is either $m_0(t)$ or $m_1(t)$, both continuous, corresponding to the two sure signals. Denote the sample function of the process by $x(t)$ and the threshold by $c > 0$. Then Grenander[1] shows that if the integral equation

$$\int_0^1 r(s,t)g(s)\,ds = m_1(t) - m_0(t) \tag{2}$$

has a square-integrable solution, the optimum decision scheme under the Neyman-Pearson criterion is the following:

$$\begin{aligned} &\text{choose} \quad m_0(t) \quad \text{if} \quad \int_0^1 x(t)g(t)\,dt < c, \\ &\text{choose} \quad m_1(t) \quad \text{if} \quad \int_0^1 x(t)g(t)\,dt \geqq c. \end{aligned} \tag{3}$$

Suppose the two sure signals in the above problem are replaced by two *stochastic* (Gaussian) signals and the additive noise is included in these signals so that the decision between two sure signals becomes now the decision between two Gaussian signals. Furthermore, suppose the optimality criterion is changed from the Neyman-Pearson's to the error-probability minimization. Then, the problem becomes optimum reception of Gaussian signals under the minimum error-probability criterion. More precisely, let $\{x_t, 0 \leqq t \leqq 1\}$ be a Gaussian process whose mean function is identically zero and whose covariance function is either $r_0(s,t)$ or $r_1(s,t)$, continuous and positive-definite, with the accompanying a priori probabilities α and $1 - \alpha$ respectively. Then what are the counterparts of (2) and (3)? That is, under what conditions can the optimum decision scheme be specified in terms of a correlation integral involving the sample function, and what is the decision scheme itself?

1.2 *Review of Literature*

Despite momentous foundations laid by Grenander in 1950, little progress was made toward rigorous solution of the above problem during the succeeding decade, due primarily to restrictions of the mathematical scope to elementary probability theory. The majority of the work is characterized by two features: (*i*) use (and misuse) of the classical method of likelihood ratio and (*ii*) attempts to specify the decision scheme in terms of some integrals involving the sample function. In order to use the classical method, however, the continuous (parameter) process must first be "represented" by a (finite) sequence of random variables. Thus Middleton[2] and Price [3] sample $\{x_t, 0 \leqq t \leqq 1\}$ to obtain the representing sequence $x_{t_1}, \cdots, x_{t_n}$ and form their likelihood ratio l_n :

$$l_n(x_{t_1}, \cdots, x_{t_n}) = | R_0^{(n)}(R_1^{(n)})^{-1} |^{\frac{1}{2}} \exp\left\{\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}[(R_0^{(n)})^{-1} - (R_1^{(n)})^{-1}]_{ij}x_{t_i}x_{t_j}\right\}, \tag{4}$$

where $R_0^{(n)}$ and $R_1^{(n)}$ are two alternative covariance matrices of $x_{t_1}, \cdots, x_{t_n}$ given respectively by $(R_0^{(n)})_{ij} = r_0(t_i, t_j)$ and $(R_1^{(n)})_{ij} = r_1(t_i, t_j)$; $i, j = 1, \cdots, n$. Then, as $n \to \infty$ and each sampling interval becomes infinitesimal, the decision scheme is specified in terms of the limits of the exponent and the factor before the exponential in (4), provided these limits exist. Middleton argues on a formal basis that the exponent of (4) becomes an integral

$$\int_0^1 x_t y_t \, dt,$$

where the new process $\{y_t , 0 \leqq t \leqq 1\}$ is given as one of the solutions of a pair of certain simultaneous "stochastic integral equations." Price also formally argues that the exponent converges to an integral

$$\int_0^1 \int_0^1 x_s[g_1(s,t) - g_0(s,t)]x_t \, ds \, dt,$$

where g_0 and g_1 are given as solutions of a certain pair of ordinary integral equations.

Davis,[4] Bello[5] and Turin,[6] on the other hand, make orthonormal expansions of the process and use the Fourier coefficients as the representing sequence. However, the formulation of Davis and Bello is based upon a ratio of probability density functions of *two* sequences of Fourier coefficients corresponding to two separate orthonormal expansions, which is *not* a likelihood ratio; while the fundamental notion in Turin's formulation is "probability density functions of processes," which are unbounded functions in general.

One difficulty common among all the papers is the total absence of convergence proofs for series of random variables. As mentioned in Section 1.1, the sense of convergence must be "for almost all sample functions." Yet, for example, it is not clear on what ground the exponent of (4) should converge for almost all sample functions to those stochastic integrals, nor is the existence of the integrals themselves shown.

The other common difficulty, of a more fundamental nature, is the lack of optimality proofs. Considering the process as an ensemble of "well behaved" functions of time, it is intuitively plausible that such an ensemble should be "adequately" described by the distributions of the "infinitely densely" sampled values of the member functions or by the distributions of the Fourier coefficients of some orthogonal expansions in $\mathfrak{L}_2$ (the space of square-integrable functions). Namely, the continuous (parameter) process should somehow be "representable" by a sequence (infinite in general) of random variables. However, the optimality of the resultant decision scheme should in general be affected by selection of the representing sequence. Obviously, there are innumerable ways of sampling the process, resulting in innumerable decision schemes. Similarly, there are as many sequences of Fourier coefficients to represent the process as orthonormal bases of $\mathfrak{L}_2$. Yet, should all the representing sequences eventually yield the decision schemes with the same error probability, the minimum? If not, which sequences are the best representations? Even if the best sequence is chosen, on what grounds will the error probability remain minimum in the limit as $n \to \infty$, since, after all, the classical method is valid only for a finite n?

Note that there is no a priori need for the use of either likelihood ratios or representations, so long as the proposed decision scheme is shown to have the minimum error probability. In fact, Slepian[7] shows interesting special examples (of the singular case) where minimality of the error probability is explicitly proved. From a different point of view, Parzen recently restores Grenander's basic formulation, where what is called the Radon-Nikodym derivative plays the role of the likelihood ratio in the classical theory, and puts the sampling method on a more rigorous basis.

1.3 *Summary of Main Results and Remarks*

Solution of the problem of optimum reception stated in Section 1.1 rests on the following two fundamental (measure theoretical) facts:

(a) If P_0 and P_1 are two Gaussian (probability) measures, they must be either (*i*) "equivalent," i.e., $P_0 \equiv P_1$, or (*ii*) "orthogonal" (or "singular"), i.e., $P_0 \perp P_1$.

(*b*) If P_0 and P_1 are equivalent, there exists a certain nonnegative random variable $f(\omega)$, called the Radon-Nikodym derivative of P_1 with respect to P_0, and a set of ω points in Ω such that $f(\omega) \geqq \alpha/(1 - \alpha)$ can be taken as the desired critical region, denoted by Λ_α in Section 1.1. On the other hand, if P_0 and P_1 are orthogonal, there exists a set H of ω points in Ω such that $P_0(H) = 0$ and $P_1(H) = 1$, and the critical region can be taken to be such a set H. In short, the following set S_α serves as the critical region:

$$S_\alpha = \begin{cases} \{f(\omega) \geqq \alpha/(1 - \alpha)\} & \text{if} \quad P_0 \equiv P_1 , \\ H & \text{if} \quad P_0 \perp P_1 . \end{cases} \tag{5}$$

Thus, the problem of determining the critical region now becomes the problem of finding such a random variable and a set H.

Next, through the use of theory of martingales, the following facts can be established:

For almost all sample functions,

(*i*) if (and only if)

$$\lim_{n\to\infty} \operatorname{tr}\,[(R_0^{(n)})^{-1} R_1^{(n)} - 2I - R_0^{(n)} (R_1^{(n)})^{-1}] < \infty, \tag{6}$$

then

$$\lim_{n\to\infty} l_n(x_{t_1}, \cdots, x_{t_n}) = f(\omega) \quad \text{under both hypotheses;*} \tag{7}$$

(*ii*) if (and only if) (6) is not satisfied, then

$$\lim_{n\to\infty} l_n(x_{t_1}, \cdots, x_{t_n}) = \begin{cases} 0 & \text{under null hypothesis,} \\ \infty & \text{under alternative hypothesis,} \end{cases} \tag{8}$$

provided that the sequence $\{t_i\}$ is dense in the interval $0 \leqq t \leqq 1$, where "tr" stands for "trace" and the likelihood ratio l_n, together with $R_0^{(n)}$ and $R_1^{(n)}$, is previously defined in (4).†

Examination of (7) and (8) in conjunction with (5) immediately leads to the conclusion that, irrespective of the hypotheses,

$$S_\alpha = \{\lim_{n\to\infty} l_n(x_{t_1}, \cdots, x_{t_n}) \geqq \alpha/(1-\alpha)\}. \tag{9}$$

Thus, if $x(t_1), \cdots, x(t_n)$ are the values of the sample function (the received waveform) $x(t)$, $0 \leqq t \leqq 1$, sampled arbitrarily but with the restriction that each sampling interval becomes infinitesimal as $n \to \infty$, then the optimum decision scheme becomes the following:

$$\begin{aligned} &\text{choose} \quad r_0(s,t) \quad \text{if} \quad \lim_{n\to\infty} l_n[x(t_1), \cdots, x(t_n)] < \alpha/(1-\alpha), \\ &\text{choose} \quad r_1(s,t) \quad \text{if} \quad \lim_{n\to\infty} l_n[x(t_1), \cdots, x(t_n)] \geqq \alpha/(1-\alpha). \end{aligned} \tag{10}$$

Furthermore, according to (*i*), if the given covariance functions $r_0(s,t)$ and $r_1(s,t)$ are such that (6) is satisfied by their covariance matrices $R_0^{(n)}$ and $R_1^{(n)}$ obtained through sampling, then, regardless of whether $r_0(s,t)$ or $r_1(s,t)$ is the true covariance function, the above limit is finite for almost all sample functions, and the error probability associated with the decision scheme (10) is minimum. According to (*ii*), on the other hand, if $r_0(s,t)$ and $r_1(s,t)$ are such that $R_0^{(n)}$ and $R_1^{(n)}$ do not satisfy (6), then for almost all sample functions the limit vanishes if $r_0(s,t)$ is true, while the limit diverges if $r_1(s,t)$ is true; and, independent of the given a priori probabilities, the associated error probability simply vanishes, resulting in the case of "perfect reception."

* Recall that the null hypothesis is the hypothesis that $r_0(s,t)$ is the true covariance function of the process while the alternative is the hypothesis that $r_1(s,t)$ is the true covariance function.

† (6) and (7) are also found in Parzen.[8]

It should be noted, first of all, that the sequence of sampled values is not used to represent the continuous process but to obtain the crucial random variable f and set H through formation of the likelihood ratio. Secondly, under the assumption of the covariance functions being continuous, it can be proved that, regardless of the sampling manner, the limit of the likelihood ratio satisfies either (7) or (8), thus yielding the same error probability, so long as each sampling interval becomes infinitesimal as $n \to \infty$.* Lastly, negation of condition (6) can be regarded as a necessary and sufficient condition for perfect reception.

Having obtained the optimum decision scheme (10), the question of possible simplification naturally arises next. Examination of the form of the likelihood ratio (4) suggests that, if the limits of the exponent and the factor before the exponential exist separately, decision scheme (10) may be rewritten in terms of these limits. Such an attempt already appears in the literature, as mentioned in Section 1.2. However, the crucial mathematical consideration hinges upon the condition under which such a procedure can be justified. Here, the following condition is shown to be necessary and sufficient:

$$\begin{aligned} &\lim_{n\to\infty} \operatorname{tr} [(R_0^{(n)})^{-1} R_1^{(n)} - I] < \infty \\ &\lim_{n\to\infty} \operatorname{tr} [R_0^{(n)} (R_1^{(n)})^{-1} - I] < \infty . \end{aligned} \tag{11}$$

Note that this condition implies (6), as it should, and excludes the case of perfect reception. In fact, condition (11) states not only that the sum of two traces converge as condition (6) requires, but also that the two traces converge individually. In conclusion: If condition (11) is satisfied, then there exist a positive constant β and a random variable θ such that

$$\beta = \lim_{n\to\infty} | R_0^{(n)} (R_1^{(n)})^{-1} |, \tag{12}$$

$$\theta = \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} [(R_0^{(n)})^{-1} - (R_1^{(n)})^{-1}]_{ij}\, x_{t_i} x_{t_j} \tag{13}$$

for almost all sample functions under both hypotheses; and the optimum decision scheme (10) is reduced to the following:

$$\begin{aligned} &\text{choose} \quad r_0(s,t) \quad \text{if} \quad \theta(x) < \log (1/\beta)[\alpha/(1-\alpha)]^2, \\ &\text{choose} \quad r_1(s,t) \quad \text{if} \quad \theta(x) \geqq \log (1/\beta)[\alpha/(1-\alpha)]^2, \end{aligned} \tag{14}$$

* This does not imply that two different decision schemes yield the same decision for every sample function; rather, a set N of sample functions, for which two decisions differ, give no contribution to the error probability, i.e., $P_0(N) = 0 = P_1(N)$.

where $\theta(x)$ is the value of θ for the sample function $x(t)$, which is obtained by simply replacing x_{t_i} and x_{t_j} in (13) by $x(t_i)$ and $x(t_j)$.

Although the above decision scheme is certainly a step toward simplification compared with (10), it is still inconvenient, if not unfeasible, for physical application, since it requires limit operations for each received waveform. Yet, so long as the likelihood ratio is formed in terms of the sampled values, elimination of the limit operation appears to be impossible. Recall, however, the problem of optimum detection of sure signals in noise mentioned in Section 1.1. There, the likelihood ratio is formed in terms of the Fourier coefficients of the so-called Karhunen-Loéve expansion of the process instead, thus resulting in the decision scheme specified in terms of an *integral* in place of an *infinite series*, as shown by (3). Needless to say, in the present problem where there are two covariance functions instead of one, additional mathematical complications should be inevitable. Nevertheless, an optimum decision scheme which is essentially comparable to (3) can be obtained, as will now be shown.

Let $\lambda_1 \geqq \lambda_2 \geqq \cdots$ and $\psi_1(t), \psi_2(t), \cdots$ be the eigenvalues and the orthonormal eigenfunctions associated with the covariance function $r_0(s,t)$, and, similarly, let $\mu_1 \geqq \mu_2 \geqq \cdots$ and $\varphi_1(t), \varphi_2(t), \cdots$ be those associated with $r_1(s,t)$. Then, it can be shown that, under the assumption of $r_0(s,t)$ and $r_1(s,t)$ being continuous and positive-definite, the integrals

$$\xi_i = \int_0^1 x_t \psi_i(t)\, dt, \qquad i = 1, 2, \cdots, \tag{15}$$

exist for almost all sample functions under both hypotheses, and are Gaussian distributed with means zero. Furthermore, the covariance matrix determining the joint distribution of $\xi_1, \cdots, \xi_n$ is given by either

$$\begin{aligned} (Q_0^{(n)})_{ij} &= \lambda_i \delta_{ij}, \quad \text{or} \\ (Q_1^{(n)})_{ij} &= a_{ij} = \sum_{k=1}^{\infty} \mu_k u_{ki} u_{kj}, \qquad u_{ij} = \int_0^1 \varphi_i(t)\psi_j(t)\, dt, \end{aligned} \tag{16}$$

depending upon which one of $r_0(s,t)$ and $r_1(s,t)$ is the true covariance function of the process.

Thus the likelihood ratio of $\xi_1, \cdots, \xi_n$ becomes

$$\hat{l}_n = | Q_0^{(n)} (Q_1^{(n)})^{-1} |^{\frac{1}{2}} \exp \left\{ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} [(Q_0^{(n)})^{-1} - (Q_1^{(n)})^{-1}]_{ij} \xi_i \xi_j \right\}, \tag{17}$$

which corresponds to (4). It turns out that, under the previous assumption on the covariance functions, there is a complete parallel between the two formulations, one based upon $x_{t_1}, \cdots, x_{t_n}$ and the other based upon $\xi_1, \cdots, \xi_n$. Thus, for almost all sample functions,

(*i*) if (and only if)

$$\lim_{n\to\infty} \operatorname{tr} [(Q_0^{(n)})^{-1} Q_1^{(n)} - 2I - Q_0^{(n)} (Q_1^{(n)})^{-1}] < \infty, \tag{18}$$

then

$$\lim_{n\to\infty} \hat{l}_n(\xi_1, \cdots, \xi_n) = f(\omega) \quad \text{under both hypotheses;} \tag{19}$$

(*ii*) if (and only if) (18) is not satisfied, then

$$\lim_{n\to\infty} \hat{l}_n(\xi_1, \cdots, \xi_n) = \begin{cases} 0 & \text{under null hypothesis,} \\ \infty & \text{under alternative hypothesis.} \end{cases} \tag{20}$$

Then, the optimum decision scheme corresponding to (10) becomes:

$$\begin{aligned} &\text{choose} \quad r_0(s,t) \quad \text{if} \quad \lim_{n\to\infty} \hat{l}_n[\xi_1(x), \cdots, \xi_n(x)] < \alpha/(1-\alpha), \\ &\text{choose} \quad r_1(s,t) \quad \text{if} \quad \lim_{n\to\infty} \hat{l}_n[\xi_1(x), \cdots, \xi_n(x)] \geqq \alpha/(1-\alpha), \end{aligned} \tag{21}$$

where $\xi_i(x)$, $i = 1, \cdots, n$, are the values of the random variables ξ_i for the sample function $x(t)$, namely,

$$\xi_i(x) = \int_0^1 x(t)\psi_i(t)\, dt.$$

Again, note first the role of $\{\xi_i\}$, which is not a representing sequence of the process but a means for obtaining the crucial random variable f and set H by forming the likelihood ratio. Secondly, it can be shown that, under the assumption of the two covariance functions being continuous and positive-definite, $\{\varphi_i(t)\}$ can be used in place of $\{\psi_i(t)\}$ to form $\{\xi_i\}$, but *not* any orthonormal basis of $\mathfrak{L}_2$. Lastly, as before, negation of (18) can be interpreted as a necessary and sufficient condition for perfect reception. Completing the parallel, if (and only if)

$$\begin{aligned} &\lim_{n\to\infty} \operatorname{tr} [(Q_0^{(n)})^{-1} Q_1^{(n)} - I] < \infty, \\ &\lim_{n\to\infty} \operatorname{tr} [Q_0^{(n)} (Q_1^{(n)})^{-1} - I] < \infty, \end{aligned} \tag{22}$$

then there exist $\hat{\beta}$ and $\hat{\theta}$ such that

$$\hat{\beta} = \lim_{n\to\infty} | Q_0^{(n)} (Q_1^{(n)})^{-1}|, \tag{23}$$

$$\hat{\theta} = \lim_{n \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{n} [(Q_0^{(n)})^{-1} - (Q_1^{(n)})^{-1}]_{ij} \, \xi_i \xi_j \tag{24}$$

for almost all sample functions under both hypotheses; and decision scheme (21) is reduced to the following:

$$\begin{aligned} &\text{choose} \quad r_0(s,t) \quad \text{if} \quad \hat{\theta}(x) < \log (1/\hat{\beta})[\alpha/(1 - \alpha)]^2, \\ &\text{choose} \quad r_1(s,t) \quad \text{if} \quad \hat{\theta}(x) \geqq \log (1/\hat{\beta})[\alpha/(1 - \alpha)]^2. \end{aligned} \tag{25}$$

Returning to the original goal of eliminating the limit operation, examination of (24) immediately suggests the possibility of rewriting $\hat{\theta}$ as a quadratic form in x_t. That is, if one defines

$$h^{(n)}(s,t) = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{(n)} \psi_i(s) \psi_j(t), \tag{26}$$

where

$$h_{ij}^{(n)} = [(Q_0^{(n)})^{-1} - (Q_1^{(n)})^{-1}]_{ij},$$

then, from (15),

$$\hat{\theta} = \lim_{n \to \infty} \int_0^1 \int_0^1 x_s h^{(n)}(s,t) x_t \, ds \, dt, \tag{27}$$

and $h_{ij}^{(n)}$; $i,j = 1, \cdots, n$, can be given as a solution of the matrix equation

$$Q_0^{(n)} \, (h_{ij}^{(n)}) \, Q_1^{(n)} = Q_1^{(n)} - Q_0^{(n)},$$

or, more directly, $h^{(n)}(s,t)$ can be given as a solution of the integral equation

$$\int_0^1 \int_0^1 r_0^{(n)}(s,u) h^{(n)}(u,v) r_1^{(n)}(v,t) \, du \, dv = r_1^{(n)}(s,t) - r_0^{(n)}(s,t), \tag{28}$$

where

$$r_0^{(n)}(s,t) = \sum_{i=1}^{n} \lambda_i \psi_i(s) \psi_i(t), \quad r_1^{(n)}(s,t) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \psi_i(s) \psi_j(t). \tag{29}$$

Then, the following conjecture should be imminent:

$$\hat{\theta} = \int_0^1 \int_0^1 x_s h(s,t) x_t \, ds \, dt, \tag{30}$$

where $h(s,t)$ is a solution of

$$\int_0^1 \int_0^1 r_0(s,u)h(u,v)r_1(v,t)\, du\, dv = r_1(s,t) - r_0(s,t), \tag{31}$$

which are formally the limits of (27) and (28) respectively. The essential part of the above conjecture can be shown to be correct. That is, if (31) has a solution $h(s,t)$ such that

$$\begin{aligned} &\int_0^1 \int_0^1 h^2(s,t)\, ds\, dt < \infty, \quad \text{then} \\ &\int_0^1 \int_0^1 x_s h(s,t) x_t\, ds\, dt = \lim_{n\to\infty} \sum_{i=1}^n \sum_{j=1}^n h_{ij}^{(n)} \xi_i \xi_j \end{aligned} \tag{32}$$

for almost all sample functions under both hypotheses, provided that, for all $i, j = 1, 2, \cdots,$

$$a_{ii} < 1, \qquad a_{ii} > \sum_{j=1}^{\infty}{}' \mid a_{ij} \mid, \qquad \frac{\left| \dfrac{a_{ij}}{\lambda_i} - \delta_{ij} \right|}{1 - \sum_{k=1}^{\infty} \mid \delta_{jk} - a_{jk} \mid} \leqq K, \tag{33}$$

where K is a positive constant independent of i and j.
Then the optimum decision scheme (25) is immediately reduced to the following desired form:

$$\begin{aligned} &\text{choose} \quad r_0(s,t) \quad \text{if} \quad \int_0^1 \int_0^1 x(s)h(s,t)x(t)\, ds\, dt < \log \frac{1}{\hat{\beta}} \left(\frac{\alpha}{1-\alpha} \right)^2, \\ &\text{choose} \quad r_1(s,t) \quad \text{if} \quad \int_0^1 \int_0^1 x(s)h(s,t)x(t)\, ds\, dt \geqq \log \frac{1}{\hat{\beta}} \left(\frac{\alpha}{1-\alpha} \right)^2. \end{aligned} \tag{34}$$

Difficulty of the proof lies mainly in the fact that, as n increases, the coefficients $h_{ij}^{(n)}$ themselves vary with n as well as the number of the terms of the sum, yet $h^{(n)}(s,t)$ must approach $h(s,t)$ in such a way that

$$\lim_{n\to\infty} \int_0^1 \int_0^1 x_s h^{(n)}(s,t) x_t\, ds\, dt = \int_0^1 \int_0^1 x_s h(s,t) x_t\, ds\, dt$$

for almost all sample functions under both hypotheses. This accounts for need of the auxiliary conditions (33). The first condition is not a restriction in physical application since

$$\sum_{i=1}^{\infty} a_{ii} = \sum_{j=1}^{\infty} \mu_i = \int_0^1 r_1(t,t)\, dt$$

is the average energy of the waveform in the interval $0 \leqq t \leqq 1$, which can always be normalized to assure $a_{ii} < 1$. Although the remaining two conditions are restrictive, current knowledge of infinite systems of equations does not seem to allow their removal. Thus this calls for a future investigation of the degree of restriction imposed by them in physicial application.

As anticipated, there is an apparent correspondence between the classical case of sure signals in noise mentioned in Section 1.1 and the present case of stochastic signals, namely, between (2)–(3) and (31)–(34), except for the fact that the constituent functions in the latter case are functions of two variables instead of one. As the integral of the decision scheme (3) has a simple physical interpretation (the output of a linear filter with $g(t)$ as its impulse response), so does the integral in (34). Namely, it is the output of a quadratic filter whose impulse response is $h(s,t)$. The advantage of this scheme over the others — namely, (10, (14), (21) and (25) — is obvious. Given two covariance functions, the impulse response of the filter is uniquely determined by the integral equation (31) if a solution exists, and decision is made by comparing, with a preassigned threshold, the appropriately sampled output of the filter with the received waveform as its input, instead of having to perform the limit operation for each received waveform.

Finally, it should be remarked that the optimum decision scheme above differs *formally* from those previously obtained by others.* A further, and more significant, distinction lies in the assured optimality of this scheme, inherent in its derivation, while the optimality of the others has yet to be proved separately.†

II. MATHEMATICAL THEORY

2.1 *Gaussian Processes*

Let $\{x_t, t \in T\}$ be a real Gaussian process with a parameter set $T = [0,1]$ and a finite dimensional distribution function $F_{t_1,\cdots,t_n}$, which is determined by given mean function and covariance function where $t_1, \cdots, t_n$ are an arbitrary finite subset of T. It is assumed that the mean function is identically zero on T while the covariance function is positive-definite and continuous on $T \times T$. In the present problem it is desirable to have an explicit representation of the given process $\{x_t, t \in T\}$ on a function space.‡

Let Ω be a space of real-valued functions of $t \in T$. Let $x_s(\omega)$ be the

* Although their work is briefly reviewed in Section 1.2, their decision schemes are not stated explicitly in this paper.

† This excludes Parzen's[8] case where the decision scheme is essentially (10).

‡ The next paragraph follows closely Example 2.3 in Supplement, Doob,[9] pp. 609–610.

ω function with the value $\zeta(s)$ if ω is the function $\zeta(\cdot)$, so that $x_s(\omega) = \zeta(s)$. If the t function ω has values $\zeta(t_1), \cdots, \zeta(t_n)$ at $t_1, \cdots, t_n$, the condition

$$\zeta(t_1) \leqq \rho_1, \cdots, \zeta(t_n) \leqq \rho_n$$

defines an ω set, which is denoted by

$$\{x_{t_i}(\omega) \leqq \rho_i, \qquad i = 1, \cdots, n\} \tag{35}$$

where $\rho_1, \cdots, \rho_n$ are arbitrary real numbers. Next, let $\mathfrak{F}$ be the class of all ω sets obtained in this way for arbitrary n, $t_1, \cdots, t_n$, and let $\mathcal{B}_T$ be the Borel field generated by $\mathfrak{F}$, and lastly let P be a probability measure defined on the sets of $\mathcal{B}_T$ whose value is given by

$$P\{x_{t_i}(\omega) \leqq \rho_i, i = 1, \cdots, n\} = F_{t_1,\cdots,t_n}(\rho_1, \cdots, \rho_n). \tag{36}$$

Then, $\{x_t(\omega), t \in T\}$ is a representation of the given process $\{x_t, t \in T\}$ on the function space Ω, and $(\Omega, \mathcal{B}_T, P)$ is the explicit probability measure space for the representation.*

(Remark) By virtue of the choice of representation space, the general elements of the space Ω coincide with the general sample functions of the process $\{x_t(\omega), t \in T\}$. Thus, the phrases, "almost everywhere (or almost surely)" and "for almost all sample functions," have the same meaning.

The assumption of continuous covariance function has the following significant consequences:

(*i*) $\{x_t(\omega), t \in T\}$ has an equivalent (with respect to P) separable and measurable process on the same ω space.† Hence, so long as the almost-everywhere valid properties of a given process are of interest, as in the case of this paper, the given process may as well be taken to be separable and measurable. Therefore, the Gaussian process $\{x_t(\omega), t \in T\}$ is henceforth assumed to be separable and measurable.

(*ii*) $\{x_t(\omega), t \in T\}$ is sample (Lebesgue) square-integrable on T almost everywhere with respect to P.‡

This immediately implies that a Lebesgue integral

* Symbolic distinction between the given process and its representation on the function space is made by explicitly writing the argument ω for the latter.

† Note continuity of the covariance function of a process is equivalent to continuity in quadratic mean of the process (Loéve,[10] p. 470), and hence it implies continuity in probability of the process. Then, according to Theorem 2.6 in Doob,[9] pp. 61–62, there exists an equivalent separable and measurable process on the same space.

‡ See Loéve,[10] pp. 520–521.

$$\xi(\omega) = \int_T x_t(\omega)\psi(t)\,dt$$

exists almost everywhere, in which $\psi(t)$ is any continuous function on T. Furthermore, since the sample Lebesgue integral of a process coincides almost everywhere with the Riemann integral in quadratic mean criterion,* and also the Riemann integral in quadratic mean criterion of a Gaussian process is a Gaussian (random) variable,† $\xi(\omega)$ is a Gaussian variable.

2.2 *Formulation of Problem*

Let $F_{0;t_1,\cdots,t_n}$ and $F_{1;t_1,\cdots,t_n}$ be two alternative Gaussian finite dimensional distribution functions of a real separable and measurable process $\{x_t(\omega),\ t \in T\}$, whose mean functions are identically zero and whose covariance functions, denoted respectively by $r_0(s,t)$ and $r_1(s,t)$, are positive-definite and continuous on $T \times T$. Let P_0 and P_1 be the Gaussian probability measures defined respectively by $F_{0;t_1,\cdots,t_n}$ and $F_{1;t_1,\cdots,t_n}$ on the Borel field $\mathcal{B}_T$ of subsets of Ω as defined previously. It is well known that P_0 and P_1 are either equivalent, $P_0 \equiv P_1$, or orthogonal, $P_0 \perp P_1$.‡

Define a set function P_e by

$$P_e(\Lambda) = \alpha P_0(\Lambda) + (1 - \alpha)P_1(\Omega - \Lambda), \qquad \Lambda \in \mathcal{B}_T, \quad (37)$$

where α is a constant, $0 < \alpha < 1$.§ Let $\Lambda_\alpha \in \mathcal{B}_T$ be such a set that

$$P_e(\Lambda_\alpha) \leqq P_e(\Lambda) \qquad \text{for all } \Lambda \in \mathcal{B}_T. \quad (38)$$

Then, the problem of interest is to specify such a set Λ_α in terms of $x_t(\omega)$.‖

Now, if $P_0 \equiv P_1$, let $f(\omega)$ be a Radon-Nikodym derivative of P_1 with respect to P_0; while, if $P_0 \perp P_1$, let $H \in \mathcal{B}_T$ be a set such that $P_0(H) = 0$ and $P_1(H) = 1$. Then, it can be shown that the following

---

* Henceforth, the "sample Lebesgue integral of a process" will simply be called the "integral of a process," unless otherwise specified. A definition of Riemann integral in quadratic mean criterion is in Loéve,[10] pp. 471–474.

† See Loéve,[10] p. 485.

‡ See Hajek.[11,12]

§ P_e is the so-called error probability. Although $0 \leqq P_e \leqq 1$ for all $\Lambda \in \mathcal{B}_T$, P_e is not a probability measure, and its full meaning is given in Section 1.1.

‖ Equivalence between this problem and that of "optimum reception of binary Gaussian processes" is discussed in detail in Section I.

set S_α satisfies condition (38):*

$$\begin{aligned} &\text{if } P_0 \equiv P_1, \quad S_\alpha = \{f(\omega) \geqq \alpha/(1-\alpha)\}, \\ &\text{if } P_0 \perp P_1, \quad S_\alpha = H. \end{aligned} \tag{39}$$

Thus the above stated problem is reduced to that of finding

(i), if $P_0 \equiv P_1$, a function of $x_t(\omega)$ equal to $f(\omega)$ almost everywhere with respect to P_0 and P_1, and

(ii), if $P_0 \perp P_1$, some such set H expressible in terms of $x_t(\omega)$.

2.3 *Solutions — I*

2.3.1 *General Solution*

Let $\{\tau_k\}$ be a sequence of points in $T = [0,1]$, which is dense in T. Let $\mathcal{B}_n$ be a Borel field generated by a class of ω sets of the form

$$\{x_{\tau_i}(\omega) \leqq \rho_i, \quad i = 1, \cdots, n\}, \tag{40}$$

and let $\mathcal{B}_\infty$ be the minimal Borel field containing $\bigcup_{n=1}^{\infty} \mathcal{B}_n$. Obviously,

$$\mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots \subset \mathcal{B}_\infty \subset \mathcal{B}_T. \tag{41}$$

Then, since $\{x_t(\omega), t \in T\}$ is a separable process, continuous in probability (with respect to P_0 and P_1), and the sequence $\{\tau_k\}$ is dense in T, it follows that, for an arbitrary set $\Lambda \in \mathcal{B}_T$, there exists a set $\Lambda' \in \mathcal{B}_\infty$ such that

$$P_0(\Lambda\Delta\Lambda') = 0 = P_1(\Lambda\Delta\Lambda'). \tag{42}†$$

Now, through the use of the covariance functions $r_0(s,t)$ and $r_1(s,t)$ and the fact that the mean functions are identically zero, the density functions p_0 and p_1 of the random variables $x_{\tau_i}(\omega)$, $i = 1, \cdots, n$, corresponding to P_0 and P_1 respectively, are obtained as follows:

$$\begin{aligned} p_m(\nu_1, \cdots, \nu_n) &= (2\pi)^{-(n/2)} \mid R_m^{(n)} \mid^{-\frac{1}{2}} \\ &\times \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} [(R_m^{(n)})^{-1}]_{ij}\,\nu_i\nu_j\right\}, \qquad m = 0, 1, \end{aligned} \tag{43}$$

where the τ_i, $i = 1, \cdots, n$, are a finite subset of $\{\tau_k\}$, and $R_m^{(n)}$, $m = 0$, 1, are $n \times n$ symmetric, positive-definite matrices defined by

$$(R_m^{(n)})_{ij} = r_m(\tau_i, \tau_j); \qquad m = 0, 1; \qquad i,j = 1, \cdots, n. \tag{44}$$

* See Appendix A. The first assertion of (39) follows from Corollary 1 in this appendix, while the second assertion is self-evident.

† See Doob,[9] pp. 51–55; in particular, Theorem 2.2 (i).

Then define a random variable $l_n(\omega)$ by

$$l_n(\omega) = \frac{p_1[x_{\tau_1}(\omega), \cdots, x_{\tau_n}(\omega)]}{p_0[x_{\tau_1}(\omega), \cdots, x_{\tau_n}(\omega)]} = | R_0^{(n)}(R_1^{(n)})^{-1} |^{\frac{1}{2}} \times \exp\left\{\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}[(R_0^{(n)})^{-1} - (R_1^{(n)})^{-1}]_{ij} x_{\tau_i}(\omega) x_{\tau_j}(\omega)\right\}. \tag{45}$$

Note that

$$l_n(\omega) \geqq 0 \quad \text{for all } n. \tag{46}$$

Furthermore, since $R_m^{(n)}$, $m = 0, 1$, are positive-definite, $p_1 = 0$ whenever $p_0 = 0$ and vice versa. Then, it can be shown that the processes $\{l_n(\omega), n \geqq 1\}$ and $\{1/l_n(\omega), n \geqq 1\}$ are martingales with respect to P_0 and P_1 respectively.*

(*i*) $P_0 \equiv P_1$: Let $E_0\{f(\omega) | \mathcal{B}_n\}$, $n = 1, 2, \cdots$, be a conditional expectation of $f(\omega)$, given $\mathcal{B}_n$, with respect to P_0. Namely,

$$\int_\Lambda E_0\{f(\omega) \mid \mathcal{B}_n\}\, dP_0 = \int_\Lambda f(\omega)\, dP_0 \quad \text{for any } \Lambda \in \mathcal{B}_n .$$

Then,

$$l_n(\omega) = E_0\{f(\omega) | \mathcal{B}_n\}, \qquad \text{a.e. } (P_0),† \tag{47}$$

and, from (41)‡

$$\lim_{n\to\infty} E_0\{f(\omega) | \mathcal{B}_n\} = E_0\{f(\omega) | \mathcal{B}_\infty\}, \qquad \text{a.e. } (P_0). \tag{48}$$

Yet, from the definition of $E_0\{f(\omega) | \mathcal{B}_\infty\}$ and (42),

$$E_0\{f(\omega) | \mathcal{B}_\infty\} = f(\omega), \qquad \text{a.e. } (P_0). \tag{49}$$

Hence,

$$\lim_{n\to\infty} l_n(\omega) = f(\omega), \qquad \text{a.e. } (P_0). \tag{50}$$

Since $P_0 \equiv P_1$, the above implies

$$\lim_{n\to\infty} l_n(\omega) = f(\omega), \qquad \text{a.e. } (P_1). \tag{51}$$

Thus, the desired function, which is equal to $f(\omega)$, a.e. (P_0, P_1), is

* See Doob,[9] pp. 91–93.

† "a.e. (P_m)," $m = 0, 1$, is used as a shorthand notation of "almost everywhere with respect to P_m." Similarly, "a.e. (P_0, P_1)" will be used to denote "almost everywhere with respect to both P_0 and P_1."

‡ See Doob,[9] p. 331.

$l_\infty(\omega)$, which is defined by

$$l_\infty(\omega) = \lim_{n\to\infty} l_n(\omega). \tag{52}$$

(*ii*) $P_0 \perp P_1$: From (46), $\lim_{n\to\infty} l_n(\omega) < \infty$, a.e. (P_0).* In fact, it can be shown that

$$\lim_{n\to\infty} l_n(\omega) = 0, \qquad \text{a.e. } (P_0).\dagger \tag{53}$$

By using the same argument, it follows that

$$\lim_{n\to\infty} [1/l_n(\omega)] = 0, \qquad \text{a.e. } (P_1). \tag{54}$$

Hence, for an arbitrary constant $c > 0$,

$$P_0\{\lim_{n\to\infty} l_n(\omega) \geqq c\} = 0, \qquad P_1\{\lim_{n\to\infty} l_n(\omega) \geqq c\} = 1.$$

Thus, the desired set H, with $P_0(H) = 0$ and $P_1(H) = 1$, is

$$H = \{\lim_{n\to\infty} l_n(\omega) \geqq \alpha/(1-\alpha)\}. \tag{55}$$

In summary, upon combination of (52) and (55) in conjunction with (39), the desired set S_α is

$$S_\alpha = \{\lim_{n\to\infty} l_n(\omega) \geqq \alpha/(1-\alpha)\}, \tag{56}$$

irrespective of whether $P_0 \equiv P_1$ or $P_0 \perp P_1$.

2.3.2 *Special Solutions and Summary*

Under certain restrictive conditions, the set S_α can be specified in terms of well defined functions of $x_t(\omega)$. It is the purpose of this subsection to obtain such specifications as well as the accompanying conditions in terms of the given covariance functions $r_0(s,t)$ and $r_1(s,t)$.

(*i*) If $P_0 \equiv P_1$, it has already been shown that

$$S_\alpha = \{l_\infty(\omega) \geqq \alpha/(1-\alpha)\}.$$

Thus, it is of interest to obtain a condition for $P_0 \equiv P_1$.‡

Define

$$\eta_n(\omega) = [l_n(\omega) - 1] \log l_n(\omega), \qquad n = 1, 2, \cdots. \tag{57}$$

---

* See Doob,[9] p. 319; Theorem 4.1 (*i*).

† See Doob,[9] pp. 345–346.

‡ Such conditions are already available (e.g., Parzen,[8] Shepp[13]). For more detail, see Yaglom.[14]

Then, since $(\rho - 1) \log \rho$, $\rho > 0$, is a real, continuous and convex function of ρ and $E_0\{|(l_n(\omega) - 1) \log l_n(\omega)|\} < \infty$, $n = 1, 2, \cdots$; $\{\eta_n(\omega), n \geqq 1\}$ constitute a semi-martingale (with respect to P_0).* Hence, $E_0\{\eta_n(\omega)\}$, $n = 1, 2, \cdots$, forms a monotone nondecreasing sequence

$$E_0\{\eta_1(\omega)\} \leqq E_0\{\eta_2(\omega)\} \leqq \cdots, \tag{58}$$ †

which must either converge or diverge. Then, according to (53),

if $P_0 \perp P_1$, then

$$\lim_{n\to\infty} E_0\{\eta_n(\omega)\} = \infty. \tag{59}$$

Hence, since P_0 and P_1 can be either equivalent or orthogonal, it follows that

$P_0 \equiv P_1$, if

$$\lim_{n\to\infty} E_0\{\eta_n(\omega)\} < \infty. \tag{60}$$

It can be shown that the converse of (59) is also true,‡ i.e.,

$$\text{if} \quad \lim_{n\to\infty} E_0\{\eta_n(\omega)\} = \infty, \qquad \text{then } P_0 \perp P_1. \tag{61}$$

This implies that the condition of (60) is also necessary. Thus, through substitution of (45) into (57) and application of (43) for expectation calculation,§

$P_0 \equiv P_1$, if and only if

$$\lim_{n\to\infty} \operatorname{tr} [(R_0^{(n)})^{-1} R_1^{(n)} - 2I + R_0^{(n)} (R_1^{(n)})^{-1}] < \infty. \tag{62}$$ ||

where $R_0^{(n)}$ and $R_1^{(n)}$ are defined in terms of $r_0(s,t)$ and $r_1(s,t)$ by (44).

(*ii*) Examination of (45), (50) and (51) indicates that, in addition to condition (62), if

$$\lim_{n\to\infty} | R_0^{(n)} (R_1^{(n)})^{-1} | = \beta, \qquad 0 < \beta < \infty, \tag{63}$$

then

* See Doob,[9] pp. 295–296, Theorem 1.1 (iii). "E_0" denotes expectation with respect to P_0, namely, an integration over Ω with respect to P_0.
† See Doob,[9] p. 324, Theorem 4.1s.
‡ See Hajek;[12] in particular, Lemma 2.1.
§ For this calculation, use the following equality: $E_0\{\eta_n(\omega)\} = E_1\{\log l_n(\omega)\} - E_0\{\log l_n(\omega)\}$, $n = 1, 2, \cdots$.
|| "tr" denotes "trace," and I is the $n \times n$ identity matrix.

$$\lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} [(R_0^{(n)})^{-1} - (R_1^{(n)})^{-1}]_{ij}\, x_{\tau_i}(\omega) x_{\tau_j}(\omega) < \infty, \qquad \text{a.e. } (P_0, P_1). \tag{64}$$

Thus, by defining $\theta(\omega)$ as the above limit, i.e.,

$$\theta(\omega) = \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} [(R_0^{(n)})^{-1} - (R_1^{(n)})^{-1}]_{ij} x_{\tau_i}(\omega) x_{\tau_j}(\omega), \tag{65}$$

the set $S_\alpha \in \mathcal{B}_T$ can be specified as follows:

$$S_\alpha = \{\theta(\omega) \geqq \log (1/\beta)(\alpha/(1-\alpha))^2\}. \tag{66}$$

It will now be shown that two conditions (62) and (63), required for the above specification of S_α, are equivalent to the following pair of conditions:

$$\lim_{n\to\infty} \operatorname{tr} [(R_0^{(n)})^{-1} R_1^{(n)} - I] < \infty,$$

and (67)

$$\lim_{n\to\infty} \operatorname{tr} [R_0^{(n)} (R_1^{(n)})^{-1} - I] < \infty.$$

Define

$$\zeta_n(\omega) = -\log l_n(\omega),$$
$$\zeta_n'(\omega) = l_n(\omega) \log l_n(\omega) \qquad n = 1, 2, \cdots. \tag{68}$$

Thus,

$$\eta_n(\omega) = \zeta_n'(\omega) + \zeta_n(\omega), \qquad n = 1, 2, \cdots. \tag{69}$$

Again, just as in the case of $\eta_n(\omega)$, both $\{\zeta_n(\omega), n \geqq 1\}$ and $\{\zeta_n'(\omega), n \geqq 1\}$ are semi-martingales with respect to P_0, and

$$E_0\{\zeta_1(\omega)\} \leqq E_0\{\zeta_2(\omega)\} \leqq \cdots,$$
$$E_0\{\zeta_1'(\omega)\} \leqq E_0\{\zeta_2'(\omega)\} \leqq \cdots. \tag{70}$$

Furthermore, from (53),

$$\text{if } P_0 \perp P_1, \qquad \text{then } \lim_{n\to\infty} E_0\{\zeta_n(\omega)\} = \infty. \tag{71}$$

However, from (69) and (70), divergence of $E_0\{\zeta_n(\omega)\}$ implies that of $E_0\{\eta_n(\omega)\}$. Hence, according to (61), the converse of (71) holds. Then, again from (70) and the equivalence-or-orthogonality dichotomy of P_0 and P_1,

$$P_0 \equiv P_1 \quad \text{if and only if} \quad \lim_{n\to\infty} E_0\{\zeta_n(\omega)\} < \infty. \tag{72}$$

Thus, upon substitution of (45) into (68) and application of (43) for expectation calculation, an alternative necessary and sufficient condition for $P_0 \equiv P_1$ is obtained as follows:

$$\lim_{n\to\infty} \{\log | (R_0^{(n)})^{-1} R_1^{(n)} | + \operatorname{tr} [R_0^{(n)} (R_1^{(n)})^{-1} - I]\} < \infty. \tag{73}$$

Now, under the condition (63), the above condition implies that

$$\lim_{n\to\infty} \operatorname{tr} [R_0^{(n)} (R_1^{(n)})^{-1} - I] < \infty. \tag{74}$$

Then, upon combination of conditions (62) and (74), condition (67) immediately follows.

The result of this section may be summarized as follows:

(i) In general,

$$S_\alpha = \{\lim_{n\to\infty} l_n(\omega) \geqq \alpha/(1 - \alpha)\},$$

where $l_n(\omega)$ is defined by (45).

(ii) If $P_0 \equiv P_1$, which is true if and only if

$$\lim_{n\to\infty} \operatorname{tr} [(R_0^{(n)})^{-1} R_1^{(n)} - 2I + R_0^{(n)} (R_1^{(n)})^{-1}] < \infty,$$

then $\lim_{n\to\infty} l_n(\omega) = f(\omega)$, a.e. (P_0, P_1); thus by defining $l_\infty(\omega) = \lim_{n\to\infty} l_n(\omega)$,

$$S_\alpha = \{l_\infty(\omega) \geqq \alpha/(1 - \alpha)\}.$$

(iii) if

$$\lim_{n\to\infty} \operatorname{tr} [(R_0^{(n)})^{-1} R_1^{(n)} - I] < \infty,$$

$$\lim_{n\to\infty} \operatorname{tr} [R_0^{(n)} (R_1^{(n)})^{-1} - I] < \infty,$$

then

$$S_\alpha = \{\theta(\omega) \geqq \log (1/\beta)(\alpha/(1 - \alpha))^2\}$$

where $\theta(\omega)$ and β are defined by (65) and (63) respectively.

2.4 *Solutions — II*

2.4.1 *General and Special Solutions*

Let $\lambda_1 \geqq \lambda_2 \geqq \cdots$ and $\psi_1(t), \psi_2(t), \cdots$ be the eigenvalues and the corresponding orthonormal eigenfunctions associated with the covariance function $r_0(s,t)$.* Similarly, let $\mu_1 \geqq \mu_2 \geqq \cdots$ and $\varphi_1(t), \varphi_2(t), \cdots$ be such eigenvalues and eigenfunctions associated with $r_1(s,t)$. Then, according to the discussion in 2.1 (*ii*), continuity on T of each $\psi_i(t)$ implies that the integrals

$$\xi_i(\omega) = \int_T x_t(\omega)\psi_i(t)\,dt, \qquad i = 1, 2, \cdots, \tag{75}$$

exist a.e. (P_0, P_1), and are Gaussian random variables. In fact, it can be shown that the density functions $\hat{p}_0$ and $\hat{p}_1$ of $\xi_1(\omega), \cdots, \xi_n(\omega)$ corresponding to P_0 and P_1 are given by†

$$\hat{p}_m(\nu_1, \cdots, \nu_n) = (2\pi)^{-(n/2)} \, | Q_m^{(n)} |^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} [(Q_m^{(n)})^{-1}]_{ij}\, \nu_i\nu_j\right\}, \qquad m = 0, 1, \tag{76}$$

where $Q_m^{(n)}$, $m = 0, 1$, are $n \times n$ symmetric and positive-definite matrices defined by

$$(Q_0^{(n)})_{ij} = \lambda_i\delta_{ij}, \qquad (Q_1^{(n)})_{ij} = \sum_{k=1}^{\infty} \mu_k u_{ki} u_{kj}, \tag{77}$$

where

$$u_{ij} = \int_T \varphi_i(t)\psi_j(t)\,dt. \tag{78}$$

Let $\hat{\mathfrak{B}}_n$ be a Borel field generated by a class of ω sets of the form

$$\{\xi_i(\omega) \leqq \rho_i, \; i = 1, \cdots, n\}, \tag{79}$$

and let $\hat{\mathfrak{B}}_\infty$ be the minimal Borel field containing $\bigcup_{n=1}^{\infty} \hat{\mathfrak{B}}_n$. Obviously,

$$\hat{\mathfrak{B}}_1 \subset \hat{\mathfrak{B}}_2 \subset \cdots \subset \hat{\mathfrak{B}}_\infty \subset \hat{\mathfrak{B}}_T. \tag{80}$$

It can be shown that, for an arbitrary $\Lambda \in \mathfrak{B}_T$, there exists some $\hat{\Lambda} \in \hat{\mathfrak{B}}_\infty$ such that

$$P_0(\Lambda\Delta\hat{\Lambda}) = 0. \tag{81‡}$$

Now define a random variable $\hat{l}_n(\omega)$ by

* More precise definitions of these eigenvalues and eigenfunctions are given in Appendix B.
† See Appendix C.
‡ See Appendix D.

$$\hat{l}_n(\omega) = \frac{\hat{p}_1[\xi_1(\omega), \cdots, \xi_n(\omega)]}{\hat{p}_0[\xi_1(\omega), \cdots, \xi_n(\omega)]}$$

$$= |Q_0^{(n)}(Q_1^{(n)})^{-1}|^{\frac{1}{2}} \exp\left\{\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}[(Q_0^{(n)})^{-1} - (Q_1^{(n)})^{-1}]_{ij}\xi_i(\omega)\xi_j(\omega)\right\}, \tag{82}$$

where (76) is substituted for the second equality. Again, note that $\hat{l}_n(\omega)$ is nonnegative for all n and also the fact that $\hat{p}_1 = 0$ whenever $\hat{p}_0 = 0$ and vice versa since $Q_m^{(n)}$, $m = 0, 1$, are positive-definite. Thus, again the processes $\{\hat{l}_n(\omega), n \geqq 1\}$ and $\{1/\hat{l}_n(\omega), n \geqq 1\}$ are martingales with respect to P_0 and P_1 respectively.

By following step-by-step the same procedure as the one in the preceding section,* the following results are obtained:

(*i*) In general,

$$S_\alpha = \{\lim_{n\to\infty} \hat{l}_n(\omega) \geqq \alpha/(1 - \alpha)\}. \tag{83}$$

(*ii*) If $P_0 \equiv P_1$, which is true if and only if

$$\lim_{n\to\infty} \operatorname{tr}\, [(Q_0^{(n)})^{-1}\, Q_1^{(n)} - 2I + Q_0^{(n)}\, (Q_1^{(n)})^{-1}] < \infty, \tag{84}$$

then

$$\lim_{n\to\infty} \hat{l}_n(\omega) = f(\omega), \qquad \text{a.e. } (P_0, P_1); \tag{85}$$

thus by defining

$$\hat{l}_\infty(\omega) = \lim_{n\to\infty} \hat{l}_n(\omega), \tag{86}$$

$$S_\alpha = \{\hat{l}_\infty(\omega) \geqq \alpha/(1 - \alpha)\}. \tag{87}$$

(*iii*) If

$$\begin{aligned} &\lim_{n\to\infty} \operatorname{tr}\, [(Q_0^{(n)})^{-1}\, Q_1^{(n)} - I] < \infty, \\ &\lim_{n\to\infty} \operatorname{tr}\, [Q_0^{(n)}\, (Q_1^{(n)})^{-1} - I] < \infty, \end{aligned} \tag{88}$$

then there exists a constant $\hat{\beta}$, $0 < \hat{\beta} < \infty$, such that

* In effect, it amounts to replacing $\mathcal{B}_n$ and $l_n(\omega)$, $n = 1, 2, \cdots n$, by $\hat{\mathcal{B}}$ and $\hat{l}_n(\omega)$ respectively.

$$\lim_{n\to\infty} | Q_0^{(n)} (Q_1^{(n)})^{-1} | = \hat{\beta}; \tag{89}$$

and, from (85) and (82), it follows that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} [(Q_0^{(n)})^{-1} - Q(_1^{(n)})^{-1}]_{ij}\xi_i(\omega)\xi_j(\omega) < \infty, \quad \text{a.e. } (P_0, P_1); \tag{90}$$

thus, by defining $\hat{\theta}(\omega)$ as the above limit,

$$S_\alpha = \{\hat{\theta}(\omega) \geqq \log (1/\hat{\beta})(\alpha/(1 - \alpha))^2\}. \tag{91}$$

2.4.2 *Integral Expression for* $\hat{\theta}(\omega)$

For the purpose of physical application, it is desirable to express the random variable $\hat{\theta}(\omega)$ as a simpler function of $x_t(\omega)$, in particular, without involving limit operation. Examination of the definition of $\hat{\theta}(\omega)$, i.e.,

$$\hat{\theta}(\omega) = \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} [(Q_0^{(n)})^{-1} - (Q_1^{(n)})^{-1}]_{ij}\xi_i(\omega)\xi_j(\omega), \tag{92}$$

indicates that $\hat{\theta}(\omega)$ might be expressible as a quadratic form in $x_t(\omega)$, i.e.,

$$\int_T\int_T x_s(\omega)h(s,t)x_t(\omega)\, ds\, dt$$

if such a square-integrable function $h(s,t)$, $(s,t) \in T \times T$, exists and can be determined uniquely. It is the purpose of this subsection to make the above statement more definite and precise.

Define an $n \times n$ symmetric matrix $H^{(n)}$ by

$$H^{(n)} = (Q_0^{(n)})^{-1} - (Q_1^{(n)})^{-1}.$$

Then,

$$Q_0^{(n)}H^{(n)}Q_1^{(n)} = Q_1^{(n)} - Q_0^{(n)},$$

or, through (77), the equation for the i-jth element becomes

$$\sum_{k=1}^{n} \lambda_i(H^{(n)})_{ik}(Q_1^{(n)})_{kj} = (Q_1^{(n)}{}_{ij} - \lambda_i\delta_{ij}; \qquad i, j = 1, \cdots, n.$$

In other words, every ith row of $H^{(n)}$ satisfies the following system of equations:*

$$\sum_{k=1}^{n} a_{jk}h_{ik}^{(n)} = b_j(i) \qquad j = 1, \cdots, n,$$

* Note that the solution is unique, since the matrix (a_{ij}) is positive-definite.

where

$$a_{ij} = \sum_{k=1}^{\infty} \mu_k u_{ki} u_{kj}, \qquad i, j = 1, 2, \cdots, \qquad (93)^*$$
$$b_j(i) = (a_{ij}/\lambda_i) - \delta_{ij},$$

or its standard form

$$h_{ij}^{(n)} = \sum_{k=1}^{n} c_{jk} h_{ik}^{(n)} + b_j(i), \qquad j = 1, \cdots, n, \qquad (94)$$

where

$$c_{ij} = \delta_{ij} - a_{ij}.$$

Now, for each $i = 1, 2, \cdots$, consider the following infinite system of equations:

$$h_{ij} = \sum_{k=1}^{\infty} c_{jk} h_{ik} + b_j(i), \qquad j = 1, 2, \cdots. \qquad (95)$$

According to the theory of infinite systems of equations,† if (95) has a solution $(h_{i1}, h_{i2}, \cdots)$ for each $i = 1, 2, \cdots$, such that

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij}^2 < \infty, \qquad (96)$$

then $(h_{i1}, h_{i2}, \cdots)$ is unique and

$$h_{ij} = \lim_{n \to \infty} h_{ij}^{(n)}, \qquad j = 1, 2, \cdots, \qquad (97)$$

for each $i = 1, 2, \cdots$, where $(h_{i1}^{(n)}, \cdots, h_{in}^{(n)})$, $i = 1, \cdots, n$, is the solution of (94); provided that (95) satisfies the following conditions: for each $i = 1, 2, \cdots$,

$$\sum_{j=1}^{\infty} | c_{ij} | < 1, \qquad (98)$$

and there exists a constant $K_i > 0$, independent of j, such that

$$| b_j(i) | \leqq K_i\left(1 - \sum_{k=1}^{\infty} | c_{jk} |\right), \qquad j = 1, 2, \cdots. \qquad (99)$$

On the other hand, if $(h_{i1}, h_{i2}, \cdots)$ is a solution of (95) for each $i = 1, 2, \cdots$, satisfying (96), then the following integral equation

$$\int_T \int_T r_0(s,u) h(u,v) r_1(v,t) \, du \, dv = r_1(s,t) - r_0(s,t) \qquad (100)$$

* Note that $(Q_1^{(n)})_{ij} = a_{ij}$; $i, j = 1, \cdots, n$.
† See Kantorovich and Krylov,[15] pp. 20–33.

has a square-integrable solution $h(s,t)$,

$$\int_T \int_T h^2(s,t)\, ds\, dt < \infty, \tag{101}$$

such that

$$h(s,t) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij} \psi_i(s) \psi_j(t), \qquad \text{in the mean.}^* \tag{102}$$

Conversely, if $\tilde{h}(s,t)$ is a square-integrable solution of (100), then (95) has a unique solution $(\tilde{h}_{i1}, \tilde{h}_{i2}, \cdots)$ for each $i = 1, 2, \cdots$, satisfying $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \tilde{h}_{ij}^2 < \infty$, such that

$$\tilde{h}_{ij} = \int_T \int_T \psi_i(s) \tilde{h}(s,t) \psi_j(t)\, ds\, dt.\dagger \tag{103}$$

Now, extend the definition of $h_{ij}^{(n)}$, $i = 1, 2, \cdots, n$,‡ by adding

$$h_{ij}^{(n)} = 0; \qquad i, j = n + 1, n + 2, \cdots . \tag{104}$$

Then, (90) and (92) can be rewritten as

$$\lim_{n \to \infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij}^{(n)}\, \xi_i(\omega) \xi_j(\omega) < \infty, \qquad \text{a.e. } (P_0, P_1), \tag{105}$$

and

$$\hat{\theta}(\omega) = \lim_{n = \infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij}^{(n)}\, \xi_i(\omega) \xi_j(\omega). \tag{106}$$

According to the theory of coordinate and projective limits in sequence spaces,§ (97) and (105) imply that

$$\hat{\theta}(\omega) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij} \xi_i(\omega) \xi_j(\omega), \qquad \text{a.e. } (P_0, P_1), \tag{107}$$

since

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \xi_i^2(\omega) \xi_j^2(\omega) < \infty, \qquad \text{a.e. } (P_0, P_1). \tag{108}$$

On the other hand, from (102) and square-integrability, a.e. (P_0, P_1),

---

* See Appendix E.1.
† See Appendix E.2.
‡ Namely, $(h_{i1}, \cdots, h_{in})$ is the solution of (94) for each $i = 1, \cdots, n$.
§ See Cooke,[16] pp. 282–289; in particular, Theorem (10.3, II), extended to the case of double sequences.

of $\{x_t(\omega),\ t \in T\}$,

$$\int_T\int_T x_s(\omega)h(s,t)x_t(\omega)\ ds\ dt = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} h_{ij}\xi_i(\omega)\xi_j(\omega), \qquad \text{a.e. } (P_0, P_1). \tag{109}$$

Hence,

$$\hat{\theta}(\omega) = \int_T\int_T x_j(\omega)h(s,t)x_t(\omega)\ ds\ dt, \qquad \text{a.e. } (P_0, P_1). \tag{110}$$

2.4.3 *Discussion and Summary*

Recall that, in order to specify the set $S_\alpha \in \mathcal{B}_T$ for given α as (91), it is sufficient to assume (88), which assures existence of $\hat{\theta}(\omega)$ and $\hat{\beta}$ defined by (89) and (92) respectively. Moreover, in order to express $\hat{\theta}(\omega)$ as (110), it requires the additional assumptions that (*i*) the integral equation (100) have a square-integrable solution and (*ii*) the conditions (98) and (99) be satisfied.

It can be shown, however, under the assumptions (*i*) and the following:

$$a_{ii} < 1, \qquad i = 1, 2, \cdots, \tag{111}$$

the conditions (*ii*) and (88) can be replaced by the following:

$$a_{ii} > \sum_{j=1}^{\infty}{}' \mid a_{ij} \mid, \qquad i = 1, 2, \cdots, \tag{112}^*$$

and that there exists a constant $K > 0$, independent of $i, j = 1, 2, \cdots$, such that

$$\mid (a_{ij}/\lambda_i) - \delta_{ij} \mid \leqq K(a_{ij} - \sum_{k=1}^{\infty}{}' \mid a_{jk} \mid), \tag{113}$$

where a_{ij} is defined by (93).† It is quite possible that, once the condition (*i*) is assumed, the conditions (111), (112) and (113) may be superfluous. That is to say, in some special cases, if the integral equation (100) admits a square-integrable solution $h(s,t)$ it may be possible to prove directly that

$$h(s,t) = \lim_{n\to\infty} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} h_{ij}^{(n)}\psi_i(s)\psi_j(t), \tag{114}$$

in the mean, which immediately implies (97) and (105), thus establish-

* The prime on the summation sign symbolizes omission of the term $j = i$.
† See Appendix D.

ing (107) and leading to (110). However, in the general case, establishment of (114) does not seem possible, nor does finding a sufficient condition for (114), without making the resultant condition excessively implicit and complex.

2.5 *Summary*

If

$$\int_T\int_T r_0(s,u)h(u,v)r_1(v,t)\ du\ dv = r_1(s,t) - r_0(s,t)$$

has a solution $h(s,t)$,

$$\int_T\int_T h^2(s,t)\ ds\ dt < \infty,$$

then the set $S_\alpha \in \mathcal{B}_T$, given α $(0 < \alpha < 1)$, can be specified as

$$S_\alpha = \left\{\int_T\int_T x_s(\omega)h(s,t)x_t(\omega)\ ds\ dt \geqq \log\frac{1}{\hat{\beta}}\left(\frac{\alpha}{1-\alpha}\right)^2\right\}, \tag{115}$$

where

$$\hat{\beta} = \lim_{n\to\infty} |\ Q_0^{(n)}\ (Q_1^{(n)})^{-1}\ |,$$

and

$$(Q_0^{(n)})_{ij} = \lambda_i\delta_{ij}, \qquad (Q_1^{(n)})_{ij} = a_{ij}; \qquad i, j = 1, \cdots, n,$$

and

$$a_{ij} = \sum_{k=1}^{\infty} \mu_k u_{ki} u_{kj}; \qquad i,j = 1, 2, \cdots,$$

$$u_{ij} = \int_T \varphi_i(t)\psi_j(t)\ dt;$$

where $\lambda_1 \geqq \lambda_2 \geqq \cdots$; $\psi_1(t)$, $\psi_2(t)$, $\cdots$, and $\mu_1 \geqq \mu_2 \geqq \cdots$; $\varphi_1(t)$, $\varphi_2(t)$, $\cdots$, are the eigenvalues and the corresponding orthonormal eigenfunctions associated with the given covariance functions $r_0(s,t)$ and $r_1(s,t)$, which are positive-definite and continuous on $T \times T$; provided that

(1) $a_{ii} < 1$, $i = 1, 2, \cdots$,

(2) $a_{ii} > \sum_{j=1}^{\infty}{}' \,|\ a_{ij}\ |$, $\quad i = 1, 2, \cdots$,

(3) the following is bounded uniformly in $i, j = 1, 2, \cdots$:

$$\frac{\left| \dfrac{a_{ij}}{\lambda_i} - \delta_{ij} \right|}{1 - \sum_{k=1}^{\infty} | \delta_{jk} - a_{jk} |} \leqq K.$$

III. ACKNOWLEDGMENTS

The author gratefully acknowledges his deep indebtedness to D. Slepian and L. A. Shepp, without whose encouragement and assistance this work could not have been completed. Acknowledgment should be made also to I. W. Sandberg, I. Jacobs and W. L. Nelson for their valuable suggestions.

APPENDIX A

Theorem on Optimality

Let P_0 and P_1 be probability measures defined on a Borel field $\mathcal{B}$ of subsets of an abstract space Ω. Through the use of Lebesgue decomposition theorem and Radon-Nikodym theorem:* for a nonempty set $H \in \mathcal{B}$ with $P_0(H) = 0$, there exists a nonnegative function $f(\omega)$ integrable over Ω with respect to P_0 such that

$$P_1(\Lambda) = \int_\Lambda f(\omega) dP_0 + P_1(\Lambda \cap H) \tag{116}$$

for an arbitrary $\Lambda \in \mathcal{B}$.†

Theorem: For an arbitrary constant $k > 0$, define a set $S \in \mathcal{B}$ by

$$S = \{f(\omega) \geqq k\} \cup H. \tag{117}$$

Then,

$$kP_0(S) + P_1(S^c) - kP_0(\Lambda) - P_1(\Lambda^c) \leqq 0 \tag{118}$$

for an arbitrary set $\Lambda \in \mathcal{B}$ where S^c and Λ^c are the complements of S and Λ with respect to Ω.

Proof:
Put $\rho = kP_0(S) + P_1(S^c) - kP_0(\Lambda) - P_1(\Lambda^c)$. By adding and subtracting $kP_0(S \cap \Lambda)$ and $P_1(S^c \cap \Lambda^c)$,

* See Loéve,[10] pp. 130–132.
† This paragraph closely parallels Grenander,[1] pp. 209–210.

$$\begin{aligned}\rho &= k[P_0(S) - P_0(S \cap \Lambda)] + P_1(S^c) - P_1(S^c \cap \Lambda^c) \\ &\qquad - k[P_0(\Lambda) - P_0(S \cap \Lambda)] - P_1(\Lambda^c) + P_1(S^c \cap \Lambda^c) \\ &= kP_0(S \cap \Lambda^c) - P_1(S \cap \Lambda^c) + P_1(S^c \cap \Lambda) \\ &\qquad - kP_0(S^c \cap \Lambda).\end{aligned} \tag{119}$$

From (116) and (117), with $k > 0$ and $P_0(H) = 0$,

$$\begin{aligned}P_1(S \cap \Lambda^c) &- P_1(S^c \cap \Lambda) \\ &= \int_{S \cap \Lambda^c} f(\omega)dP_0 - \int_{S^c \cap \Lambda} f(\omega)dP_0 \\ &\qquad + P_1(S \cap \Lambda^c \cap H) - P_1(S^c \cap \Lambda \cap H) \\ &\geqq kP_0(S \cap \Lambda^c) - kP_0(S^c \cap \Lambda),\end{aligned} \tag{120}$$

since

$$P_1(S^c \cap \Lambda \cap H) = P_1(\{f(\omega) < k\} \cap H^c \cap \Lambda \cap H) \leqq P_1(H^c \cap H) = 0.$$

By substituting (120) into (119),

$$\rho \leqq 0,$$

which proves (118). (Q.E.D.)

Corollary 1. Suppose $P_0 \equiv P_1$, and let $k = [\alpha/(1 - \alpha)]$, $0 < \alpha < 1$. Then, a set S_α defined by

$$S_\alpha = \{f(\omega) \geqq \alpha/(1 - \alpha)\} \tag{121}$$

has the property expressed by (118), i.e.,

$$\alpha P_0(S_\alpha) + (1 - \alpha)P_1(S_\alpha{}^c) \leqq \alpha P_0(\Lambda) + (1 - \alpha)P_1(\Lambda^c) \tag{122}$$

for an arbitrary $\Lambda \in \mathcal{B}$.

Proof:

Note that $P_0 \equiv P_1$ implies $P_1(H) = 0$. Hence, in (118),

$$\begin{aligned}kP_0(S) + P_1(S^c) &= [\alpha/(1 - \alpha)]P_0(S_\alpha \cup H) + P_1(S_\alpha \cup H) \\ &= [\alpha/(1 - \alpha)]P_0(S_\alpha) + P_1(S_\alpha).\end{aligned}$$

Thus, substitution of the above into (118) and multiplication by $1 - \alpha$ proves (122).

Corollary 2. Take Ω to be R_n, an n-dimensional Euclidean space, and $\mathcal{B}$

to be Borel field of right semi-closed, semi-infinite intervals in R_n, *denoted by* $\mathfrak{F}_n$. *Let* $p_m(x_1, \cdots, x_n)$, $m = 0, 1$ *and* $(x_1, \cdots, x_n) \in R_n$, *be Baire density functions corresponding to* P_m, $m = 0, 1$; *i.e.,*

$$P_m\{x_i \leqq \rho_i, i = 1, \cdots, d\} = \int_{-\infty}^{\rho_1} \cdots \int_{-\infty}^{\rho_n} dx_1 \cdots dx_n \, p_m(x_1, \cdots, x_n). \tag{123}$$

Suppose $p_1(x_1, \cdots, x_n) = 0$ *whenever* $p_0(x_1, \cdots, x_n) = 0$. *Then* $S_{\alpha,n}$ *defined by*

$$S_{\alpha,n} = \left\{\frac{p_1(x_1, \cdots, x_n)}{p_0(x_1, \cdots, x_n)} \geqq \frac{\alpha}{1-\alpha}\right\}$$

has the property expressed by (122).*

Proof:

Note that $P_0 \equiv P_1$, thus $P_1(H) = 0$. Then, from (116),

$$f(x_1, \cdots, x_n) = \frac{p_1(x_1, \cdots, x_n)}{p_0(x_1, \cdots, x_n)}, \qquad \text{a.e. } (P_0, P_1)$$

Hence, apply Corollary 1.

APPENDIX B

Preliminaries on Integral Operators†

Let L be an integral operator with a real, symmetric, continuous and positive-definite kernel $r(s,t)$ defined on the rectangle $T \times T$ where T is the closed interval [0,1]. That is,

$$Lf(t) \equiv \int_T r(s, t) f(s) \, ds, \tag{124}$$

where $f(t)$ is an arbitrary real-valued function in the space of square-integrable functions on T, which is symbolically denoted by $\mathfrak{L}_2(0,1)$, or simply by $\mathfrak{L}_2$.

Then, according to the theory of linear operators, all the eigenvalues of L are positive, of finite multiplicity, and finite or denumerably infinite in number. Thus, counting each eigenvalue as many times as its multiplicity, we can construct an ordered sequence of eigenvalues,

* This replaces the Neyman-Pearson theorem in the classical theory of testing simple hypotheses when the criterion changes from the Neyman-Pearson's to the minimum error probability. See Cramér,[18] pp. 529–530.

† See Riesz-Nagy,[17] pp. 227–246.

$$\lambda_1 \geqq \lambda_2 \geqq \cdots , \tag{125}$$

and the corresponding sequence of orthonormal eigenfunctions (using the Gram-Schmidt orthonormalization process if necessary),

$$\psi_1(t), \psi_2(t), \cdots . \tag{126}$$

Then, according to Mercer's theorem,

$$r(s,t) = \sum_{i=1}^{\infty} \lambda_i \psi_i(s) \psi_i(t), \tag{127}$$

where the series converges uniformly on T. Consequently, $\psi_i(t)$ is continuous on T for all i, and

$$\sum_{i=1}^{\infty} \lambda_i = \int_T r(t,t)\, dt < \infty , \tag{128}$$

namely, the sum of all eigenvalues is finite.

Furthermore, because of the positive definiteness of the kernel $r(s,t)$, the set of the eigenfunctions $\{\psi_i(t)\}$ forms an orthonormal basis of $\mathfrak{L}_2$. Let $\{\varphi_i(t)\}$ be another orthonormal basis of $\mathfrak{L}_2$. Then,

$$\psi_j(t) = \sum_{i=1}^{\infty} u_{ij} \varphi_i(t), \quad \text{in the mean}, \tag{129}$$

where

$$u_{ij} = \int_T \varphi_i(t) \psi_j(t)\, dt, \tag{130}$$

which satisfies the following orthogonality conditions:

$$\begin{aligned} \sum_{k=1}^{\infty} u_{ik} u_{jk} &= \int_T \varphi_i(t) \varphi_j(t)\, dt = \delta_{ij} , \\ \sum_{k=1}^{\infty} u_{ki} u_{kj} &= \int_T \psi_i(t) \psi_j(t)\, dt = \delta_{ij} . \end{aligned} \tag{131}$$

APPENDIX C

Density Functions of $\xi_i(\omega)$, $i = 1, \cdots , n$

It has been established in Section 2.4.1 that the random variables defined by (75), i.e.,

$$\xi_i(\omega) = \int_T x_t(\omega)\, \psi_i(t)\, dt, \qquad i = 1, 2, \cdots, \tag{132}$$

are Gaussian variables with respect to P_0 and P_1, where

$$E_m\{x_t(\omega)\} = 0, t \in T, m = 0, 1, \quad \text{and} \quad \int_T x_t^2(\omega)\, dt < \infty,$$

$$\text{a.e. } (P_0, P_1) .$$

C.1 *With respect to P_0*

Through repeated use of Fubini's theorem,

$$E_0\{\xi_i(\omega)\} = \int_T E_0\{x_t(\omega)\}\, \psi_i(t)\, dt = 0, \qquad i = 1, 2, \cdots, \tag{133}$$

and

$$\begin{aligned} E_0\{\xi_i(\omega)\ \xi_j(\omega)\} &= \int_T\int_T E_0\{x_s(\omega)x_t(\omega)\}\psi_i(s)\psi_j(t)\, ds\, dt \\ &= \int_T\int_T r_0(s,t)\psi_i(s)\ \psi_j(t)\ ds\, dt \\ &= \lambda_i\delta_{ij}\,; \quad i,j = 1, 2, \cdots, \end{aligned} \tag{134}$$

where Mercer's theorem is used for the third equality. Then, since $\xi_i(\omega)$, $i = 1, \cdots, n$, are Gaussian variables, (133) and (134) immediately give (76) and (77) with $m = 0$.

C.2 *With respect to P_1*

By substituting (129) into (132),

$$\xi_i(\omega) = \sum_{k=1}^{\infty} u_{ki}\eta_k(\omega), \qquad \text{a.e. } (P_0, P_1) \tag{135}$$

where

$$\eta_i(\omega) = \int_T x_t(\omega)\varphi_i(t)\, dt, \qquad i = 1, 2, \cdots,^* \tag{136}$$

which exist a.e. (P_0, P_1), and Gaussian variables just as $\xi_i(\omega)$, $i = 1, 2, \cdots$, are. Then, the results in C.1 imply that

$$E_1\{\eta_i(\omega)\eta_j(\omega)\} = \mu_i\delta_{ij}\,; \qquad i,j = 1, 2, \cdots . \tag{137}$$

Define

$$\xi_j^{(m)}(\omega) = \sum_{k=1}^{m} u_{kj}\eta_k(\omega), \qquad j = 1, \cdots, n, \tag{138}$$

* Note $\eta_i(\omega)$ here must not be confused with the one in Section 2.3.2.

and let $F_1^{(m)}$ be the distribution function of $\xi_1^{(m)}(\omega), \cdots, \xi_n^{(m)}(\omega)$, and let $f_1^{(m)}(\tau_1, \cdots, \tau_n)$, $-\infty < \tau_j < \infty$, $j = 1, \cdots, n$, be its characteristic function with respect to P_1, i.e.,

$$f_1^{(m)}(\tau_1, \cdots, \tau_n) = E_1\left\{\exp\left[i \sum_{j=1}^{n} \tau_j \xi_j^{(m)}(\omega)\right]\right\}. \tag{139}$$

Then, according to Levy's continuity theorem,* $\lim_{m\to\infty} F_1^{(m)}$ exists if and only if $\lim_{m\to\infty} f_1^{(m)}(\tau_1, \cdots, \tau_n)$ exists for every τ_j, $-\infty < \tau_j < \infty$, and continuous at $\tau_j = 0$, $j = 1, \cdots, n$; and, furthermore, when $\lim_{m\to\infty} F_1^{(m)} = F_1$ exists, its characteristic function $f_1(\tau_1, \cdots, \tau_n)$ is equal to $\lim_{m\to\infty} f_1^{(m)}(\tau_1, \cdots, \tau_n)$ for all τ_j, $-\infty < \tau_j < \infty$, $j = 1, \cdots, n$. Hence, it suffices to obtain $\lim_{m\to\infty} f_1^{(m)}(\tau_1, \cdots, \tau_n)$, namely, the limit of (139) as $m \to \infty$, and to assure its continuity at the origin.

By substituting (138) into (139),

$$\begin{aligned} f_1^{(m)}(\tau_1, \cdots, \tau_n) &= E_1\left\{\exp\left[i \sum_{j=1}^{n} \tau_j \sum_{k=1}^{m} u_{kj}\eta_k(\omega)\right]\right\} \\ &= E_1\left\{\exp\left[\sum_{k=1}^{m} i\eta_k(\omega) \sum_{j=1}^{n} \tau_j u_{kj}\right]\right. \\ &= \prod_{k=1}^{m} \exp\left[-\frac{1}{2}\mu_k\left(\sum_{j=1}^{n} \tau_j u_{kj}\right)^2\right] \\ &= \exp\left[-\frac{1}{2}\sum_{k=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{n} \tau_i\tau_j\mu_k u_{ki}u_{kj}\right]. \end{aligned}$$

Note that

$$\sum_{k=1}^{\infty} |\mu_k u_{ki} u_{kj}| = \sum_{k=1}^{\infty} \mu_k |u_{ki}u_{kj}| \leqq \sum_{k=1}^{\infty} \mu_k < \infty, \tag{138}$$

since

$$\begin{aligned} |u_{ki}u_{kj}| &= \left|\int \varphi_k(t)\psi_i(t)\,dt\right|\left|\int \varphi_k(t)\psi_j(t)\,dt\right| \\ &\leqq \int \varphi_k^2(t)\,dt\left[\int \psi_i^2(t)\,dt \int \psi_j^2(t)\,dt\right]^{\frac{1}{2}} \\ &= 1. \end{aligned}$$

* See Cramér,[18] p. 102.

Hence,

$$\sum_{k=1}^{\infty}\sum_{i=1}^{n}\sum_{j=1}^{n} \tau_i\tau_j\mu_k u_{ki}u_{kj} = \sum_{i=1}^{n}\sum_{j=1}^{n} \tau_i\tau_j \sum_{k=1}^{\infty} \mu_k u_{ki}u_{kj}.$$

Then, putting

$$(Q_1^{(n)})_{ij} = \sum_{k=1}^{\infty} \mu_k u_{ki}u_{kj}\,; \qquad i, j = 1, \cdots, n,$$

continuity of exponential functions implies

$$\lim_{m\to\infty} f^{(m)}(\tau_1, \cdots, \tau_n) = \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} (Q_1^{(n)})_{ij}\tau_i\tau_j\right],$$

which is obviously continuous at $\tau_i = 0$, $i = 1, \cdots, n$. Note that the right-hand side above is the characteristic function of the Gaussian distribution function with the density function (76) and (77) with $m = 1$.

APPENDIX D

P_0 — Equivalence between $\mathcal{B}_T$ and $\hat{\mathcal{B}}_\infty$

It is to be proved that, for an arbitrary set $\Lambda \in \mathcal{B}_T$, there exists a nonempty set $\hat{\Lambda} \in \hat{\mathcal{B}}_\infty$ such that $P_0(\Lambda\Delta\hat{\Lambda}) = 0$. Note, however, that the above statement is equivalent to the following:

> Let $\mathfrak{F}_T \subset \mathcal{B}_T$ be a class of all sets $\Lambda \in \mathcal{B}_T$ such that $\Lambda \in \mathfrak{F}_T$ implies existence of a nonempty set $\hat{\Lambda} \in \hat{\mathcal{B}}_\infty$ with $P_0(\Lambda\Delta\hat{\Lambda}) = 0$. Then, $\mathfrak{F}_T = \mathcal{B}_T$.*

The second statement will be proved.

D.1 For every $t \in T$,

$$x_t(\omega) = \sum_{k=1}^{\infty} \xi_k(\omega)\psi_k(t), \qquad \text{a.e.}(P_0). \tag{139}$$

Proof:

According to the discussion in Section 2.1, (*ii*), $\xi_k(\omega)$, $k = 1, 2, \cdots$, are equal, a.e. (P_0), to the Riemann integrals in quadratic mean criterion of $x_t(\omega)\psi_k(t)$ on T. Hence, from the proper orthogonal decomposi-

* It must be proved first that such an $\mathfrak{F}_T$ is not empty. This will be done in Section D.2.

tion theorem,* the series of (139) converges in quadratic mean with respect to P_0 to $x_t(\omega)$ uniformly on T. Furthermore, $\xi_k(\omega)$, $k = 1, 2, \cdots$, are mutually independent Gaussian variables with means zero and variances λ_k, $k = 1, 2, \cdots$, with respect to P_0.† Hence, the series converges, a.e. (P_0), to a limit for every $t \in T$ since the series of its variances converges for every $t \in T$, i.e.,

$$\sum_{k=1}^{\infty} E_0\{\xi_k^2(\omega)\}\psi_k^2(t) = \sum_{k=1}^{\infty} \lambda_k\psi_k(t)\psi_k(t) = r_0(t,t) < \infty,$$

from Mercer's theorem. Yet, since both the convergence in quadratic mean and the convergence almost everywhere imply the convergence in probability measure, this limit must be equal, a.e. (P_0), to $x_t(\omega)$ for every $t \in T$. (Q.E.D.)

D.2 Let $\Lambda_T \in \mathcal{B}_T$ be defined by

$$\Lambda_T = \{x_{t_i}(\omega) \leqq \rho_i,\ i = 1, \cdots, n\}. \tag{140}$$

Then there exists a nonempty set $\hat{\Lambda}_T \in \hat{\mathcal{B}}_\infty$ such that

$$P_0(\Lambda_T \Delta \hat{\Lambda}_T) = 0.$$

Proof:

Consider a set defined by

$$\hat{\Lambda}_T = \left\{\sum_{k=1}^{\infty} \xi_k(\omega)\psi_k(t_i) \leqq \rho_i,\ i = 1, \cdots, n\right\}. \tag{141}$$

Clearly, $\hat{\Lambda}_T \in \hat{\mathcal{B}}_\infty$. Define $\Gamma_t \in \mathcal{B}_T$ by

$$\Gamma_t = \left\{x_t(\omega) = \sum_{k=1}^{\infty} \xi_k(\omega)\psi_k(t)\right\}, \qquad t \in T. \tag{142}$$

Note that (139) implies

$$P_0(\Gamma_t) = 1, \qquad t \in T. \tag{143}$$

Then it is self-evident that, for $t_i \in T$, $i = 1, \cdots, n$,

$$\begin{aligned} \Lambda_T &= \Lambda_T \cap \left(\bigcap_{i=1}^{n} \Gamma_{t_i}\right) + \Lambda_T \cap \left(\bigcup_{i=1}^{n} \Gamma_{t_i}^c\right), \\ \hat{\Lambda}_T &= \hat{\Lambda}_T \cap \left(\bigcap_{i=1}^{n} \Gamma_{t_i}\right) + \hat{\Lambda}_T \cap \left(\bigcup_{i=1}^{n} \Gamma_{t_i}^c\right), \end{aligned} \tag{144}$$

where $\Gamma_{t_i}^c$ is the complement of Γ_{t_i}. Note that, from (142),

* See Loéve,[10] pp. 478–479.
† See Appendix C.1.

$$\Lambda_T \cap \left(\bigcap_{i=1}^{n} \Gamma_{t_i}\right) = \hat{\Lambda}_T \cap \left(\bigcap_{i=1}^{n} \Gamma_{t_i}\right), \tag{145}$$

and, from (143),

$$P_0\left[\Lambda_T \cap \left(\bigcup_{i=1}^{n} \Gamma_{t_i}{}^c\right)\right] = 0 = P_0\left[\hat{\Lambda}_T \cap \left(\bigcup_{i=1}^{n} \Gamma_{t_i}{}^c\right)\right]. \tag{146}$$

Hence, upon combination of 144, 145, and (146),

$$P_0(\Lambda_T \Delta \hat{\Lambda}_T) = 0. \qquad \text{(Q.E.D.)}$$

D.3 $\mathfrak{F}_T = \mathfrak{B}_T$.

Proof:

First, it is easily seen that the class $\mathfrak{F}_T$ is a field. Moreover, it will now be shown that $\mathfrak{F}_T$ is a Borel field. Let $\Lambda_i \in \mathfrak{F}_T$, $i = 1, 2, \cdots$. Then, from the definition of $\mathfrak{F}_T$, there exists $\hat{\Lambda}_i \in \hat{\mathfrak{B}}_\infty$ such that

$$P_0(\Lambda_i \Delta \hat{\Lambda}_i) = 0, \qquad i = 1, 2, \cdots . \tag{147}$$

Define two sequences of null sets M_i and N_i, $i = 1, 2, \cdots$, by

$$M_i = \Lambda_i - \hat{\Lambda}_i, \qquad N_i = \hat{\Lambda}_i - \Lambda_i . \tag{148}$$

Then,

$$\hat{\Lambda}_i - N_i \subset \Lambda_i \subset \hat{\Lambda}_i \cup M_i, \qquad i = 1, 2, \cdots .$$

Hence,

$$\bigcup_{i=1}^{\infty} \hat{\Lambda}_i - \bigcup_{i=1}^{\infty} N_i \subset \bigcup_{i=1}^{\infty} \Lambda_i \subset \left(\bigcup_{i=1}^{\infty} \hat{\Lambda}_i\right) \cup \left(\bigcup_{i=1}^{\infty} M_i\right),$$

which implies

$$\bigcup_{i=1}^{\infty} \hat{\Lambda}_i - \bigcup_{i=1}^{\infty} \Lambda_i \subset \bigcup_{i=1}^{\infty} N_i,$$

$$\bigcup_{i=1}^{\infty} \Lambda_i - \bigcup_{i=1}^{\infty} \hat{\Lambda}_i \subset \bigcup_{i=1}^{\infty} M_i.$$

Thus,

$$P_0\left[\left(\bigcup_{i=1}^{\infty} \Lambda_i\right) \Delta \left(\bigcup_{i=1}^{\infty} \hat{\Lambda}_i\right)\right] = 0,$$

namely,

$$\bigcup_{i=1}^{\infty} \Lambda_i \in \mathfrak{F}_T .$$

Furthermore, since

$$\bigcap_{i=1}^{\infty} \Lambda_i = \left(\bigcup_{i=1}^{\infty} \Lambda_i^c \right)^c$$

and also $\mathfrak{F}_T$ is a field,

$$\bigcap_{i=1}^{\infty} \Lambda_i \in \mathfrak{F}_T .$$

Hence, $\mathfrak{F}_T$ is a Borel field.

Secondly, note that $\mathfrak{F}_T$ contains the generating class of $\mathcal{B}_T$ as shown by (140) and (35). Hence,

$$\mathfrak{F}_T \supset \mathcal{B}_T .$$

Yet, from the definition of $\mathfrak{F}_T$,

$$\mathfrak{F}_T \subset \mathcal{B}_T .$$

Therefore,

$$\mathfrak{F}_T = \mathcal{B}_T . \qquad \text{(Q.E.D.)}$$

APPENDIX E

Equivalence between Two Equations

E.1 *Preliminary*

Through Mercer's theorem,

$$r_0(s,t) = \sum_{i=1}^{\infty} \lambda_i \psi_i(s) \psi_i(t) , \tag{149}$$

uniformly.

$$r_1(s,t) = \sum_{k=1}^{\infty} \mu_k \varphi_k(s) \varphi_k(t), \tag{150}$$

Then,

$$\int_T \int_T r_1(s,t) \psi_i(s) \psi_j(t) \, ds \, dt = \sum_{k=1}^{\infty} \mu_k u_{ki} u_{kj} = a_{ij} .$$

Hence,

$$r_1(s,t) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} \psi_i(s) \psi_j(t), \qquad \text{in the mean.} \qquad (151)^*$$

* This is a trivial extension of well-known results in the case of functions of one variable. A special case of (151) is found in Courant and Hilbert,[19] pp. 73–74.

E.2 *Equivalence between Two Equations*

Equation (95) can be rewritten as

$$\sum_{k=1}^{\infty} \lambda_i h_{ik} a_{kj} = a_{ij} - \lambda_i \delta_{ij}, \qquad j = 1, 2, \cdots, \tag{152}$$

where $i = 1, 2, \cdots$. Repeating (100),

$$\int_T \int_T r_0(s,u) h(u,v) r_1(v,t)\, du\, dv = r_1(s,t) - r_0(s,t). \tag{153}$$

(a) If $(h_{i1}, h_{i2}, \cdots)$ is a solution of (152) for each $i = 1, 2, \cdots$, with

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij}^2 < \infty, \tag{154}$$

then a square-integrable function $h(s,t)$ with

$$h(s,t) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij} \psi_i(s) \psi_j(t), \qquad \text{in the mean}, \tag{155}$$

satisfies (153).

Proof:

The left-hand side of (153) is clearly square-integrable. Hence, it has the following expansion:

$$\int_T \int_T r_0(s,u) h(u,v) r_1(v,t)\, du\, dv = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \lambda_i h_{ik} a_{kj} \psi_i(s) \psi_j(t), \quad \text{in the mean}, \tag{156}$$

since, through substitution of (149), (150), and (155),

$$\int_T \int_T \left[\int_T \int_T r_0(s,u) h(u,v) r_1(v,t)\, du\, dv \right] \psi_i(s) \psi_j(t)\, ds\, dt = \sum_{k=1}^{\infty} \lambda_i h_{ik} a_{kj}; \qquad i,j = 1, 2, \cdots. \tag{157}$$

Yet, by virtue of $(h_{i1}, h_{i2}, \cdots)$ being a solution of (152) for each $i = 1, 2, \cdots$, the right-hand side of (156) becomes

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \lambda_i h_{ik} a_{kj} \psi_i(s) \psi_j(t) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (a_{ij} - \lambda_i \delta_{ij}) \psi_i(s) \psi_j(t), \tag{158}$$

the right-hand side of which in turn becomes, from (149) and (150),

$$r_1(s,t) - r_0(s,t)$$

$$= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (a_{ij} - \lambda_i \delta_{ij}) \psi_i(s) \psi_j(t), \qquad \text{in the mean.} \qquad (159)$$

Thus, upon combination of (156), (158), and (159),*

$$\int_T \int_T r_0(s,u) h(u,v) r_1(v,t)\, du\, dv = r_1(s,t) - r_0(s,t). \qquad (160)$$

(Q.E.D.)

(b) If $h(s,t)$ is a square-integrable solution of (153), then $(h_{i1}, h_{i2}, \cdots)$ satisfies (152) for each $i = 1, 2, \cdots$, where

$$h_{ij} = \int_T \int_T h(s,t) \psi_i(s) \psi_j(t)\, ds\, dt. \qquad (161)$$

Proof:

Since $h(s,t)$ is square-integrable, it has the expansion of (155) where h_{ij}; $i,j = 1, 2, \cdots$, are defined by (161); thus (157) is established. Meanwhile, from (149) and (150),

$$\int_T \int_T [r_1(s,t) - r_0(s,t)] \psi_i(s) \psi_j(t) = a_{ij} - \lambda_i \delta_{ij}; \qquad i,j = 1, 2, \cdots. \qquad (162)$$

Then, combination of (160), (157) and (162) establishes

$$\sum_{k=1}^{\infty} \lambda_i h_{ik} a_{kj} = a_{ij} - \lambda_i \delta_{ij}. \qquad \text{(Q.E.D.)}$$

APPENDIX F

Alternative Conditions

Assume

$$a_{ii} < 1, \qquad i = 1, 2, \cdots, \qquad (163)$$

and the integral equation (100)

$$\int_T \int_T r_0(s,u) h(u,v) r_1(v,t)\, du\, dv = r_1(s,t) - r_0(s,t) \qquad (164)$$

has a square-integrable solution.† Then, the conditions that

* Note that, if a sequence of functions converges in the mean to two limits, the limits are equal almost everywhere. Furthermore, if the limits are continuous, they are equal everywhere. Note also that continuity of the left-hand side of (156) can easily be seen through the use of the Schwartz inequality.

† Recall from Appendix E that this implies $\sum_{i=1} \sum_{j=1} h_{ij}^2 < \infty$.

$$a_{ii} > \sum_{j=1}^{\infty}{}' \, | a_{ij} | , \tag{165}*$$

and there exists a constant $K > 0$, independent of $i,j = 1, 2, \cdots$, such that

$$| (a_{ij}/\lambda_i) - \delta_{ij} | \leqq K \left(a_{jj} - \sum_{k=1}^{\infty}{}' \, | a_{jk} | \right), \tag{166}$$

imply the conditions (98), (99) and (88); namely, for each $i = 1, 2, \cdots$,

$$\sum_{j=1}^{\infty} | c_{ij} | < 1, \tag{167}$$

and there exists a constant $K_i > 0$ such that

$$b_j(i) \leqq K_i \left(1 - \sum_{k=1}^{\infty} | c_{jk} | \right), \qquad j = 1, 2, \cdots , \tag{168}$$

and finally

$$\begin{aligned} &\lim_{n\to\infty} \operatorname{tr} [(Q_0^{(n)})^{-1} Q_1^{(n)} - I] < \infty, \\ &\lim_{n\to\infty} \operatorname{tr} [Q_0^{(n)} (Q_1^{(n)})^{-1} - I] < \infty. \end{aligned} \tag{169}$$

Proof:

First, note that

$$a_{ii} > 0, \qquad i = 1, 2, \cdots , \tag{170}$$

and

$$\sum_{i=1}^{\infty} a_{ii} = \sum_{i=1}^{\infty} \mu_i . \tag{171}$$

For, from (93) and the fact that $\mu_k > 0$, $k = 1, 2, \cdots$,

$$a_{ii} = \sum_{k=1}^{\infty} \mu_k u_{ki}^2 > 0,$$

and, from (131),

$$\sum_{i=1}^{\infty} a_{ii} = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \mu_k u_{ki}^2 = \sum_{k=1}^{\infty} \mu_k \sum_{i=1}^{\infty} u_{ki}^2 = \sum_{k=1}^{\infty} \mu_k . \dagger$$

---

* The prime symbolizes omission of the term $j = i$ in the summation.
† For justification of interchange of order of summation, see Apostol,[20] pp. 374–375.

Second, through (93) and (94) with (163), (165) and (166),

$$1 - \sum_{j=1}^{\infty} | c_{ij} | = 1 - \sum_{j=1}^{\infty} | \delta_{ij} - a_{ij} |$$

$$= 1 - | 1 - a_{ii} | - {\sum_{j=1}^{\infty}}' | a_{ij} |$$

$$= a_{ii} - {\sum_{j=1}^{\infty}}' | a_{ij} | > 0,$$

and

$$\frac{b_j(i)}{1 - \sum_{k=1}^{\infty} | c_{jk} |} = \frac{| (a_{ij}/\lambda_i) - \delta_{ij} |}{a_{jj} - {\sum_{k=1}^{\infty}}' | a_{jk} |} \leqq K; \qquad i,j = 1, 2, \cdots ,$$

which prove (167) and (168).

Last, note from the definition of $h_{ij}^{(n)}$, $i,j = 1, 2, \cdots$,*

$$\mathrm{tr}\, [(Q_0^{(n)})^{-1} Q_1^{(n)} - I] = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{(n)} a_{ij} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij}^{(n)} a_{ij} ,$$

$$- \mathrm{tr}\, [Q_0^{(n)} (Q_1^{(n)})^{-1} - I] = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij}^{(n)} \lambda_i \delta_{ij} = \sum_{i=1}^{\infty} h_{ii}^{(n)} \lambda_i .$$

Yet, according to the theory of infinite systems of equations, for each $i = 1, 2, \cdots$,

$$| h_{ij}^{(n)} | \leqq K_i , \qquad j = 1, 2, \cdots .\dagger$$

By putting $K_i = K$, $i = 1, 2, \cdots$,

$$| h_{ij}^{(n)} | \leqq K; \qquad i,j = 1, 2, \cdots .$$

Then,

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} | h_{ij}^{(n)} a_{ij} | \leqq K \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} | a_{ij} | < \infty ,$$

$$\sum_{i=1}^{\infty} | h_{ij}^{(n)} \lambda_i | \leqq K \sum_{i=1}^{\infty} \lambda_i < \infty ,$$

since

* Recall:

$$(Q_0^{(n)})^{-1} Q_1^{(n)} - I = [(Q_0^{(n)})^{-1} - (Q_1^{(n)})^{-1}] Q_1^{(n)} ,$$

$$h_{ij}^{(n)} = \begin{cases} [(Q_0^{(n)})^{-1} - (Q_1^{(n)})^{-1}]_{ij} ; & i,j = 1, \cdots , n, \\ 0; & i,j = n + 1, n + 2, \cdots . \end{cases}$$

† See Kantorovich and Krylov,[15] pp. 26–27.

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} | a_{ij} | = \sum_{i=1}^{\infty} \left(a_{ii} + \sum_{j=1}^{\infty}{}' | a_{ij} | \right) < 2 \sum_{i=1}^{\infty} a_{ii} = 2 \sum_{i=1}^{\infty} \mu_i < \infty .$$

Hence, from (97),

$$\begin{aligned}
\lim_{n\to\infty} | \operatorname{tr} [(Q_0{}^{(n)})^{-1} Q_1{}^{(n)} - I] | &= \lim_{n\to\infty} \left| \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij}{}^{(n)} a_{ij} \right| \\
&= \left| \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij} a_{ij} \right| \\
&\leqq \left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij}{}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij}{}^2 \right)^{\frac{1}{2}} \\
&< \infty ,
\end{aligned}$$

$$\begin{aligned}
\lim_{n\to\infty} | \operatorname{tr} [Q_0{}^{(n)} (Q_1{}^{(n)})^{-1} - I] | &= \lim_{n\to\infty} \left| \sum_{i=1}^{\infty} h_{ii}{}^{(n)} \lambda_i \right| \\
&= \left| \sum_{i=1}^{\infty} h_{ii} \lambda_i \right| \\
&\leqq \left(\sum_{i=1}^{\infty} h_{ii}{}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{\infty} \lambda_i{}^2 \right)^{\frac{1}{2}} \\
&< \infty .
\end{aligned}$$

(Q.E.D.)

REFERENCES

1. Grenander, U., Stochastic Processes and Statistical Inference, Arkiv för Matematik, **17**, No. 1, 1950, pp. 195–277.
2. Middleton, D., On the Detection of Stochastic Signals in Additive Normal Noise, I.R.E. Trans., **IT-3**, pp. 86–121.
3. Price, R., Optimum Detection of Stochastic Signals in Noise with Applications to Scatter-Multipath Communications, I.R.E. Trans., **IT-2**, pp. 125–135.
4. Davis, R. C., The Detectability of Random Signals in the Presence of Noise, I.R.E. Trans., **PGIT-3**, pp. 52–62.
5. Bello, P., Some Results on the Problem of Discriminating between Two Gaussian Processes, I.R.E. Trans., **IT-7**, pp. 224–233.
6. Turin, G. L., On Optimal Diversity Reception, I.R.E. Trans., **IT-7**, pp. 154–166.
7. Slepian, D., Some Comments on the Detection of Gaussian Signals in Gaussian Noise, I.R.E. Trans., **IT-4**, pp. 65–68.
8. Parzen, E., Probability Density Functionals and Reproducing Kernel Hilbert Spaces, Proc. of Symposium on Time Series Analysis, John Wiley, New York, 1963, pp. 155–169.
9. Doob, J. L., *Stochastic Processes*, John Wiley, New York, 1953.
10. Loeve, M., *Probability Theory*, 2nd ed., Van Nostrand, Princeton, 1960.
11. Hajek, J., A Property of J-Divergence of Marginal Probability Distributions, Czechoslovak Math. J., **83**, 8, 1958, pp. 460–463.
12. Hajek, J., On a Property of Normal Distribution of Any Stochastic Process (in Russian), Czechoslovak Math., J., **83**, 8, 1958, pp. 610–618. (A translation

appears in Selected Translations in Mathematical Statistics and Probability, Institute of Mathematical Statistics and American Mathematical Society, **1**, pp. 245–252.)
13. Shepp, L. A., The Singularity of Gaussian Measures in Function Space, Proc. Natl. Acad. Sci., **52**, 1964, pp. 430–433.
14. Yaglom, A. M., On the Equivalence and Perpendicularity of Two Gaussian Probability Measures in Function Space, Proc. Symp. on Time Series Analysis, John Wiley, New York, 1963, pp. 327–348.
15. Kantorovich, L. V., and Krylov, V. I., *Approximate Methods of Higher Analysis*, Interscience, New York, 1958.
16. Cooke, R. G., *Infinite Matrices and Sequence Spaces*, MacMillan, London, 1950.
17. Riesz, F., and Sz-Nagy, B., *Functional Analysis*, Frederick Ungar Publishing Co., New York, 1955.
18. Cramer, H., *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1945.
19. Courant, R., and Hilbert, D., *Methods of Mathematical Physics*, vol. I, Interscience, New York, 1953.
20. Apostol, T. M., *Mathematical Analysis*, Addison-Wesley, Reading, Mass., 1957.