STABILITY OF THE LIKELIHOOD RATIO IN SIGNAL DETECTION USING PERTURBATION OPERATORS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

ANDRES C. SALAZAR

1967

This is to certify that the

thesis entitled

STABILITY OF THE LIKELIHOOD
RATIO IN SIGNAL DETECTION USING
PERTURBATION OPERATORS

presented by

ANDRES CLARENCE SALAZAR

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in E. E.

[signature]
Major professor

Date May 12, 1967

O-169

ABSTRACT

STABILITY OF THE LIKELIHOOD RATIO
IN SIGNAL DETECTION USING PERTURBATION OPERATORS

by Andres C. Salazar

Attention has been focused in recent years on the continuity of a signal detection scheme with respect to a small change in the noise power. This continuity may correspond to the stability or robustness of the stochastic decision-making hypothesis test.

The case for stability investigated here is that of detecting a sure signal sent through zero-mean Gaussian noise with continuous autocorrelation $R(s, t)$ on the finite interval $0 \leq t \leq T$ with a maximum likelihood test. Once the entire detection scheme is set on the $L^2[0, T]$ mathematical platform, the operator R_0 corresponding to the original noise autocorrelation is perturbed with ϵR_1 where ϵ is a small real parameter and R_1 is a positive semidefinite operator of norm less than R_0. The likelihood ratio, an integral part of the detection plan, has a form

$$\gamma = \Sigma \frac{a_k w_k}{\lambda_k}$$

where a_k and w_k are Fourier coefficients of the sent signal and received signal respectively relative to the eigenfunctions ϕ_k of the noise autocorrelation and where λ_k are the eigenvalues of the autocorrelation kernel. With R_0 in force the ratio has a certain variance and with chosen threshold completes the detection

scheme. Now that R_0 has been perturbed what changes are instituted in γ ?

A brief summary of the work of William L. Root in this area is followed by an explanation of the difference between that work and the one contained in this thesis. Static perturbation is a term applied to Root's work since it deals with changes in the likelihood ratio stochastic properties keeping the parameters a_k and λ_k the same before and after perturbation.

Simple examples of dynamic perturbation, a term used to denote consideration of all changes of the likelihood ratio after perturbation, are followed by a lengthy discussion of changes both stochastic and function-wise of the ratio when an arbitrarily small norm perturbing operator is used. The method for finding the perturbed ratio coefficients is similar to the one used in perturbing the Hamiltonian operator in quantum mechanics. That is, $R_0 + \epsilon R_1$ is the perturbed noise operator while $\phi_k = \phi_{ko} + \epsilon \phi_{k1} + \epsilon^2 \phi_{k2} + \dots$ and $\lambda_k = \lambda_{ko} + \epsilon \lambda_{k1} + \epsilon^2 \lambda_{k2} + \dots$ are the perturbed k^{th} eigenfunction and eigenvalue. The parameter ϵ is small enough so that first order terms are assumed the significant ones; hence, the solutions to the first set of recursive equations (formed for each power of ϵ)

$$(R_0 - \lambda_{ko} I) \phi_{k1} = (\lambda_{k1} I - R_1) \phi_{ko}$$

$$(\phi_{k1} | \phi_{ko}) = 0$$

$$\lambda_{k1} = (\phi_{ko} | R_1 \phi_{ko})$$

are sufficient to determine $\gamma = \gamma_o + \epsilon\,\gamma_1$ and $\mathrm{Var}\,\gamma = \mathrm{Var}\,\gamma_o + \ldots + \epsilon\, 2\,\mathrm{Cov}\,(\gamma_o, \gamma_1)$ the new liklihood ratio form and variance where γ_1 and $\mathrm{Cov}\,(\gamma_o, \gamma_1)$ are rather involved expressions. Upon determining these new quantities we can account for the way the detection test is kept optimum while an ϵ "distance" away from the original one (γ_o, $\mathrm{Var}\,\gamma_o$ and threshold t_{γ_0}). In this way a continuity or stability can be shown for a finite variance detection scheme.

STABILITY OF THE LIKELIHOOD RATIO IN SIGNAL DETECTION USING PERTURBATION OPERATORS

By

Andres C. Salazar

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1967

646142
9/27/67

ACKNOWLEDGEMENT

por la corina, el día se pasa pero la mañana se acuerda.

The topic for this thesis was an outcome of the author's all too brief acquaintance with William L. Root, a dedicated scholar and teacher.

My gratitude for a warm and fulfilled educational experience goes to Professor Richard Dubes, thesis advisor, teacher and good friend, as well as to the secondary members of the doctoral committee: J. J. Masterson (Mathematics), H. Salehi (Statistics), W. Kilmer (Electrical Engineering), J. Strelzoff (Electrical Engineering).

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The detection of a finite energy sure signal in additive Gaussian noise is our major concern in the first stages of our problem. We shall deal briefly in this chapter with the background material necessary to set up a plausible detection scheme for such a signal on a finite interval.

I. DETECTION OF A KNOWN SIGNAL IN NOISE

The known signal will henceforth be denoted $a(t)$, an integrable square function on the interval $[0, T]$. The total noise process present in the communication channel as well as in the receiver is additive and will be described by a zero mean real second order stochastic process, $x(t)$, continuous in the mean square sence, i.e., with continuous autocorrelation function, $R(s,t)$.

$$R(s,t) = E\, x(t)\, x(s) \qquad 0 \leq s,t \leq T \tag{1.1}$$

The symbol E is the expectation operation relative to the probability measure imposed upon the model. From this measure two hypotheses can be distinguished.

H_0 : signal is not present in the received waveform.

H_1 : signal is present in the received waveform.

The received waveform can then be accordingly:

$$w(t) = x(t) \qquad \text{for } H_0$$

$$w(t) = a(t) + x(t) \quad \text{for } H_1 .$$

Autocorrelation functions, $R(x, y)$, form kernels in L_2 space if they are L_1 integrable. If they are also wide sense stationary they have a Fourier transform. This type of co-variance function has a relative maximum at the origin which is unsurpassed at any other point but can be equalled and is an even positive semidefinite function. Second order stationarity permits by means of the Fourier transform convenient spectral analysis of the detection problem. Studies of noise through Fourier spectra, emphasized by N. Wiener, help in determining in what frequency range the noise power is somewhat attenuated so that the desired signal power strength can be concentrated there for a better overall signal to noise ratio.[1]

The ergodic hypothesis to which the importance of second order stationarity owes its importance implies stationarity which claims that the noise energy level and structure will not change relative to time. The concept of ergodicity and to a lesser degree the weaker condition of wide sense stationarity is based on homogeneity and an isotropic media. The situation is random but the prime assumption is that the "natural course of fluctuation" is allowed at all times. This assumption may be unreasonable for a macroscopic experiment where several factors causing significant change in the environment may prevail for different sections in the communication channel causing truly an unstable media. It may

---

[1] Norbert Wiener, Time Series, Chapter III (Cambridge, Mass.: MIT Press, 1949).

therefore be presumptuous in some cases to try and measure time varying noise energy levels and fit them to an ergodic or weaker wide sense stationary formula for use in signal detection schemes. No claim should be made concerning the magnitude of variation of the noise energy structure. It would be just as presumptuous to assume the noise is never ergodic. On the contrary, the tendency toward ergodicity should always be recognized.

Indeed, the approach in this thesis is to recognize the small but significant variations about the ergodic value of the noise energy. This variation, it is maintained, may be the cause of either detection test instability and/or type I and type II errors whose magnitudes may not conform to estimates given by the wide sense stationary model.

II. AUTOCORRELATION FUNCTIONS

Before we set up the detection scheme we are going to use it is best to describe the properties of the autocorrelation function. First, when we speak of autocorrelations we do not implicity mean "time average" or the wide sense stationary statistical average

$$E\, x(t)\, x(t+\tau) = R(\tau) \,.$$

Specifically, a second order random function $x(t)$ on J, an index set,with properly defined probability space $(\Omega, \mathcal{B}, \mu)$ is a family of second order random variables such that

$$E\, |x(t)|^2 < \infty \qquad \text{for every } t \in J \,.$$

We will always assume the random functions have zero mean.

The second order moments are called variances and the function defined on the set $J \times J$ is termed the autocorrelation function or covariance function.

$$R(t, s) = E\, x(t)\, x(s) \tag{1.1}$$

It can be shown by the CBS inequality that the function exists and is finite almost everywhere (Loève p. 465). The reverse is also true; $R(s, t)$ finite implies

$$E\,|x(t)|^2 = R(t, t) < \infty \qquad t \in J .$$

Hence, covariances or autocorrelation are a manifest nature of the second order random function.

The properties of the autocorrelation can be summarized with a few notes (Loève pp. 466-8).

1. A function $R(s, t)$ on $J \times J$ is an autocorrelation if and only if it is nonnegative definite (sometimes termed positive semidefinite).
2. The class of autocorrelations is closed under additions, multiplications and passages to the limit.
3. Since every nonnegative number is an autocorrelation it follows then that polynomials of covariances with positive coefficients and their limits are also autocorrelations.
4. Given a continuous autocorrelation on a closed interval J the stochastic process $x(t)$ is normal or Gaussian if and only if the random variables

$$x_n = \int_J x(t)\ \phi_n(t)\ dt \qquad (1.2)$$

are normal or Gaussian. $\{\phi_n(t)\}_{n=1}^{\infty}$ here are the eigenfunctions corresponding to the discrete spectra of the autocorrelation function via the Karhunen-Loève expansion (section I. 3).

If the autocorrelation is not positive definite, its eigenfunction sequence is <u>not</u> necessarily a generating set for all $L_2(J)$ where J is the closed interval of the continuously indexed random process. Later we will see that this point implies that the detection procedure may be singular, i.e. there exists a signal in $L_2(J)$ which can be detected without error.

III. $L_2(J)$ KERNELS AND THE KARHUNEN-LOÈVE EXPANSION

A spectral theorem for stochastic processes known as the Karhunen-Loève Expansion plays a central role in the signal detection scheme of interest here. It permits the noise process $x(t)$ to be expressed as

$$x(t) = \sum_{n=1}^{\infty} x_n\ \phi_n(t) \qquad 0 \leq t \leq T \ (= J) \qquad (1.3)$$

where the equal sign really represents convergence in a stochastic as well as an $L_2(J)$ mean square norm. Here the $\{\phi_k\}_{k=1}^{\infty}$ are the eigenfunctions of the autocorrelation $R(x, y)$, an $L_2(J)$ kernel, and <u>per se</u> have no statistical properties. The Fourier coefficients x_n assume then the entire stochastic characteristic of $x(t)$. We have the Karhunen-Loève Expansion then splitting the noise process into two orthogonal sequences (Appendix D), the functions $\{\phi_k\}_{k=1}^{\infty}$

and the stochastic set $\{x_k\}_{k=1}^{\infty}$. Each set can now be dealt with more clearly.

We now see the importance of dealing with $L_2(J)$ kernels and indeed the whole $L_2(J)$ space structure. From Appendix C we see that an $L_2(J)$ kernel is a bivariate function $K(x, y)$ for which both functions

$$A(x) = [\int_0^T K^2(x, y)\, dy]^{1/2} \qquad B(y) = [\int_0^T K(x, y)\, dx]^{1/2} \tag{1.4}$$

exist almost everywhere in J and belong to class $L_2(J)$. It can be shown that every symmetric and nonnull $L_2(J)$ kernel has at least one eigenvalue. Further, any nonnull symmetric $L_2(J)$ kernel has an infinite number of eigenvalues or is a PG kernel, i.e., a Pincherle-Goursat kernel whose expression is generally put as

$$K(x, y) = \sum_{n=1}^{N} \lambda_n \phi_n(x) \phi_n(y) \qquad N < \infty, \; 0 \leq x, y \leq T$$

where $\{\phi_n\}_{n=1}^{N}$ and $\{\lambda_n\}_{n=1}^{N}$ are the eigenvectors and eigenvalues respectively. A Hilbert-Schmidt kernel on the other hand is one for which the following is true:

$$\int_0^T \int_0^T |K(x, y)|^2 \, dxdy < \infty$$

This indicates a PG kernel is Hilbert-Schmidt but not necessarily vice versa.

In our case the autocorrelation function, in view of its properties, forms an $L_2(J)$ kernel and hence submits to eigen-function expansions. The Karhunen-Loève Expansion is such an expansion. Hence, the second order characteristics of the random process will form the properties of the covariance function.

If the $L_2(J)$ kernel, $R(x, y)$, is a positive definite function then the stochastic process is of the proper type; but if the kernel is only positive semidefinite then it is improper. These kernels will correspond to positive $L_2(J)$ and nonnegative $L_2(J)$ operators. From Hilbert space terms we understand positive operators will yield eigenfunction sets which generate all of $L_2(J)$ space while nonnegative operators will have a null space which does not merely contain the zero vector.

The common assumption, due in part to the fact that natural disturbances are commonly of Gaussian distributions, is that the noise process is of Gaussian character as is the k^{th} Fourier coefficient of the noise relative to the k^{th} eigenfunction of $R(x, y)$. This is true for $k = 1, 2, 3, ..$

The following will partially summarize the background needed for the detection scheme to be discussed in the next section.

Let $x(t)$, $0 \leq t \leq T \doteq J$, be a real second order Gaussian random process continuous in mean square with mean zero. $R(s, t) = E\, x(t)\, x(s)$ is then a symmetric, nonnegative definite continuous function in $[0, T] \times [0, T]$ and an integral operator's kernel in $L_2(J)$.

$$[Rf](t) = \int_0^T R(t, s)\, f(s)\, ds = g(t) \quad \epsilon \quad L_2(J) \qquad (1.5)$$

R is self adjoint, nonnegative definite and Hilbert-Schmidt. If R is positive definite its eigenvectors will form a total sequence in $L_2(J)$.

Denote the eigenvectors and eigenvalues of R by $\{\phi_n\}_{n=1}^{\infty}$ and $\{\lambda_n\}_{n=1}^{\infty}$ respectively. By the spectral theorem for stochastic processes:

$$x(t) = \sum_{n=1}^{\infty} x_n \phi_n(t) \tag{1.5}$$

using the stochastic mean square as the metric where

$$x_n = \int_0^T x(t)\, \phi_n(t)\, dt \; . \tag{1.6}$$

Let

$$a_n = \int_0^T a(t)\, \phi_n(t)\, dt \tag{1.7}$$

where a(t) is a known waveform in $L_2(J)$. If $w(t) = a(t) + x(t)$ then $w_n = a_n + x_n$ are uncorrelated random variables for all n. Since the stochastic process is Gaussian w_n, $n = 1, 2, \ldots,$ are also independent.

IV. HYPOTHESIS TESTING FOR SIGNAL DETECTION

For distinguishing between two hypotheses which in our case is between signal and no signal present in the received waveform the maximum likelihood test is an optimum Bayes procedure. The integral part of this test is the likelihood ratio of the two distributions assumed for two hypotheses.

$$\frac{\text{Probability density (observation taken}/H_1\text{ true)}}{\text{Probability density (observation taken}/H_0\text{ true)}} = \frac{f_1}{f_0} = \Lambda \tag{1.8}$$

Use of the ratio in various detection procedures varies only in regard to selection of the threshold t for deciding the boundary between the two hypothesis regions. More knowledge or differing detection penalties of error would influence the selection of a threshold.

For each experiment $\Lambda > t$ implies H_1 is true while $\Lambda < t$ is the result of H_0 predominant. The ratio Λ is a random variable with two distributions defined on the sample space of possible observations. It then has possibly two means and variance corresponding to its two distributions.

$$P_1 (\Lambda > t) = \int_{A_1} d\, F_1(x) \qquad (1.9a)$$

A_1 = region of sample space for which $\Lambda > t$

$$P_0(\Lambda < t) = \int_{A_1^c} d\, F_0(x) \qquad (1.9b)$$

Two kinds of errors may evolve from this simple hypothesis test and they are called respectively errors of the first (type I) kind and second (type II) kind.

Error type I: $\quad$ Prob (H_1 chosen/H_0 true) = e_I $\qquad$ (1.10a)

Error type II: $\quad$ Prob (H_0 chosen/H_1 true) = e_{II} $\qquad$ (1.10b)

In our case we consider the likelihood ratio for Gaussian noise studied under second order variations. Several auto-correlations with different operator spectra and corresponding eigenfunctions may have Gaussian statistical natures. Second order properties of a stochastic process have to do with the rules joining the familial members, this fact corresponding to "energy" for an engineer and first order properties corresponding to the first order distribution of each random variable $x(t)$. The situation we want to investigate is then the energy or power variations of the noise in a communication channel and how these variations affect the general signal detection procedure as exemplified by a general likelihood ratio. We insist on first order stationarity at all times so that the general assumption, namely that of a Gaussian distribution for each random variable in the noise stochastic process, is valid.

Continuing with the Gaussian example we form a detection test which is a form of likelihood ratio test when a sure signal is sent. Denote with f_0 the probability (conditional) density function of no signal and f_1, signal (Root, 1964).

$$f_0 \{w_1, w_2, w_3, \ldots, w_N | 0\} = \frac{1}{\sqrt{2\pi \lambda_1 \lambda_2 \ldots \lambda_N}} \exp\left(-\frac{1}{2} \sum_{n=1}^{N} \frac{w_n^2}{\lambda_n}\right) \tag{1.11a}$$

$$f_1 \{w_1, w_2, \ldots, w_N | 1\} = \frac{1}{\sqrt{2\pi \lambda_1 \lambda_2 \ldots \lambda_N}} \exp\left(-\frac{1}{2} \sum_{n=1}^{N} \frac{(w_n - a_n)^2}{\lambda_n}\right) \tag{1.11b}$$

$$\Lambda = \frac{f_1}{f_0} = \exp\left[\sum^{N} \frac{w_n a_n}{\lambda_n} - \frac{1}{2} \sum^{N} \frac{a_n^2}{\lambda_n}\right] \tag{1.12}$$

$$\eta = \ln \Lambda = \Sigma \frac{w_n a_n}{\lambda_n} - \frac{1}{2} \Sigma \frac{a_n^2}{\lambda_n} \tag{1.13}$$

The latter term is simply a constant ψ so γ can be defined as

$$\gamma = \eta + \psi = \eta + \frac{1}{2} \Sigma \frac{a_n^2}{\lambda_n} \tag{1.14}$$

The decision rule is then to compare γ with a threshold t_γ

$$\gamma(w) = \Sigma \frac{w_n a_n}{\lambda_n} \gtrless t_\gamma \tag{1.15}$$

We compute the means and variances according to the different hypotheses.

$$E_o \gamma = 0 \qquad E_1 \gamma = \Sigma \frac{a_k^2}{\lambda_k} = \beta^2 \tag{1.16a}$$

$$\text{Var}_o \gamma = \text{Var}_1 \gamma = \Sigma \frac{a_k^2}{\lambda_k} = \beta^2 \tag{1.16b}$$

If β^2 is finite no mistake free decision can be made. An example of this will be shown in Chapter II. We call $\beta^2 < \infty$ the non-singular case and $\beta^2 = \infty$ the singular case. Actually singularity as such is usually defined to be the mistake free case if realizable.

Grenander (1950) has shown that the problem can be singular in two ways. First, the integral operator kernel $R(s,t)$ may have a nonzero null space whereas $a(t)$ has a nonzero projection onto this null space. Hence, there exists an element, ρ,

$$\rho \in L_2(J) \ni (\rho \mid \phi_n) = 0 \quad n = 1, 2, \ldots \tag{1.17}$$

where $\{\phi_n\}_{n=1}^{\infty}$ is the set of eigenfunctions for R but

$$(\rho \mid a(t)) \neq 0 \tag{1.18}$$

The operation $(\rho \mid w(t))$ where $w(t) = a(t) + x(t)$ will distinguish between the two hypotheses with probability one. Second, the series

$$\sum_{k=1}^{N} \frac{a_k^2}{\lambda_k} \quad \text{if} \quad N = \infty$$

may diverge so from certain Grenander theorems there exists a test for distinguishing both hypotheses with probability one.[2]

The question arises as to whether a singular test ever exists in a linear realizable receiver. One claim is that the case is inherently prevented in the equipment if not in the incongruence of the mathematical model to the real noise process.

V. OBJECT OF THESIS

This then is the problem. Fixing the threshold t_γ for the assumed stochastic structure of the noise $x(t)$ and its autocorrelation function we wish to know how badly the effect of either not knowing the precise autocorrelation function or of varying it slightly from the presumed value can affect the likelihood ratio of the detection test. We wish to know how the likelihood ratio will be affected first, in its functional form, second, in its statistical properties and third, in the detection test apparatus. The problem is then one of perturbation. Distrubing the kernel $R(x, y)$ of an $L_2(J)$ operator does what to the eigenvalues, eigenfunctions which are an integral part of the ratio? The perturbation is restricted to an additive and small operator, ϵR_1.

$$R(\epsilon) = R_0 + \epsilon R_1 \tag{1.19}$$

---

[2] U. Grenander, "Stochastic Processes and Statistical Inference," Ark. Mat. 1:195-277, 1950.

The parameter ϵ here is a small positive number and R_1 is bounded operator with positive semidefinite symmetric kernel. Property 3 of the autocorrelation (Section II) permits the kernel of $R(\epsilon)$ to be an autocorrelation also.

In Chapter II William L. Root's work in the analysis of stability of the likelihood ratio will be reviewed.

The method for finding the perturbed likelihood ratio coefficients in Chapter III is similar to the one used in perturbing the Hamiltonian operator in quantum mechanics. That is, $R_0 + \epsilon R_1$ is the perturbed noise operator while

$$\phi_k = \phi_{ko} + \epsilon\, \phi_{k1} + \epsilon^2\, \phi_{k2} + \dots$$

and

$$\lambda_k = \lambda_{ko} + \epsilon\, \lambda_{k1} + \epsilon^2\, \lambda_{k2} + \dots$$

are the perturbed k^{th} eigenfunction and eigenvalue. The parameter ϵ is small enough so that first order terms are assumed the significant ones; hence, the solutions to the first set of recursive equations (formed for each power of ϵ)

$$\begin{aligned} (R_0 - \lambda_{ko} I)\, \phi_{k1} &= (\lambda_{k1} I - R_1)\, \phi_{ko} \\ (\phi_{k1} | \phi_{ko}) &= 0 \\ \lambda_{k1} &= (\phi_{ko} | R_1\, \phi_{ko}) \end{aligned}$$

are sufficient to determine $\gamma = \gamma_o + \epsilon\gamma_1$ and $\text{Var}\ \gamma = \text{Var}\ \gamma_o + \dots + \epsilon\, 2\, \text{Cov}(\gamma_o, \gamma_1)$ the new likelihood ratio form and variance where γ_1 and $\text{Cov}\,(\gamma_o, \gamma_1)$ are rather involved expressions. Upon determining these new quantities we can account for the way the detection test is kept optimum while an ϵ "distance" away from the original one (γ_o, $\text{Var}\ \gamma_o$, and threshold t_{γ_0}). In this way a continuity or stability can be shown for a finite variance detection scheme.

CHAPTER II

STATIC STABILITY ANALYSIS

In this chapter we wish to review one method used for investigating stability of the likelihood ratio in signal detection (Root 1964) which sheds light into how the problem may be approached and studied. The summary of Root's work presented here will help establish an attitude towards the problem which might not come about easily since the references may not be readily accessible.

The method consists of perturbing the central or assumed autocorrelation function with a symmetric PG kernel. It then will be shown that the likelihood ratio will be subject to a larger variance than in the unperturbed case. This fact will force larger type I and type II errors for a fixed threshold. Finally, a summary of the operations used in such a stability examination will be put into a Hilbert space trio of equations for the special case in which $\Sigma \left(\frac{a_k}{\lambda_k}\right)^2 < \infty$.

I. PINCHERLE-GOURSAT PERTURBATION

We wish to know what effect a perturbation of the autocorrelation function with a PG kernel will have on the likelihood ratio. We are still considering a sure signal $a(t)$ on $[0, T]$ corrupted by an additive Gaussian, zero mean noise process with continuous positive semidefinite autocorrelation function $R_0(t, s)$ whose operator has a discrete $L_2[0, T]$ spectrum. The corresponding R_0 normalized eigenfunction sequence denoted by $\{\phi_k(t)\}_{k=1}^{\infty}$ may or may not be a generative basis for $L_2[0, T]$. The family or random variables representing the noise indexed by the interval $[0, T]$ is Gaussian or normal with zero mean before and after perturbation.

Let $R_0(t, s)$ be perturbed by $R_1(t, s)$, the kernel of operator R_1.

$$(R_1 \phi_k(t) \mid \phi_j(t)) \triangleq c_k c_j \qquad k, j = 1, 2, \ldots N \qquad (2.1)$$

Otherwise, R_1 is the zero operator for all other k, j. The ϕ_k's are eigenfunctions of $R_0(t, s)$. The autocorrelation of the second order perturbed noise process is

$$R(t, s) = R_0(t, s) + R_1(t, s) \qquad 0 \leq t, s, \leq T . \qquad (2.2)$$

Defining R_1 in this manner implies it is Hilbert Schmidt, self-adjoint, real, positive semidefinite and continuous since the set members $\{\phi_k(t)\ \phi_j(s)\}$ are continuous. For the sake of perturbation we insist R_1's operator norm is equal to ϵ, a small real constant.

$$\| R_1 \|^2 = \sum^N \sum^N |(R_1 \phi_k \mid \phi_j)|^2 = \Sigma \Sigma\, c_j^2\, c_k^2 = \left| \sum^N |c_k|^2 \right|^2 \triangleq \epsilon^2 \qquad (2.3)$$

Since $R_1(t, s)$ is a positive semidefinite function it is a covariance function. The sum of two covariance functions is a covariance function.

The signal's expansion

$$a(t) = \sum_{k=1}^{\infty} a_k \phi_k(t) \qquad (2.4a) \qquad a_k = \int_0^T a(t)\, \phi_k(t)\, dt \qquad (2.4b)$$

makes sense if $a(t)$ belongs to the space $[\{\phi_k\}]$. For a real noise process with $R_0(t, s)$ as its autocorrelation,

$$x_n = \int_o^T x(t)\,\phi_n(t)\,dt \quad (2.5a) \qquad E\, x_n x_k = \lambda_n\, \delta_{nk}\,. \quad (2.5b)$$

II. PERTURBED LIKELIHOOD RATIO

Of interest is the change in variance of the likelihood ratio when the autocorrelation has been perturbed. From equation 1.15 the ratio is

$$\gamma(w) = \Sigma \frac{w_k\, a_k}{\lambda_k} \quad (2.6a) \qquad w_k = \int_o^T w(t)\,\phi_k(t)\,dt \quad (2.6b)$$

with decision rule:

accept H_1 (signal present) if $\gamma(s) > t_\gamma$

reject H_1 (signal absent) if $\gamma(w) < t_\gamma$

where t_γ is the threshold.

In what follows the subscript indicates hypothesis 0 or 1 while the prime mark means the expectation is with respect to the measure of the perturbed model.

$$E'_o\ \gamma(w) = 0 \qquad E'_1\ \gamma(w) = \sum_{n=1}^{\infty} \frac{a_n^2}{\lambda_n}$$

$$\mathrm{Var}_o\ \gamma(w) = E'_o\ |\gamma(w)|^2 \tag{2.7}$$

Expanding the latter equation,

$$E'_o\,|\gamma(w)|^2 = E'_o\ \Sigma \frac{a_n w_n}{\lambda_n}\ \Sigma \frac{a_k w_k}{\lambda_n} = \Sigma\Sigma \frac{a_n a_k}{\lambda_n \lambda_k}\ E'_o\ w_n w_k \tag{2.8}$$

However,

$$E'_o\, w_n w_k = E'_o \int_o^T w(t)\, \phi_n(t)\, dt \int_o^T w(s)\, \phi_k(s)\, ds$$

$$= \int_o^T \int_o^T \phi_k(s)\, \phi_n(t)\; E' w(t)\, w(s)\, dt\, ds \qquad (2.9)$$

claiming for the moment the measures commute in the latter step. For hypothesis 0 we have $w(t) = x(t)$ so

$$E'_o\, w(t)\, w(s) = R(t, s)$$

where

$$R(t, s) = R_0(t, s) + R_1(t, s)\ .$$

Expanding equation 2.9:

$$E'_o\, w_n\, w_k = \int_o^T \int_o^T \{ R_0(t, s) + R_1(t, s)\}\, \phi_n(t)\, \phi(s)\; dt\, ds$$

$$= \lambda_n\, \delta_{nk} + \int_o^T \int_o^T R_1(t, s)\, \phi_n(t)\, \phi_k(s)\, dt\, da$$

$$= \lambda_k\, \delta_{nk} + (R_1 \phi_n \,|\, \phi_k) \qquad (2.10)$$

where δ_{nk} is the kronecker delta

$$\delta_{nk} = \begin{cases} 1 & n = k \\ 0 & n \neq k\ . \end{cases}$$

We conclude that

$$\mathrm{Var}'_o\, \gamma(w) = \Sigma\Sigma \frac{a_n a_k}{\lambda_n \lambda_k} \{ \lambda_n\, \delta_{nk} + (R_1 \phi_n \,|\, \phi_k)\} \qquad (2.11a)$$

$$\triangleq \Sigma \frac{a_n^2}{\lambda_n} + \Delta^2 \triangleq \beta^2 + \Delta^2 \qquad (2.11b)$$

where

$$\Delta^2 = \Sigma\Sigma \frac{a_n a_k}{\lambda_n \lambda_k} (R_1 \phi_k | \phi_n) \tag{2.12}$$

$$\Delta^2 = |\Sigma \frac{a_n c_n}{\lambda_n}|^2 \qquad |\Delta| = |\sum^{N} \frac{a_n c_n}{\lambda_n}|$$

Now

$$\mathrm{Var}_1' \gamma(w) = E_1' |\gamma(w) - \Sigma \frac{a_n^2}{\lambda_n}|^2$$

$$= E_1' |\gamma(w)|^2 - |E(\gamma(w))|^2 \tag{2.13}$$

where

$$E_1' |\gamma(w)|^2 = E_1' \Sigma\Sigma \frac{a_n a_k}{\lambda_n \lambda_k} w_n w_k$$

$$= \Sigma\Sigma \frac{a_n a_k}{\lambda_n \lambda_k} E_1' (x_n + a_n)(x_k + a_k) \tag{2.14}$$

$$E_1' |\gamma(w)|^2 = \Sigma\Sigma \frac{a_n a_k}{\lambda_n \lambda_k} E_1' (x_n x_k + a_n a_k)$$

$$= \Sigma\Sigma \frac{a_n a_k}{\lambda_n \lambda_k} (\lambda_n \delta_{nk} + (R_1 \phi_k \;\; \phi_n) + a_n a_k)$$

$$\Sigma \frac{a_k^2}{\lambda_k} + \Sigma\Sigma \frac{a_n^2 a_k^2}{\lambda_n \lambda_k} + \Sigma\Sigma \frac{a_n a_k}{\lambda_n \lambda_k} (R_1 \phi_n | \phi_k)$$

$$= \Sigma \frac{a_n^2}{\lambda_k} + |\Sigma \frac{a_n c_n}{\lambda_n}|^2 + (\Sigma \frac{a_n^2}{\lambda_k})^2 \tag{2.15}$$

so

$$\mathrm{Var}_1' (\gamma(w)) = \Sigma \frac{a_n^2}{\lambda_n} + (\Sigma \frac{a_n^2}{\lambda_n})^2 + |\Sigma \frac{a_n c_n}{\lambda_n}|^2 - (\Sigma \frac{a_k^2}{\lambda_k})^2$$

$$= (\Sigma \frac{a_n^2}{\lambda_n}) + |\Sigma \frac{a_n c_n}{\lambda_n}|^2 \tag{2.16}$$

Finally,

$$\mathrm{Var}_1' \ \gamma(w) = \beta^2 + \Delta^2 = \mathrm{Var}_0' \ \gamma(w) \ . \tag{2.17}$$

For the unperturbed model, the error probabilities are:

$$e_I = \frac{1}{\sqrt{2\pi}} \int_{\frac{t_\gamma}{\beta}}^{\infty} e^{-u^2/2} \, du \tag{2.18a}$$

$$e_{II} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{t_\gamma - \beta^2}{\beta}} e^{-u^2/2} \, du \ . \tag{2.18b}$$

The error probabilities for the perturbed model are:

$$e_I' = \frac{1}{\sqrt{2\pi}} \int_{\frac{t_\gamma}{\sqrt{\beta^2+\Delta^2}}}^{\infty} e^{-u^2/2} \, du \tag{2.19a}$$

$$e_{II}' = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{t_\gamma - \beta^2}{\sqrt{\beta^2+\Delta^2}}} e^{-u^2/2} \, du \ . \tag{2.19b}$$

For a fixed threshold note that $e_I' > e_I$ and $e_{II}' > e_{II}$. Since the integrals are continuous the error probabilities are continuous relative to the perturbation parameter Δ^2. That is, as $|\Delta|$ goes to zero, e_I' approaches e_I and e_{II}' reduces to e_{II}.

If $t_\gamma = k\,\beta^{3/2}$ is chosen we have

$$\begin{aligned} e_I &= \frac{1}{\sqrt{2\pi}} \int_{k\beta^{1/2}}^{\infty} e^{-u^2/2} \, du \\ e_{II} &= \frac{1}{\sqrt{2\pi}} \int_{\infty}^{k\beta^{1/2} - \beta} e^{-u^2/2} \, du \end{aligned} \tag{2.20}$$

As β aproaches ∞ , e_I and e_{II} approach 0. Thus, singularity is implied. Later, we will see that if t_γ is not taken in this manner this case is unstable in that Δ^2 could be finite or infinite for comparably small ϵ .

III. SINGULARITY AND NONSINGULARITY

A few remarks about both the singular and nonsingular cases of detection in regard to the gross magnitude of β^2 are in order here. Examples of both cases will be given.

It may be that R_0 as well as R_1 are not positive operators so that they have a nontrivial nullspace within $L_2[0, T]$. In this case the singular case of detection exists. That is, there exists a signal $a(t)$ in $L_2[0, T]$ which is independent of the R_0 eigenfunction sequence. It may be shown then that this signal is mistake-free detectable (Grenander (1950) or Root (1963)). However, this case of singularity is discounted (Davenport and Root) by the argument that a linear system is always used to receive the waveform $w(t) = a(t) + x(t)$. Hence, $w(t)$ represents the signal corrupted by the transmission media noise and any introduced by the receiver. The important idea here is that the detection does not occur until <u>after</u> the receiver whether it introduces noise or not. Be it as it may, we shall consider this singular case and its ramifications only briefly.

We will now proceed with specific cases for β^2 but first must consider a most important mathematical lemma.

Landau's Lemma: Given a series $\sum_{n=1}^{\infty} q_n = \infty$ then for arbitrary $\epsilon > 0$ there exists a series $\Sigma z_n^2 = \epsilon$ such that $\sum_{n=1}^{\infty} q_n z_n = \infty$.

This lemma will be used extensively in the following discussion of β^2 examples in detection.

1. singular case

$$\mu^2 = \Sigma \left(\frac{a_n}{\lambda_n}\right)^2 = \infty$$

(a) N finite

$$\sum_{n=1}^{N} \frac{a_n c_n}{\lambda_n} = \Delta < \infty$$

(b) For $N = \infty$ and any ϵ we have a possibility

$$\sum_{n=1}^{\infty} \frac{a_n c_n}{\lambda_n} < \infty \quad \text{or} \quad = \infty$$

by Landau's Lemma. So regardless of the smallness in perturbation described by ϵ we may still have $\Delta^2 = \infty$ or $\Delta^2 < \infty$. For fixed t_γ this may mean unbounded errors as seen in equations 2.19a, 2.19b.

2. nonsingular $\beta^2 < \infty$

(a) For finite N and $\sum_{n=1}^{N} c_n^2 = \epsilon$ or for

$$\sum_{n=1}^{\infty} \frac{a_n^2}{\lambda_n} = \beta^2 < \infty$$

we have

$$\sum_{n=1}^{N} \frac{a_n c_n}{\lambda_n} < \infty$$

(b) For $N = \infty$ and $\sum_{n=1}^{\infty} c_n^2 = \epsilon$ it is possible $\{\frac{a_n}{\lambda_n}\}_{n=1}^{\infty}$ is unbounded while $\Sigma \frac{a_k^2}{\lambda_k}$ is convergent so $\Sigma \frac{a_n c_n}{\lambda_n} = \infty$ is not forbidden by Landau's Lemma.

3. nonsingular

$$\Sigma \frac{a_n^2}{\lambda_n^2} = \mu^2 < \infty$$

(a) N finite

$$\sum_{n=1}^{N} \frac{a_n c_n}{\lambda_n} < \infty$$

(b) N infinite, $\sum_{n=1}^{\infty} c_n^2 = \epsilon$, then by CBS inequality

$$\left|\frac{a_n c_n}{\lambda_n}\right| \leq \sqrt{\Sigma \left(\frac{a_n}{\lambda_n}\right)^2} \sqrt{\Sigma c_n^2} < \infty$$

4. singular $\beta^2 = \infty$

(a) N finite

$$\left| \Sigma \frac{a_n c_n}{\lambda_n} \right| = \Delta < \infty$$

(b) N infinite $\Sigma c_n^2 = \epsilon$. Then

$$\frac{1}{\lambda_{max}} \Sigma \frac{a_n^2}{\lambda_n} < \Sigma \frac{a_n^2}{\lambda_n^2}$$

so the $\Delta = \infty$ possibility exists.

The preceding cases show that a great part of the stability question depends on whether $\sum_{\lambda=1}^{\infty} \left(\frac{a_k}{\lambda_k}\right)^2 = \mu^2$ is finite or not.

Smaller variance indicates stability (Root 1963) in that Δ^2 can not have value range from a finite magnitude to an infinite one.

IV. REFORMULATION OF PROCEDURE IN SPECIAL CASE

A formulation of the approach presented in this chapter to detection stability for the special case of $\Sigma \left(\frac{a_k}{\lambda_k}\right)^2 < \infty$ can be put in terms of the unbounded operator R_0^{-1} when R_0 is CC.

If $\sum_{k=1}^{\infty} \left(\frac{a_k}{\lambda_k}\right)^2 < \infty$ then a(t) is a member of the R_0^{-1} domain. That is,

$$\int_o^T R_0(t, s)\, g(s)\, ds = a(t) \qquad 0 \leq t \leq T \qquad (2.21)$$

has a solution g(t) in $L_2[0, T]$.

$$R_o^{-1} a = g$$

Acutally, $g(t) = \sum_{k=1}^{\infty} \frac{a_k \phi_k(t)}{\lambda_k}$ since

$$\int_o^T R(t, s)\, g(t) = \sum_{k=1}^{\infty} \frac{a_n}{\lambda_n} \int_o^T R(t, s)\, \phi_n(t)\, dt = \sum_{n=1}^{\infty} a_n \phi_n(s) = a(s) \qquad (2.22)$$

$$0 \leq s \leq T$$

$$R_0\, g = a$$

Recalling

$$\gamma(w) = \sum_{k=1}^{\infty} \frac{a_k w_k}{\lambda_k}$$

we see readily

$$(g(t) \,|\, w(t)) = \gamma(w)$$

because

$$(g|w) = (\Sigma \frac{a_n \phi_n}{\lambda_n} \mid \Sigma \frac{w_k \phi_k}{\lambda_k}) = \Sigma\Sigma \frac{a_n w_k}{\lambda_n} \delta_{nk}$$

$$(g|w) = \sum_{k=1}^{\infty} \frac{a_k w_k}{\lambda_k} = \gamma(w) \tag{2.23}$$

Further, we claim $\Delta = (R_1 g/g)$. This can be seen when we put Δ^2 from equation 2.12 in the following form:

$$\begin{aligned}\Delta^2 &= \Sigma\Sigma \frac{a_n a_k}{\lambda_n \lambda_k} (R_1 \phi_k | \phi_n) \\ &= (R_1 \Sigma \frac{a_k \phi_k}{\lambda_k} \Big| \Sigma \frac{a_n \phi_n}{\lambda_n}) \\ &= (R_1 g | g) \end{aligned} \tag{2.24}$$

We then have a trio of equations in compact form describing the entire detection and perturbation procedures.

$$\begin{aligned} &\gamma(w) = (g(t) \mid w(t)) &&\sim &&\gamma = (g|w) \\ &[Rg](t) = a(t) &&\sim &&Rg = a \\ &\Delta^2 = (Rg(t) \mid g(t)) &&\sim &&\Delta^2 = (R \mid g) \end{aligned} \tag{2.25}$$

We have thus formulated the detection procedure into a series of integral equations in $L_2[0, T]$. This likelihood ratio's variance is augmented with Δ^2 which approaches zero in mean square as R_1 approaches the zero operator in operator norm.

CHAPTER III

DYNAMIC PERTURBATION

Dynamic perturbation considers the changes which will occur to the parameters in the likelihood ratio expression as well as the changes in the stochastic properties such as variance and mean. Error probabilities should then point up a "dynamic" change in magnitude since we are using the optimum hypothesis test in both the unperturbed and perturbed models as differentiated from the models in the static case. The continuity property can be studied under the condition that a true likelihood ratio is being used before and after perturbation.

The key to this dynamic approach is operator perturbation theory. If some type of "small" operator is added to the original what changes in the eigenvalues and vectors are instituted? Some theorems have been developed to answer this question but are entrenched in a mathematical mire of notation and at times in quantum theoretic arguments and so must be modified for our pruposes here.

Besides considering simple examples of dynamic perturbation, this chapter outlines the assumptions needed to establish for the likelihood ratio a perturbation procedure similar to the one used to perturb the Hamiltonian operator in quantum mechanics. Work by Franz Rellich (1953) will help in formulating the procedure needed to obtain the first order equation set from the groups of recursive equations set up by this type of operator perturbation. Once the changes in the likelihood ratio parameters are found the stochastic aspect of the ratio is investigated after perturbation. It will be shown that both stochastic and functional facets of the ratio will change proportionately to a small perturbation (ϵ) parameter. Continuity of the detection scheme relative to a small change in noise energy structure will then be implied through stability of the finite-variance likelihood ratio hypothesis test.

I. COMMENTS ON STATIC PERTURBATION

In the previous chapter the changes in detection error probabilities were examined under the hypothesis that the perturbation kernel $R_1(s,t)$ would affect only the variance of the likelihood ratio. This is to say, the same eigenvalues and eigenfunctions were used in the likelihood ratio while its variance changed and duly affected the detection error probability integrals. This approach to the perturbation problem might be called the "static" one and its basis warrants further discussion and closer examination before more interpretation is devoted to it. In the following brief review of the Chapter II approach it is hoped that the need for a more comprehensive perturbation approach will become apparent. This second approach is termed the "dynamic" one for reasons which will become apparent in due time.

In the perturbation example found in the preceding chapter the autocorrelation function, R(s, t), is presumed known at first.

If Gaussian noise is assumed its eigenvalues and eigenfunctions establish probability density functions as well as a likelihood ratio and threshold for the hypothesis test. The variance and means of the likelihood test statistic are determined. Now if a second order perturbation is assumed on the autocorrelation function, the change in the original likelihood ratio comes in second order form, namely in the variance of the ratio under either hypothesis. The error probabilities are made greater by the increase in variance. If the proper threshold is chosen it can be shown, however, that as the perturbation approaches zero in magnitude the errors will approach their unperturbed magnitudes. The error probabilities are hence "continuous " at a chosen threshold point with respect to the second order perturbation parameter.

Now the question raised here is why the original ratio was used with parameters, a_k and λ_k, made obsolete by the perturbation itself. Root (1964) maintains that this procedure yields an indication of "stability" of the hypothesis test relative to the second order noise statistics. An objection to this reasoning is that it really is unfair to use obsolete parameters to show the continuity of the test despite the fact that this procedure simulates the change in error probabilities actually experienced by a receiver-processor. Although a nonadaptive receiver processor would indeed use obsolete parameters when the noise structure is perturbed, the question raised here is not that the decision making apparatus used obsolete parameters but that a true measure of changes in error probabilities is not obtained. That is, it is unfair to claim that the change in error probabilities is shown entirely by comparing the variances of the likelihood test function in the original and perturbed states and by the effect these variances have on the error integrals. Consequently, it may be claimed improper continuity and untrue error changes are the result of the Chapter II procedure.

II. DYNAMIC PERTURBATION--SIMPLE EXAMPLES

Following are some rather simple examples of dynamic perturbation of a likelihood ratio. We will forego the statistical analysis of the perturbation and will only demonstrate models of continuity in the "function" aspect of the likelihood ratio. In essence we will additively perturb the autocorrelation $R(s, t)$ will kernels which range in simplicity from a diagonal to a general commuting one and then see what happens to the likelihood ratio parameters.

Example 1. Perturbation relative to a diagonal PG operator

Let $R_1(t, s) = \sum_{k=1}^{N} c_k^2 \, \phi_k(s) \, \phi_k(t) \qquad 0 \leq t, s \leq T$

be the perturbing kernel. We have that $\Sigma c_k^2 = \epsilon^2$ and that $R(t, s) + A(t, s)$ have eigenvalues $\lambda_k + c_k^2$ and the same eigen-vectors $\{\phi_k(t)\}_{k=1}^{N}$ up to the index N while $R(t, s)$ continues with $\{\phi_k(t)\}_{k=1}^{\infty}$; a_k will not change in the likelihood detection ratio γ.

$$\hat{\gamma}(w) = \sum_{k=1}^{N} \frac{a_k w_k}{\lambda_k + c_k^2} + \sum_{k=N+1}^{\infty} \frac{a_k w_k}{\lambda_k} \qquad (3.1)$$

where ^ denotes the perturbed models.

$$\mathrm{Var}\ \hat{\gamma}(w) = \sum_{k=1}^{N} \frac{a_k^2}{\lambda_k + c_k^2} + \sum_{k=N+1}^{\infty} \frac{a_k^2}{\lambda_k} < \mathrm{Var}\ \gamma(w)$$

$\mathrm{Var}\ \gamma(w) > \mathrm{Var}\ \hat{\gamma}(w)$ and we certainly have

$$\lim_{\epsilon^2 \to 0} \mathrm{Var}\ \hat{\gamma}(w) = \mathrm{Var}\ \gamma(w)$$

Example 2. Perturbation relative to a CC diagonal operator

This example is really an extension of the second in case $\sum_{k=1}^{\infty} c_k^2 = \epsilon^2 < \infty$ and

$$\hat{\gamma}(w) = \sum_{k=1}^{\infty} \frac{a_k w_k}{\lambda_k + c_k^2} \quad \text{and} \quad \mathrm{Var}\ \hat{\gamma}(w) = \sum_{k=1}^{\infty} \frac{a_k^2}{\lambda_k + c_k^2}$$

Example 3. Perturbation relative to a general commuting operator

In the previous examples the pervading quality of the perturbing operators was that they commuted with R. These foregoing cases can really be included in the general case of operators which commute with R. It is this type of operator which permits the eigenvalue set to be perturbed while the eigenfunction set remains the same.

Let us assume we have reindexed the eigenfunctions $\{\phi_k\}_{k=1}^{\infty}$ so that $R_0 \phi_k = \lambda_k \phi_k$ and $R_1 \phi_k = \lambda_k \phi_k$ $k = 1, 2, \ldots$ with equal multiplicity and same order. If

$$(R_0 + \epsilon R_1)\phi_k = \lambda_k (1 + \epsilon)\phi_k$$

then

$$\hat{R}\hat{\phi}_k = \hat{\lambda}_k \hat{\phi}_k \qquad \hat{\lambda}_k = \lambda_k (1 + \epsilon)$$

$$\hat{R} = R_0 + \epsilon R_1$$

with say ϵ a small parameter. Since $\hat{\phi}_k = \phi_k$, a_k remains the same in the likelihood ratio γ.

$$\hat{\gamma} = \sum_{k=1}^{\infty} \frac{a_k w_k}{\lambda_k (1+\epsilon)}$$

$$\hat{\beta}^2 = \operatorname{Var} \hat{\gamma} = \Sigma \frac{a_k^2}{\lambda_k (1+\epsilon)} \tag{3.2}$$

Case 1. If $\epsilon > 0$ then the variance $\hat{\beta}^2 < \infty$ if $\beta^2 < \infty$; $\hat{\beta}^2 < \beta^2 < \infty$. The other possible cases have already been discussed.

As noted above, the eigenfunctions do not change in this type of perturbation. Hence the parameters a_k remain the same and the only "perturbed parameters" are the eigenvalues in the operator spectrum. As recalled, this was not the case in the Chapter II example. When the operator was perturbed no guarantee was made that the other parameters which include eigenfunctions and eigenvalues changed as well as the variance.

The next step would be to broaden the class of perturbing operators to either a general CC class or a bounded class.

So far we have considered the continuity of the likelihood ratio relative to changes in the eigenvalue set only. We accomplished this by examining perturbing operators which commuted with R. The eigenvector set and thus the a_k sequence, signal Fourier coefficient relative to $\{\phi_k(t)\}_{k=1}^{\infty}$, the R eigenvector set, remained untouched by the fluctuation. This type of perturbation can intuitively be seen as an inward or introverted disturbance since the operator's own structure, namely the eigenvalue or eigenvector sets, is being used for describing the varying behavior. This is the simplest type of perturbation which can occur. Nevertheless, it does seem plausible that the noise structure as described by the autocorrelation kernel and corresponding operator will in fact vary about a center functional pattern with additive variations which are themselves describable by the central function pattern.

The time has come however to consider the more general case which includes both perturbed eigenvectors and eigenvalues. Hence, we must consider now the change in a_k (w_k), the Fourier coefficient of the sent (received) signal, as well in the likelihood ratio which leads to the hypothesis test. We will need to delve a little into operator perturbation theory in order to extract the needed mathematical results and formulations. However, these more general cases will give the broader view into kernel perturbations needed to deal with the effect second order noise fluctuation has on a signal detection scheme with likelihood ratio hypothesis test.

III. AN INTRODUCTION TO FIRST ORDER PERTURBATION

Before we proceed to investigate the signal detection problem it would perhaps be appropriate to pause here to consider what uses perturbation theory has had in other scientific areas and point out how this use has prompted the study manifested in this paper. There is indeed more than a passing analog between its use elsewhere (especially in quantum mechanics) and its purpose here.

The following analysis of the signal detection perturbation has been instigated by work in quantum mechanics and systems whose behavior is described by operators, eigenvectors and eigenvalues. Corson (1950), and Rellich (1953) and Powell-Craseman (1961) references all deal with Q-M (quantum mechanics) problems of perturbing Hamiltonian operators or the differential operators appearing in the Schroedinger equation. These references, for the most part, deal with differential operators while the operator of interest in signal detection has integral kernel representations. The analogy between differential and integral operators theories will be exploited and the theorems developed for differential operators will be adopted to our purposes.

In the foregoing chapters it has been pointed out that the signal detection problem is being split into two subproblems. In one, the noise autocorrelation function is known or is unchanged in form. The second investigates the effect of an unknown additive perturbation to the autocorrelation or a sudden change in the autocorrelation. This perturbation incurs change in eigenvalues and eigenvectors of the unperturbed problem structure. In Q-M the operators perturbed are Hamiltonians K and H, i.e., system energy representations. These operators are analogous to the noise energy representation in the time domain, the autocorrelation function. The wave function or eigenfunction of the H and K operator depict the state behavior of the system, whether atomic or grossly mechanical, while the eigenfunction of R portrays the function basis or distinct "function states" of the noise, i.e. $\{\phi_k\}_{k=1}^{\infty}$.

In Q-M, disturbance of eigenvectors and eigenvalues resulting from operator perturbation will reduce in a continuous way to the original unperturbed operator and corresponding eigen-structure as the perturbing operator reduces to zero norm. This continuity hypothesis is hard to understand if the spectra of the unperturbed and perturbed models are discrete (continuous) and continuous (discrete) respectively since a discontinuous change occurs in the vanishing perturbing operator nature. As will be stated later in a fuller context, singularities might appear in the sense discrete spectra may become continuous.

Consider a brief look into first order perturbation for a system using the Hamiltonian operator H; λ_n is the nth energy level while ψ_n is the nth state function.

$$H(\epsilon) = H_0 + \epsilon H_1$$

$$\psi_n(\epsilon) = \psi_{n0} + \epsilon \psi_{n1}$$

$$\lambda_n(\epsilon) = \lambda_{n0} + \epsilon \lambda_{n1}$$

where the second subscripts correspond to the order of perturbation. Suppose

$$H(\epsilon)\, \psi_n(\epsilon) = \lambda_n(\epsilon)\, \psi_n(\epsilon) \quad .$$

Then substituting, we have

$$H_0 \psi_{n0} + \epsilon (H_1 \psi_{n0} + H_0 \psi_{nl}) = \lambda_{n0} \psi_{n0} + \epsilon (\lambda_{nl} \psi_{n0} + \lambda_{n0} \psi_{nl}) + O(\epsilon^2)$$

Equating equal powers of ϵ :

$$\begin{aligned} &H_0 \psi_{n0} = \lambda_{n0} \psi_0 && \text{zeroeth order} \\ &H_0 \psi_{nl} + H_1 \psi_{n0} = \lambda_{nl} \psi_{n0} + \lambda_{n0} \psi_{nl} && \text{first order} \end{aligned} \tag{3.3}$$

Taking the Hilbert space product on both sides with ψ_{n0} we have, since H_0 is self adjoint

$$(\psi_{n0} | H_0 \psi_{nl}) = (H_0 \psi_{n0} | \psi_{nl})$$

then,

$$(\psi_{n0} | H_1 \psi_{n0}) = \lambda_{nl} (\psi_{n0} | \psi_{n0})$$

$$\lambda_{nl} = \frac{(\psi_{n0} | H_1 \psi_{n0})}{(\psi_{n0} | \psi_{n0})} \quad . \tag{3.4}$$

The increase of energy level of the nth state is then the expectation of the perturbing operator relative to the nth original state vector.

The foregoing brief use of perturbation techniques in Q-M should demonstrate the analogy with our problem at hand. Although not done in this thesis, it is hoped that this analogy may be carried further in due time so that extraction of results on one side may be adapted to the other.

Finally, we turn to a more general formulation of operator perturbation so that we can consider the case of signal detection. The notation followed is that used in the Appendices A-D. For engineers and especially those in finite state systems the z transform or generating function concept is well known and is pointed out where applicable. The procedure followed is aligned well with that found in Rellich, Powell-Crasemann and similar references but is specifically altered and augmented where necessary for signal detection perturbation.

Consider the following operator and vector equation.

$$(R_0 + \epsilon R_1)\, \hat{\phi}_k = \hat{\lambda}_k \hat{\phi}_k \tag{3.5}$$

where $\hat{\phi}_k$ and $\hat{\lambda}_k$ are the new eigenvector and eigenvalue relative to the new and perturbed operator $R_0 + \epsilon R_1$.

Assuming ϵ is small we suppose $\hat{\lambda}_k$ and $\hat{\phi}_k$ can be put into the following power series of ϵ, a form of the z transform or generating function of ϵ parameter.

$$(R_0 + \epsilon R_1)(\phi_{ko} + \epsilon\phi_{k1} + \epsilon^2\phi_{k2} + \ldots) = (\lambda_{ko} + \epsilon\lambda_{k1} + \epsilon^2\lambda_{k2} + \ldots)(\phi_{ko} + \epsilon\phi_{k1} + \epsilon^2\phi_{k2} + \ldots$$

Of course,

$$\phi_{ko} = \phi_k, \quad \lambda_{ko} = \lambda_k \ .$$

$$(R_0 + \epsilon R_1)\ \Phi(\epsilon) = \lambda(\epsilon)\ \Phi(\epsilon)$$

where $\Phi(\epsilon)$ and $\lambda(\epsilon)$ are generating functions with ϵ parameter and $\lambda(\epsilon)$ a scalar model and $\Phi(\epsilon)$ actually a function model, i.e., assuming convergence

$$\Phi(\epsilon) = \phi_{ko}(t) + \epsilon\,\phi_{k1}(t) + \ \ldots \tag{3.6}$$

If the convergence of $\Phi(\epsilon)$ and $\lambda(\epsilon)$ were not possible the above formulation is presumptuous. Suppose, however, the convergence of $\Phi(\epsilon)$ and $\lambda(\epsilon)$ are sufficiently "well behaved" so that term by term multiplication is permitted.

$$
\begin{aligned}
(R_0+\epsilon R_1)(\phi_{ko}+\epsilon\phi_{k1}+..) &= (R_0\phi_{ko}+\epsilon R_0\phi_{k1}+\epsilon^2 R_0\phi_{k2}+..)+\epsilon R_1\phi_{ko}+\epsilon^2 R_1\phi_{k1}+\ldots \\
&= R_0\phi_{ko}+\epsilon(R_0\phi_{k1}+R_1\phi_{ko})+\epsilon^2(R_0\phi_{k2}+R_1\phi_{k1})+\ldots \\
&= (\lambda_{ko}+\epsilon\lambda_{k1}+\epsilon^2\lambda_{k2}+..)(\phi_{ko}+\epsilon\phi_{k1}+\ldots) \\
&= \lambda_{ko}\phi_{ko}+\epsilon(\lambda_{k1}\phi_{ko}+\lambda_{ko}\phi_{k1})+\epsilon^2(\lambda_{k2}\phi_{ko}+\lambda_{k1}\phi_{k1} \\
&\qquad +\lambda_{ko}\phi_{k2})+\ldots
\end{aligned}
$$

By equating equal powers of ϵ we can obtain the following recursive equations for any k integer:

$$
\begin{aligned}
R_0\phi_{ko} &= \lambda_{ko}\phi_{ko} \\
R_0\phi_{k1}+R_1\phi_{ko} &= \lambda_{k1}\phi_{ko}+\lambda_{ko}\phi_{k1} \qquad (3.7) \\
&\vdots \\
R_0\phi_{kn}-\lambda_{ko}\phi_{kn} &= \lambda_{kn}\phi_{ko}+\lambda_{kn-1}\phi_{k1}+\ldots+\lambda_{k1}\phi_{kn-1}-R_1\phi_{kn-1} \qquad (3.8)
\end{aligned}
$$

for arbitrary n, a positive integer. To find exact values for the unknowns indexed by $n \neq 0$ we revert to Hilbert space methods.

The first equation

$$R_0\phi_{ko}-\lambda_{ko}\phi_{ko} = 0$$

is really

$$R_0 \phi_k - \lambda_k \phi_k = 0$$

which is the unperturbed equation and yields no new information. The general equation, however, is of the form

$$R_0 \phi_{kn} - \lambda_{ko} \phi_{kn} = \alpha_{kn} \qquad \text{a vector also.} \tag{3.9}$$

$$(R_0 \phi_{kn} - \lambda_{ko} \phi_{kn} | \phi_{ko}) = (\alpha_{kn} | \phi_{ko})$$

Recall our inner product is

$$\int_o^T a(t)\, \overline{b(t)}\, dt = (a | b) \quad .$$

Since R_0 and R_1 are symmetric we have

$$(R_0 \phi_{kn} | \phi_{ko}) = (\phi_{kn} | R_0 \phi_{ko}) = \lambda_{ko} (\phi_{kn} | \phi_{ko})$$

$$(\alpha_{kn} | \phi_{ko}) = (R_0 \phi_{kn} | \phi_{ko}) - \lambda_{ko} (\phi_{kn} | \phi_{ko})$$

$$= \lambda_{ko} (\phi_{kn} | \phi_{ko}) - \lambda_{ko} (\phi_{kn} | \phi_{ko}) = 0$$

Thus α_{kn} is perpendicular to ϕ_{ko} for n arbitrary $\neq 0$

$$(\alpha_{kn} | \phi_{ko}) = \lambda_{kn} (\phi_{ko} | \phi_{ko}) + \lambda_{kn-1} (\phi_{ko} | \phi_{k1}) + \ldots + \lambda_{k1} (\phi_{ko} | \phi_{kn-1}$$

$$- (\phi_{ko} | R_1 \phi_{kn-1})$$

$$= 0$$

For $n = 1$

$$\lambda_{k1} (\phi_{ko} | \phi_{ko}) - (\phi_{ko} | R_1 \phi_{ko}) = (\alpha_{kn} | \phi_{ko}) = 0$$

$$\lambda_{k1} (\phi_{ko} | \phi_{ko}) = (\phi_{ko} | R_1 \phi_{ko})$$

$$\lambda_{k1} = \frac{(\phi_{ko} | R_1 \phi_{ko})}{(\phi_{ko} | \phi_{ko})} \tag{3.10}$$

If λ_{ko} had its corresponding eigenvector ϕ_{ko} normalized then this would be further simplified to

$$\lambda_{k1} = (\phi_{ko} | R_1 \phi_{ko}) \ .$$

Referring back to our recursive equation (3.7) we have the right side of

$$R_0 \phi_{k1} - \lambda_{ko} \phi_{k1} = \lambda_{k1} \phi_{ko} - R_1 \phi_{ko}$$

known. Now to put a further restriction on the eigenfunction series. Let us continue the normalization of eigenfunctions as follows

$$\text{arbitrary } k: \quad ||\phi_k(\epsilon)||^2 = 1 = (\phi_{ko} + \epsilon\phi_{k1} + .. | \phi_{ko} + \epsilon\phi_{k1} + ..)$$

Again, using the generating function assumption which permits term by term multiplication and addition;

$$\begin{aligned} 1 &= (\phi_{ko} | \phi_{ko}) \\ 0 &= (\phi_{ko} | \phi_{k1}) + (\phi_{k1} | \phi_{ko}) \\ &\cdot \\ &\cdot \\ &\cdot \\ 0 &= (\phi_{ko} | \phi_{kn}) + (\phi_{k1} | \phi_{kn-1}) + \dots + (\phi_{kn} | \phi_{ko}) \end{aligned} \tag{3.11a}$$

Since R_0 and R_1 are symmetric we must have real perturbing eigenfunctions so

$$(\phi_{kj} | \phi_{k\ell}) = (\phi_{k\ell} | \phi_{kj})$$

for all ℓ, k integers > 0.

We conclude that for $n = 1$

$$2(\phi_{ko} | \phi_{k1}) = 0 \qquad \text{or} \qquad (\phi_{ko} | \phi_{k1}) = 0$$

We have then sufficient conditions to determine uniquely ϕ_{k1} $k = 1, 2, \ldots$ from the second order equation

$$(R_0 - \lambda_{ko} I)\phi_{k1} = (\lambda_{k1} I - R_1)\phi_{ko}$$

and the first order equation

$$(\phi_{ko} | \phi_{k1}) = 0 \qquad (3.11b)$$

Similar equations are set for other powers of ϵ:

$$R_0 \phi_{kn} - \lambda_{ko}\phi_{kn} = \lambda_{kn}\phi_{ko} + \lambda_{kn-1}\phi_{k1} + \ldots + \lambda_{k1}\phi_{kn-1} - R_1\phi_{kn-1}$$

$$0 = (\phi_{ko} | \phi_{kn}) + (\phi_{k1} | \phi_{kn-1}) + \ldots + (\phi_{kn} | \phi_{ko})$$

It is readily seen from the preceding argument that the higher orders of approximating eigenvalues and eigenfunctions are determined in a sequence which begins with λ_{ko} and Φ_{ko} and proceeds to as high an index n as is desired.

Although the disturbing operator R_1 was in the beginning assumed to be such that the generating functions $\phi(\epsilon)$, $\lambda(\epsilon)$ were convergent it need not be necessarily the case especially if no other restrictions are placed on R_1 besides symmetry. It can happen, for example, that R_1 is an unbounded operator and this may cause havoc in the a priori generating function convergence. Such an example is given in Rellich's notes which perhaps was a motivating factor in his investigation and eventual theorem claiming a "boundedness condition" is necessary for the perturbing operator. An unbounded operator's perturbation may cause an eigenvalue to move from a finite position on the real line to an infinite one regardless of how small an ϵ is used. Consequently, the likelihood ratio γ which is our main concern in this paper in relation to the distrubance of its parameters becomes unwieldy in the sense of having infinite components (infinite eigenvalue in the denominator of a series term).

The use of an unbounded operator as a perturbation may be dismissed easily by some workers in signal detection by arguing the noise autocorrelation function is at all times bounded and hence yields a bounded operator as well. Also, the Karhunen-Loève expansion leads one to believe pure point spectra are the rule for autocorrelation functions. The perturbing unbounded operators beside causing unbounded spectra at times can cause the disappearance of point spectra and replacement with continuous spectra. Hence, we cannot state that a small perturbation parameter indicates a small perturbation in the operator's eigenvector structure. Neither may only the first order perturbation parameter equations be the significant ones. Also, the sign of the parameter ϵ may have a large effect on how the perturbed operator behaves. There is no question that careless generalizations such as "a perturbation of a bounded operator does not boundlessly change the eigenvector structure" is unwarranted. For our case of likelihood ratios there should exist a doubt of whether perturbations of bounded autocorrelations and associated operators can indeed cause not small but very significant changes in the operator's structure and the likelihood ratio which is heavily dependent upon it.

Another facet of the perturbation procedure of the first order here concerns the eigenvector sequence of the perturbed operator. There is no guarantee (Rellich, p. 153) that the new eigenvector sequence is a basis for $L_2[0, T]$. Hence, the perturbed operator is to be assumed a positive one at all times.

IV. PERTURBATION OF A DISCRETE SPECTRUM

In the previous discussion we have written the form

$$\phi_k(\epsilon) = \phi_{ko}(t) + \epsilon \phi_{k1}(t) + \epsilon^2 \phi_{k2}(t) + \ldots \tag{3.12}$$

$$\lambda_k(\epsilon) = \lambda_{ko} + \epsilon \lambda_{k1} + \epsilon^2 \lambda_{k2} + \ldots \tag{3.13}$$

$$k = 1, 2, 3, \ldots \qquad 0 \leq t \leq T \quad \text{(J interval)}$$

for small ϵ real in the sense of a generating function. We neglected to specify a possible region of convergence relative to ϵ and more seriously omitted mentioning the type of norm we are to consider for the convergence. Since 3.12 is a function series while $\lambda_k(\epsilon)$ is a scalar series we need to specify appropriate convergence criteria for both.

From 3.13 we may use the normal mathematical rule for convergence of a power series since it is a scalar series. Since we are speaking of $\mathcal{L}_c$ operators, i.e., bounded operators, we have

$$\lambda_k(\epsilon) \leq M \sum_{h=0}^{\infty} \epsilon^n$$

for some scalar $M < \infty$. We have then that the least upper bound (sup) of ϵ such that 3.13 converges is any ϵ for which

$$|\epsilon| < 1$$

is true for any k. Note the norm for scalars used here is the absolute value. Analogously, we use the $L_2(J)$ pseudonorm instead of the absolute value and apply the Cauchy-Hadamard theorem (p. 382 Fulks) to 3.12. The radius of convergence for 3.12 is

$$\rho_k = \frac{1}{\lim_n \sup ||\phi_{kn}||^{1/n}} \tag{3.14}$$

so that $\phi_k(\epsilon)$ converges for $|\epsilon| < \rho_k$. Then for $\{\phi_k\}_{k=1}^{\infty}$ the radius of convergence is $\rho = \inf_k \{\rho_k\}$.

An added point we can make here is the existence of

$$R(\epsilon) = R_0 + \epsilon R_1 + \epsilon^2 R_2 + \dots$$

and its convergence region. Define an element $f(\epsilon)$ in $L_2(J)$ for $|\epsilon| < \rho$ as <u>regular</u> if a power series in ϵ exists in terms of a sequence $\{f_0, f_1, \dots\}$ all in $L_2(J)$. Define an operator R as <u>regular</u> if there exists a $\rho_0 \neq 0$ and $\rho_0 > 0$ real such that for arbitrary $f \in L_2(J)$, $R(\epsilon)f$ is a generating function sequence or equivalently is a regular element in $L_2[0, T]$ for $|\epsilon| < \rho_0$. Accordingly, if

$$||R_n f|| \leq d^{n+1} ||f|| \qquad n = 0, 1, 2, \dots$$

then $\rho = \frac{1}{d}$ if such a d exists. This criterion is in terms of a majorant series. Another rule (Riesz-Nagy p. 373) is

$$\| R_n f \| \leq \frac{M}{r^{n-1}} (\| f \| + \| R_0 f \|) \qquad n = 1, 2, \ldots$$

with $M = \sup_n \{ \| R_n \| \}$ and r real positive providing r exists. Our concern is autocorrelation functions and $L_2(J)$ kernels so

$$R(x, y, \epsilon) = R_0(x, y) + \epsilon R_1(x, y) + \epsilon^2 R_2(x, y) + \ldots$$

is convergent for $|\epsilon| < \rho$ and uniform in $0 \leq x, y, \leq T$ (= J interval) with $R_n(x, y)$ continuous in J. If

$$| R_n(x, y) | \leq d^{n+1}$$

for $\epsilon < \frac{1}{d}$ with such a d existent $R(x, y, \epsilon)$ is a generating function or power series in ϵ.

This background has then set up the proper attitude for understanding the theorems (Riesz Nagy p. 373-9, Rellich p. 76, 99-, 153-) which in a following brief summary can give the conditions necessary for 3.12, 2.13 and the following equation for a finite multiplicity proper value (indexed by k) and corresponding eigenfunction power series in ϵ to exist.

$$R(\epsilon)\, \phi_k(\epsilon) = \lambda_k(\epsilon)\, \Phi_k(\epsilon) \qquad |\epsilon| < r \qquad (3.15)$$

$$k = 1, 2, \ldots$$

Note first that we have neglected to state how the domain $\mathcal{D}_R(\epsilon)$ is reacting relative to ϵ. In the beginning, we assumed R was self adjoint or hypermaximal (closed and symmetric and defined over all $L_2(J)$). A Hermitian operator is symmetric but defined only on a dense subspace of $L_2(J)$. For different ϵ's the domain $\mathcal{D}_R(\epsilon)$ might enlarge or reduce relative to $L_2(J)$. For our consideration it is desirable and indeed necessary that the domain $\mathcal{D}_R(\epsilon)$ remain stable with respect to the parameter ϵ. There is no reason to consider an unstable domain in the perturbation of an autocorrelation function since we are discussing them in the light of being defined over all $L_2(J)$.

Let us then speak of $R(\epsilon) \in \mathcal{L}_c$ defined on $L_2(J)$, convergent in a nonnull ϵ neighborhood and symmetric. If λ_k is the k^{th} isolated eigenvalue of $R(0)$ with m_k multiplicity then there exists generating functions or power series $\lambda_k^{(1)}(\epsilon), \lambda_k^{(2)}(\epsilon), \ldots$ $\lambda_k^{(m)}(t)$ and corresponding $\phi_k^{(1)}(\epsilon), \phi_k^{(2)}(\epsilon), \ldots, \phi_k^{(m)}(\epsilon)$ all convergent in their respective norms in the same ϵ neighborhood.[4] Let λ_k be isolated in $\lambda_k - d < \lambda < \lambda_k + d_2$ for nonzero d_1, d_2. The following then holds.

(a) $$R(\epsilon)\, \phi_k^{(i)}(\epsilon) = \lambda_k^{(i)}(\epsilon)\, \phi_k^{(i)}(\epsilon) \qquad (3.16)$$

(b) $$\lambda_k^{(i)}(0) = \lambda_k$$

(c) $$(\phi_k^{(i)}(\epsilon) \,|\, \phi_k^{(j)}(\epsilon)) = \delta_{ij}$$

(d) in $\lambda_k - d_1' < \lambda < \lambda_k + d_2'$ the spectrum of $R(\epsilon)$ has been split into $\lambda_k^{(1)}(\epsilon), \ldots, \lambda_k^{(m)}(\epsilon)$ for some existent d_1', d_2' with $d_1' < d_1$, $d_2' < d_2$.

---

[4] Convergence of each of the three types of power series is with respect to the metric of each one respectively (scalar, $L_2(J)$, or operator norm).

The important hypothesis here is that $R(\epsilon) \in \mathcal{L}_c$ and that it has a domain independent of ϵ and is convergent in operator norm in a ϵ neighborhood. Our previous example of perturbation $R_0 + \epsilon R_1$ treated in section 2 of this chapter was certainly such a $R(\epsilon)$. This condition or hypothesis may be weakened to $R(\epsilon)$ hermitian and regular with domain $\mathcal{D}_R(\epsilon)$ valid for the eigenvector set $\{\phi_k^{(i)}(\epsilon)\}_{i=1}^{m}$ (Rellich Theorem 3, p. 100). Note the weakened condition has a varying $\mathcal{D}_R(\epsilon)$. We treat next a criterion for the completeness of $\{\phi_k(\epsilon)\}_{k=1}^{\infty}$.

If the operator $R(0)$ is positive definite and has a discrete spectrum we wish rules to guarantee the operator may be perturbed by $R(\epsilon)$ and still remain positive with discrete spectrum. Nonsingularity in the detection problem, consequently, would be preserved. This doubt is not to be taken lightly for if nonsingularity leads to singularity through perturbation (or vice versa) then the detection or hypothesis test perturbation problem is not well posed. Equivalently, how strong are the conditions which must be met in order to preserve positivity for the operator or completeness of the eigenvector $\{\phi_k(\epsilon)\}_{k=1}^{\infty}$ sequence?

Criterion: (Rellich p. 153-162) Consider $\{R_k\}_{k=0}^{\infty}$ with R_k all hermitian in $L_2[0,T]$ with R_0 CC and sequence bounded. Let

$$\|R_n f\| \leq k^n \|R_0 f\| \tag{3.17}$$

for some positive constant k and every n index, $f \in L_2[0,T]$.

$$R(\epsilon) = R_0 + \epsilon R_1 + \epsilon^2 R_2 + \dots$$

is hermitian bounded, regular for $|\epsilon| < \frac{1}{k}$ with eigenvalues $\{\lambda_k(\epsilon)\}_{k=1}^{\infty}$ and eigenfunctions $\{\phi_k(\epsilon)\}_{k=1}^{\infty}$ all convergent power series in $|\epsilon| < \frac{1}{2k}$ and

$$\begin{aligned} &\text{(a)} \qquad R(\epsilon)\,\phi_n(\epsilon) = \lambda_n(\epsilon)\,\phi_n(\epsilon) \\ &\text{(b)} \qquad \{\phi_k(\epsilon)\}_{k=1}^{\infty} \text{ complete in } L_2[0,T] \qquad\qquad (3.18) \\ &\text{(c)} \qquad |\lambda_n(\epsilon)| \to 0 \text{ as } n \to \infty . \end{aligned}$$

As long as the domain $\mathcal{D}_R(\epsilon)$ is independent of ϵ and $R(\epsilon)$ is regular for some ϵ neighborhood and $R(0)$ is self adjoint with discrete spectrum then $R(\epsilon)$ has a regular discrete spectrum.

For an integral operator 3.17 may be put as (Rellich p. 155)

$$|(u|R_n u)| \leq k^n |(u|R_0 u)| ,$$

$$\int_0^T \left|\int_0^T R(x,y,\epsilon)\,u(y)\,dy\right|^2 dx \leq M^2 \int_0^T \left|\int_0^T R(x,y,0)\,u(y)\,dy\right|^2 \qquad (3.19)$$

or

$$\left|\int_0^T\int_0^T R(x,y,\epsilon)\,u(x)\,\overline{u(y)}\,dx\,dy\right| \leq M \int_0^T\int_0^T R(x,y,0)\,u(x)\,\overline{u(y)}\,dx\,dy .$$

V. STATISTICAL PROPERTIES OF THE LIKELIHOOD EXPANSION

In the preceding section we have explained the conditions for the existence of the forms $R(\epsilon)$, $\phi_k(\epsilon)$ and $\lambda_k(\epsilon)$ and some of their important properties such as completeness of $\{\phi_k(\epsilon)\}_{k=1}^{\infty}$. Now we wish to consider changes in the statistical instead of the function aspects of the likelihood ratio with perturbation.

From 2.6 the likelihood ratio is

$$\gamma = \sum_{k=1}^{\infty} \frac{a_k w_k}{\lambda_k}$$

where a_k is the coefficient of the sure signal relative to the kth eigenfunction of $R(x, y)$, the autocorrelation function of the noise process; w_k is the coefficient of the received waveform relative to the same function, $\phi_k(t)$, while λ_k is the statistical variance of n_k the Fourier coefficient of the noise process relative to $\phi_k(t)$. The following expression has meaning when taken in light of what we have considered previously in this chapter. The prime will signify, for the moment, the perturbed likelihood ratio.

$$\gamma'_{(\epsilon)} = \sum_{k=1}^{\infty} \frac{a_k(\epsilon)\, w_k(\epsilon)}{\lambda_k(\epsilon)} \tag{3.20}$$

We want to know how this perturbed ratio compares with the unperturbed one either in function parameters or statistically. Since we have dealt with the first order perturbation of the operator kernel it is only consistent that we consider first order approximations to the eigenvalues and eigenvectors in view of the fact that convergences for the three perturbed parameters may be comparable. It is indeed impractical to assume one series will converge much faster than another when dealing with our assumed small ϵ parameter. Expanding from 3.13, 3.12:

$$\gamma' = \sum_{k=1}^{\infty} \frac{\sum_{j=0}^{\infty} \epsilon^j a_{kj} \sum_{\ell=0}^{\infty} \epsilon^{\ell} w_{k\ell}}{\sum_{m=0}^{\infty} \lambda_{km} \epsilon^m} \tag{3.21}$$

$$\doteq \Sigma \frac{a_{ko} w_{ko}}{\lambda_{ko} + \epsilon \lambda_{k1}} + \epsilon \Sigma \frac{a_{ko} w_{k1} + w_{ko} a_{k1}}{\lambda_{ko} + \epsilon \lambda_{k1}} + O(\epsilon^2)$$

$$\doteq \Sigma \frac{a_{ko} w_{ko}}{\lambda_{ko}} \left(\frac{1}{1 - (-\epsilon \frac{\lambda_{k1}}{\lambda_{ko}})} \right) + \epsilon \Sigma \frac{a_{ko} w_{k1} + w_{ko} a_{k1}}{\lambda_{ko} + \epsilon \lambda_{k1}} \tag{3.22}$$

$$\gamma' \doteq \sum_{k=1}^{\infty} \frac{a_{ko} w_{ko}}{\lambda_{ko}} \left(\sum_{n=0}^{\infty} (-\epsilon \frac{\lambda_{k1}}{\lambda_{ko}})^n \right) + \epsilon \Sigma \frac{a_{ko} w_{k1} + w_{ko} a_{k1}}{\lambda_{ko} + \epsilon \lambda_{k1}} + O(\epsilon^2)$$

$$\doteq \Sigma \frac{a_{ko} w_{ko}}{\lambda_{ko}} (1 - \epsilon \frac{\lambda_{k1}}{\lambda_{ko}}) + \epsilon \Sigma \frac{a_{ko} w_{k1} + w_{ko} a_{k1}}{\lambda_{ko} + \epsilon \lambda_{k1}} + O(\epsilon^2)$$

$$\doteq \Sigma \frac{a_{ko} w_{ko}}{\lambda_{ko}} + \epsilon \left\{ \Sigma \frac{a_{ko} w_{k1}}{\lambda_{ko}} (1 - \epsilon \frac{\lambda_{k1}}{\lambda_{ko}}) + \Sigma \frac{a_{k1} w_{ko}}{\lambda_{ko}} (1 - \epsilon \frac{\lambda_{k1}}{\lambda_{ko}}) + \dots \right.$$

$$\left. \dots + - \Sigma \frac{a_{ko} w_{ko}}{\lambda_{ko}} (\frac{\lambda_{k1}}{\lambda_{ko}}) \right\} + O(\epsilon^2)$$

$$\doteq \gamma_o + \epsilon \left[\Sigma \frac{a_{ko} w_{k1}}{\lambda_{ko}} + \Sigma \frac{a_{k1} w_{ko}}{\lambda_{ko}} - \Sigma \frac{a_{ko} w_{ko}}{\lambda_{ko}} \frac{\lambda_{k1}}{\lambda_{ko}} \right] + \dots$$

$$\dots + O(\epsilon^2) \tag{3.23a}$$

$$\gamma' \doteq \gamma_o + \epsilon \gamma_1 + O(\epsilon^2) \tag{3.23b}$$

$$\text{where} \quad \gamma_o = \sum_{k=1}^{\infty} \frac{a_{ko} w_{ko}}{\lambda_{ko}} \tag{3.23c}$$

Dropping the reference to the terms of order ϵ^2 or greater we have

$$\gamma' \doteq \gamma_o + \epsilon\,\gamma_1$$

Let us consider γ_1 alone for the moment:

$$\gamma_1 = \left[\sum_{k=1}^{\infty} \frac{a_{ko} w_{k1}}{\lambda_{ko}} + \sum_{k=1}^{\infty} \frac{a_{k1} w_{ko}}{\lambda_{ko}} - \sum_{k=1}^{\infty} \frac{a_{ko} w_{ko}}{\lambda_{ko}} \frac{\lambda_{k1}}{\lambda_{ko}}\right] . \tag{3.24}$$

The minus sign and corresponding term will not make γ' negative if $\epsilon \frac{\lambda_{k1}}{\lambda_{ko}} < 1$ even when the first two terms of γ_1 vanish. The question remains as to how important this assumption is. Since we have assumed convergence of the $\lambda_k(\epsilon)$ series we must have $\epsilon\,\lambda_{k1} < \lambda_{ko}$ for all k especially when the perturbation of the spectrum is assumed small corresponding to small $\epsilon\,R_1$. But we shall discuss this later in more detail.[5] Terms of the first and zero order are present in γ_1 and the first order perturbation terms are in the numerator. From 3.7, 3.10, 3.11b:

$$(R_0 - \lambda_{ko} I)\,\phi_{k1} = (\lambda_{k1} I - R_1)\,\phi_{ko}$$

$$(\phi_{k1} | \phi_{ko}) = 0 \qquad\qquad k = 1, 2, 3, \ldots$$

$$\lambda_{k1} = (\phi_{ko} | R_1\,\phi_{ko}) \tag{3.25}$$

Solving for ϕ_{k1} for $k = 1, 2, 3, \ldots$ will yield a_{k1} for $k = 1, 2, 3, \ldots$ needed for γ_1 through

$$a_{k1} = \int_o^T a(t)\ \phi_{k1}(t)\ dt$$

$$w_{k1} = \int_o^T w(t)\ \phi_{k1}(t)\ dt \quad . \qquad\qquad k = 1, 2, 3, \ldots$$

---

[5] See discussion after equation 3.38.

Hence, we can find the entries for γ_1 from the above 3.25 equations. Note here that the first order perturbation operator is ϵR_1

$$R' = R_0 + \epsilon R_1 .$$[6]

Compute now[7]

$$\mathrm{Var}\,(\gamma_o + \epsilon\gamma_1) = \mathrm{Var}\,\gamma_o + \epsilon\, 2\,\mathrm{Cov}(\gamma_o, \gamma_1) \tag{3.26}$$

Consider for the moment $\mathrm{Cov}(\gamma_o, \gamma_1)$ using 3.23c, 3.24

$$\mathrm{Cov}(\gamma_o, \gamma_1) = \mathrm{Cov}\left[\Sigma\, \frac{a_{ko} w_{ko}}{\lambda_{ko}},\ \left(\Sigma\, \frac{a_{ko} w_{kl}}{\lambda_{ko}} + \Sigma\, \frac{a_{kl} w_{ko}}{\lambda_{ko}} - \Sigma\, \frac{a_{ko} w_{ko}}{\lambda_{ko}} \frac{\lambda_{kl}}{\lambda_{ko}}\right)\right] \tag{3.27}$$

For the zero hypothesis or no sure signal in the received waveform $w(t)$ we have $w_{kl} = n_{kl}$ and $w_{k0} = n_{k0}$ for $k = 1, 2, \ldots.$ where n_{kl} and n_{k0} have zero means. Equation 3.27 then becomes

$$\mathrm{Cov}(\gamma_o, \gamma_1) = \Sigma\, \gamma_o \gamma_1 = E(\Sigma\, \frac{a_{ko} n_{ko}}{\lambda_{ko}})(\Sigma\, \frac{a_{ko} n_{kl}}{\lambda_{ko}} + \Sigma\, \frac{a_{kl} n_{ko}}{\lambda_{ko}} - \Sigma\, \frac{a_{ko} n_{ko}}{\lambda_{ko}} \frac{\lambda_{kl}}{\lambda_{ko}}) \tag{3.28}$$

When we discussed the perturbation technique on a function and kernel basis we neglected to affirm that its statistical property via the

---

[6] The effect of this second order perturbation on the first order statistical nature of $n(t)$, the noise process, is not one of an additive stochastic process $y(t)$ necessarily, i.e., $\hat{n}(t) = n(t) + y(t)$ is not necessarily true.

[7] $\mathrm{Cov}\,(\gamma_j, \gamma_k)$ $j, k > 1$ can be shown to be less or of equal magnitude to that of $\mathrm{Var}\,\gamma_o$, $\mathrm{Cov}\,(\gamma_o, \gamma_1)$ so that the ϵ^2 or higher orders of ϵ coefficients will diminish their significance.

Karhunen-Loève expansion has not changed. That is, for the perturbed process, regardless of the parameter ϵ:

$$E' \, n_k(\epsilon) \, n_j(\epsilon) = \delta_{kj} \, \lambda_k(\epsilon) \ . \tag{3.29}$$

Expanding,

$$E' \, n_k(\epsilon) \, n_j(\epsilon) = E' \sum_{\ell=1}^{\infty} n_{k\ell} \epsilon^{\ell} \sum_{m=0}^{\infty} n_{jm} \epsilon^{m}$$

but $E' \, n_k(\epsilon) \, n_j(\epsilon)$ is equal to $\lambda_k(\epsilon) \, \delta_{kj}$

$$E' \, n_k(\epsilon) \, n_j(\epsilon) = \lambda_{ko} + \epsilon \, \lambda_{k1} + \epsilon^2 \lambda_{k2} + \epsilon^3 \lambda_{k3} + \ldots$$

For proper convergence of the ϵ generating functions or power series we may equate like powers of ϵ so:

$$\begin{aligned}
&E' \, n_{ko} \, n_{jo} = \lambda_{ko} \, \delta_{kj} \\
&E' \, n_{k1} \, n_{jo} + E' \, n_{ko} \, n_{j1} = \lambda_{k1} \, \delta_{kj} \\
&E' \, n_{k1} \, n_{j1} + E' \, n_{k2} \, n_{jo} + E' \, n_{ko} \, n_{j2} = \lambda_{k2} \, \delta_{kj} \\
&E' \, n_{k1} \, n_{j2} + E' \, n_{k2} \, n_{j1} + E' \, n_{ko} \, n_{j3} + E' \, n_{k3} \, n_{jo} = \lambda_{k3} \, \delta_{kj} \\
&\vdots
\end{aligned} \tag{3.30}$$

simplifying, this becomes[8]

$$\begin{aligned}
&\mathrm{Var} \, n_{ko} = \lambda_{ko} \\
&2 \, \mathrm{Cov} \, n_{k1} \, n_{ko} = \lambda_{k1} \\
&2 \, \mathrm{Cov} \, n_{k2} \, n_{ko} + \mathrm{Var} \, n_{k1} = \lambda_{k2} \\
&2 \, \mathrm{Cov} \, n_{k1} \, n_{k2} + 2 \, \mathrm{Cov} \, n_{ko} \, n_{k3} = \lambda_{k3} \\
&\vdots
\end{aligned}$$

---

[8]We will drop the prime superscript after this point.

For reference, we have the following:

$$\mathrm{Var}\,(n_{ko} + \epsilon\, n_{k1} + \epsilon^2 n_{k2} + \epsilon^3 n_{k3} + \ldots) =$$

$$= \mathrm{Var}\, n_{ko} + \epsilon\, 2\,\mathrm{Cov}\,(n_{k1}, n_{ko}) + \epsilon^2 \begin{bmatrix} \mathrm{Var}\, n_{k1} + \\ + 2\,\mathrm{Cov}\,(n_{k2}, n_{ko}) \end{bmatrix} + \epsilon^3 \begin{bmatrix} 2\,\mathrm{Cov}\,(n_{k1}, n_{k2}) + \\ + 2\,\mathrm{Cov}\,(n_{k3}, n_{ko}) \end{bmatrix}$$

$$\ldots + \epsilon^4 \begin{bmatrix} \mathrm{Var}\, n_{k2} + \\ + 2\,\mathrm{Cov}\,(n_{k3}, n_{k1}) \\ + 2\,\mathrm{Cov}\,(n_{k4}, n_{ko}) \end{bmatrix} + \epsilon^5 \begin{bmatrix} 2\,\mathrm{Cov}\, n_{k5}\, n_{ko} + \\ + 2\,\mathrm{Cov}\, n_{k4}\, n_{k1} + \\ + 2\,\mathrm{Cov}\, n_{k3}\, n_{k2} \end{bmatrix} + \epsilon^6 \begin{bmatrix} \mathrm{Var}\, n_{k3} + \\ + 2\,\mathrm{Cov}\, n_{k6}\, n_{ko} + \\ + 2\,\mathrm{Cov}\, n_{k5}\, n_{k1} + \\ + 2\,\mathrm{Cov}\, n_{k4}\, n_{k2} \end{bmatrix}$$

$$+ \epsilon^7 \begin{bmatrix} 2\,\mathrm{Cov}\, n_{k7}\, n_{ko} \\ + 2\,\mathrm{Cov}\, n_{k6}\, n_{k1} \\ + 2\,\mathrm{Cov}\, n_{k5}\, n_{k2} \\ + 2\,\mathrm{Cov}\, n_{k4}\, n_{k3} \end{bmatrix} + \ldots$$

the coefficient of ϵ^{ℓ} is $\sum_{j=0}^{\ell} \mathrm{cov}\, n_{k,\ell-j}\, n_{kj}$ so we can express

$$\mathrm{Var}\,(n_k(t)) = \sum_{\ell=0}^{\infty} \{ \sum_{j=0}^{\ell} \mathrm{cov}\, n_{k,\ell-j}\, n_{kj} \} \, \epsilon^{\ell} \qquad (3.32)$$

Thus, $\mathrm{cov}\,(n_{kj}, n_{k\ell})$ may not equal zero, although $\mathrm{cov}\,(n_{kj}, n_{m\ell}) = 0$ for $k \neq m$; $j, \ell = 1, 2, \ldots$

We are still considering first order perturbation here so

$$\mathrm{Var}(w_{ko} + \epsilon w_{k1}) = \mathrm{Var}\, n_{ko} + \epsilon\, 2\, \mathrm{cov}(n_{k1}, n_{ko})$$

$$= \mathrm{Var}\, n_{ko} + \epsilon\, 2\, E\, n_{k1}\, n_{ko} = \lambda_{ko} + \epsilon \lambda_{k1}$$

under hypothesis 0:

$$E\, w_{ko} = E\, n_{ko} = 0 .$$

Conclude

$$E\, n_{k1}\, n_{ko} = \frac{1}{2} \lambda_{k1} = \mathrm{cov}\,(n_{ko}, n_{k1}) \qquad (3.35)$$

Returning to the original calculation of $\mathrm{cov}\,(\gamma_o, \gamma_1)$ of 3.28 we have

$$\mathrm{Cov}_o(\gamma_o, \gamma_1) = E\,\gamma_o\gamma_1 = E\left[\sum_j \sum_k \left(\frac{a_{ko} n_{ko}}{\lambda_{ko}}\right)\left(\frac{a_{jo} n_{j1}}{\lambda_{jo}} + \frac{a_{j1} n_{jo}}{\lambda_{jo}} - \frac{a_{jo} n_{jo}}{\lambda_{jo}} \frac{\lambda_{j1}}{\lambda_{jo}}\right)\right]$$

$$= \sum_j \sum_k \frac{a_{ko} a_{jo}}{\lambda_{ko}\lambda_{jo}} E\, n_{ko} n_{j1} + \frac{a_{ko} a_{j1}}{\lambda_{ko}\lambda_{jo}} E\, n_{ko} n_{jo} - \frac{a_{ko} a_{jo}}{\lambda_{ko}\lambda_{jo}} \frac{\lambda_{j1}}{\lambda_{jo}} E\, n_{ko} n_{jo}$$

$$= \sum_j \sum_k \frac{a_{ko} a_{jo}}{\lambda_{ko}\lambda_{jo}} \delta_{jk}\, \mathrm{cov}(n_{ko}, n_{j1}) + \frac{a_{ko} a_{j1}}{\lambda_{ko}\lambda_{jo}} \lambda_{ko} \delta_{jk} - \frac{a_{ko} a_{jo}}{\lambda_{ko}\lambda_{jo}} \frac{\lambda_{j1}}{\lambda_{jo}} \lambda_{ko} \delta_{jk}$$

$$= \sum_k \frac{a_{ko}^2}{\lambda_{ko}\lambda_{ko}} \mathrm{cov}(n_{ko}, n_{k1}) + \frac{a_{ko} a_{k1}}{\lambda_{ko}} - \frac{a_{ko}^2}{\lambda_{ko}} \left(\frac{\lambda_{k1}}{\lambda_{ko}}\right)$$

But from 3.35 we have then

$$\mathrm{Cov}(\gamma_o, \gamma_1) = \sum_k \frac{a_{ko}^2}{\lambda_{ko}^2} \left(\frac{1}{2}\lambda_{k1}\right) + \frac{a_{ko} a_{k1}}{\lambda_{ko}} - \frac{a_{ko}^2}{\lambda_{ko}} \frac{\lambda_{k1}}{\lambda_{ko}} \qquad (3.36)$$

But simplifying 3.36 reduces to

$$\mathrm{Cov}(\gamma_o, \gamma_1) = \sum_k \frac{a_{ko}a_{k1}}{\lambda_{ko}} - \frac{1}{2} \frac{a_{ko}^2}{\lambda_{ko}} \left(\frac{\lambda_{k1}}{\lambda_{ko}}\right) \tag{3.37}$$

We can now write 3.26 in its true form using 3.23c and 3.37

$$\mathrm{Var}(\gamma_o + \epsilon\,\gamma_1) = \sum_k \left(\frac{a_{ko}^2}{\lambda_{ko}}\right) + \epsilon\, 2\left\{\sum_k \frac{a_{ko}a_{k1}}{\lambda_{ko}} - \frac{1}{2} \frac{a_{ko}^2}{\lambda_{ko}} \left(\frac{\lambda_{k1}}{\lambda_{ko}}\right)\right\}$$

$$= \sum_k \frac{a_{ko}^2}{\lambda_{ko}} \left(1 - \epsilon \frac{\lambda_{k1}}{\lambda_{ko}}\right) + \epsilon \left(\frac{2\, a_{k1}a_{ko}}{\lambda_{ko}}\right)$$

$$\mathrm{Var}(\gamma_o + \epsilon\,\gamma_1) = \sum_{k=1}^{\infty} \frac{a_{ko}^2}{\lambda_{ko}} \left(1 - \epsilon \frac{\lambda_{k1}}{\lambda_{ko}}\right) + \epsilon \sum_{k=1}^{\infty} \frac{2\, a_{k1}a_{ko}}{\lambda_{ko}} \tag{3.38}$$

Note that as $\epsilon \to 0$ $\mathrm{Var}(\gamma_o + \epsilon\,\gamma_1) \to \mathrm{Var}\,\gamma_o$ independent of $\|R_1\|$ (whose influence is manifested in λ_{k1} and a_{k1}). Further, $\mathrm{Var}(\gamma_o + \epsilon\gamma_1) \to \mathrm{Var}\,\gamma_o$ independent of ϵ if $\|R_1\| \to 0$ since then $\lambda_{k1} \to 0$ as well as $a_{k1} \to 0$. Therefore, this result 3.38 is consistent with the zeroeth order variance since it collapses into it if either $\|R_1\|$ or $\epsilon \to 0$.

Since ϵ is small $\frac{\epsilon\,\lambda_{k1}}{\lambda_{ko}} < 1$ certainly in $|\epsilon| < 1$ since we must have $\frac{\lambda_{kj+1}}{\lambda_{kj}} < 1$ for convergence if $\frac{\lambda_{kj+1}}{\lambda_{kj}} \to 1$. So $1 - \epsilon \frac{\lambda_{k1}}{\lambda_{ko}} > 0$ for infinitely many k if $|\epsilon| < 1$. Yet this is not enough to guarantee

$$\sum \frac{a_{ko}^2}{\lambda_{ko}} \left(1 - \epsilon \frac{\lambda_{k1}}{\lambda_{ko}}\right) > 0$$

but it does say the greatest number of series terms are positive. The series will be positive if $\epsilon\,\lambda_{k1} < \lambda_{ko}$ however for $k = 1, 2, \ldots$ This requirement is manifested by $R_o > \epsilon\, R_1$ or that R_o is a larger operator than $\epsilon\, R_1$ which is obviously the case since the perturbation is indeed assumed small, much smaller than the original operator. Note the order of eigenvalues here $\epsilon \lambda_{k1} < \lambda_{ko}$ and that we are <u>not</u> implying λ_{k1} are eigenvalues of R_1.

$$\mathrm{Var}(\gamma_o + \epsilon \gamma_1) = \Sigma \frac{a_{ko}^2}{\lambda_{ko}} (1 - \epsilon \frac{\lambda_{k1}}{\lambda_{ko}}) + \epsilon \, \Sigma \frac{2 a_{k1} a_{ko}}{\lambda_{ko}} \tag{3.40}$$

Could the conclusion be drawn that the new variance is smaller or larger than the original one? That is, can the following hold true?

$$\Sigma \frac{a_{ko}^2}{\lambda_{ko}} \frac{\lambda_{k1}}{\lambda_{ko}} - \Sigma \frac{2 a_{k1} a_{ko}}{\lambda_{ko}} >$$

$$\Sigma \frac{a_{ko}^2}{\lambda_{ko}} \left(\frac{\lambda_{k1}}{\lambda_{ko}} - 2 \frac{a_{k1}}{a_{ko}}\right) \overset{?}{>} 0 \tag{3.41}$$

For $\mathrm{Var}(\gamma_o + \epsilon \gamma_1)$ to be smaller than $\mathrm{Var}\, \gamma_o$ we need for $k \neq 0$

$$\frac{\lambda_{k1}}{\lambda_{ko}} - 2 \frac{a_{ko}}{a_{k1}} > 0$$

such that the whole sum 3.41 is positive. It can be seen that if $\frac{\lambda_{k1}}{\lambda_{ko}}$ $k = 1, 2, \ldots$ are very small for all k this may be impossible. Otherwise, the restrictions we have placed on our perturbation technique could permit

$$\mathrm{Var}\left(\gamma_o + \epsilon \gamma_1\right) < \mathrm{Var}\, \gamma_o \, .$$

Let us try to compute bounds for $\mathrm{Var}(\gamma_o + \epsilon \gamma_1)$ in terms of $\mathrm{Var}\, \gamma_o$ and a few parameters we may derive now. Let $a(t)$ be a signal integrable square on $[0, T] = J$. Let perturbation R_1 be such that $R_o > R_1$, and $\lambda_{ko} > \lambda_{k1}$ $k = 1, 2, \ldots$ Let

$$0 \leq \cancel{\varpi} = \inf_k \left\{\frac{\lambda_{k1}}{\lambda_{ko}}\right\} < 1$$

$$0 < \mathfrak{R} = \sup_k \left\{\frac{\lambda_{k1}}{\lambda_{ko}}\right\} \leq H < \infty, \quad H \text{ a constant.} \tag{3.43}$$

$$m = \sup_j \left\{ K_j : a_{k1} \geq K_j \; a_{ko} \quad k = 1, 2, \ldots \right\} \tag{3.44}$$

m finite since a_k's are finite

$$M = \inf_j \left\{ K_j : |a_{k1}| \leq K_j \, |a_{ko}| \quad k = 1, 2, \ldots \right\} \tag{3.45}$$

M finite since a_k's are finite

then

$$\Sigma \frac{a_{ko}^2}{\lambda_{ko}} - \epsilon \, \Sigma \frac{a_{ko}^2}{\lambda_{ko}} \frac{\lambda_{k1}}{\lambda_{ko}} + \epsilon \, \Sigma \frac{2\, a_{k1} a_{ko}}{\lambda_{ko}} \leq$$

$$\leq \operatorname{Var} \gamma_o - \epsilon \, (\operatorname{Var} \gamma_o) \varpi + 2 \, \epsilon \, (\operatorname{Var} \gamma_o) \, M \tag{3.46}$$

$$\operatorname{Var} \gamma_o - \epsilon \, \mathfrak{R} \operatorname{Var} \gamma_o + 2 \epsilon \, m \operatorname{Var} \gamma_o \leq \Sigma \frac{a_{ko}^2}{\lambda_{ko}} - \epsilon \, \Sigma \frac{a_{ko}^2}{\lambda_{ko}} \frac{\lambda_{k1}}{\lambda_{ko}} +$$

$$+ \epsilon \, \Sigma \frac{2\, a_{k1} a_{ko}}{\lambda_{ko}}$$

so

$$\max \left(0, \operatorname{Var} \gamma_o \left(1 - \epsilon(\mathfrak{R} - 2m)\right)\right) \leq \operatorname{Var}(\gamma_o + \epsilon \gamma_1) \leq \operatorname{Var} \gamma_o (1 + \epsilon(2M - \varpi)) \tag{3.47}$$

$$\theta \leq \operatorname{Var}(\gamma_o + \epsilon \, \gamma_1) \leq \Theta$$

For ϵ small enough we need not worry about limiting the left hand side of the inequality so we have

$$- \epsilon \, (\mathfrak{R} - 2m) \leq \Delta \operatorname{Var} \leq \epsilon \, (2M - \varpi) \tag{3.48}$$

Let $G = \max(\Omega - 2m, 2M - \omega)$ then

$$-\epsilon G \leq \Delta \mathrm{Var} \leq \epsilon G$$

$$|\Delta \mathrm{Var}| \leq \epsilon G \tag{3.49}$$

VI. RESULTS RELATIVE TO THE THRESHOLD

In previous sections of this chapter we have compared what we have called the nuclear likelihood ratio γ_o with a perturbed model $\gamma(\epsilon)$. It perhaps should have been stressed that this comparison is seemingly unjust if a fixed threshold is to be involved. As we recall from 1.16 the threshold was chosen for γ simply because η in 1.15 was different from the expression by a constant.

$$\eta = \ln \frac{f_1}{f_o} = \Sigma \frac{w_n a_n}{\lambda_n} - \frac{1}{2} \Sigma \frac{a_n^2}{\lambda_n} \tag{1.15}$$

$$\gamma = \eta + \frac{1}{2} \Sigma \frac{a_n^2}{\lambda_n} = \eta + \psi \tag{1.16a}$$

$$\gamma = \Sigma \frac{w_n a_n}{\lambda_n} \gtrless t_\gamma \tag{1.16b}$$

The constant, of course, supposedly is ψ. Yet if perturbation of the second order properties of the noise is considered this quantity is not a constant relative to the disturbance parameter ϵ.

$$\psi_o = \frac{1}{2} \Sigma \frac{a_{no}^2}{\lambda_{no}} = \frac{1}{2} \beta_o^2 \tag{3.46}$$

$$\psi(\epsilon) = \frac{1}{2} \Sigma \frac{a_n^2(\epsilon)}{\lambda_n(\epsilon)} = \frac{1}{2} \beta^2(\epsilon) \tag{3.47}$$

Thus, in order to consider the changes in the detection structure relative to a threshold, we will have to explicate the changes occurring to ψ as well. To do this, it is important to use the expression for the likelihood ratio in η form. We must keep in mind, however, that γ_o and η_o as well as $\gamma(\epsilon)$ and $\eta(\epsilon)$ have the same stochastic variance and properties except mean.

$$\eta(\epsilon) = \Sigma \frac{w_n(\epsilon)\, a_n(\epsilon)}{\lambda_n(\epsilon)} - \frac{1}{2} \Sigma \frac{a_n^2(\epsilon)}{\lambda_n(\epsilon)} \tag{3.48}$$

$$= \gamma(\epsilon) - \psi(\epsilon)$$

From 3.23c,

$$\psi(\epsilon) = \frac{1}{2} \Sigma \frac{(a_{no} + \epsilon\, a_{n1} + \epsilon^2\, a_{n2} + \ldots)^2}{\lambda_{no} + \epsilon \lambda_{n1} + \epsilon^2 \lambda_{n2} + \ldots}$$

Approximating to the first few orders of ϵ we have

$$\psi(\epsilon) \doteq \frac{\frac{1}{2} \Sigma\, a_{no}^2 + \epsilon\,(2\, a_{no} a_{n1}) + \epsilon^2 (a_{n1}^2 + 2\, a_{n2}\, a_{no}) + \ldots}{\lambda_{no}\, (1 + \epsilon \frac{\lambda_{n1}}{\lambda_{no}} + \epsilon^2 \frac{\lambda_{n2}}{\lambda_{no}})}$$

$$\doteq \frac{1}{2} \Sigma \frac{a_{no}^2 + (2\, a_{no} a_{n1}) + O(\epsilon^2)}{\lambda_{no}(1 - (-\epsilon \frac{\lambda_{n1}}{\lambda_{no}}))}$$

$$\doteq \frac{1}{2} \Sigma \frac{\{a_{no}^2 + \epsilon(2\, a_{no} a_{n1}) + O(\epsilon^2)\}\ (1 - \epsilon \frac{\lambda_{n1}}{\lambda_{no}} + O(\epsilon^2)}{\lambda_{no}}$$

$$\doteq \frac{1}{2} \Sigma \frac{a_{no}^2}{\lambda_{no}} + \epsilon \{\Sigma \frac{a_{no} a_{n1}}{\lambda_{no}} - \frac{1}{2} \Sigma \frac{a_{no}^2}{\lambda_{no}} (\frac{\lambda_{n1}}{\lambda_{no}})\} + O(\epsilon^2)$$

$$\psi(\epsilon) = \psi_o + \epsilon\,\psi_1 \tag{3.50}$$

$$\psi_1 = \frac{1}{2}\,\Sigma\,\frac{a_{no}}{\lambda_{no}}\,(2\,a_{nl} - a_{no}\,\frac{\lambda_{nl}}{\lambda_{no}}) \tag{3.51}$$

If $\frac{\lambda_{nl}}{\lambda_{no}}$ is small for $n = 1, 2, \ldots$ ψ_1 may be positive if it is convergent. However, there is no apparent reason to rule out $\psi_1 < 0$.

Nevertheless, $\psi(\epsilon)$ is still positive since $\epsilon\,\frac{\lambda_{nl}}{\lambda_{no}} << 1$

$$\psi(\epsilon) = \frac{1}{2}\,\Sigma\,\frac{a_{no}^2}{\lambda_{no}}\,(1 - \epsilon\,\frac{\lambda_{nl}}{\lambda_{no}}) + \epsilon\,\frac{1}{2}\,\Sigma\,\frac{2\,a_{no}a_{nl}}{\lambda_{no}} \tag{3.52}$$

so $\psi(\epsilon) > 0$.

There is a strong resemblance between 3.38 and 3.52. This means that the unperturbed relations 3.46 is kept after perturbation in the since that

$$\psi(\epsilon) = \frac{1}{2}\,\mathrm{Var}\,\gamma(t) = \beta^2(\epsilon)\;.$$

Now we can compare both expression for η before and after perturbation:

$$\eta_o = \gamma_o - \psi_o = \Sigma\,\frac{a_{no}w_{no}}{\lambda_{no}} - \frac{1}{2}\,\Sigma\,\frac{a_{no}^2}{\lambda_{no}} \tag{3.53a}$$

$$\eta(\epsilon) = \Sigma\,\frac{a_n(\epsilon)\,w_n(\epsilon)}{\lambda_n(\epsilon)} - \frac{1}{2}\,\Sigma\,\frac{a_n^2(\epsilon)}{\lambda_n(\epsilon)} \tag{3.53b}$$

$$= \gamma(\epsilon) - \psi(\epsilon)$$

But to the first order:

$$\eta(\epsilon) = (\gamma_o + \epsilon\,\gamma_1) - (\psi_o + \epsilon\,\psi_1) \tag{3.54a}$$

$$= (\gamma_o - \gamma_o) + \epsilon(\gamma_1 - \psi_1)$$

$$\eta(\epsilon) = \eta_o + \epsilon\eta_1 \tag{3.54c}$$

From 3.24 and 3.51,

$$\eta_1 = \{\Sigma \frac{a_{ko}w_{kl}}{\lambda_{ko}} + \Sigma \frac{a_{kl}w_{ko}}{\lambda_{ko}} - \Sigma \frac{a_{ko}w_{ko}}{\lambda_{no}} (\frac{\lambda_{kl}}{\lambda_{ko}})\} + \ldots \tag{3.55}$$

$$- \frac{1}{2} \Sigma \frac{a_{ko}}{\lambda_{ko}} (2 a_{kl} - a_{ko} \frac{\lambda_{kl}}{\lambda_{ko}})$$

We can see from 1.15 that $\text{Var}\ \eta = = \text{Var}\ \gamma$.

$$\text{Var}\ \eta_o = \text{Var}\ \gamma_o \tag{3.56a}$$

$$\text{Var}\ \eta(\epsilon) = \text{Var}\ \gamma(\epsilon) \tag{3.56b}$$

Hence,

$$\Delta\ \text{Var}\ \eta = |\text{Var}\ \eta(\epsilon) - \text{Var}\ \eta_o| = |\text{Var}\ \gamma(\epsilon) - \text{Var}\ \gamma_o|$$

$$\Delta\ \text{Var}\ \eta = \Delta\ \text{Var}\ \gamma$$

The change in variance then is nil relative to consideration of the ψ expression.

Say the threshold for η is t_η and for γ is t_γ

$$\gamma \gtrless t_\gamma \qquad\qquad \eta \gtrless t_\eta$$

are the respective likelihood tests. The differences between thresholds are

$$\Delta t = t_{\gamma o} - t_{\eta o} = \frac{1}{2} \Sigma \frac{a_{no}^2}{\lambda_{no}}$$

$$\Delta t(\epsilon) = t_{\gamma(\epsilon)} - t_{\eta(\epsilon)} = \frac{1}{2} \Sigma \frac{a_n^2(\epsilon)}{\lambda_n(\epsilon)}$$

$$\Delta t_o = \frac{1}{2} \Sigma \frac{a_{no}^2}{\lambda_{no}} = \psi_o \tag{3.55}$$

$$\Delta t(\epsilon) = \psi(\epsilon) = \psi_o + \epsilon \psi_1 \tag{3.56}$$

$$\text{change in } \Delta t = \Delta t(\epsilon) - \Delta t_o = \epsilon \psi_1 \tag{3.57}$$

Equation 3.57 is noteworthy. It gives the magnitude of the change in threshold needed to preserve a true maximum likelihood test (or modified) quaranteed by the Neyman Pearson lemma to be the most powerful test for determining H_0 or H_1.

Thus, in order to consider energy changes manifested by ϵR_1 and their effect on the detection test used, namely the likelihood ratio, the first order approximation to the threshold dispersion between η and γ required to maintain it is less than the 3.57 value. To calculate the actual threshold change required to maintain the likelihood ratio test we need the rule for deciding what the new threshold will be. Suppose for example the threshold t_η chosen is midway between the means of η under hypothesis 1 and 0. From 3.53a

$$\begin{aligned} E_o \eta_o &= -\frac{1}{2} \frac{a_{no}^2}{\lambda_{no}} \\ E_1 \eta_o &= \frac{1}{2} \Sigma \frac{a_{no}^2}{\lambda_{no}} \end{aligned} \tag{3.58}$$

then $\quad t_{\eta o} = 0$

From 3.54c and 3.55

$$E_o \eta(\epsilon) = E_o \eta_o + \epsilon E_o \eta_1$$

$$E_o\,\eta(\epsilon) = -\frac{1}{2}\,\Sigma\,\frac{a_{no}^2}{\lambda_{no}} + \frac{1}{2}\epsilon\,\Sigma\,\frac{a_{ko}}{\lambda_{ko}}(a_{ko}\,\frac{\lambda_{kl}}{\lambda_{ko}} - 2\,a_{kl})$$

$$E_1\,\eta(\epsilon) = \frac{1}{2}\,\Sigma\,\frac{a_{no}^2}{\lambda_{no}} + \epsilon\,\{2\,\Sigma\,\frac{a_{kl}a_{ko}}{\lambda_{ko}} - \Sigma\,\frac{a_{kl}a_{ko}}{\lambda_{ko}} - \Sigma\,\frac{a_{ko}^2}{\lambda_{ko}}\,\frac{\lambda_{kl}}{\lambda_{ko}} + \ldots$$

$$\ldots + \frac{1}{2}\,\Sigma\,\frac{a_{ko}^2}{\lambda_{ko}}\,\frac{\lambda_{kl}}{\lambda_{ko}}\,\}$$

$$E_1\,\eta(\epsilon) = \frac{1}{2}\Sigma\,\frac{a_{no}^2}{\lambda_{no}} + \epsilon\,\{\Sigma\,\frac{a_{kl}a_{ko}}{\lambda_{ko}} - \frac{1}{2}\,\Sigma\,\frac{a_{ko}^2}{\lambda_{ko}}\,(\frac{\lambda_{kl}}{\lambda_{ko}})\,\}$$

$$t_{\eta(\epsilon)} = \frac{1}{2}\,\{\,E_o\eta(\epsilon) + E_1\,\eta(\epsilon)\}$$

$$= \frac{1}{2}\{\frac{1}{2}\,\epsilon\{\,\Sigma\,\frac{a_{ko}^2}{\lambda_{ko}}\,(\frac{\lambda_{kl}}{\lambda_{ko}}) - 2\,\Sigma\,\frac{a_{ko}a_{kl}}{\lambda_{ko}}\} + \epsilon\,\{\Sigma\,\frac{a_{kl}a_{ko}}{\lambda_{ko}} - \frac{1}{2}\,\frac{a_{ko}^2}{\lambda_{ko}}\,\frac{\lambda_{kl}}{\lambda_{ko}}\,\}\,\}$$

$$t_{\eta(\epsilon)} = 0 \tag{3.59}$$

$$\Delta t_\eta = t_{\eta(\epsilon)} - t_{\eta o} = 0 \tag{3.60}$$

For the γ test:

$$\left.\begin{aligned}
E_o\,\gamma_o &= E_o\,\eta_o + E_o\,\psi_o = E_o\,\eta_o + \psi_o\\
E_1\,\gamma_o &= E_1\,\eta_o + E_1\,\psi_o = E_1\,\eta_o + \psi_o\\
E_o\,\gamma(\epsilon) &= E_o\,\eta(\epsilon) + E_o\,\psi(\epsilon) = E_o\,\eta(\epsilon) + \psi(\epsilon)\\
E_1\,\gamma(\epsilon) &= E_1\,\eta(\epsilon) + E_1\,\psi(\epsilon) = E_1\,\eta(\epsilon) + \psi(\epsilon)
\end{aligned}\right\} \tag{3.61}$$

$$t_{\gamma o} = \frac{1}{2}\,\{\psi_o + E_o\,\eta_o + E_1\,\eta_o\} = t_{\eta o} + \frac{1}{2}\,\psi_o$$

$$t_{\gamma(\epsilon)} = \frac{1}{2}\,\psi(\epsilon) + t_{\eta(\epsilon)}$$

$$\triangle t_\gamma = t_{\gamma(\epsilon)} - t_{\gamma o} = \triangle t_\eta + \frac{1}{2}(\psi(\epsilon) - \psi_o) \tag{3.62}$$

For example, we have from 3.51 and 3.59,

$$\triangle t_\gamma = t_{\eta(\epsilon)} + \frac{1}{2}\epsilon\,\psi_1 \tag{3.63}$$

$$= \frac{1}{2}\epsilon\left[\left\{\Sigma\ \frac{a_{ko}a_{kl}}{\lambda_{ko}} - \frac{1}{2}\Sigma\ \frac{a_{no}^2}{\lambda_{no}}\ \frac{\lambda_{nl}}{\lambda_{no}}\right\}\right]$$

$$\triangle t_\gamma = \triangle t_\gamma = \frac{1}{2}\epsilon\,\psi_1 \tag{3.64}$$

<u>The 3.64 result is an example of the change in threshold required to maintain the modified maximum likelihood test in a detection problem with second order perturbation of the form ϵR_1.</u>

What we have said about convergence in the past still holds throughout this section. If convergence occurs then a singular case is in play. As mentioned in chapter one, we are acutely interested in the nonsingular case for perturbation. It is the nonsingular case in which $\Sigma \frac{a_{ko}^2}{\lambda_{ko}} < \infty$ is generally found. We see that in this case 3.64 is convergent if $\frac{\lambda_{kl}}{\lambda_{ko}} \leq 1$ for $k = 1, 2, 3, \ldots$ or at least for infinitely many k. It is significant to note the direct dependence on ϵ of 3.64. If the series is convergent, for small parameter ϵ the quantity of changes in threshold necessary is small also.

CHAPTER IV

AN EXAMPLE AND THESIS CONCLUSIONS

The ideas and theory expounded in Chapter III will be put into more tangible form in this chapter with an example taken in this case of perturbing a wide sense stationary autocorrelation with a non wide sense stationary one. The signal chosen will be a sine wave. After a few comments on this example a general summary of the thesis will follow in which its results will be coordinated and explicitly outlined.

I. AN EXAMPLE

In order to avoid tedious mathematical detail the example to be taken here will be one for which the unperturbed and perturbed kernels are in simplest terms. The procedure for extracting the eigenfunctions of the unperturbed kernel is outlined in Davenport and Root (pp. 99-101 and Appendix p. 371) so that only the final results of that extraction will be shown here.

The autocorrelation function of the original or nuclear noise process is

$$R_0(t, s) = R_0(|t-s|) = e^{-|t-s|}, \quad -T \leq s, t \leq T,$$

whose eigenfunctions and eigenvalues can be shown to be solutions to the differential equation

$$\lambda \, \phi''(t) + \frac{2-\lambda}{\lambda} \, \phi(t) = 0 \tag{4.1}$$

with $0 < \lambda < 2$ only.

The eigenfunctions and eigenvalues of the unperturbed kernel are known to be in a split form:

$\phi_k(t) = c_k \cos b_k t$	(4.2a)	$\hat{\phi}_k(t) = \hat{c}_k \sin \hat{b}_k t$	(4.2b)
$b_k \tan t_k T = 1$	(4.3a)	$\hat{b}_k \cot \hat{b}_k T = 1$	(4.3b)
$\lambda_k = \dfrac{2}{b_k^2 + 1}$	(4.4a)	$\lambda_k = \dfrac{2}{\hat{b}_k^2 + 1}$	(4.4b)
$c_k = \dfrac{1}{\sqrt{T + \dfrac{\sin 2 b_k T}{2b_k}}}$	(4.5a)	$c_k = \dfrac{1}{\sqrt{T - \dfrac{\sin 2 \hat{b}_k T}{2 \hat{b}_k}}}$	(4.5b)

By choosing $R_1(t, s) = e^{(-|t| - |s|)}$ a non wide sense stationary covariance function the three important equations of first order perturbation,

$$(R_0 - \lambda_{ko} I)\,\phi_{k1} = (\lambda_{k1} I - R_1)\,\phi_{ko} \qquad (3.25a)$$

$$(\phi_{k1} \,|\, \phi_{ko}) = 0 \qquad (3.25b)$$

$$\lambda_{k1} = (\phi_{ko} \,|\, R_1\, \phi_{ko}) \qquad (3.25c)$$

$$k = 1, 2, \ldots$$

can now be put into integral forms. Equations 3.25c, 4.2 are used to find

$$\lambda_{k1} = \frac{4\, c_k^2}{(1 + b_k^2)^2} = c_k^2 \lambda_{ko}^2 \qquad (4.6)$$

$$\frac{1}{\sqrt{2T}} \leq c_k \leq \frac{1}{\sqrt{T}} \qquad k = 1, 2, \ldots \qquad (4,7)$$

The dual nature of the eigenstructure of R_0 incurs a similar calculation for $\hat{\lambda}_{k1}$.

$$\hat{\lambda}_{k1} = 0 \qquad\qquad k = 1, 2, \ldots$$

To find ϕ_{k1} equations 3.25a, 3.26c are used along with the new found knowledge of 4.6, 4.2 and 4.4. The resulting calculations are lengthy but involve only algebraic manipulations.

$$\begin{aligned}
\phi_{k1} &= B_k \, t \sin b_k t + A_k \cos b_k t \\
B_k &= c_k^2 \, \lambda_{ko}^2 \\
A_k &= \frac{c_k^3}{b_k^2} \left(\frac{1}{2} - T c_k^2 \cos^2 b_k T \right) \\
\hat{\phi}_{k1} &= 0 \, .
\end{aligned} \qquad (4.9)$$

A few characteristics of this particular example are examined before Chapter III equations involving the change in variance are brought in.

First, we note that $\lambda_{k1} < \lambda_{ko}$ for infinitely many k since $\lambda_{ko} \to 0$ as $b_k \to \infty$ and eventually $c_k^2 \lambda_{ko} < 1$. From 4.6 $\lambda_{k1} = c_k^2 \lambda_{ko}^2$ where c_k is bounded above and below

$$\frac{1}{\sqrt{2T}} \leq c_k = \frac{1}{\sqrt{T + \dfrac{\sin 2 b_k T}{2 \, bk}}} \leq \frac{1}{\sqrt{T}}$$

Second, a_{k1} is related to a_{k0} in the following manner from 4.9:

$$a_{k1} = c_k^2 \lambda_{ko}^2 (t \sin b_k t | \, a(t)) + \frac{c_k^2}{b_k^2} \left(\frac{1}{2} - T c_k^2 \cos^2 b_k T \right) a_{ko} \qquad (4.13)$$

Recalling $\hat{\lambda}_{k1} = 0$ and $\hat{\phi}_{k1} = 0$, we have $\hat{a}_{k1} = 0$. (4.14)

The nuclear autocorrelation function $R_0(t, s)$ has the split nature of eigenfunction sets but only one of the sets will be perturbed by the kernel $R_1(s,t) = e^{-|t| - |s|}$.

If the signal $a(t) = A \sin \omega_m t$, $-T \leq t \leq T$ the following coefficients result:

$$\hat{a}_{ko} = 0$$

$$a_{k0} = A\, c_k \left[\frac{\sin (b_k - \omega_m)T}{(b_k - \omega_m)} - \frac{\sin (b_k + \omega_m)T}{(b_k + \omega_m)} \right] \tag{4.15}$$

$$a_{k1} = \frac{c_k^2}{b_k^2} \left\{\frac{1}{2} - T c_k^2 \cos^2 b_k T\right\} c_k A \left\{ \frac{\sin (b_k - \omega_m)T}{b_k - \omega_m} - \frac{\sin (b_k + \omega_m)T}{b_k + \omega_m} \right\} \tag{4.16}$$

There exists ω_m such that

$$a_{\bar{k}o} = 0 = a_{\bar{k}1}\,, \quad \text{e.g. } \omega_m = \frac{1}{3} b_{\bar{k}} \text{ for some } \bar{k} > 0.$$

Finally, the variance of the perturbed likelihood ratio can be computed from 3.38 for this example.

$$\frac{\lambda_{k1}}{\lambda_{ko}} = \frac{c_k^2 \lambda_{ko}^2}{\lambda_{ko}} = c_k^2 \lambda_{ko} \leq \frac{\lambda_{ko}}{T} \tag{4.17}$$

$$\operatorname{Var}(\gamma_o + \epsilon \gamma_1) = \Sigma \frac{a_{ko}^2}{\lambda_{ko}} \left(1 - \epsilon \frac{\lambda_{k1}}{\lambda_{ko}}\right) + \epsilon\, \Sigma \frac{2 a_{k1} a_{ko}}{\lambda_{ko}} \tag{3.38}$$

For large b_k,

$$c_k^2 \sim \frac{1}{T} \tag{4.18}$$

so that the series term of 3.38 approaches

$$\sim \frac{a_{ko}^2}{\lambda_{ko}} \left(1 - \epsilon \frac{2}{T}\right) + \epsilon \left\{ 2 a_{ko}^2 \frac{\frac{1}{T}}{2 b_k^2} \left(\frac{1}{2} - \cos^2 b_k T \right) \right\} \tag{4.19}$$

$$\sim \frac{a_{ko}^2}{\lambda_{ko}} \left(1 - \epsilon \frac{\lambda_{ko}}{T}\right) \tag{4.20}$$

II. CONCLUSIONS

An attempt has been made in this thesis to underscore the importance of the stability of a signal detection scheme. An example of a scheme, namely one involving the likelihood ratio for a sure signal in zero-mean Gaussian noise with continuous autocorrelation on a finite interval J, has been perturbed by a slight change in the second order noise statistics. After reviewing previous work and setting up simple examples, a more general case of perturbation is formulated in additively disturbing the autocorrelation function with an $L_2(J)$ positive semidefinite kernel. The change in form and variance of the perturbed likelihood ratio is found to be of the order of ϵ, a small positive parameter. Using the parameter strategically, continuity of the signal detection scheme is shown through stability of a finite-variance likelihood ratio. The concept kept in mind throughout Chapters III and IV (Section I) which encompass the new results is that the detection test was to be kept the optimum Bayes procedure both before and after perturbation.

Changes in the form of the likelihood ratio is illustrated by equation 3.23a while change in variance is shown by 3.38. Threshold dispersion between the true Gaussian likelihood ratio and the one used predominantly in the thesis is demonstrated by 3.57. In selecting the point midway between the means of the ratio under either hypothesis (before and after perturbation), the change in threshold necessary for an optimum signal detection scheme is shown to vary directly also with ϵ parameter (equation 3.64).

In Chapter IV, Section I, kernels corresponding to the R_0 and R_1 operators in equation 3.25 are chosen, the former a wide sense stationary one and the latter a non-wide sense stationary one, so that the forms of the series in the results of Chapter III involving the changes in form and variance of the likelihood ratio can be seen more easily. Choosing the sine wave as the sure signal, the perturbed forms are seen to vary directly and simply with parameter ϵ, most clearly shown for large index as demonstrated by 4.20.

The relevance of showing continuity of a detection scheme is made clear by the fact that there is no guarantee the hypothesis test for detecting the presence of a sure signal in additive noise is robust or stable. If not stable relative to small changes in the noise energy, the error probabilities would be liable to drastic change and the detection model would then be little more than useless. What this thesis has <u>not</u> done is shown how good the additive Gaussian noise with continuous autocorrelation function model is in relation to noise experienced in actual signal detection equipment. We have not tried to defend the model as a good or true one, but have tried to show its stable character. In this way perhaps, an indirect defense for its "goodness" may be implied.

APPENDICES

A. HILBERT SPACE

Many concepts in signal detection theory can be put into simpler mathematical forms when the notion of a hilbert space is introduced. In this appendix section are outlined some important hilbert space concepts which are used throughout the text.

I. PREHILBERT SPACE

A prehilbert space (PHS) is a complex vector space P with a scalar product of any two vectors, $x, y \in P$, denoted by $(x|y)$, defined with the following properties:

1. $(x|y) = \overline{(y|x)}$ (where bar denotes complex conjugate)
2. $(x + y|z) = (x|z) + (y|z)$
3. $(\lambda x|y) = \lambda (x|y)$, λ a scalar.
4. $(x|y) > 0$ when $x \neq 0$, the zero vector.

The above four statements imply the following elementary theorems:

<u>Theorem</u>:

1. $(x|y + z) = (x|y) + (x|z)$
2. $(x|\lambda y) = \lambda (x|y)$
3. $(\theta|y) = (y|\theta) = 0$
4. $(x - y|z) = (x|z) - (y|z)$; $(x|y - z) = (x|y) - (x|z)$
5. If $(x|z) = (y|z)$ for all z then $x = y$.

Every PHS is a metric space with the distance function defined through the inner product thus:

$$d^2(x, y) = (x-y|\, x-y) = \| x-y \|^2 .$$

Hence, the NORM of a vector in a PHS is given by

$$\| x \| = \sqrt{(x|\, x)} .$$

Properties of the norm are expressed through the following:

Theorem:

1. $\| \lambda x \| = | \lambda | \; \| x \|$
2. $\| x \| > 0$ if $x \neq \theta$, the zero vector.
3. $\| x + y \|^2 + \| x-y \|^2 = 2 \| x \|^2 + 2 \| y \|^2$ Parallelogram law
4. $|(x|\, y)| = 1/4 \{ \| x+y \|^2 - \| x-y \|^2 + i \| x+iy \|^2 - i \| x-iy \|^2 \}$

   Polarization identity
5. $| (x|\, y) | \leq \| x \| \; \| y \|$ Cauchy-Schwartz-Boniakovsky Inequality (CBS Inequality)
6. $\| x+y \| \leq \| x \| + \| y \|$
7. $\| x-y \| \geq 0$; $\| x-y \| = 0 \leftrightarrow x = y$
8. $\| x-y \| = \| y-x \|$
9. $\| x-z \| \leq \| x-y \| + \| y-z \|$

Metric character (6–9)

II. HILBERT SPACE

Once the norm is defined, it is possible to speak of the convergence of a sequence in a prehilbert space and only then will the concept of a hilbert space naturally follow.

Definitions. Let x_m be a sequence of vectors in a PHS.

1. The sequence of vectors converges to a limit vector x if $\| x_n - x \| \to 0$ as $n \to \infty$, i.e., give $\epsilon > 0$, $\exists N \ni n > N$ implies $\| x_n - n \| < \epsilon$ and x is unique in the PHS.

2. The sequence of vectors is <u>Cauchy</u> if $\| x_n - x_m \| \to 0$ as $m, n \to \infty$, i.e., given $\epsilon > 0$ $\exists N \ni \| x_m - x_n \| < \epsilon$ if $m, n > N$.
3. A PHS is <u>Complete</u> if every Cauchy sequence converges.

   A PHS which is complete is called a <u>hilbert space</u>.

To find a set of vectors which will generate the entire hilbert space H is a problem demanding a few more specialized definitions. By the word "generate" we mean that a sequence $\{x_k\}_{k=1}^{\infty}$ generates H if for any vector x in H, scalars $\{a_k\}_{k=1}^{\infty}$ exist such that

$$x = \sum_{k=1}^{\infty} a_k x_k \ .$$

<u>Definitions</u>.

1. For x, y in a PHS, x is <u>orthogonal</u> to y iff $(x | y) = 0$, that fact denoted by $x \perp y$. If $x \perp x_j$, $j = 1, 2, 3, \ldots$ then $x \perp \{x_j\}_{j=1}^{\infty}$ and $x \perp [\{x_j\}]$.
2. An <u>orthogonal sequence</u> $\{x_j\}$ has $x_k \perp x_j$, $j \neq k$.
3. An <u>orthonormal sequence</u> $\{x_j\}$ has $x_k \perp x_j$, $j \neq k$ $\| x_k \| = 1$ for $k = 1, 2, 3, \ldots$
4. A set S of vectors in a PHS P is <u>total</u> if the only vector z of P orthogonal to every vector of S is the zero vector $z = \theta$. A <u>total sequence</u> is defined accordingly.
5. An <u>orthonormal basis</u> for a hilbert space H is a set $\{x_n\}_{n=1}^{\infty}$ which is both total and orthonormal.
6. A hilbert space is separable if it possesses a total set in V, the vector space it is defined over.

B. ℓ^2, L^2 AND OPERATORS

Two prime examples of hilbert spaces are those denoted by L^2 and ℓ^2. Besides being the best known and most utilized spaces they provide exceptional insight into general hilbert space theory. In the following we will try to give an informal definition of these two hilbert spaces after which we will proceed to give a short introduction into operators and their character in L^2 and ℓ^2.

I. L^2 AND ℓ^2

The space ℓ^2 is the set of all square summable complex sequences. Each sequence is a countably infinite dimensional vector. The inner product associated with ℓ^2 is

$$(a \mid b) = \sum_{j=1}^{\infty} a_j \, b_j \qquad \text{where } a = (a_1, a_2, \ldots) \text{ and } b = (b_1, b_2, \ldots) \text{ and both belong to } \ell^2.$$

The space L^2 is the set of all square Lebesque-summable complex functions over the open infinite plane. An analogous space, $L^2[a, b]$ is the set of all square Lebesque-summable complex functions over the finite interval $[a, b]$. The inner product used in $L^2[a, b]$ is demonstrated by the following:

$$\int_a^b f(t) \, \overline{g(t)} \, dt = (f \mid g)$$

where $f(t)$ and $g(t)$ both belong to $L^2[a, b]$.

II. OPERATORS

An operator or linear transformation defined on a hilbert space H is a continuous linear mapping T such that

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty$$

is a bona fide vector in H for any x, y in H and any scalars α, β in the scalar field defined with the vector space.

A bounded operator is one for which there exists a constant M greater than zero for which the norm of any vector in its range is less than M times the vector's norm. That is, T is bounded if and only if for all x in H there exists a $M > 0$ such that $\|Tx\| \leq M \|x\|$. The smallest such constant M such that the preceding holds is called the <u>operator norm</u>, denoted by $\|T\|$

$$\|T\| = \inf_{M} \{ M \text{ a constant} > 0 : \|Tx\| \leq M\|x\| \text{ all } x \text{ in } H. \}$$

The collection of all bounded operators is denoted by $\mathcal{L}_c$. It is easily seen that a bounded operator will carry bounded sets into bounded sets. If it also carries any bounded set into a convergent set than it is termed CC, Completely Continuous or Compact. A CC operator is sometimes referred to as a "small" operator since it somewhat "compresses" a vector set which is bounded into a convergent form.

An operator's adjoint operator is defined by the following equation:

$$(Tx | y) = (x | T^{*}y) \quad \text{for all } x, y \text{ in } H, \text{ a hilbert space.}$$

We can now define more operator types using this "adjoint" notion.

Definitions. Let T be an operator defined on H, a hilbert space.

1. T is isometric if $T^*T = I$ where I is the identity operator, i.e., $Ix = x$ for all x in H.
2. T is unitary if $T^*T = TT^* = I$.
3. T is self adjoint if $T^* = T$.
4. T is normal if $T^*T = TT^*$.

An operator T has as a proper value or eigenvalue and x as its corresponding eigenvector if $Tx = x$. It will be seen in appendix C that self adjoint operators are easily characterized by their eigenvalues and eigenvectors.

Operators in L^2 are bivariate kernels of either two complex variables if L^2 is complex function space or two real variables if L^2 is restricted to be a real function space. The space ℓ^2 has infinite dimensional matrices as its operators. Both types of operators are characterized by the following:

1. $$(Tx)(s) = \int_0^T T(t, s)\, x(t)\, dt = y(s) \qquad 0 \leq s \leq T$$

   ($Tx = y$ in $L_2[0, T]$)

2. $$T = \begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ & & \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \end{bmatrix}$$

   $$Tx = \begin{bmatrix} \Sigma\, a_{1j}\, x_j \\ \Sigma\, a_{2j}\, x_j \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = y$$

   ($Tx = y$ in ℓ^2).

C. L_2 KERNELS

The space of functions which provides a platform for the integral equations we are going to consider is $L^2[0, T]$. Integral operators on this space, denoted by capital roman letters R, K, ..., have corresponding kernels $R(x, y)$, $K(x, y)$ some of whose properties are explained in the following discussion.

If $\alpha(t)$ and $\beta(t)$ belong to $L^2[0, T]$ the general integral equation of interest can be written

$$\int_0^T R(t, s)\, \alpha(t)\, dt = \beta(s) \qquad 0 \leq s \leq T \tag{C.1}$$

Abstractly, $R\alpha = \beta$. If for every $\alpha \in L^2[0, T]$, $R\alpha$ belongs to $L^2[0, T]$ also then we say R is an $L^2[0, T]$ kernel and is "defined" over all $L^2[0, T]$. Otherwise R is simply an operator and $R(t,s)$ is its kernel.

For any element, $f(t) \in L^2[0, T]$, its norm, $\| f \|$, is defined by the following:

$$\| f \| = \sqrt{\int_0^T | f(t) |^2 \, dt} \; . \tag{C.2}$$

Convergence in the space $L^2[0, T]$ is based on this norm.

$$\lim_{N \to \infty} \int_0^T | g(t) - \Sigma \, a_k f_k(t) |^2 \, dt = 0 \tag{C.3}$$

$$\longleftrightarrow g(t) = \lim_{N \to \infty} \sum_{k=1}^{N} a_k f_k(t)$$

$$\longleftrightarrow \| g(t) - F_N \| \to 0 \text{ as } N \to \infty \quad \text{where}$$

$$F_N = \sum_{k=1}^{N} a_k f_k(t)$$

A kernel $K(s,t)$ is <u>self adjoint</u> or <u>hypermaximal</u> if it is defined over all $L^2[0,T]$ and

$$K(s,t) = \overline{K(t,s)} \ .$$

If the latter equation holds but $K(s,t)$ is defined only on a dense subspace of $L^2[0,T]$ then it is <u>hermitian</u>. Symmetric kernels are usually thought of as real $L^2[0,T]$ kernels which have the following property:

$$R(t,s) = R(s,t) \ . \tag{C. 4}$$

Autocorrelation functions form symmetric kernels if they are a product of a second order real stochastic process. Also, autocorrelations are either nonnegative or positive kernels (the latter implies the former):

$$\left.\begin{array}{ll} \int_0^T K(s,t)\, g(s)\, \overline{g(t)}\, ds\, dt \geq 0 & \text{nonnegative definite} \\ \int_0^T K(s,t)\, g(s)\, \overline{g(t)}\, ds\, dt > 0 & \text{positive definite} \end{array}\right. \tag{C. 5}$$

where $K(s,t)$ is the kernel in question and $g(t)$ is any member of $\mathcal{D}_k$, the domain of operator K which has $K(s,t)$ as its kernel.

Symmetric kernels have at least one eigenvalue λ and corresponding eigenvector ϕ. That is,

$$\int_0^T K(x,y)\, \phi(x)\, dx = \lambda\, \phi(y) \qquad 0 \leq y \leq T \tag{C. 6}$$

holds for some λ and $\phi(y)$ where $K(s,t)$ is a symmetric kernel. An extension of this is given by the following theorem.

<u>Theorem.</u> Every nonzero symmetric $L^2[0, T]$ kernel either has an infinite number of eigenvalues or is a PG kernel. (Tricomi, p. 105).

A Pincherle-Goursat kernel (PG) or a kernel of finite rank has a finite number of eigenvalues and can be represented by

$$K(s, y) = \sum_{j=1}^{\infty} \lambda_j \phi_j(x) \overline{\phi_j(y)} \tag{C.7}$$

where $\{\lambda_j\}_{j=1}^{N}$ and $\{\phi_j\}_{j=1}^{N}$ are its eigenvalues and eigenvectors, respectively.

For any symmetric kernel, if $\sum_{j=1}^{\infty} \lambda_j \phi_j(s) \overline{\phi_j(t)}$ converges uniformly then

$$K(s, t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(s) \overline{\phi_j(t)} \tag{C.8}$$

where $\{\lambda_j\}_{j=1}^{\infty}$ and $\{\phi_j\}_{j=1}^{\infty}$ are $K(s, t)$'s eigenvalues and eigenfunctions as above. Whether $\{\phi_k\}_{k=1}^{\infty}$ is complete or not the following is true for any symmetric kernel; $K(x, y)$:

$$K(x, y) = \lim_{N \to \infty} \sum_{j=1}^{N} \lambda_j \phi_j(x) \phi_j(y) \tag{C.9}$$

Mercer's theorem claims uniform as well as mean square convergence for symmetric, continuous, positive definite kernels. That is, both C.8 and C.9 hold for such kernels.

Orthogonality of a function in $L^2[0, T]$ relative to the kernel $K(s, t)$ is expressed by

$$\int_0^T K(s, t) \; y(s) \, ds = 0 \; . \tag{C.10}$$

If C. 10 holds for a symmetric kernel than $y(s)$ is orthogonal to K's eigenfunctions (Tricomi p. 109).

Picard's theorem gives a criterion for the expansion of a function in terms of the kernel's eigenfunction set:

Picard Theorem. The equation for symmetric kernel $R(t, s)$,

$$\int_o^T R(t, s)\, x(t)\, dt = y(s) \quad \text{for} \quad 0 \le s \le T,$$

has a solution in $L^2[0, T]$ if and only if

(a) $y(t) = \text{l.i.m.} \sum_{n=1}^{\infty} \phi_n(t)$

(b) $y_n = \int_o^T y(t)\, \phi_n(t)\, dt$

(c) $\sum_{n=1}^{\infty} \left| \frac{y_n}{\lambda_n} \right|^2 < \infty$ where $\{\lambda_n\}_{n=1}^{\infty}$ are $R(t, s)$ eigenfunctions.

(d) $x(t) = \text{l.i.m.} \sum_{n=1}^{N} \frac{y_n\, \phi_n(t)}{\lambda_n} \qquad 0 \le t \le T$

The expansion expressed in (d) is unique if $\{\phi_n(t)\}_{n=1}^{\infty}$ is complete.

D. KARHUNEN LOEVE EXPANSIONS

Of singular importance in the study of vector spaces or stochastic processes is a spectral theorem which breaks down into simpler components either operators defined on the vector space or the stochastic process. We will give a version of the spectral theoremin hilbert space (CC self adjoint operators) and will follow that with the analogous spectral theorem in stochastic processes known as the Karhunen-Loève Expansion theorem.

Spectral theorem for self adjoint CC operators.

A nonnull CC self adjoint operator T in a hilbert space H has:

1. an eigenvector set $\{\phi_n\}_{n=1}^{\infty}$ corresponding to the existent eigenvectors $\{\lambda_n\}_{n=1}^{\infty}$ of T.
2. $\lambda_j \to 0$ as $j \to \infty$.
3. $\{\phi_j\}_{j=1}^{\infty}$ generate all H if T has no null space except θ.
4. $Tx = \sum_{j=1}^{\infty} \lambda_j (\phi_j | x) \phi_j$ for any x in H.

Karhunen-Loève Expansion theorem.

A stochastic process $x(t)$ with zero mean on $[0, T]$ and continuous autocorrelation function $R(t, s) = E\, x(t) \overline{x(s)}$ has:

1. a series expansion for $x(t)$ relative to $\{\phi_j\}_{j=1}^{\infty}$ the orthonormal set of $R(t, s)$ eigenfunctions.

$$x(t) = \sum_{n=1}^{\infty} x_n \phi_n(t) \quad \text{where} \quad x_n = \int_0^T x(t)\, \phi_n(t)\, dt$$

2. $Ex_n x_k = 0$ for $n \neq k$ and $E|x_n|^2 = \lambda_n$.

3. $E \left| x(t) - \sum_{n=1}^{N} x_n \phi_n(t) \right|^2 \to 0$ as $N \to \infty$ for all t in $[0, T]$.

4. the above expansion of $x(t)$ unique and

$$\int_o^T R(t, s)\, \phi_n(s)\, ds = \lambda_n \phi_n(t) \qquad 0 \leq t \leq T$$

BIBLIOGRAPHY

Berberian, S. K., Introduction to Hilbert Space. New York: Oxford, 1961. 206 pp.

Corson, E. M., Perturbation Methods in Quantum Mechanics of n Electron Systems. New York: Hafner, 1950. 308pp.

Davenport, W. S., Root, W. L., Introduction to Random Signals and Noise. New York: McGraw Hill, 1958. 393pp.

Fulks, W., Advanced Calculus. New York: Wiley, 1961. 521pp.

Grenander, U., Stochastic Processes and Statistical Inference. Ark. Mat.1: 195-277 (1950).

Masterson, J. J., Notes for Hilbert Space Theory Course, Michigan State Univ. Winter-Spring terms, 1966.

Pitcher, T. S., "Likelihood Ratios for Gaussian Processes," Ark. Mat. 4:5; 35-44, 1959.

Powell, J. L., Crasemann, B., Quantum Mechanics. Reading, Mass.: Addison Wesley, 1961. 495pp.

Rellich, Franz, Perturbation Theory of Eigenvalue Problems, New York Univ. bound lecture notes, 1953.

Riesz, F., Sz.-Nagy, B., Functional Analysis. New York: Ungar, 1955. 468pp.

Root, W. L., Kelly, E. J., "A representation of vector valued random processes," Journal of Math and Physics, 39: 211-16, 1960.

Root, W. L., "Singular Gaussian Measures in Detection Theory," Proc. Symposium Time Series Analysis, Brown Univ. New York: Wiley, 1963. 497pp.

Root, W. L., "Stability in Signal Detection Problems," Proc. Symposium Applied Math, Vol. XVI, 247-263, 1964.

Root, W. L., Detection Theory Course Notes, Michigan State Univ. Spring term, 1966.

Slepian, D., "Some Comments on Detection of Gaussian Signals in Gaussian Noise," IRE Trans. Information Theory, 4:65-8, No. 4, 1958.

Tricomi, F. G., Integral Equations. New York: Interscience, 1957. 238pp.

Wainstein, L. A., Zubakov, V. D., Extraction of Signals from Noise. Englewood Cliffs, N. J.: Prentice Hall, 1962. 382pp.

Wiener, N., Time Series. Cambridge, Mass.: MIT Press. 1949.

Wiener, N., Fourier Integral and Certain of Its Applications. New York: Dover. 1933.