SPHERICALLY INVARIANT PROCESSES:
THEIR NONLINEAR STRUCTURE, DISCRIMINATION, AND ESTIMATION*

by

Steel T. Huang
Department of Mathematics
University of Cincinnati
Cincinnati, Ohio 45221

and

Stamatis Cambanis
Department of Statistics
University of North Carolina
Chapel Hill, North Carolina 27514

ABSTRACT

The structure of the nonlinear space of a spherically invariant process is studied and the problem of discriminating between two spherically invariant processes as well as the problem of nonlinear estimation in spherically invariant processes are solved.

Approved for public release; distribution unlimited.

* This research was supported by the Air Force Office of Scientific Research under Grant AFOSR-75-2796.
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)

NOTICE OF TRANSMITTAL TO DDC

This technical report has been reviewed and is approved for public release IAW AFR 190-12 (7b). Distribution is unlimited.

A. D. BLORE
Technical Information Officer
**Title:** Spherically Invariant Processes: Their Nonlinear Structure, Discrimination and Estimation

**Authors:** Steel T., Huang Stanatis, Cambanis

**Performing Organization Name and Address:**
University of North Carolina
Department of Statistics
Chapel Hill, NC 27514

**Contract or Grant Number(s):** AF-AFOSR-75-2796-75

**Program Element, Project, Task Area & Work Unit Numbers:** 61102F 2304/A5

**Report Date:** 1977

**Number of Pages:** 39

**Distribution Statement:** Approved for public release; distribution unlimited.

**Security Classification:** UNCLASSIFIED

**Supplementary Notes:**
- The structure of the nonlinear space of a spherically invariant process is studied and the problem of discriminating between two spherically invariant processes as well as the problem of nonlinear estimation in spherically invariant processes are solved.

**Keywords:** Spherically invariant processes; Gaussian processes; nonlinear space; discrimination between two spherically invariant processes; nonlinear estimation

**Abstract:**

The structure of the nonlinear space of a spherically invariant process is studied, and the problem of discriminating between two spherically invariant processes as well as the problem of nonlinear estimation in spherically invariant processes are solved.
0. INTRODUCTION

This paper attempts a systematic study of spherically invariant processes (SIP's), i.e., processes whose finite dimensional distributions are mixtures of Gaussian distributions. Section 1 contains the basic properties of SIP's, including their representation in terms of Gaussian processes, which are used throughout. The structure of the nonlinear space of a second order SIP is considered in Section 2. Section 3 solves the problem of discriminating between two second order SIP's, and Section 4 the nonlinear estimation problem for second order SIP's and in particular for Gaussian processes.

Our basic notation and terminology is as follows. \( X = (X_t, t \in T) \) is a stochastic process defined on a probability space \((\Omega, \mathcal{B}, \mathbb{P})\). \( T \) is an arbitrary index set; sometimes it is taken to be a real line interval, but since this is clear from the context it is not emphasized. \( \mathcal{B} \) is usually taken to be \( \mathcal{B}(X) \), the \( \sigma \)-field generated by the random variables of the process \( X \), or \( \overline{\mathcal{B}}(X) \), the completion of \( \mathcal{B}(X) \) with respect to \( \mathbb{P} \). For each \( \omega \in \Omega \), \( X(\omega) \) is the corresponding sample path of the process \( X \), which is an element of \( \mathbb{R}^T \), the space of all functions defined on \( T \). \( X \) induces a probability measure \( \mu = \mathbb{P} \cdot X^{-1} \) on \((\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))\), where \( \mathcal{B}(\mathbb{R}^T) \) is the \( \sigma \)-field generated by the cylinder sets of \( \mathbb{R}^T \). \( X \) is a coordinate process if \((\Omega, \mathcal{B}, \mathbb{P}) = (\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), \mu) \) and \( X_t(\omega) = \omega(t) \). The nonlinear space of \( X \), \( L^2_2(X) = L^2_2(\Omega, \mathcal{B}(X), \mathbb{P}) \), is the set of all \( \mathcal{B}(X) \)-measurable random variables with finite second moment which are called (nonlinear) \( L^2_2 \)-functionals of \( X \). \( X \) is a second order process if \( \mathbb{E}X^2_t < \infty \) for all \( t \in T \). The linear space of a second order process \( X \), \( H(X) \), is the closed subspace of \( L^2_2(X) \) spanned by \( X_t, t \in T \), and its elements are called linear \( L^2_2 \)-functionals of \( X \).
1. SPHERICALLY INVARIANT PROCESSES (SIP's)

It is well known that all mean square estimation problems on Gaussian processes have linear solutions and that Gaussian processes are closed under linear operations. Vershik (1964) showed that these two properties do not uniquely characterize the Gaussian processes. They do, however, characterize the class of SIP's.

Let \( X = (X_t, t \in T) \) be a second order process with mean \( m(t) \) and covariance function \( r(t,s) \). Then \( X \) is said to be an SIP if all r.v.'s in \( H(X-m) \) having the same variance have the same distribution. Also, \( X \) is called degenerate if \( H(X-m) \) is finite dimensional. Since the nonlinear structure of a non-degenerate and a degenerate SIP are considerably different, we will restrict our present investigation to the nondegenerate case; the finite dimensional case will be treated elsewhere. A SIP \( X \) is a mixture of Gaussian processes. It can be determined by its mean \( m(t) \), a covariance function \( R(t,s) \), and a probability distribution \( F(\alpha) \) on \( \mathbb{R}^+ \); the characteristic function of

\[
X_{t_1}, \ldots, X_{t_k} (t_1, \ldots, t_k \in T)
\]

is given by

\[
\exp \left( - \frac{\alpha}{2} \sum_{i,j} R(t_i, t_j)(u_i-m(t_i))(u_j-m(t_j)) \right) dF(\alpha)
\]

(Vershik (1964) and Nagornyi (1970)). In order to avoid the trivial case where \( X \) is a constant process, we will assume throughout that \( F(0^+) < 1 \).

Such an SIP determined by \( m, R \), and \( F \) will be denoted, in short, by SIP\((m,R;F)\).

A probability measure \( \mu \) on the sample space \( (\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T)) \) is said to be a spherically invariant measure (SIM) if it is induced by a SIP, or, equivalently, if the coordinate process on \( (\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), \mu) \) is a SIP. A SIM induced by a SIP \( (m, R, F) \) will be denoted by SIM\((m,R;F)\). When \( F \) puts all its mass at the point 1 the SIP \( (m,R;F) \) and the SIM \( (m,R;F) \) become
Gaussian with mean $m$ and covariance $R$; we will use the notation $\text{GP}(m, R)$ and $\text{GM}(m, R)$ respectively.

**Lemma 1.1** Let $\mu$ be an SIM $(m, R; F)$ and for each $\alpha \geq 0$ let $\mu_\alpha$ be a $\text{GM}(m, \alpha R)$. Then

\begin{equation}
\mu(E) = \int \mu_\alpha(E) dF(\alpha) \quad \text{for all } E \in B(\mathbb{R}^T).
\end{equation}

Furthermore, for any measurable function $\theta$ on $(\mathbb{R}^T, B(\mathbb{R}^T), \mu)$ which is non-negative or integrable we have

\begin{equation}
E\theta = \int E_\alpha \theta dF(\alpha)
\end{equation}

where $E\theta = \int \theta d\mu$ and $E_\alpha \theta = \int \theta d\mu_\alpha$.

**Proof:** For each fixed cylinder set $E$, $\mu_\alpha(E)$ is a measurable function of $\alpha$;
and the family of sets $E$ such that $\mu_\alpha(E)$ is $\alpha$-measurable is a $\sigma$-field. Therefore, $\mu_\alpha(E)$ is $\alpha$-measurable for every $E \in B(\mathbb{R}^T)$. Thus $\int \mu_\alpha(E) dF(\alpha)$ is well-defined and is easily checked to be a probability measure. The characteristic functions of its finite dimensional distributions are given by (1.1), and since they uniquely determine a probability measure on $(\mathbb{R}^T, B(\mathbb{R}^T))$, we have (1.2).

(1.3) holds for $\theta = 1_E (E \in B(\mathbb{R}^T))$ by (1.2). Hence it holds for simple functions $\theta$, and thus for $\theta$ nonnegative by considering a sequence of simple functions increasing to $\theta$. For $\theta$ integrable, consider the positive and negative parts separately.

It follows from Lemma 1.1 that if $X$ is a coordinate SIP $(m, R; F)$ then, under each $\mu_\alpha$, $X$ is a $\text{GP}(m, \alpha R)$. The following theorem generalizes this fact.

**Theorem 1.2.** Let $X$ be a SIP $(m, R; F)$ on $(\Omega, B(X), P)$. Then for each $\alpha \geq 0$ there exists a probability measure $P_\alpha$ on $(\Omega, B(X))$ such that $\mu_\alpha = P_\alpha \cdot X^{-1}$ is a $\text{GM}(m, \alpha R)$. Hence under each $P_\alpha$, $X$ is a $\text{GP}(m, \alpha R)$ and the corresponding
formulae (1.2) and (1.3) hold.

Proof: Let \( \mu \) be the SIM induced on \((\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))\) by \( X \) and let \( \mu_\alpha, \alpha \geq 0 \), be \( \text{GM}(m, R) \) as in Lemma 1.1. Let \( \mathcal{C}(\mathbb{R}^T) \) be the field of all cylinder sets in \( \mathcal{B}(\mathbb{R}^T) \) and \( \mathcal{C}(X) = X^{-1} \cdot \mathcal{C}(\mathbb{R}^T) \subset \mathcal{B}(X) \). For each \( D \in \mathcal{C}(X) \), define \( P_\alpha(D) = \mu_\alpha(E) \) if \( D = X^{-1}(E) \). We claim that \( P_\alpha \) is well defined. It is not hard to see (by looking at their characteristic functions) that when restricted to a finite dimensional space \( \mu_\alpha \) is absolutely continuous with respect to \( \mu \). Thus if \( D = X^{-1}(E) = X^{-1}(E') \), \( E, E' \in \mathcal{C}(\mathbb{R}^T) \), then \( \mu(E) = P(D) = \mu(E_1) \) and hence \( \mu_\alpha(E) = \mu_\alpha(E') \). Consequently, \( P_\alpha \) is a probability measure on the field \( \mathcal{C}(X) \) and hence it has a unique extension to a probability measure on \( \mathcal{B}(X) \). Now it follows that \( \mu_\alpha = P_\alpha \circ X^{-1} \) since the two measures coincide on \( \mathcal{C}(\mathbb{R}^T) \). \( \square \)

(1.2) in Theorem 1.2 was proved by Gualtierotti (1974) for \( P \) an SIM on a separable Hilbert space. Note that since \( X \) is of second order, it follows from
\[
\omega > E X_t^2 = \int E X_t^2 dF(\alpha) = \int [m^2(t) + \alpha R(t, t)] dF(\alpha)
\]
that
\[
\alpha_1 = \int_\alpha dF(\alpha) < \infty .
\]

**Theorem 1.3.** Let \( X \) be a SIP \((0, R; F)\) and \( \{\xi_n\} \) a sequence in \( H(X) \). Then there exist versions of the \( \xi_n \)'s such that under each \( P_\alpha \) the r.v.'s \( \{\xi_n\} \) are jointly Gaussian with zero mean and covariance
\[
E_{\alpha_n} \xi_m = \frac{\alpha}{\alpha_1} E \xi_n \xi_m = \alpha E \xi_1 \xi_n \xi_m ,
\]
and, moreover, for any measurable function \( g \) on \((\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty)) \) we have
\[
g_{\alpha}(\xi_1, \xi_2, \ldots) = \xi_1 g(\sqrt{\alpha} \xi_1, \sqrt{\alpha} \xi_2, \ldots)
\]
whenever one of these expectations exists.
Proof: In view of Theorem 1.2 and the fact that every measurable function \( \theta(\omega) \) on \((\Omega, \mathcal{B}(\Omega))\) is of the form \( f(X(\omega)) \) for some measurable function \( f \) on \((\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))\), it suffices to prove the assertion for \( X \) a coordinate process. For each r.v. \( \xi_n(x) \), \( x \in \mathbb{R}^T \), there is a sequence \( \xi_i^{(n)}(x) \) of finite linear combinations of \( \{f_t(x)=x(t), t \in T\} \) such that \( \xi_n(x) = \lim_i \xi_i^{(n)}(x) \) a.s. \([\nu]\).

Let \( C_n = \{x \in \mathbb{R}^T: \xi_n^{(n)}(x) \to \cdot \text{ as } i \to \infty\} \) and let \( C = \cap C_n^c. \) \( C \) is clearly a measurable linear space having \( \mu \) probability 1. Now take each \( \xi_n \) to be \( \lim_i \xi_i^{(n)}(x) \) for \( x \in C \) and 0 for \( x \notin C \). Since \( 1 = \mu(C) = \int \mu_\alpha(C) d\nu(\alpha) \) by Lemma 1.1, and \( F(0^+) < 1 \), there is \( \alpha_0 > 0 \) such that \( \mu_{\alpha_0}(C) = 1 \). Let \( Y \) be a \( G(0, \mathbb{R}) \) on some probability space \((\Omega_0, \mathcal{B}_0, \mathbb{Q})\). Then since \( C \) is a linear space we have for all \( \alpha > 0 \)

\[
\mu_{\alpha}(C) = Q(\sqrt{\alpha} Y \in C) = Q(\sqrt{\alpha_0} Y \in C) = \mu_{\alpha_0}(C) = 1.
\]

Thus \( \xi_n = \lim_i \xi_i^{(n)} \) a.e. \([\mu_\alpha]\), for all \( \alpha > 0 \). Since under each \( \mu_\alpha \), \( \{\xi_i^{(n)}, n \geq 1, i \geq 1\} \) is a Gaussian family, it follows that \( \{\xi_n\} \) are jointly Gaussian under each \( \mu_\alpha \).

To show (1.4) we recall that for a Gaussian family a.s. convergence is equivalent to mean square convergence. Thus

\[
E \xi_i \xi_j = \lim_i E \xi_i^{(n)} \xi_j^{(m)} = \lim_{i,j} \alpha_i \xi_i^{(n)} \xi_j^{(m)} = \alpha \xi_i \xi_j,
\]

and hence \( E \xi_i \xi_j = \int E \xi_i \xi_j d\nu(\alpha) = \alpha \int \xi_i \xi_j d\nu(\alpha) = \frac{\alpha_1}{\alpha} E \xi_i \xi_j. \)

It is easy to see that \( \{\xi_n\} \) under \( \mu_\alpha \) has the same probability law as \( \{\sqrt{\alpha} \xi_n\} \) under \( \mu_1 \). (1.5) is now evident.

We now introduce a r.v. \( A \) which will play a central role in the representation of a SIP as well as in the study of its nonlinear space.
Pick an orthogonal sequence \( \{\xi_n\} \) from \( \mathcal{H}(X) \) with \( E\xi_n^2 = \alpha_1 \). Then by Theorem 1.3 we may assume that under each \( \mathbb{P}_\alpha \), \( \{\xi_n\} \) is a sequence of independent zero mean Gaussian r.v.'s with \( E\xi_n^2 = \alpha \). Define

\[
A_n = \frac{1}{n} \sum_{i=1}^{n} \xi_i^2.
\]

By the law of large numbers, we have \( \lim A_n = \alpha \) a.s. \([\mathbb{P}_\alpha]\). Let \( \mathcal{C}^* = \{\omega \in \Omega : A_n(\omega) \to \alpha\} \). Then

\[
P(\mathcal{C}^*) = \int \mathcal{P}_\alpha(\mathcal{C}^*)d\mathbb{P}(\alpha) = \int d\mathbb{P}(\alpha) = 1.
\]

Thus \( A_n \) converges a.s. \([\mathbb{P}]\). Let

\[
(1.6) \quad A(\omega) = \begin{cases} 
\lim A_n(\omega) & \text{if } \omega \in \mathcal{C}^* \\
0 & \text{if } \omega \notin \mathcal{C}^*.
\end{cases}
\]

Then

\[
(1.7) \quad A = \alpha \quad \text{a.s. } [\mathbb{P}_\alpha],
\]

and the distribution function of \( A \) is \( F \) since

\[
(1.8) \quad P(A \leq \alpha) = \int \mathcal{P}_\alpha(A \leq \alpha)d\mathbb{P}(\alpha) = \int_{[0,\alpha]}(\alpha)d\mathbb{P}(\alpha) = F(\alpha).
\]

If we put \( \mathcal{C}_\alpha = \{\omega \in \Omega : A(\omega) = \alpha\} \) then \( \mathcal{P}_\alpha(\mathcal{C}_\alpha) = 1 \), and thus the probability measures \( \mathcal{P}_\alpha, \alpha \geq 0 \), are mutually singular (which is of course well known).

We now arrive at the main theorem of this section.

**Theorem 1.4.** A nondegenerate second order process \( X \) on \((\Omega, \mathcal{B}(X), \mathbb{P})\) is a SIP \((0; R; F)\) if and only if it has the representation

\[
(1.9) \quad X_t = A^{1/2}Y_t \quad \text{a.s. for all } t \in T
\]

where \( A \) is a nonnegative r.v. with distribution function \( F \) and, conditioned on \( A > 0 \), \( Y \) is a nondegenerate GP(0, R) independent of \( A \).
Proof: The "if" part is clear. To show the "only if" part, let $A$ be the r.v. defined by (1.6) and let

$$
Y = \begin{cases} \frac{X}{A^{1/2}} & \text{if } A > 0 \\ 0 & \text{if } A = 0. \end{cases}
$$

(1.10)

Then it follows from Theorem 1.2 and (1.7) that $P(X_t = A^{1/2}Y_t) = \int p_\alpha(x) = 1$; and that $Y$ is a GP($0, R$) under each $p_\alpha$ for $\alpha > 0$ and thus also under $\bar{F}(\cdot) = (1-F(0^+))^{-1} \int_{(0,\infty)} p_\alpha(\cdot) dF(\cdot)$, which is just the conditional probability of $P$ given $A > 0$. Now for $D \in \mathcal{B}(\mathbb{R})$ and $E \in \mathcal{B}(\mathbb{R}^T)$ we have

$$
\bar{P}(A \in D, Y \in E) = \frac{1}{1-F(0^+)} \int_{(0,\infty)} 1_D(x)p_\alpha(Y \in E) dF(x)
$$

$$
= \frac{1}{1-F(0^+)} P(A \in D) \cdot \bar{P}(Y \in E)
$$

$$
= \bar{P}(A \in D) \bar{P}(Y \in E)
$$

since $p_\alpha(Y \in E) = \bar{P}(Y \in E)$ for $\alpha > 0$. Hence $A$ and $Y$ are conditionally independent given $A > 0$.

The representation (1.9) of a SIP was first noted by Besson (1974) who, without constructing the r.v. $A$, showed its existence by employing a result in Bretagnolle et. al (1966) concerning symmetrically dependent (exchangeable) r.v.'s. Our approach seems more elementary and direct, and also yields further results.

(1.9) reveals that a SIP is merely a conditional Gaussian process. (More specifically, given $A = \alpha$, the SIP $(0, R; F)$ is a GP $(0, \alpha R)$.) By requiring $F(0^+) = 0$, Theorem 1.4 may be stated in a more appealing way.
THEOREM 1.5. A nondegenerate second order process $X$ on $(\Omega,\mathcal{B}(X),\mathbb{P})$ is a SIP $(0,R;F)$ satisfying $F(0^+) = 0$ if and only if it has the representation

$$X_t = A_1^{1/2}Y_t \quad \text{a.s., } t \in T$$

where $Y$ is a nondegenerate GP $(0,R)$ and $A$ is a positive r.v. independent of $Y$ having distribution function $F$.

Theorem 1.4 enables one to read off many properties of SIP's immediately. For instance a SIP $(0,R;F)$ $X$ is of order $p > 0$, i.e., $E|X_t|^p < \infty$, if and only if $E A_1^{p/2} = \int_0^\infty A_\alpha^{p/2} dF(\alpha) < \infty$. If $X$ is of order 2, then continuity in probability of $X$ is equivalent to mean square continuity [Besson (1974)]. All the usual (local and global) analytic properties of the sample functions of $X$ depend only on $R$ and not on $F$, while properties of maxima and crossings depend on both $R$ and $F$. Kallianpur's zero-one law and Slepian's lemma take the following form for SIP's. (The proofs are straightforward and are thus omitted.)

COROLLARY 1.6. If $\mu$ is an SIM $(0,R;F)$ and $L$ a $\mathcal{B}(\mathbb{R}_T^\infty)$-measurable linear subspace of $\mathbb{R}_T^\infty$, then $\mu(L) = F(0^+)$ or 1. Furthermore, $\mu(L) = 1$ if and only if for some $\alpha > 0$, $L$ is $\overline{\mathcal{B}}_\alpha(\mathbb{R}_T^\infty)$-measurable and $\mu_\alpha(L) = 1$.

In Corollary 1.6 $\mu_\alpha$ are as in Lemma 1.1 and $\overline{\mathcal{B}}(\mathbb{R}_T^\infty)$, resp. $\overline{\mathcal{B}}_\alpha(\mathbb{R}_T^\infty)$, denotes the completion of $\mathcal{B}(\mathbb{R}_T^\infty)$ with respect to $\mu$, resp. $\mu_\alpha$.

COROLLARY 1.7. If $X_i$ is a separable SIP $(0,R_i;F_i)$, $i = 1,2$, and if

$$R_1(t,t) = R_2(t,t) \quad , \quad R_1(t,s) \leq R_2(t,s) \quad \text{for all } t,s \in T,$$

$$F_1(0^+) = F_2(0^+) \quad , \quad F_1(\alpha) \leq F_2(\alpha) \quad \text{for all } \alpha > 0,$$

then for all $\alpha$,
\[
P\{\sup_{t \in T} X_1(t) < u\} \leq P\{\sup_{t \in T} X_2(t) < u\}.
\]

We now give an interesting example of a sample continuous martingale whose family of \(\sigma\)-fields is not continuous. Let \(W = \{W_t, 0 \leq t < \infty\}\) be a Wiener process, and let \(X = A^{1/2}W\) where \(A\) is a nonnegative r.v. independent of \(W\) and whose \(\sigma\)-field is nontrivial. It is easily checked that \(X\) is a sample continuous martingale. Let \(\mathcal{B}_t = \mathcal{B}(X_s, 0 \leq s \leq t), 0 \leq t < \infty\). We will show that \(\mathcal{B}_t\) is not continuous at \(t = 0\). Fix \(t > 0\) and consider the quadratic variation \(M(t)\) of \(X\) over the interval \([0,t]\). We have

\[
M(t) = \lim_{n \to \infty} \sum_{j=1}^{2^n-1} \left[ X\left(\frac{j+1}{2^n} t\right) - X\left(\frac{j}{2^n} t\right) \right]^2
\]

\[
= A \lim_{n \to \infty} \sum_{j=1}^{2^n-1} \left[ W\left(\frac{j+1}{2^n} t\right) - W\left(\frac{j}{2^n} t\right) \right]^2
\]

\[= At \text{ a.s.} \]

where the last equality is a theorem of Lévy [Doob (1953)]. This implies that \(A\) is \(\mathcal{B}_t\)-measurable for all \(t > 0\).
In this section we study the structure of the nonlinear space \( L_2(X) \) of a SIP(\( O, R; F \)) \( X \), using the canonical representation of \( X \) in Theorem 1.4 and the well known properties of the nonlinear space of a GP. When \( X \) is second order, the relation between the linear space \( H(X) \) and the nonlinear space \( L_2(X) \) is shown in Theorem 2.4 and complete orthonormal sets (CONS's) in \( L_2(X) \) are given in Theorems 2.2 and 2.3. When \( F \) has a moment generating function, Theorem 2.5 shows that \( L_2(X) \) has the orthogonal decomposition
\[
\sum_{p \geq 0} H_p(X),
\]
where \( H_p(X) \) is the \( p \)-th homogeneous chaos of \( X \); and Theorem 2.6 shows the relation between \( H_p(X) \) and \( H(X) \) and gives CONS's in each \( H_p(X) \).

Theorems 2.2 to 2.4 are based on the following property.

\textbf{Lemma 2.1} Let \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \) be two independent \( \sigma \)-fields on a probability space \((\Omega, \mathcal{B}, P)\) such that \( \mathcal{B} \) is generated by \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \). Then
\[
L_2(\Omega, \mathcal{B}, P) \cong L_2(\Omega, \mathcal{B}_1, P) \otimes L_2(\Omega, \mathcal{B}_2, P)
\]
under the correspondence \( fg \leftrightarrow f \circ g \).

\textbf{Proof:} We will write \( L_2(\mathcal{B}_1) \) for \( L_2(\Omega, \mathcal{B}_1, P) \). Consider the mapping taking \( f \circ g \) to \( fg \) for all \( f \in L_2(\mathcal{B}_1), g \in L_2(\mathcal{B}_2) \). From the independence of \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \) it follows that \( fg \in L_2(\mathcal{B}) \) and
\[
\langle f_1 \circ g_1, f_2 \circ g_2 \rangle_{L_2(\mathcal{B}_1) \otimes L_2(\mathcal{B}_2)} = \langle f_1 g_1, f_2 g_2 \rangle_{L_2(\mathcal{B})}
\]
Thus the mapping preserves inner products and it has a unique extension to an isomorphism between the closed subspace spanned by \( \{fg\} \), which is in fact \( L_2(\mathcal{B}_1) \otimes L_2(\mathcal{B}_2) \), and the closed subspace of \( L_2(\mathcal{B}) \) spanned by \( \{fg\} \). The assertion is proved if we show that \( M = \{fg: f \in L_2(\mathcal{B}_1), g \in L_2(\mathcal{B}_2)\} \) is a complete set in \( L_2(\mathcal{B}) \).
Consider the family \( B^* \) of sets of the form \( \cup_{i=1}^n (E_i \cap F_i), E_i \in B_1 \) and \( F_i \in B_2 \). Since \( B_1 \) and \( B_2 \) are fields and they generate \( B \), it is readily seen that \( B^* \) is a field generating \( B \). We may even assume, after a moment's reflection, that the sets \( E_i \cap F_i, 1 \leq i \leq n \), are disjoint. Now a standard monotone class argument shows that the set \( \{ \cup_{i=1}^n (E_i \cap F_i) \} = \cup_{i=1}^n E_i \cap F_i \) is complete in \( L_2(B) \). Thus \( M \) is complete.

Let \( X = \{ X_t, t \in T \} \) be a second order nondegenerate SIP\((O,R;F)\) and \( X = A^Y \) its canonical representation of Theorem 1.4. Let \( \{ \xi_y, y \in \Gamma \} \) \((\Gamma \) linearly ordered) and \( \{ e_n, 1 \leq n \leq N \} \) \((N \) may be infinite) be CONS's in \( H(X) \) and \( L_2(dF) \) respectively.

**THEOREM 2.2** If \( F(O^+) = 0 \) then the family

\[
(2.1) \quad e_n(A) \cdot \left( \frac{1}{p_1 \cdots p_k} \right)^{\frac{1}{2}} \prod_{i=1}^k H \left( \frac{\alpha_i}{A} \right) \xi_{y_1} \cdots H \left( \frac{\alpha_i}{A} \right) \xi_{y_k}
\]

\[
= e_n(A) \cdot \left( \frac{1}{p_1 \cdots p_k} \right)^{\frac{1}{2}} \prod_{i=1}^k \frac{\alpha_i}{A} \xi_{y_1} \cdots \prod_{i=1}^k \frac{\alpha_i}{A} \xi_{y_k}
\]

where \( 1 \leq n \leq N, k \geq 1, p_1 \cdots p_k = p > 0, y_1 < \cdots < y_k \) in \( \Gamma \), is a CONS in \( L_2(X) \).

**Proof:** Assume \( F(O^+) = 0 \). Then \( B(A) \) and \( B(Y) \) are independent \( \sigma \)-fields generating \( B(X) \). By Lemma 2.1 we have

\[
(2.2) \quad L_2(X) = L_2(A) \otimes L_2(Y).
\]

It is easily verified that \( L^2(Y) = H(X) \) and \( \{ \eta_y = \left( \frac{1}{A} \right)^{\frac{1}{2}} \xi_{y}, y \in \Gamma \} \) is complete in \( H(Y) \). (Recall that \( Y \) is a GP.) Indeed \( \{ \eta_y, y \in \Gamma \} \) is a CONS in \( H(Y) \) since

\[
E \left( \frac{\alpha_i}{A} \xi_{y_1} \xi_{y_k} \right) = \int \frac{\alpha_i}{A} E \xi_{y_1} \xi_{y_k} dF(\alpha) = E \xi_{y_1} \xi_{y_k}
\]

(by Theorem 1.3). Now (2.2) and the celebrated theorem of Cameron and Martin yield the result. \( \Box \)
We remark that in dealing with a SIP, the case \( F(0^+) = 0 \)
(i.e. \( X \) cannot be zero with positive probability) is of main interest.

We now show how to obtain the counterpart of Theorem 2.2 when
\( 0 < F(0^+) < 1 \). Let \( \Omega_1 = \{ \omega \in \Omega : A(\omega) = 0 \} \), \( \Omega_2 = \{ \omega \in \Omega : A(\omega) > 0 \} \), and
consider the restrictions of \((\Omega, B(X), P)\) to \( \Omega_1 \) and \( \Omega_2 \) respectively,
\( (\Omega_1, B_{\Omega_1}(X), P_{\Omega_1}) \) and \( (\Omega_2, B_{\Omega_2}(X), P_{\Omega_2}) \). Then
\[
L_2(X) = L_2(\Omega_1, B_{\Omega_1}(X), P_{\Omega_1}) \oplus L_2(\Omega_2, B_{\Omega_2}(X), P_{\Omega_2}) .
\]
It is easily seen that
\[
L_2(\Omega_1, B_{\Omega_1}(X), P_{\Omega_1}) \cong \mathbb{R} ,
\]
where \( \mathbb{R} \) is equipped with the usual inner product \( \langle x, y \rangle_{\mathbb{R}} = xy \). Also \( P_{\Omega_2} = Q_{\Omega_2} \)
where \( Q \) is defined on \( B(X) \) by \( Q(\omega) = \int_{(0, \infty)} p(\omega) F(\alpha) dF(\alpha) \), and thus
\[
L_2(\Omega_2, B_{\Omega_2}(X), P_{\Omega_2}) \cong L_2(\Omega, B(X), \bar{P} \frac{Q}{1-F(0^+)})
\]
under the correspondence \((1-F(0^+))^{-\frac{1}{2}} 1_{\Omega_2} f \leftrightarrow f \). But under \( \bar{P} \), \( X \) is a
SIP(0; R; G) where \( G(\alpha) = (1-F(0^+))^{-\frac{1}{2}} (F(\alpha) - F(0^+)) \), and Theorem 2.2 yields
the following CONS.

**Theorem 2.3** If \( 0 < F(0^+) < 1 \) and if \( \{e_n, 1 \leq n \leq N\} \) is a CONS in \( L_2(d\{F-F(0^+)\}) \)
then the family
\[
\begin{align*}
&\left(\frac{1}{F(0^+)}\right)^{\frac{1}{2}} 1_{\{0\}}(A), \\
&\left(\frac{1}{1-F(0^+)}\right)^{\frac{1}{2}} 1_{(0, \infty)}(A) \ e_n(A) \left(\frac{1}{p_1 \cdots p_k}\right)^{\frac{1}{2}} \frac{A}{A_1} \ P \frac{A}{A_1} \ \cdots \ \frac{A}{A_k} \ A(\xi_1) \ \cdots \ A(\xi_k)
\end{align*}
\]
where \( 1 \leq n \leq N, k \geq 1, p_1 \cdots p_k = p \geq 0, \gamma_1 < \cdots < \gamma_k \) in \( \Gamma \), is a CONS in
\( L_2(X) \).
Consider the correspondence

\[ e_n(A) \left( \frac{1}{p_1^{\lambda_1} \ldots p_k^{\lambda_k}} \right)^{\frac{1}{2}} \left( \frac{A}{H_1^{\lambda_1}} \right) \rightarrow (\xi_{\gamma_1} \ldots \xi_{\gamma_k}) \]

(2.4)

when \( F(0^+) = 0 \), and when \( 0 < F(0^+) < 1 \) the correspondence

\[ 1 \leftrightarrow \left( \frac{1}{F(0^+)} \right)^{\frac{1}{2}} 1_{(0,\infty)}(A), \]

(2.5)

where \( k_1, p_1 + \ldots + p_k = p \geq 0 \), \( \gamma_1 < \ldots < \gamma_k \), form a CONS in \( H(X) \), we have

**Theorem 2.4**

\[ L_2(X) = L_2(dF) \otimes (\xi_{\gamma_1} \ldots \xi_{\gamma_k}) \]

Since \( \left( \frac{p_1^{\lambda_1} \ldots p_k^{\lambda_k}}{p_1^{\lambda_1} \ldots p_k^{\lambda_k}} \right)^{\frac{1}{2}} \left( \xi_{\gamma_1} \ldots \xi_{\gamma_k} \right) \),

Suppose that every \( \xi \in H(X) \) has all moments finite. Let \( P(X) \) be the linear space of all polynomials in elements of \( H(X) \) and let \( P_0(X) \) (\( p \geq 0 \)) be the linear space of all polynomials in \( P(X) \) of degree at most \( p \); hence \( P_0(X) \) is the set of all constants. Let \( Q_0(X) = P_0(X) \) and for \( p \geq 1 \) let \( Q_p(X) \) be the set of all polynomials in \( P_p(X) \) orthogonal to \( P_{p-1}(X) \). Denote
by $H_p(X)$ the closure of $Q_p(X)$ in $L_2(X)$. $Q_p(X)$ is called the $p$-th polynomial chaos and $H_p(X)$ is called the $p$-th homogenous chaos.

When $X$ is a SIP($O,R;F$), in order to have all moments of $\xi \in H(X)$ finite we introduce the following "moment" condition:

(M) The moment generating function of $F$ exists, i.e.

$$\int e^{at} dF(\alpha) < \infty \quad \text{for all } t \in \mathbb{R}.$$  

Under the condition (M) we have for $\xi \in H(X)$ (by (1.5)),  

$$E[|\xi|^p] = \int E_\alpha |\xi|^p dF(\alpha) = \int \alpha^{p/2} E_1 |\xi|^p dF(\alpha) = \alpha_p/2 E_1 |\xi| < \infty$$

where $\alpha_p = \int \alpha^p dF(\alpha)$, $p \geq 0$.

**Theorem 2.5** If (M) holds,  

$$L_2(X) = \bigoplus H_p(X).$$

**Proof:** It is well-known (see Neveu (1968)) that if $e^{|\xi|} \in L_2(X)$ for every $\xi \in H(X)$ then $L_2(X) = \bigoplus H_p(X)$ (for arbitrary process $X$). Thus it suffices to show that $e^{|\xi|}$ is integrable if (M) holds. Since under each $P_\alpha$, $\xi$ is a zero mean Gaussian variable with variance $\sigma^2 = \frac{\alpha}{\alpha_1} E_{\xi^2}$, we have

$$E_\alpha e^{\gamma |\xi|} = \frac{2}{\sqrt{2\pi} \sigma} \int_0^\infty e^{-x^2/(2\sigma^2)} \, dx$$

$$= \frac{2e^{\sigma^2/2}}{\sqrt{2\pi} \sigma} \int_{-\sigma^2}^{\sigma^2} e^{-x^2/2\sigma^2} \, dx \leq 2e^{\sigma^2/2}.$$
Thus, under (M),
\[ E \alpha \geq |\xi| = \int E \alpha e^{i\xi} \, dF(\alpha) \leq 2 \int e^{-2\alpha} \, dF(\alpha) < \infty. \]

We now establish the relationship between the decomposition of Theorem 2.5 and the representation of Theorem 2.4. Assume the condition (M). Then

\[ L_2(X) = \oplus H_r(X), \text{ and } \{A^n, 0 \leq n < \infty\} \text{ is a complete set of } L_2(A). \]

For each \( p \geq 0 \),

Fixed \( q \geq 0 \), applying the Gram-Schmidt onthonormalization procedure, to \( \{A^n\} \) with respect to the inner product \( \langle A^m, A^n \rangle_q = E A^{m+n+q} \) we obtain the set \( \{e_n^q(A), 0 \leq n < \infty\} \) which is complete in \( L_2(A) \) and orthonormal relative to \( \langle *, * \rangle_q \). Note that \( e_n^q(A) \) is a polynomial in \( A \) with degree \( n \).

Now assume \( F(0+) = 0 \). By Theorem 2.4 we have \( L_2(A) \oplus (\oplus \tilde{H}_p(X)) = L_2(X) \).

Denote this isomorphism by \( \phi \). Let

\[ H_{n,q}(X) = \phi\{e_n^q(A)^{\frac{A_1}{A}})^{q/2} \oplus \tilde{H}_q^p(X)\}. \]

It is not difficult to see from (2.4) that

\[ H_{n,q}(X) \subset \bar{P}_{2n+q}(X) \setminus \bar{P}_{2n+q-1}(X) \]

since \( A \in \bar{P}_2 \), from the definition of \( A \).

**Theorem 2.6** Under conditions (M) and \( F(0+) = 0 \),

\[ H_p(X) = \bigoplus H_{n,q}(X) \]

and a \textit{CONS in } \( H_p(X) \) is given by

\[ \left( \frac{A_q^1}{A_1} \right) \cdots \left( \frac{A_q^k}{A_k} \right) e_n^q(A)^{\frac{A_q^1}{A_1}} \cdots \frac{A_q^1}{A_1} (\xi_1^{q_1}) \cdots \frac{A_q^k}{A_k} (\xi_k^{q_k}) \]

where \( 2n+q = p, k \geq 1, q_1 + \cdots + q_k = q, y_1 < \cdots < y_k \).
Proof: First we show that \( H_{n,q}(X) \perp H_{n',q}(X) \) if \((n,q) \neq (n',q')\). It is clear from (2.6) that \( H_{n,q}(X) \perp H_{n',q}(X) \) if \( q \neq q' \). Suppose now that \( q = q' \) and \( n \neq n' \). For \( \theta, \theta' \in H^{p,q}(X) \) we have

\[
E\{ \phi(e_n^q(A)\left(\frac{A}{\alpha_1}\right)^{q/2} \otimes \theta) \phi(e_{n'}^q(A)\left(\frac{A}{\alpha_1}\right)^{q/2} \otimes \theta') \} = \frac{1}{q}E_{\theta, \theta'}\{ \alpha_{n}^q e_n^q \langle A^q \rangle \} H_{q}(X) = 0
\]

since \( \langle e_n^q, e_{n'}^q \rangle_q = 0 \) and thus \( H_{n,q}(X) \perp H_{n',q}(X) \).

In order to show (2.8) it suffices to show that \( \overline{\mathbb{P}}(X) = \bullet \ H_{n,q}(X) \)

because of (2.7). We need to show that \( \xi_1 \ldots \xi_k \in \bullet \ H_{n,q}(X) \) for all \( \xi_1, \ldots, \xi_k \in H(X) \) and \( r_1 + \ldots + r_k \leq p \). For ease of exposition we show this for \( \xi^r \) only. Write \( \xi = \sum \eta, \eta \in H(Y) \). Then

\[
\xi^r = \alpha_1^r \eta = \alpha_1^r \eta = \eta \left( \frac{H_{r,2n+q}}{r,2n+q,4} + \text{Const.} \right) + \text{Const.} \left( \frac{H_{r,2n+q}}{r,2n+q,4} \right)
\]

Note that \( \alpha_1^m \eta \in H_{n/2m,2n+q}(X) \) and (2.8) is proved.

(2.9) follows from (2.4), (2.6), (2.8) and \( ||e_n^q(A)\left(\frac{A}{\alpha_1}\right)^{q/2}||^2 = \alpha_1^q \).

Of course \( H_1(X) = H(X) \) and, for instance \( H_2(X) = H_{1,0}(X) \bullet H_{0,2}(X) \) where

\[
H_{1,0}(X) = \overline{\text{sp}}(A_{-\alpha_1}) \text{ and } H_{0,2}(X) = \overline{\text{sp}}(\left(\frac{\alpha_1}{\alpha_2}\right)^{p_1^2}H(A_{-\alpha_1} \xi_{1,2} H(A_{-\alpha_1} \xi_{2,2}) : p_1 + p_2 = 2, y_1 < y_2) \).
3. EQUIVALENCE AND SINGULARITY OF SIP's

In this section we combine the representation of SIP's (Theorem 1.4) and the dichotomy of GP's to the problem of discriminating between two SIP's. The discrimination problem is completely solved by identifying the Lebesgue decomposition of the distributions of the two processes and the Radon-Nikodym derivative of the absolutely continuous part. Two processes are called equivalent (~), resp. singular (1), if their induced measures on $\mathcal{B}(\mathbb{R}^T)$ are equivalent (~), resp. singular (1). The discrimination problem for two second order SIP's: SIP(0,m;F) and SIP(m,S;G) is fully resolved by first noting that without loss of generality we may assume that either GP(0,R) ~ GP(m,S) or else GP(0,\alpha R) \perp GP(m,\beta S) for all \alpha, \beta > 0. When GP(0,R) ~ GP(m,S), the Lebesgue decomposition of SIM(m,S;G) with respect to SIM(0,R;F) is given in Theorem 3.1 and the Radon-Nikodym derivative of the absolutely continuous part in Theorem 3.4. In particular, if GP(0,R) ~ GP(m,S), then dF ~ dG implies SIP(0,R;F) ~ SIP(m,S;G), and dF \perp dG implies SIP(0,R;F) \perp SIP(m,S;G). Theorem 3.2 shows that if GP(0,\alpha R) \perp GP(m,\beta S) for all \alpha, \beta > 0 then SIP(0,R;F) \perp SIP(m,S;G).

Finally, if dF ~ dG then SIP(0,R;F) and SIP(m,R;G) are either equivalent or singular, and necessary and sufficient conditions for equivalence along with an expression for the Radon-Nikodym derivative are given in Theorem 3.5.

For reasons of clarity in this section we attach the underlying probability measure to the usual notation for expectation, variance, linear space, etc.

We first state the general theorem concerning the equivalence and the singularity of two GP's (see for instance Pang (1973)). Fix \alpha > 0. Let $X = (X_t, t \in T)$ be a GP(0,\alpha R) on the probability space $(\Omega, \mathcal{B}(X), \mathbb{P}_\alpha)$. Let $Q_\alpha$ be a second probability on $(\Omega, \mathcal{B}(X))$ under which $X$ is a GP(m,\alpha S). Then either
\( P_\alpha \sim Q_\alpha \) or \( P_\alpha \perp Q_\alpha \). \( P_\alpha \sim Q_\alpha \) if and only if the following conditions are satisfied:

\[ (*) \] There exist positive constants \( K_1 \) and \( K_2 \) such that

\[ K_2 \text{Var}_p \xi \leq \text{Var}_Q \xi \leq K_1 \text{Var}_p \xi \]

for every \( \xi \in L(X) \), the linear space of all finite linear combinations \( \sum_{n=1}^{N} t_n \xi t_n \in T \).

\[ (**) \] There exists \( \xi \in H_p(\alpha) \) such that \( m(t) = E_p(\xi \chi_t) \), \( t \in T \); i.e. \( m \in \mathcal{H}(R) \),
the reproducing kernel Hilbert space of the covariance \( R \).

\[ (***) \] If \( B : H_p(\alpha) \rightarrow H_p(\alpha) \) is the positive self-adjoint operator defined
by \( \text{Cov}_q(\xi, \eta) = \text{Cov}_p(\xi, \eta) \) for all \( \xi, \eta \in L(X) \), then \( B - I \) is Hilbert-Schmidt.
Moreover if conditions \( (*) \), \( (**) \) and \( (***) \) hold true and if \( \{ \lambda_n \} \) and \( \{ \xi_n \} \)
denote the sets of eigenvalues and corresponding normalized eigenvectors of \( B_\alpha \), then

\[ (3.1) \quad \frac{dQ_\alpha}{dp} = \exp \left\{ B^{-1/2}_\alpha \xi - \frac{1}{2} \text{Var}_p \left( B^{-1/2}_\alpha \xi \right) - \frac{1}{2} \! \sum \! \left( \xi_n^2 - \lambda_n^{-1} \right) + \log \lambda_n \right\} \]

From now on \( X = (X_t, t \in T) \) will be a second order nondegenerated SIP
\((0, R; F) \) on \( (\Omega, \mathcal{B}(X), P) \). Recall that \( R \) is a covariance function and \( F \) is a
distribution function on \( \mathbb{R}^+ \) with finite first moment \( \alpha_1 \). \( P \) is a mixture
of Gaussian measures, \( P(E) = \int P_\alpha(E) dF(\alpha) \), and under each \( P_\alpha \), \( X \) is a
GP\((0, \alpha R) \). Let \( A \) be the r.v. associated with \( X \) and \( P \) as in Section 1.

Now consider a second probability measure \( Q \) on \( (\Omega, \mathcal{B}(X)) \) under which \( X \)
is a second order nondegenerate SIP \((m, S; G) \). Then \( Q(E) = \int Q_\alpha(E) dG(\alpha) \) and
under each \( Q \), \( X \) is a GP\((m, \alpha S) \). Denote the first moment of \( G \) by \( \alpha_1' \). We
are interested in the equivalence and mutual singularity of the measures \( P \) and \( Q \). Since \( P_0 \perp Q_0 \) if \( m \neq 0 \), we shall assume throughout this section that
\( F(0^+) = 0 \). Also we may assume without loss of generality that either \( P_1 \sim Q_1 \)
or \( P_\alpha \perp Q_\beta \) for all \( \alpha, \beta > 0 \), since clearly

\( \Box.2) \) a SIP \((m, R; F(\alpha))\), is also a SIP \((m, cR; F(c\alpha))\) for every \( c > 0 \).

**THEOREM 3.1** Let \( P_1 \sim Q_1 \) and let \( dG = dG' + dG'' \) be the Lebesgue decomposition of \( dG \) with respect to \( dF \) with \( dG' \ll dF \) and \( dG'' \perp dF \). Then

\[
Q(E) = \int Q_\alpha(E) dG'(\alpha) + \int Q_\alpha(E) dG''(\alpha)
\]
is the Lebesgue decomposition of \( Q \) with respect to \( P \). Hence if \( P_1 \sim Q_1 \) and \( dF \sim dG \) then \( P \sim Q \); if \( P_1 \sim Q_1 \) and \( dF \perp dG \) then \( P \perp Q \).

**Proof:** Note that \( P_1 \sim Q_1 \) implies \( P_\alpha \sim Q_\alpha \) for all \( \alpha > 0 \). Let \( P(E) = 0 \). Then

\[
\int P_\alpha(E) dF(\alpha) = 0 \quad \text{and} \quad P_\alpha(E) = 0 \quad \text{a.e.} \quad [dF].
\]
Since \( P_\alpha \sim Q_\alpha \) and \( dG' \ll dF \), we have \( Q_\alpha(E) = 0 \) a.e. \([dG']\), and thus \( \int Q_\alpha(E) dG'(\alpha) = 0 \). This implies \( \int Q_\alpha(\cdot) dF(\alpha) \ll P(\cdot) \).

Since \( dG'' \perp dF \), there exists \( E \in B(\mathbb{R}) \) such that \( \int_E dG''(\alpha) = \int_E dF(\alpha) = 0 \).

Note that \( A = \alpha \) a.e. \([Q_\alpha]\) because \( A = \alpha \) a.e. \([P_\alpha]\) and \( P_\alpha \sim Q_\alpha \). Thus we have

\[
P(A \in E') = \int P_\alpha(A \in E') dF(\alpha) = \int_E dF(\alpha) = 0,
\]

\[
\int Q_\alpha(A \in E) dG''(\alpha) = \int_E dG''(\alpha) = 0,
\]
which imply \( \int Q_\alpha(\cdot) dG''(\alpha) \perp P(\cdot) \).  

The second assertion of Theorem 3.1 was first stated in Gualtierotti (1974) for \( P \) and \( Q \) SIM's on a separable Hilbert space.

**THEOREM 3.2** If \( P_\alpha \perp Q_\beta \) for all \( \alpha, \beta > 0 \) then \( P \perp Q \).

**Proof:** The proof is adapted from Pang (1973).

The following remarks will be used without further comment. For \( \xi, \eta \in L(X) \), we have
\[ E_{p_{\alpha}} \xi = E_{p_{\alpha}} \xi = 0, \quad \text{Cov}_{p_{\alpha}} (\xi, \eta) = \frac{\alpha}{\alpha_1} \text{Cov}_{p} (\xi, \eta), \]

(3.3)

\[ E_{q_{\alpha}} \xi = E_{q_\alpha} \xi, \quad \text{Cov}_{q_{\alpha}} (\xi, \eta) = \frac{\alpha}{\alpha_1} \text{Cov}_{q} (\xi, \eta). \]

If \( P_1 \) and \( Q_1 \) satisfy (*) then for every sequence \( \{ \xi_n \} \) in \( H_{p_{\alpha}} (X) \) there exist versions of \( \xi_n \)'s such that \( \{ \xi_n \} \) is also a sequence in \( H_{p_{\alpha}} (X), H_{p_{\alpha}} (X), H_Q (X), \) \( H_{Q_{\alpha}} (X) \) for all \( \alpha > 0 \), and satisfies (3.3). (The proof of this is similar to that of Theorem 1.3).

Since \( P_1 \perp Q_1 \), one of the conditions (*) , (**) and (***) must be violated. First suppose that (*) is not satisfied; for instance, suppose that there exists no constant \( K_1 \) such that \( \text{Var}_{Q_1} \xi \leq K_1 \text{Var}_{P_{\alpha}} \xi \). Then there exists a sequence \( \{ \xi_n \} \) in \( L(X) \) such that \( \text{Var}_{P_1} \xi_n = 1 \) and \( \text{Var}_{Q_1} \xi > n^2 \). Then, as \( n \to \infty \), we have

\[ P(\{|\xi_n| > \sqrt{n}\} \leq \frac{\alpha_1}{n} \to 0, \]

\[ Q(\{|\xi_n| > \sqrt{n}\} = \int_{Q_{\alpha}} \{|\xi_n| > \sqrt{n}\} dG(\alpha) \]

\[ \geq \int (1 - \int_{-(\alpha n)}^{-(\alpha n)} - \frac{x^2}{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx) dG(\alpha) \]

\[ \to 1 \]

which imply \( P \perp Q \).

Next suppose (*) holds but (**) does not hold. Then, for each \( n \), there exists \( \xi_n \in L(X) \) such that

\[ E_{Q_1} \xi_n > (n \text{Var}_{Q_1} \xi_{1/2}) \geq (k_2n \text{Var}_{P_1} \xi_{1/2}) \]

(see Pang (1973)). Consequently, as \( n \to \infty \)

\[ P(\xi_n > \frac{1}{2} E_{Q_1} \xi_n) \leq P(\xi_n > \frac{1}{2} (k_2n \text{Var}_{P_1} \xi_{1/2}) \leq \frac{4\alpha_1}{k_2n} \to 0, \]

\[ Q(\xi_n > \frac{1}{2} E_{Q_1} \xi_n) \geq Q(|\xi_n - E_{Q_1} \xi_n| < \frac{1}{2} E_{Q_1} \xi_n) \]

\[ \geq Q(|E_{Q_1} \xi_n| < \frac{1}{2} (n \text{Var}_{Q_1} \xi_{1/2}) \]

\[ \geq 1 - \frac{4\alpha_1}{n} \to 1, \]
and thus $P \perp Q$.

Finally, suppose that (*) and (**) are satisfied, but (***) is not. We may assume in this case that $m = 0$. We claim that there exists a sequence $\{\xi_n\} \in H_p(X)$ such that $\text{Cov}_p(\xi_i, \xi_j)$

\begin{equation}
\text{Cov} \quad \delta_{ij} \quad \text{and} \quad \text{Cov}_{Q_1}(\xi_i, \xi_j) = \mu_1 \delta_{ij}
\end{equation}


\begin{equation}
\sum (1-\mu_n)^2 = \infty \quad \text{and} \quad \frac{1}{\mu_n} - \mu_n > 0 \quad \text{for all } n \quad \text{or} \quad < 0 \quad \text{for all } n.
\end{equation}

Given this, consider the events

$$E_m = \left\{ \sum_{j=1}^m \frac{2}{1} \frac{1}{\mu_n^2} - 1 \leq \frac{1}{2} \sum_{j=1}^m \frac{1}{\mu_n} - \mu_n \right\}.$$ 

We shall deal with the case $\frac{1}{\mu_n} - \mu_n > 0$ only. Note that $K_2 \leq \mu_n \leq K_1$ by (*).

For $\alpha \geq 1$ we have

$$P_{\alpha}(E_m) = P_1\left\{ \sum_{j=1}^m \frac{1}{\mu_n^2} - 1 \leq \frac{1}{2} \sum_{j=1}^m \frac{1}{\mu_n} - \mu_n \right\} \leq P_1\left\{ \sum_{j=1}^m (\xi_n - 1) (\frac{1}{\mu_n} - 1) \geq \frac{1}{2} \sum_{j=1}^m (1-\mu_n)^2 \right\} \leq \frac{K_1}{k_2^2} \frac{1}{k_1^2 (1-\mu_n)^2} \to 0$$

and for $\beta \leq 1$

$$Q_{\beta}(E_m) = Q_1\left\{ \sum_{j=1}^m \frac{1}{\mu_n^2} - 1 \leq \frac{1}{2} \sum_{j=1}^m \frac{1}{\mu_n} - \mu_n \right\} \geq Q_1\left\{ \sum_{j=1}^m (\xi_n^2 - \mu_n) (\frac{1}{\mu_n} - 1) \leq \frac{1}{2} \sum_{j=1}^m (1-\mu_n)^2 \right\} \geq 1 - \frac{8K_1^2}{k_1^2 (1-\mu_n)^2} \to 1.$$ 

Thus if $E = \text{limsup} \ E_m$, then $P_{\alpha}(E) = 0$ for $\alpha \geq 1$ and $Q_{\beta}(E) = 1$ for $\beta \leq 1$. This implies by (3.2) that for all $a, b > 0$ there exist $E_{ab} \in \mathcal{B}(X)$ such that
\( P_\alpha(E_{ab}) = 0 \) for \( \alpha \geq a \) and \( Q_\beta(E_{ab}) = 1 \) for \( \beta \leq b \). Hence, as \( a \to 0 \) and \( b \to \infty \), we have

\[
P(E_{ab}) = \int_{[0,a]} P_\alpha(E_{ab}) dF(\alpha) \leq F(a) \to F(0^+) = 0,
\]
\[
Q(E_{ab}) \geq \int_{[0,b]} Q_\beta(E_{ab}) dG(\beta) = G(b) \to 1,
\]
and consequently \( P \perp Q \).

To complete the proof, we now verify (3.4). Two cases are to be considered. First, suppose that \( B_1 - I \) is compact but not Hilbert-Schmidt. Let \( \{\lambda_n\} \) and \( \{\eta_n\} \) be the eigenvalues and the corresponding normalized eigenvectors of \( B_1 - I \). We have \( \lambda_n \to 0 \), \( \sum \lambda_n^2 = \infty \) and \( B_1 \eta_n = (1+\lambda_n)\eta_n \).

By choosing a suitable subsequence of \( \{\eta_n\} \), we obtain the desired sequence \( \{\xi_n\} \) in (3.4). Second, suppose that \( B_1 - I \) is not compact. Being invertible (by \((*)\)), \( B_1 \) is not compact and thus its essential spectrum is not \( \{0\} \). Furthermore, since \( B_1 - I \) is not compact, there is at least one point \( \mu \neq 1 \) in the essential spectrum of \( B_1 \), and thus also in the essential numerical range of \( B \). (For a nice discussion of essential spectrum and essential numerical range see Fillmore et. al. (1972)). Now by a known result in operator theory (Lemma 2 in Anderson and Stampfli (1971) stated for a separable Hilbert space but true for a nonseparable Hilbert space as well), there exists an orthonormal sequence \( \{\eta_n\} \) in \( H_1(X) \) such that \( \text{Cov}(\eta_i, B_1 \eta_j) = \delta_{ij} \delta_{ij} \) and \( \eta_n \to \mu \). Again \( \{\xi_n\} \) in (3.4) is obtained by choosing a suitable subsequence of \( \{\eta_n\} \).

Theorem 3.2 is not true in general, without the Gaussian assumption on \( P_\alpha \) and \( Q_\beta \), since there are uncountably many measures involved. For example, if \( P_\alpha \) is the uniform measure on \([0,1]\) for each \( \alpha \), if \( Q_\beta \) is the one point mass at \( \beta \) for each \( \beta \), and if \( dF \sim dG \sim P_1 \), then \( P_\alpha \perp Q_\beta \) for all \( \alpha \) and \( \beta \), but \( P \sim Q \).
Now suppose that \( P_1 \sim Q_1 \) and let \( Q'(*) = \int Q_{\alpha}(*) dG'(\alpha) \) be the absolutely continuous part of \( Q \) given in Theorem 3.1. We will calculate the Radon-Nikodym derivative \( \frac{dQ'}{dP} \). Theorems 3.1 and 3.2, together with the expression of \( \frac{dQ'}{dP} \), provide a complete solution to the problem of discriminating between two SIP's. To this end, we prepare the following

**Lemma 3.3** Let \( \rho_\alpha = \frac{dQ_{\alpha}}{dP_\alpha} \) a.s. [\( P_\alpha \)]. If \( \rho_{A(\omega)}(\omega) \) is a measurable function, then

\[
\frac{dQ'}{dP}(\omega) = \rho_{A(\omega)} \frac{dG'}{dF}(A(\omega)) \text{ a.s. [P].}
\]

**Proof:** For every \( E \in B(X) \), we have

\[
\int E \rho_A \frac{dG'}{dF}(A)dP = \int \int E \rho_\alpha \frac{dG'}{dF}(\alpha)dP_\alpha dF(\alpha)
= \int Q_{\alpha}(E) dG'(\alpha) = Q'(E)
\]
as required. \( \square \)

Note that the measurability of \( \rho_{A(\omega)}(\omega) \) is not automatic since each \( \rho_\alpha \) can be arbitrarily changed on a set of \( P_\alpha \)-measure zero.

**Theorem 3.4** Suppose \( P_1 \sim Q_1 \). Then

(i) there exists \( \xi \in H_p(X) \) such that \( m(t) = \frac{1}{\alpha_1} C_p(\xi X_t), \) \( t \in T; \)

(ii) there exists a self-adjoint positive operator \( B \) on \( H_p(X) \) defined by

\[
\frac{1}{\alpha_1} \text{Cov} Q(\xi, \eta) = \frac{1}{\alpha_1} \text{Cov}_p(\xi, B\eta)
\]

for all \( \xi, \eta \in L(X) \), and such that \( B - I \) is Hilbert-Schmidt.

Moreover if \( \{\lambda_n\} \) and \( \{\xi_n\} \) is the set of eigenvalues of \( B \) and their corresponding eigenvectors with norms \( \alpha_1^{1/2} \), then

\[
\frac{dQ'}{dP} = \exp \left( \frac{1}{\alpha_1} B^{-1/2} \xi - \frac{1}{2\alpha_1} \text{Var}_p(B^{-1/2} \xi) \right)
- \frac{1}{2} \sum_{n=1}^{\infty} \left( \frac{\xi_n^2}{\alpha_1} \right) \cdot \frac{dG'}{dF}(A) .
\]
Proof: Let \( X = A^{1/2}Y \) be the canonical representation of \( X \), where \( Y \) is a \( \text{GP}(0,R) \) under \( P \) and under each \( P_\alpha \) (\( \alpha > 0 \)). It is clear that each \( \xi \in H_p(X) \) or \( H_p(X) \) has a version of the form \( A^{1/2} \eta \) with \( \eta \in H_p(Y) \). In this sense, we have \( H_p(X) = A^{1/2}H(Y) = H_p(X) \). Therefore every operator on \( H_p(X) \) induces an operator on \( H_p(X) \) in an obvious way, and vice versa.

Since \( P_1 \sim Q_1 \), there exists a \( \xi \in H_p(X) \) such that \( m(t) = E_{P_1}(\xi X_t), t \in T \), and a self-adjoint positive operator \( B_1 \) on \( H_p(X) \) defined by \( \text{Cov}_{Q_1}(\xi, \eta) = \text{Cov}_{P_1}(\xi, B_1 \eta), \eta \in H_p(X) \). A simple computation shows that \( \xi \) satisfies (i) and that \( B \), the operator on \( H_p(X) \) induced by \( B_1 \), satisfies (ii).

By Lemma 3.3, (3.5) will follow if we show

\[
\frac{dQ}{dP_\alpha} = \exp \left\{ \frac{1}{\alpha} B_\alpha^{-1/2} \xi - \frac{1}{2\alpha} \text{Var}_p \left( B_\alpha^{-1/2} \xi \right) - \frac{1}{2} \sum_{n=1}^\infty \left[ \frac{\xi_n}{\alpha} \left( \frac{1}{n} - 1 \right) \right] \lambda_n \right\} \text{a.s. } [P_\alpha].
\]

Let \( B_\alpha \) be the operator on \( H_p(X) \) induced by \( B \). Again, a simple computation shows that \( \alpha^{-1} \xi \) satisfies condition (**), and that \( B_\alpha \) satisfies condition (***) and has eigenvalues \( \lambda_n \) and corresponding eigenvectors \( \{\alpha^{-1/2} \xi_n\} \). Thus we have by (3.1)

\[
\frac{dQ}{dP_\alpha} = \exp \left\{ \frac{1}{\alpha} B_\alpha^{-1/2} \xi - \frac{1}{2\alpha} \text{Var}_p \left( B_\alpha^{-1/2} \xi \right) - \frac{1}{2} \sum_{n=1}^\infty \left[ \frac{\xi_n}{\alpha} \left( \frac{1}{n} - 1 \right) \right] \lambda_n \right\} \text{a.s. } [P_\alpha]
\]

which is equivalent to (3.6) since \( B_\alpha^{-1/2} \xi = B^{-1/2} \xi \) a.s. \([P_\alpha]\). \( \square \)
**Theorem 3.5** Suppose $R = S$ and $dF \sim dG$. Then either $P \sim Q$ or $P \perp Q$, and $P \sim Q$ if and only if

$$m(t) = E_p(\xi X_t), t \in T,$$

for some $\xi \in H_p(X)$. In the case of equivalence,

$$\frac{dQ}{dP} = \exp \left( \frac{\alpha_i}{A} \left( \xi - \frac{1}{2} \text{Var}_p \xi \right) \right) \frac{dG}{dF} \quad (A)$$

**Proof:** Clearly if (3.7) holds we have $P_1 \sim Q_1$, and if (3.7) does not hold we have $P_\alpha \perp Q_\beta$ for all $\alpha, \beta$. The first assertion follows then from Theorems 3.1 and 3.2. (3.8) follows from Theorem 3.4 by comparing (3.7) with (3.6) and noting that $R = S$ implies $B = I$. 

Sytaya (1969) derived (3.8) for $P$ and $Q$ SIM's on a separable Hilbert space and $F = G$. 
4. NONLINEAR ESTIMATION AND PREDICTION

Using the tensor product structure of the nonlinear space we solve the general nonlinear estimation problem for SIP's, and in particular for GP's, in the sense that we reduce the nonlinear problem to a standard linear estimation problem, the theory of which is well developed. Also we derive a lower bound for the mean square error of the nonlinear prediction for a certain class of prediction problems.

4.1 Nonlinear Estimation

Let \( X = (X_t, t \in T) \) be a second order process with zero mean. Consider the following estimation problem: we observe \( X_t \) for \( t \in S \), a subset of \( T \), and we want to estimate an \( L_2 \)-functional \( \theta \) of \( X \) based on the observations. We are interested in finding the best estimate \( \hat{\theta} \), an \( L_2 \)-functional of \( (X_t, t \in S) \) which minimizes the mean square error of estimation \( E(\hat{\theta} - \theta)^2 \), and it is well known that

\[
\hat{\theta} = E(\theta | X_t, t \in S).
\]

In general, \( \hat{\theta} \) is extremely difficult to determine. However, if \( X \) is a SIP, we have a complete solution. In formulating the main result we will use the notation of Section 2, and identify \( L_2(X) \) with \( L_2(A) \otimes (\oplus_{p \geq 0} H_{p}(X)) \) by Theorem 2.4. We let \( L_2(X;S) = L_2(X_t, t \in S) \) and \( H(X;S) = H(X_t, t \in S) \).

**Theorem 4.1** Let \( X \) be a nondegenerate SIP \((0,R;F)\) with \( F(0+) = 0 \) and let \( \theta \in L_2(X) \) have the following orthogonal development

\[
\theta = \sum_{n \geq 1} \sum_{p_1 + \ldots + p_k = p} a_{p_1 \ldots p_k} (A) \otimes (\xi_1 \otimes \ldots \otimes \xi_k).
\]
Suppose that \((X_t, t \in S)\) is nondegenerate. Then
\[
\hat{\theta} = \sum_{n \geq 1} \sum_{p_1 + \ldots + p_k = p} \sum_{\gamma_1 \ldots \gamma_k} \tilde{e}_{\gamma_1} \ldots \tilde{e}_{\gamma_k} \in (A) \otimes \tilde{e}_{\gamma_1} \ldots \tilde{e}_{\gamma_k}.
\]
where for \(\xi \in H(X)\),
\[
\hat{\xi} = \text{Proj}_{H(X;S)} \xi = \mathbb{E}(\xi | X_t, t \in S).
\]

**Proof:** Recall the definition of the r.v. \(A\) and observe that it is independent of the choice of the defining sequence \(\{\xi_i\}\). It then follows that \(A \in L_2(X;S)\) and every \(\rho \in L_2(X;S)\) has the orthogonal development
\[
\rho = \sum_{q \geq 0} \sum_{q_1 + \ldots + q_j = q} \sum_{b, \delta, \ldots, \delta_j} \tilde{e}_{\delta_1} \ldots \tilde{e}_{\delta_j} \in (A) \otimes (\eta_{\delta_1} \ldots \eta_{\delta_j})
\]
where \(\{\eta_\delta, \delta \in \Delta\} (\Delta \text{ linearly ordered})\) is a CONS in \(H(X;S)\).

We have
\[
\hat{\theta} = \mathbb{E}(\theta | X_t, t \in S) = \text{Proj}_{L_2(X;S)} \theta
\]
\[
= \sum_{p \geq 0} \sum_{p_1 + \ldots + p_k = p} \sum_{\gamma_1 \ldots \gamma_k} \tilde{e}_{\gamma_1} \ldots \tilde{e}_{\gamma_k} \in (A) \otimes (\xi_{\gamma_1} \ldots \xi_{\gamma_k}).
\]

Thus to show the theorem it suffices to show that
\[
\text{Proj}_{L_2(X;S)} \in (\xi_{\gamma_1} \ldots \xi_{\gamma_k}) = \mathbb{E}(\xi_{\gamma_1} \ldots \xi_{\gamma_k}) \in (A) \otimes (\xi_{\gamma_1} \ldots \xi_{\gamma_k}).
\]

If \(\xi_1, \ldots, \xi_p \in H(X)\) and \(\eta_1, \ldots, \eta_p \in H(X;S)\) then it follows from \(<\xi_i, \eta_j>_H(X) = <\hat{\xi}_i, \hat{\eta}_j>_H(X)\) that
\[
<\xi_1 \ldots \xi_p, \eta_1 \ldots \eta_p>_H(X) = <\hat{\xi}_1 \ldots \hat{\eta}_p>_H(X)
\]
and hence for each \(\rho \in L_2(X;S)\) we have
<e_n(A) \odot (\xi_{\gamma_1} \tilde{\circ} \ldots \tilde{\circ} \xi_{\gamma_k}), p>_{L_2(X)}
= \sum_{m} b^m \delta_{1 \ldots j}^m \delta_{1 \ldots j} <e_n(A) \odot (\xi_{\gamma_1} \tilde{\circ} \ldots \tilde{\circ} \xi_{\gamma_k}), p>_{L_2(X)}
= \sum_{m} b^m \delta_{1 \ldots j} <e_n(A) \odot (\xi_{\gamma_1} \tilde{\circ} \ldots \tilde{\circ} \xi_{\gamma_k}), p>_{L_2(X)}
= <e_n(A) \odot (\xi_{\gamma_1} \tilde{\circ} \ldots \tilde{\circ} \xi_{\gamma_k}), p>_{L_2(X)}.

Since e_n(A) \odot (\xi_{\gamma_1} \tilde{\circ} \ldots \tilde{\circ} \xi_{\gamma_k}) \in L_2(X;S), the result follows.

A similar result can be obtained when F(0+) > 0. Of particular interest is the case where X is in fact a Gaussian process. In this case we have A = 1 = a_1,
L_2(X) \tilde{=} \bigoplus_{p \geq 0} H_\mathbb{P}(X), \widetilde{Q}_p \sim H_\mathbb{P}(X); and Theorem 4.1 yields the following.

**COROLLARY 4.2** Let X be a GP(0,R) and let \( \theta \in L_2(X) \) have the following orthogonal development

\[
\theta = \sum_{p \geq 0} \sum_{p_1 + \ldots + p_k = p} a_{p_{1 \ldots p_k}} \xi_{\gamma_1} \tilde{\circ} \ldots \tilde{\circ} \xi_{\gamma_k}.
\]

Then

\[
\hat{\theta} = \sum_{p \geq 0} \sum_{p_1 + \ldots + p_k = p} a_{p_{1 \ldots p_k}} \xi_{\gamma_1} \tilde{\circ} \ldots \tilde{\circ} \xi_{\gamma_k}
\]

where for \( \xi \in H(X) \),

\[
\hat{\xi} = \text{Proj}_{H(X;S)} \xi \quad (=E(\xi|X_t, t\in S)).
\]

Consequently,

\[
\text{Proj}_{\widetilde{Q}_p} = \text{Proj}_{L_2(X;S)} = \text{Proj}_{L_2(X;S)} \cdot \text{Proj}_{\widetilde{Q}_p} = \text{Proj}_{\widetilde{Q}_p}(S)
\]
where \( p(S) \) denotes the \( p \)-th homogeneous chaos of \((X_t, t \in S)\).

**COROLLARY 4.3** If \( X \) is a nondegenerate SIP \((0, R; F)\) and if \( \int e^{t \alpha} dF(\alpha) < \infty \)
for all \( t \in \mathbb{R} \), then for all \( \xi \in \mathcal{H}(X) \),

\[
E(H^{\frac{\alpha}{2}E_{t}^{2}}(\xi)|X_t, t \in S) = H^{\frac{\alpha}{2}E_{t}^{2}}(\xi)
\]

\[
E(\exp(\xi - \frac{\alpha}{2}E_{t}^{2})|X_t, t \in S) = \exp(\xi - \frac{\alpha}{2}E_{t}^{2})
\]

In particular, if \( X \) is a \( GP(0, R) \) and \( \xi \in \mathcal{H}(X) \), then

\[
E(H^{\frac{\alpha}{2}E_{t}^{2}}(\xi)|X_t, t \in S) = H^{\frac{\alpha}{2}E_{t}^{2}}(\xi)
\]

\[
E(\exp(\xi - \frac{\alpha}{2}E_{t}^{2})|X_t, t \in S) = \exp(\xi - \frac{\alpha}{2}E_{t}^{2})
\]

If \( X \) is a zero mean Gaussian martingale then

\[ H^{\frac{\alpha}{2}E_{t}^{2}}(\xi), \exp(X_t - \frac{\alpha}{2}E_{t}^{2}) \]

are martingales.

**Proof:** Only the first assertion requires proof and we shall prove it for the case \( F(0+) = 0 \) only. \( \int e^{t \alpha} dF(\alpha) < \infty \), \( t \in \mathbb{R} \), implies \( \lambda^{p/2} \in L^2(A) \). Thus

\[
\theta_1 = H^{\frac{\alpha}{2}E_{t}^{2}}(\xi) = (p!)^{\frac{1}{2}}(\frac{A}{\alpha_1})^{\frac{1}{2}} \cdot (\frac{1}{p!})^{\frac{1}{2}} (\frac{A_1}{\alpha_1})^{\frac{1}{2}} \theta_2^{p/2} = H^{\frac{\alpha}{2}E_{t}^{2}}(\xi)
\]

(by (2.4))

\[
\theta_1 = (p!)^{\frac{1}{2}}(\frac{A}{\alpha_1})^{\frac{1}{2}} \theta_2^{p/2} = H^{\frac{\alpha}{2}E_{t}^{2}}(\xi)
\]

and by Theorem 4.1

\[
\hat{\theta}_1 = (p!)^{\frac{1}{2}}(\frac{A}{\alpha_1})^{\frac{1}{2}} \hat{\theta}_2^{p/2} = H^{\frac{\alpha}{2}E_{t}^{2}}(\xi)
\]

Now write
\[ \theta_2 = \exp\{\xi - \frac{A}{\alpha_1} E_1^2\} = \sum_{p \geq 0} \frac{1}{p!} H_{p, \alpha_1} A E_1^2(\xi), \]

the convergence of the series being pointwise. It follows from the expression of \( \theta_1 \) that all terms in the above sum are mutually orthogonal and each term has square norm \( \frac{\alpha_p}{\alpha_1^p} \cdot \frac{(E_1^2)^{\frac{p}{2}}}{p!} \), where \( \alpha_p = \int \alpha^p dF(\alpha) \). Therefore the series converges also in \( L_2(X) \) and

\[ \hat{\theta}_2 = \sum_{p \geq 0} \frac{1}{p!} E_{p, \alpha_1} H A E_1^2(\xi) | X_t, t \leq s \]  
\[ = \alpha_1 \hat{\xi} - \frac{A}{\alpha_1^2} E_1^2 \] .

For \( X \) a Wiener process it is well known that \( (X_{t-t}, t \geq 0) \) and \( (\exp(X_{-t-t}), t \geq 0) \) are martingales.

If \( X \) is a zero mean Gaussian process and \( T = (-\infty, \infty) \) (or any interval)

then by Corollary 4.3 we have that for all \( s \leq t \),

\[ E(H_{p, EX_t^2}(X_t) | X_u, u \leq s) = H_{p, EX_t^2} \hat{X}_{t,s} \]

where

\[ \hat{X}_{t,s} = E(X_t | X_u, u \leq s). \]

An expression for \( \hat{X}_{t,s} \) can always be obtained via the Cramér-Hida representation of \( X_t \):

\[ x_t = \sum_{n=1}^{N} \int_{-\infty}^{t} f(n)(t,u) \, d\xi_u(n). \]

Then we have

\[ \hat{x}_{t,s} = \sum_{n=1}^{N} \int_{-\infty}^{s} f(n)(t,u) \, d\xi_u(n). \]
The case with $p = 2$, i.e. the $L_2$-functional $\chi_t^2 - \mathbb{E}X_t^2$, is considered in Hida and Kallianpur (1975) for a special class of Gaussian processes $X$. It should be clear that whenever a simple expression is available for $\hat{\chi}_{t,s}$, then a simple expression is also found for the nonlinear predictor of $H_p, \mathbb{E}X_t^2(X_t)$.

We close this section with a simple example. Let $(X_t, -\infty < t < \infty)$ be a stationary Gaussian process with $\mathbb{E}X_t = 0$, $\mathbb{E}X_t^2 = 1$, and continuous covariance $R(t)$, which is reciprocal or quasi-Markov on $[0,T]$. Then it is known [13, 5] that $R(t)$, $0 \leq t \leq T$, has one of the following three forms: $Ae^{-at} + (1-A)e^{at}$ where $a > 0$, $A > \frac{1}{2}$, $T \leq (2a)^{-1} \log(A/|\Lambda-1|)$; $\cos at$ where $a > 0$, $T \leq \pi/a$; $1-\cos t$, $0 \leq a \leq 2/T$. We want to estimate $\theta$, an $L_2$-functional of $X_t$, $0 \leq u < t < v \leq T$, based on observations $X_s$, $s \in S = (0,u)\cup(v,T)$. Since $X$ is reciprocal or quasi-Markov on $[0,T]$ we have

$$\hat{\chi}_t = E(X_t|X_s, s \in S) = \alpha X_u + \beta X_v$$

where

$$\alpha = \frac{R(u-t) - R(v-t)R(u-v)}{1-R^2(u-v)} \quad \text{and} \quad \beta = \frac{R(v-t) - R(u-t)R(u-v)}{1-R^2(u-v)}.$$

Since $\theta$ is an $L_2$-functional of $X_t$, it has the orthogonal development

$$\theta = \sum_p a_p H_p(X_t),$$

and thus the best estimate $\hat{\theta}$ of $\theta$ is given by

$$\hat{\theta} = \sum_{p \geq 0} a_p H_p(\hat{\chi}_t) = \sum_{p \geq 0} a_p H_p \chi_t^2(\hat{\chi}_t) = \sum_{p \geq 0} a_p \chi_t^2(\alpha X_u + \beta X_v).$$

4.2 Nonlinear Prediction

Consider the following prediction problem: Let $X = (X_t, t \in T)$, $T$ an interval, be a second order process and let $Y = (Y_t = \theta_t(X_t), t \in T)$ with $\theta_t$ a real function
such that $\mathbb{E}Y_t = 0$ and $\mathbb{E}Y_t^2 < \infty$ for all $t \in \mathbb{T}$. Suppose that on the basis of the past values of $Y$ up to time $t$ we want to find the best prediction of the future value $Y_{t+\tau}$, $\tau > 0$. Two predictors are of special interest: the optimal linear predictor $\hat{Y}_n(t, \tau)$ and the nonlinear predictor $\hat{Y}_{nl}(t, \tau)$. The optimality is in the sense of minimizing the mean square error within the class of all linear and nonlinear predictors respectively. It is well known that

$$\hat{Y}_{nl}(t, \tau) = \mathbb{E}(Y_{t+\tau} | Y_s, s \leq t), \quad \hat{Y}_l(t, \tau) = \text{Proj}_H(Y_s, s \leq t)Y_{t+\tau}.$$

The corresponding mean square predictor errors are denoted by

$$\sigma_{nl}^2(t, \tau) = \mathbb{E}[Y_{t+\tau} - \hat{Y}_{nl}(t, \tau)]^2, \quad \sigma_l^2(t, \tau) = \mathbb{E}[Y_{t+\tau} - \hat{Y}_l(t, \tau)]^2.$$

Now introduce a "super predictor" $\hat{Y}_s(t, \tau)$ as the nonlinear predictor of $Y_{t+\tau}$ based on $(X_s, s \leq t)$, i.e.

$$\hat{Y}_s(t, \tau) = \mathbb{E}(Y_{t+\tau} | X_s, s \leq t),$$

and denote its mean square error by $\sigma_s^2(t, \tau)$. It is clear that

$$(4.1) \quad \sigma_s^2(t, \tau) \leq \sigma_{nl}^2(t, \tau) \leq \sigma_l^2(t, \tau)$$

and thus $\sigma_s^2$ provides a lower bound for the mean square errors of linear and nonlinear prediction. If $X$ is a SIP then $\sigma_s^2$ can be obtained as in Section 4.1 by solving an estimation problem. If, in addition, $\theta_t$ is one-to-one for each $t$ then the $\sigma$-fields generated by $X_t$ and $Y_t$ coincide. In this case

$$\hat{Y}_{nl}(t, \tau) = \hat{Y}_s(t, \tau) = \mathbb{E}(Y_{t+\tau} | X_s, s \leq t),$$

and the nonlinear predictor can be again obtained by solving an estimation problem.

In the important case where $X = (X_t, t \in \mathbb{R})$ is a zero mean stationary Gaussian process with covariance function $R(t, s) = R(t-s)$ and $\theta_t = \theta$ we can calculate the lower bound $\sigma_s^2(t, \tau) = \sigma_s^2(\tau)$ as follows. Write
\[ Y_t = 0(X_t) = \sum_{p=1}^{\infty} a_p H_p \sigma^2(X_t) \]

where \( \sigma^2 = \mathbb{E}X_t^2 \). Note that for all \( \xi, \eta \in H(X) \),

\[ \mathbb{E}H_p, \mathbb{E}^2(\xi)H_p, \mathbb{E}^2(\eta) = p! < \xi^{(p)}, \eta^{(p)} > = p! (\mathbb{E}(\xi \eta))^{(p)} \]

\[ \mathbb{E}H_p, \mathbb{E}^2(\xi)H_q, \mathbb{E}^2(\eta) = 0 \quad \text{if} \ p \neq q. \]

Thus we have

\[ \mathbb{E}Y_t Y_s = \sum_{p=1}^\infty p! a_p^2 \mathbb{R}^p(t-s). \]

Let \( \hat{X}(t,\tau) = E(X_{t+\tau} | X_s, s \leq t) \) be the optimal nonlinear predictor of \( X_{t+\tau} \) (which is also the optimal linear predictor since \( X \) is Gaussian), and \( \sigma_0^2(\tau) \) be the mean square error. Then by Corollary 4.3,

\[ \hat{\gamma}_s(t,\tau) = \sum_{p=1}^\infty a_H \mathbb{E}^2(\hat{X}(t,\tau)) \]

and hence

\[ \sigma_s^2(\tau) = E(Y_{t+\tau} - \hat{\gamma}_s(t,\tau))^2 = EY_{t+\tau}^2 - E(\hat{\gamma}_s(t,\tau))^2 \]

\[ = \sum_{p=1}^\infty p! a_p^2 \sigma^2 p - \sum_{p=1}^\infty p! a_p^2 (\sigma^2 - \sigma_0^2(\tau))^p \]

\[ = \sum_{p=1}^\infty p! a_p^2 \sigma^2 p - [\sigma^2 - \sigma_0^2(\tau)]^p. \]

It is well known from the general theory of stationary processes (e.g. Doob (1953), Rozanov (1967)) that \( \sigma_0^2(\tau) \) can be obtained analytically (if not explicitly) through the Wiener-Paley factorization theorem if \( X \) is linearly regular, i.e. \( \eta H(X_t, s \leq t) = \{0\} \). When \( X \) is mean square continuous and linearly regular we now show that so is \( Y \), and hence \( \sigma_s^2(\tau) \) can be obtained analytically. \( Y \) is clearly stationary and its mean square continuity follows from the continuity of \( R \) and (4.3). The linear regularity of \( Y \) follows from the fact that
a Gaussian process is linearly regular if and only if its remote past is trivial. Here we give a purely geometric proof of this property; for a proof using Kolmogorov's zero-one law see Rosanov (1967) and Ibragimov and Linnik (1971).

**THEOREM 4.4** Let \( X = (X_t, t \in \mathbb{R}) \) be a zero mean Gaussian process, 
\[ B_t = \mathcal{B}(X_s, s \leq t), \quad t \in \mathbb{R}, \quad \text{and} \quad B_\infty = \cap B_t. \]
Then \( X \) is linearly regular if and only if \( B_\infty \) is trivial.

**Proof:** The "if" part is clear. For the "only if" part, note first that the triviality of \( B_\infty \) is equivalent to the condition
\[ \cap_{t \geq 2} (X_s, s \leq t) = \{0\}. \]
For simplicity we write \( L_t = L_2(X_s, s \leq t) \) and \( H_t = H(X_s, s \leq t) \). Thus we need to show that \( H_\infty = \cap H_t = \{0\} \) implies \( L_\infty = \cap L_t = \{0\} \). So assume that \( H_\infty = \{0\} \)
and let \( \theta \in L_\infty \). We will show that \( \theta = 0 \).

Fix a sequence \( \{t_n, n=1,2,\ldots\} \) decreasing to \( -\infty \). The family of subspaces \( H_n \) has the property that \( H_{t_m} \subseteq H_{t_n} \) for \( m < n \) and \( \cap_{n \in \mathbb{N}} H_n = \{0\} \). Thus it follows readily (see e.g. Rozanov (1967), pp. 53, 56) that \( \lim_{n \to \infty} \text{Proj}_{H_{t_n}} \xi = 0 \) for all \( \xi \in H(X) \), and hence
\[ H_{t_1} = \bigotimes_{n=1}^{\infty} D_n \]
where \( D_n \) is the orthogonal complement of \( H_{t_{n+1}} \) in \( H_{t_n} \), i.e., \( H_{t_n} = H_{t_{n+1}} \oplus D_n \).
In each \( D_n \) pick a CONS; then pool all these CONS's together to get a CONS \( \{\xi_\gamma, \gamma \in \Gamma\} \) in \( H_{t_1} \). Observe that
\[ \text{Proj}_{H_{t_n}} \xi_\gamma = 0 \text{ or } \xi_\gamma. \quad (4.6) \]
Since \( \theta \in L_\infty \subseteq L_{t_1} \), it has the orthogonal development.
and by Corollary 4.2,

$$\theta = \operatorname{Proj}_{t_n} \theta = \sum_{p \geq 0} \sum_{p_1 + \ldots + p_k = p} a_{\gamma_1 \ldots \gamma_k} \operatorname{Proj}_{t_n} \xi_{\gamma_1} \ldots \operatorname{Proj}_{t_n} \xi_{\gamma_k},$$

which together with (4.6) yields

$$E \theta^2 = \sum \sum (a_{\gamma_1 \ldots \gamma_k})^2 ||(\operatorname{Proj}_{t_n} \xi_{\gamma_1}) \ldots \operatorname{Proj}_{t_n} \xi_{\gamma_k}||^2.$$

Now let $$n \to \infty$$ in (4.7). The limit can be taken inside the summation since each summand is bounded by $$||\xi_{\gamma_1} \ldots \operatorname{Proj}_{t_n} \xi_{\gamma_k}||^2$$; and the limit of each summand is zero because of (4.6). Consequently $$E \theta^2 = 0$$, i.e. $$\theta = 0$$, and the theorem is proved. \(\square\)

Jaglom (1970) has considered the problem of comparing the performance of optimal linear and nonlinear predictors for polynomial functions of certain stationary Markov processes. Donelson and Maltz (1972) studied this problem in detail for polynomial functions of the Ornstein-Uhlenbeck process. The inequality (4.1) plays a central role in such studies. As an example, let $$X$$ be the Ornstein-Uhlenbeck process, i.e. a zero mean Gaussian process with covariance function $$R(t,s) = e^{-|t-s|}$$, and let $$Y$$ be given by (4.2). By the Markov property of $$X$$ we have

$$\hat{X}(t,t) = E(X_t | X_s, s \leq t) = e^{-t} X_t.$$

Thus it follows from (4.4) and (4.5) that

$$\hat{Y}_s(t,t) = \sum_{p \geq 1} a_{p} e^{-2t} (e^{-T} X_s) = \sum_{p \geq 1} a_{p} e^{-pT} (X_s),$$
\[ \sigma_s^2(\tau) = \sum_{p \geq 1} p!a_p^2(1 - e^{2p\tau})^2. \]

This result, with \( Y_t \) a polynomial function of \( X_t \), has been obtained by Donelson and Maltz using a different approach; they also compared \( \sigma_s^2 \) with \( \sigma_\Delta^2 \) and found that these two errors are frequently close to one another. Finally, we remark that if \( Y_t = H_p(X_t) \) then

\[ \hat{Y}_{n\lambda}(t,\tau) = \hat{Y}_s(t,\tau) = e^{-P\tau}Y_t, \]

\[ \sigma_{n\lambda}^2(\tau) = \sigma_s^2(\tau) = 1 - e^{-2p\tau}. \]
REFERENCES


