* Ph.D. dissertation under the direction of Stamatis Cambanis.

† This research was supported by the Air Force Office of Scientific Research under Grant AFOSR-75-2796.

NONLINEAR ANALYSIS OF SPHERICALLY INVARIANT PROCESSES AND ITS RAMIFICATIONS*

by

Steel T. Huang†

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1028

September, 1975

HUANG, STEEL T. Nonlinear Analysis of Spherically Invariant Processes and Its Ramifications. (Under the direction of STAMATIS CAMBANIS.)

The objective of this work is to study the structure of the nonlinear space of spherically invariant processes (which constitute a natural generalization of Gaussian processes) and to define multiple Wiener integrals for spherically invariant processes and hence for Gaussian processes, thus generalizing the notion originally introduced for a Wiener process. We obtain an orthogonal decomposition of the nonlinear space and an integral representation for its L_2-functionals, which is expressed in terms of multiple Wiener integrals.

These results are applied to nonlinear estimation and prediction theory, the nonlinear noise theory, and the equivalence of spherically invariant processes.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor Professor S. Cambanis for his constant stream of ideas and unfailing patience. He has not only given me helpful advice in my academic work, but also shown great concern for me personally.

I am grateful to Professors G. Simons and W. Hoeffding for introducing me to the theory of Probability and Statistics; and to Professors M. R. Leadbetter and C. R. Baker for leading me into the field of Stochastic Processes.

Above all I wish to acknowledge my indebtedness to my parents and my wife. It is their love, encouragement and understanding that made all this possible.

CONTENTS

I. INTRODUCTION AND SUMMARY

1. Introduction

By a *stochastic process* $X=(X_t,\ t\in T)$ we mean a family of random variables (r.v.'s) X_t, $t\in T$, defined on a probability space $(\Omega,\mathcal{B},P)$. T is an arbitrary index set but very often it is an interval on the real line; $\mathcal{B}$ is usually taken to be $\mathcal{B}(X)$, the σ-field generated by the process X, or $\overline{\mathcal{B}}(X)$, the completion of $\mathcal{B}(X)$ with respect to (w.r.t.) the measure P. For each $\omega\in\Omega$, $X(\omega)$ represents the corresponding sample path of the process, which is an element of $\mathbb{R}^T$, the space of all functions defined on T. X is called a *coordinate process* if $(\Omega,\mathcal{B},P) = \left(\mathbb{R}^T,\mathcal{B}(\mathbb{R}^T),P\right)$ and $X_t(\omega)=\omega(t)$, where $\mathcal{B}(\mathbb{R}^T)$ is the σ-field generated by cylinder sets of $\mathbb{R}^T$. X is called a *second order process* if $EX_t^2<\infty$ for all $t\in T$.

Now consider a second order process X with zero mean. There are two Hilbert spaces closely related to such a process. The first one is the *nonlinear space of* X, $L_2(X) = L_2\left(\Omega,\mathcal{B}(X),P\right)$ (consisting of all r.v.'s on $\left(\Omega,\mathcal{B}(X),P\right)$ with finite second moment) with inner product $<\xi,\eta>=E\xi\eta$. Elements of $L_2(X)$ are called *(nonlinear)* L_2*-functionals of* X. (Note that every $\mathcal{B}(X)$-measurable function θ is of the form $\theta(\omega)=\rho\left(X(\omega)\right)$ where ρ is a $\mathcal{B}(\mathbb{R}^T)$-measurable function.) The second Hilbert space is the *linear space of* X, $H(X)$, the subspace of $L_2(X)$ spanned by X_t, $t\in T$. Elements of $H(X)$ are called *linear* L_2*-functionals of* X.

It is well known that a linear problem on the process X can be cast as an optimization or a representation problem on the Hilbert space $H(X)$;

and a nonlinear problem on the process X can be cast as an optimization or a representation problem on the Hilbert space $L_2(X)$.

The linear problems on second order processes have been studied extensively over the past two decades and a satisfactory theory has been developed, especially for the stationary case (see e.g. Doob [1953], Gikhman and Skorokhod [1969]). The basic approach is to establish an isomorphism between $H(X)$ and a function Hilbert space (e.g. L_2-spaces, reproducing kernel Hilbert spaces), and then to transform the original problem to an optimazation or a representation problem on the function space, and solve it. This approach, usually called the Hilbert space method, has been fruitful for the linear theory. It is only natural to ask whether the same approach will work for nonlinear problems. The analogy does exist. However it requires knowledge of the structure of the nonlinear space $L_2(X)$. It is clear that the structure of the linear space $H(X)$ depends on the process X only through its covariance function, while the structure of the nonlinear space $L_2(X)$ is much more complex and depends on the process X through all its finite dimensional distributions.

The most useful notion in the study of the nonlinear space of a Wiener process is the Multiple Wiener Integral. This notion was first introduced by Wiener [1938], who termed it "Polynomial Chaos", and was redefined in a somewhat deeper way by Itô [1951]. Itô showed that his multiple integrals of different degree have the important property of being mutually orthogonal and also presented their connection with the celebrated Fourier-Hermite expansion of L_2-functionals of Cameron and Martin [1947]. Subsequently, Itô [1956] extended to general processes with stationary independent increments the Wiener-Itô expansion of L_2-functionals in terms of multiple Wiener integrals.

In his important work on nonlinear problems Wiener [1958] reinterpreted the multiple Wiener integrals for a Wiener process in an extremely simple and intuitive way and made some interesting applications.

Neveu [1968] and Kallianpur [1970] studied the connection between the nonlinear space of a Gaussian process and the tensor products of its linear space, which sheds new light and gives more insight on the structure of the nonlinear space.

The objective of this work is to study the structure of the nonlinear space of spherically invariant processes (which constitute a natural generalization of Gaussian processes) and to define multiple Wiener integrals for spherically invariant processes and hence for general Gaussian processes, thus generalizing the notion originally introduced for a Wiener process. These results are applied to the nonlinear estimation and prediction theory, the nonlinear noise theory, and the equivalence of spherically invariant measures.

2. Summary

In Chapter II we develop the structure of Λ_2 and λ_2-spaces, which are our choice of function Hilbert spaces in the Hilbert space method approach to the linear and nonlinear problems.

In Chapter III we obtain an orthogonal decomposition of the nonlinear space of a spherically invariant process and an integral representation for its L_2-functionals. The integral representation is expressed in terms of multiple Wiener integrals. The general properties of spherically invariant processes are also studied.

In Chapter IV, we study the equivalence of spherically invariant processes and obtain an explicit expression for the Radon-Nikodym derivative; we also derive a 0-1 law for spherically invariant processes and give some applications.

In Chapter V we solve the general nonlinear estimation problem for spherically invariant processes, and hence for Gaussian processes. As an application we derive a lower bound on the mean square prediction error for a class of nonlinear prediction problems.

In Chapter VI we introduce a new class of noises and show that they possess the same important properties as the white noise.

In the Appendix we briefly recapitulate some results on tensor products of Hilbert spaces and on Hermite polynomials which are used extensively throughout the whole work.

II. THE HILBERT SPACES $\Lambda_2(R)$ AND $\lambda_2(R)$

Throughout this chapter T will be an interval closed or open, bounded or unbounded, and $X = (X_t, t\epsilon T)$ a second order process with zero mean and covariance function $R(t,s)$. Integrals over T will be denoted by the integral sign with no subscript, and 1_E will denote the characteristic function of the set E.

It is shown in Loève [1955, p. 472] that the following two integrals

$$I(f) = R - \int f(t)dX_t$$

$$J(f) = R - \int f(t)X_t dt$$

can be defined as the mean square limits of the corresponding sequences of approximating Riemann sums if and only if the following double Riemann integrals exist

$$R - \iint f(t)f(s)d^2R(t,s)$$

$$R - \iint f(t)f(s)R(t,s)dtds \ ,$$

and then $I(f)$ and $J(f)$ are the random variables with means zero and variances the corresponding double Riemann integrals.

1. The Hilbert Spaces $\Lambda_2(R)$ and $\lambda_2(R)$

Consider the set S_I of all step functions on T $f(t) = \sum_1^N f_n 1_{(a_n,b_n]}(t)$, $(a_n,b_n] \subset T$ and define

$$\int f(t)dX_t = \sum_1^N f_n(X_{b_n} - X_{a_n}) .$$

S_I is clearly a linear space and for all $f,g \in S_I$ we have

$$E \int f(t)dX_t = 0$$

$$E \left(\int f(t)dX_t \cdot \int g(t)dX_t\right) = \iint f(t)g(s)\ d^2R(t,s)$$

where the double integral is defined in the obvious way. Two step functions f,g will be considered identical if

$$\iint \big(f(t)-g(t)\big)\big(f(s)-g(s)\big)d^2R(t,s) = 0 .$$

If we define for $f,g \in S_I$,

$$\langle f,g\rangle = \iint f(t)f(s)\ d^2R(t,s)$$

then $(S_I, \langle\circ,\circ\rangle)$ is an inner product space. Indeed $\langle f,g\rangle$ has the ordinary bilinear and symmetric properties,

$$\langle f,f\rangle = E \left(\int f dX\right)^2 \geq 0$$

and $\langle f,f\rangle = 0$ only when f is the zero element of S_I according to the convention introduced above.

Now let $\Lambda_2(R)$ be the completion of S_I, so that it is a Hilbert space with inner product denoted again by $\langle\circ,\circ\rangle$. A typical element in $\Lambda_2(R)$ is a Cauchy sequence of step functions. However, we will find it convenient to treat elements in $\Lambda_2(R)$ as "formal" functions in $t\epsilon T$ and to write $\iint f(t)g(s)d^2R(t,s)$ for the inner product $\langle f,g\rangle$ (see Theorem 1 for a partial justification).

Notice that for $f\epsilon S_I$ the integral $\int f(t)dX_t$ depends on X only through its increments. Thus we may suppose without loss of generality that there is a point $t_0 \epsilon T$ such that

$$X_{t_0} = 0 \quad \text{a.e.}$$

Under this assumption we can establish an isomorphism between $H(X)$ and $\Lambda_2(R)$ as follows. The map

$$S_I \to H(X): f \longmapsto \int fdX$$

preserves inner products and hence it can be extended to an isomorphism on $\Lambda_2(R)$ to a closed subspace of $H(X)$. But the set

$$X_t = \int 1_t(u)dX_u \ , \quad t\epsilon T \ ,$$

where $1_t = 1_{(t_0,t]}$ for $t \geq t_0$ and $= -1_{(t,t_0]}$ for $t<t_0$, generates $H(X)$ and $1_t \epsilon S_I$. It follows that the isomorphism is onto $H(X)$, i.e.

$$\Lambda_2(R) \cong H(X) \ .$$

We denote this isomorphism by I and we define the integral of $f \epsilon \Lambda_2(R)$ with respect to X (which we write as $\int f(t)dX_t$ following our convention to view elements of $\Lambda_2(R)$ as formal functions) by

$$\int f(t)dX_t = I(f) \ .$$

The properties of this integral follow from those of I and are the analogues of the properties of the integral where X has orthogonal increments (see e.g. Doob [1953]). The integral is defined for "functions" in $\Lambda_2(R)$ and thus it is of interest to identify usual functions in $\Lambda_2(R)$ besides the step functions. Two such classes of functions are identified in the following.

Under the additional assumption that $R(t,s)$ is of bounded variation on every bounded subset of $T\times T$, Cramér [1951] defined $\Lambda_2(R)$ as the completion (with respect to the same inner product) of the set S_I^* of all functions f whose double Riemann integral $R\text{-}\iint f(t)f(s)d^2R(t,s)$ exists. However, Cramér's definition is not appropriate for the general case (where R is not necessarily of bounded variation on bounded sets) since then 1_t may not be in $\Lambda_2(R)$ and thus $\Lambda_2(R)$ may not be isomorphic to $H(X)$. In this sense our definition of $\Lambda_2(R)$ is the appropriate generalization of the definition given by Cramér.

That S_I^* is always (even when R may not be of bounded variation) a subspace of $\Lambda_2(R)$ and that for $f\epsilon S_I^*$, $I(f) = R\text{-}\int f(t)dX_t$, follow immediately from the fact that $f\epsilon S_I^*$ is equivalent to the existence of $R\text{-}\int f(t)dX_t$ and the approximating Riemann sums for $R\text{-}\int f(t)dX_t$ are of the form $\int f_n dX_t$ with $f_n \epsilon S_I$. It can be also shown that when R is of bounded variation on bounded sets then the two definitions of $\Lambda_2(R)$ coincide. We show instead the following result which is more useful for our purposes.

$R(t,s)$ is said to be of bounded variation on $[a,b] \times [c,d]$ if for all N,M and points $a = t_0<t_1<\ldots<t_N = b$, $c = s_0<s_1<\ldots<s_M = d$ the sum $\sum_{n=1}^{N}\sum_{m=1}^{M} \left|\Delta_{(t_n,s_m)}^{(t_{n-1},s_{m-1})} R\right|$ is bounded, where

$$\Delta_{(t,s)}^{(t',s')} R = R(t',s')-R(t',s)-R(t,s')+R(t,s) .$$

Also R is said to be of bounded variation on every finite domain of $T\times T$ if it is of bounded variation on every $[a,b] \times [c,d] \subset T\times T$. Such an R determines uniquely a σ-finite signed measure on the Borel subsets of $T\times T$, denoted again by R, such that $R\big((t,t']\times(s,s']\big) = \Delta_{(t,s)}^{(t',s')} R$.

Let L_I be the set of all measurable functions f on T such that the following Lebesgue integrals are finite

$$\iint |f(t)f(s)| \ d^2|R|(t,s) < \infty$$

$$\iint |f(t)| \ 1_{(a,b]}(s) \ d^2|R|(t,s) < \infty$$

for all $(a,b] \subset T$, where $|R|$ is the total variation measure of R. We say that <u>*the function f in L_I represents an element in $\Lambda_2(R)$*</u> if there is a $f' \in \Lambda_2(R)$ such that for all $g \in S_I$,

$$<f',g> = \iint f(t)g(s) \ d^2R(t,s) \ .$$

Notice that if such an f' exists it is unique since S_I is dense in $\Lambda_2(R)$. We will then denote f' by f and we will write $f \in \Lambda_2(R)$. With this convention we have the following

<u>THEOREM 1</u>. *Let $R(t,s)$ be of bounded variation on every finite domain of $T \times T$. Then L_I is a dense subset of $\Lambda_2(R)$. Also if $f_1, f_2 \in L_I$ and $\iint |f_1(t)f_2(s)| d^2|R(t,s)| < \infty$ then*

$$<f_1,f_2> = \iint f_1(t)f_2(s) \ d^2R(t,s) \ .$$

<u>PROOF</u>. Let E be a bounded Borel subset of T. Then $1_E \in L_I$ and we will prove that $1_E \in \Lambda_2(R)$, i.e. there is an $f \in \Lambda_2(R)$ such that for all $g \in S_I$,

$$<f,g> = \iint 1_E(t)g(s) \ d^2R(t,s) \ .$$

Let I be a finite interval containing E (so that $|R|(I \times I) < \infty$). We can always find $I_n \subset I$, $n = 1,2,\ldots,$ with each I_n a finite union of half open intervals such that

$$|R|\big((I_n \Delta E)\times I\big) \to 0 \quad \text{as} \quad n \to \infty .$$

Since

$$|R(I_n\times I_m)-R(E\times E)| \le |R|\big((I_n\Delta E)\times E\big) + |R|\big((I_m\Delta E)\times E\big) \to 0 ,$$

it follows that $\langle 1_{I_n}, 1_{I_m}\rangle \to R(E\times E)$ and thus $\{1_{I_n}\}_{n=1}^{\infty}$ is a Cauchy sequence in $\Lambda_2(R)$. Define $f \in \Lambda_2(R)$ by $f = \lim 1_{I_n}$.

We first show that f does not depend on the approximating sequence I_n, $n = 1,2,\ldots$. Let $I_n' \subset I$, $n = 1,2,\ldots$ be another such approximating sequence and $f' = \lim 1_{I_n'}$. Then for each interval $J\subset I$ we have

$$\begin{aligned}
|\langle f-f', 1_J\rangle| &= \lim |\langle 1_{I_n} - 1_{I_n'}, 1_J\rangle| \\
&= \lim |R(I_n\times J) - R(I_n'\times J)| \\
&\le \lim \{|R|\big((I_n\Delta E)\times J\big) + |R|\big((I_n'\Delta E)\times J\big)\} \\
&\le \lim \{|R|\big((I_n\Delta E)\times I\big) + |R|\big((I_n'\Delta E)\times I\big)\} = 0 .
\end{aligned}$$

If M_I is the closed subspace of $\Lambda_2(R)$ generated by all $1_{(a,b]}$, $(a,b]\subset I$, then it is clear from the above that $f,f' \in M_I$ and $f-f' \perp M_I$. It follows that $f = f'$.

We now show that f does not depend on the finite interval I containing E. Let J and J_n, $n = 1,2,\ldots$ have the same properties as I and I_n, $n = 1,2,\ldots$. Define $I_n' = I_n\cap J$ and $J_n' = J_n\cap I$, $n = 1,2,\ldots$. Then clearly $I_n', J_n' \subset I \cap J$ and $|R|\big((I_n'\Delta E)\times(I\cap J)\big) \to 0$, $|R|(J_n'\Delta E, I\cap J) \to 0$ since $(I_n'\Delta E)\times(I\cap J) \subset (I_n\Delta E)\times I$, $(J_n'\Delta E)\times(I\cap J) \subset (J_n\Delta E)\times J$ and $|R|$ is a measure. From the result of the previous paragraph applied to I,J and $I\cap J$ we have

$$\lim 1_{I_n} = \lim 1_{I'_n} = \lim 1_{J'_n} = \lim 1_{J_n}$$

and thus f does not depend on I.

Hence f is well-defined. Now fix $J = (a,b] \subset T$ and let I be a finite interval containing $E \cup J$ and I_n, $n = 1,2,\ldots$ as before, i.e. $1_{I_n} \to f$. Then

$$\begin{aligned} |\langle f, 1_J\rangle - R(E\times J)| &= \lim |R(I_n\times J) - R(E\times J)| \\ &\le \lim |R|\big((I_n \Delta E)\times I\big) = 0 . \end{aligned}$$

Hence $\langle f, 1_J\rangle = R(E\times J)$ and since J is arbitrary it follows that for all $g \in S_I$, $\langle f,g\rangle = \iint 1_E(t)g(s)d^2R(t,s)$. Thus we have shown that $1_E \in \Lambda_2(R)$.

Now if E and F are bounded Borel subsets of T, denoting by 1_E and 1_F also the corresponding elements in $\Lambda_2(R)$ we have $\langle 1_E, 1_F\rangle = R(E\times F)$. Indeed if I is a finite interval containing $E \cup F$ and as before $1_{I_n} \to 1_E$ and $1_{J_n} \to 1_F$ we have (since R is symmetric)

$$\begin{aligned} |R(I_n\times J_m) - R(E\times F)| &\le |R|\big(I_n\times(J_m\Delta F)\big) + |R|\big((I_n\Delta E)\times F\big) \\ &\le |R|\big((J_m\Delta F)\times I\big) + |R|\big((I_n\Delta E)\times I\big) \to 0 \end{aligned}$$

and thus $\langle 1_E, 1_F\rangle = \lim R(I_n\times J_m) = R(E\times F)$.

It then follows that if ϕ is a simple function in L_I with bounded support, then $\phi \in \Lambda_2(R)$, and if ϕ_1,ϕ_2 are two such functions then

$$\langle \phi_1, \phi_2\rangle = \iint \phi_1(t)\phi_2(s)d^2R(t,s) .$$

Now let $f \in L_I$. Then there exist simple functions ϕ_n, $n=1,2,\ldots$, with bounded support such that $|\phi_n| \uparrow |f|$ on T. It follows from the

Bounded Convergence Theorem that

$$<\phi_n,\phi_m> = \iint \phi_n(t)\phi_m(s)d^2R(t,s) \to \iint f(t)f(s)d^2R(t,s) .$$

Hence ϕ_n, $n = 1,2,\ldots$, is a Cauchy sequence in $\Lambda_2(R)$ and denote its limit by f'. Then for all $g\epsilon S_I$

$$\begin{aligned}<f',g> &= \lim <\phi_n,g> = \lim \iint \phi_n(t)g(s)d^2R(t,s)\\ &= \iint f(t)g(s)d^2R(t,s)\end{aligned}$$

again by the Bounded Convergence Theorem since $f\epsilon L_I$ and $g\epsilon S_I$ imply $\iint |f(t)g(s)|d^2|R|(t,s)<\infty$. Since the values of $<f',g>$ for $g\epsilon S_I$ determine f' uniquely, the last equality implies that f' is uniquely determined by f, independently of the approximating sequence ϕ_n. It follows that $f \epsilon \Lambda_2(R)$ and thus $L_I \subset \Lambda_2(R)$. Since L_I contains S_I it is dense in $\Lambda_2(R)$.

For the last statement of the theorem, with the obvious notation, we have

$$\begin{aligned}<f_1,f_2> &= \lim <\phi_{1,n},\phi_{2,n}> = \lim \iint \phi_{1,n}(t)\phi_{2,n}(s)d^2R(t,s)\\ &= \iint f_1(t)f_2(s)d^2R(t,s)\end{aligned}$$

where the additional assumption on f_1,f_2 makes the Bounded Convergence Theorem applicable. Q.E.D.

Consider the set S_J of all functions f on T such that the Riemann integral $R - \iint f(t)f(s)R(t,s)dtds$ exists and is finite. S_J is a linear space. Two functions f and g in S_J will be considered identical if

$$R - \iint \big(f(t)-g(t)\big)\big(f(s)-g(s)\big)R(t,s)dtds = 0 .$$

For $f,g \in S_J$ we define $\int f(t)X_t dt = R\text{-}\int f(t)X_t dt$ and then we have

$$E\left(\int f(t)X_t dt \circ \int g(t)X_t dt \right) = R - \iint f(t)g(s)R(t,s)dtds \ .$$

Define for $f,g \in S_J$,

$$\langle f,g \rangle = R - \iint f(t)g(s)R(t,s)dtds \ .$$

Then $(S_J, \langle \cdot, \cdot \rangle)$ becomes an inner product space. $\lambda_2(R)$ is defined to be the completion of the inner product space S_J and so it is a Hilbert space. Again a typical element in $\lambda_2(R)$ is a sequence of functions convergent in norm. However formally we shall treat elements in $\lambda_2(R)$ as functions and write $\iint f(t)g(s)R(t,s)dtds$ as the inner product $\langle f,g \rangle$.

In order to establish an isomorphism between $H(X)$ and $\lambda_2(R)$ we shall assume that X is mean square continuous which is equivalent to the continuity of the covariance function $R(t,s)$. Consider the sequence of functions $n \circ 1_{(\tau - \frac{1}{n}, \tau]}(t)$ where τ is an interior point of T.

It is easy to show that this sequence is a Cauchy sequence in $\lambda_2(R)$, whose limit is denoted by δ_τ, and that

$$X_\tau = \text{l.i.m.} \int n \circ 1_{(\tau - \frac{1}{n}, \tau]}(t)\ X_t\ dt \ .$$

Then, the map

$$S_J \to H(X) \ ; \ f \mapsto \int f(t)\ X_t\ dt$$

preserves the inner product and its range includes X_τ for all interior τ of T (which is linearly dense in $H(X)$ by mean square continuity). Hence it can be extended to an isomorphism on $\lambda_2(R)$ onto $H(X)$. Thus $\lambda_2(R) \cong H(X)$, the isomorphism is denoted by J and for $f \in \lambda_2(R)$ we define

$$J(f) = \int f(t)X_t dt \ .$$

A useful connection between the integrals I and J and the spaces Λ_2 and λ_2 can be established as follows. Let $Z_t = \int_{t_0}^{t} X_u du = J(1_{(t_0,t]})$ where t_0 is an arbitrary but fixed point in T. ($1_{(t_0,t]} \in \lambda_2(R)$ since R is continuous.) Then,

$$\Gamma(t,s) = E\ Z_t Z_s = \int_{t_0}^{t} \int_{t_0}^{s} R(u,v)dudv \ .$$

<u>THEOREM 2.</u> *If* X *is mean square continuous then* $\lambda_2(R) = \Lambda_2(\Gamma)$ *and for all* $f \in \lambda_2(R) = \Lambda_2(\Gamma)$

$$\int f(t)X_t dt = \int f(t)dZ_t \ .$$

Hence $H(X) = H(Z)$.

<u>PROOF.</u> We will prove first that the existence of $R - \int_T \int_T f(t)f(s)R(t,s)dtds$ is equivalent to that of $R - \int_T \int_T f(t)f(s)d^2\Gamma(t,s)$. We may assume that T is a closed interval by the very definition of a Riemann integral. We also assume that $|f(t)| < M \ \forall\ t \in T$; otherwise neither Riemann integral will exist. Consider the typical Riemann sums

$$R_J = \sum \sum f(t_i)f(s_j)R(t_i,s_j)\ |A_i|\ |B_j|$$

$$R_I = \sum \sum f(t_i)f(s_j)\Gamma(A_i \times B_j)$$

where $\{A_i\}$, $\{B_j\}$ are interval partitions of T, $t_i \in A_i$, $s_j \in B_j$; and $|A_i|$, $|B_j|$ denote the lengths of these intervals. By the uniform continuity of $R(t,s)$ we have that for every $\varepsilon>0$, $|R_I - R_J| \le M^2|T|^2\varepsilon$ as $\max(|A_i|,|B_j|) \to 0$. We thus conclude that

$$R - \iint f(t)f(s)R(t,s)dtds = R - \iint f(t)f(s)d^2\Gamma(t,s)$$

and the existence of one side implies that of the other side. In short,

$$(S_I^*, \langle\cdot,\cdot\rangle_{\Lambda_2(\Gamma)}) = (S_J, \langle\cdot,\cdot\rangle_{\lambda_2(R)}) .$$

Note that S_I^* is dense in $\lambda_2(\Gamma)$ (since Γ is continuous) and S_J is dense in $\lambda_2(R)$. Thus $\lambda_2(R) = \Lambda_2(\Gamma)$.

For a step function f we clearly have

$$\int f(t)X_t dt = \int f(t)dZ_t ,$$

hence this is true for all $f \in \Lambda_2(\Gamma)$ by the continuity of I and J.

Q.E.D.

$\lambda_2(R)$ may contain interesting classes of functions larger than S_J. Let L_J be the set of all measurable functions f on T such that the following Lebesgue integrals are finite

$$\iint |f(t)f(s)R(t,s)|dtds < \infty$$

$$\iint |f(t)|\ 1_{(a,b]}(s)\ |R(t,s)|dtds < \infty$$

for all $(a,b] \subset T$. We will follow the same convention (as for Λ_2) in treating functions f in L_J as elements of $\lambda_2(R)$ if there is a $f' \in \lambda_2(R)$ such that for all g in a dense subset of $\lambda_2(R)$,

$$\langle f',g\rangle = \iint f(t)g(s)R(t,s)dtds .$$

With this convention the following is a corollary of Theorems 1 and 2.

COROLLARY 3. *Let* $R(t,s)$ *be continuous on* $T\times T$. *Then* L_J *is a dense subset of* $\lambda_2(R)$. *Also if* $f_1, f_2 \in L_J$ *and* $\iint |f_1(t)f_2(s)R(t,s)|dtds < \infty$, *then*

$$\langle f_1, f_2\rangle = \iint f_1(t)f_2(s)R(t,s)dtds \ .$$

We now remark on the relation between Λ_2, λ_2-spaces and L_2-spaces. The spaces $\Lambda_2(R)$ and $\lambda_2(R)$ are generalizations of L_2-spaces. In general they are larger than L_2-spaces. As an example, consider $R(t,s)$ a continuous covariance function on $[a,b] \times [a,b]$ and let $\Gamma(t,s)$ be defined as before. Then $\lambda_2(R) = \Lambda_2(\Gamma)$. Any function f in $L_2([a,b],dt)$ belongs to $\lambda_2(R) = \Lambda_2(\Gamma)$ since

$$\iint |f(t)f(s)R(t,s)|dtds \le \max |R(t,s)| \circ \left(\int |f(t)|dt\right)^2$$

$$\le \max |R(t,s)| \circ |b-a| \int f^2(t)dt < \infty \ .$$

However, $\delta_t \in \lambda_2(R) = \Lambda_2(\Gamma)$ is not in $L_2([a,b],dt)$ since $X_t = J(\delta_t)$ for all interior points t of T implies $R(t,s) = \iint \delta_t(u)\delta_s(v)R(u,v)dudv$.

Nevertheless, there is a special case where $\lambda_2(R)$ reduces to an L_2-space. Let X be a zero mean process with orthogonal increments. Assume $X_{t_0} = 0$ a.e. for some fixed $t_0 \in T$. Then, $R(t,s) = F\big(t_0 \vee (t\wedge s)\big) + F\big(t_0 \wedge (t \vee s)\big)$ where $F(t) = EX_t^2$ if $t \ge t_0$ and $= -EX_t^2$ if $t \le t_0$. F is non-decreasing and thus $R(t,s)$ is of bounded variation in every finite domain of $T\times T$, and the associated measure concentrates on the diagonal $t=s$ of $T\times T$. In this case $\Lambda_2(R) = L_2\big(T, F(dt)\big)$. In particular, if X is the Wiener process $\Lambda_2(R) = L_2(T,dt)$.

2. Tensor Products of $\Lambda_2(R)$ and $\lambda_2(R)$

Now let us study the tensor product spaces $\otimes^n \Lambda_2(R)$ and $\otimes^n \lambda_2(R)$. Consider the set $S_I^{(n)}$ of all step functions $K(t_1,\ldots,t_n)$ on T^n. Define the following function on $S_I^{(n)} \times S_I^{(n)}$

$$<K,G> = \iint \circ \cdots \circ \iint K(t_1,\ldots,t_n)G(s_1,\ldots,s_n)d^2R(t_1,s_1)\ldots d^2R(t_n,s_n) ,$$

and identify K with G if $<K-G,K-G> = 0$. Let $1_{I_1\times\ldots\times I_n}$, $1_{J_1\times\ldots\times J_n} \in S_I^{(n)}$ (i.e., I_i, J_j are bounded half open intervals in T). Then

$$\begin{aligned} <1_{I_1\times\ldots\times I_n}, 1_{J_1\times\ldots\times J_n}> &= \iint 1_{I_1}(t)1_{J_1}(s)d^2R(t,s) \circ \cdots \\ &\qquad \circ \iint 1_{I_n}(t)1_{I_n}(s)d^2R(t,s) \\ &= <1_{I_1},1_{J_1}>_{\Lambda_2(R)} \cdots <1_{I_n},1_{J_n}>_{\Lambda_2(R)} \\ &= <1_{I_1}\otimes\ldots\otimes 1_{I_n}, 1_{J_1}\otimes\ldots\otimes 1_{J_n}>_{\otimes^n\Lambda_2(R)} . \end{aligned}$$

This implies that $(S_I^{(n)}, <\circ,\circ>)$ is an inner product space and we shall denote by $\Lambda_2(\otimes^n R)$ the completion of $S_I^{(n)}$. Since $\{1_{I_1}\otimes\ldots\otimes 1_{I_n}\}$ is a complete set in $\otimes^n\Lambda_2(R)$, we have

$$\Lambda_2(\otimes^n R) \cong \otimes^n \Lambda_2(R) .$$

$\lambda_2(\otimes^n R)$ can be defined in a similar manner. Let $S_J^{(n)}$ be the set of functions of the form

$$K(t_1,\ldots,t_n) = \sum_{k=1}^{N} f_1^{(k)}(t_1) \ldots f_n^{(k)}(t_n)$$

where the f's belong to S_J. $S_J^{(n)}$ is a linear space. Define on $S_J^{(n)} \times S_J^{(n)}$ the function

$$\langle K,G\rangle = R - \iint \circ \ldots \circ \iint K(t_1,\ldots,t_n)G(s_1,\ldots,s_n)R(t_1,s_1) \circ \ldots$$
$$\circ\, R(t_n,s_n)dt_1ds_1 \ldots dt_nds_n$$

and identify K with G if $\langle K-G,K-G\rangle=0$. With the observation that for $f_i,g_j \in S_J$,

$$\langle f_1(t_1)\ldots f_n(t_n),\ g_1(t_1)\ldots g_n(t_n)\rangle$$
$$= R - \iint f_1(t)g_1(s)R(t,s)dtds \circ \ldots \circ R - \iint f_n(t)g_n(s)R(t,s)dtds$$
$$= \langle f_1,g_1\rangle_{\lambda_2(R)} \cdots \langle f_n,g_n\rangle_{\lambda_2(R)}$$
$$= \langle f_1\otimes\ldots\otimes f_n,\ g_1\otimes\ldots\otimes g_n\rangle_{\otimes^n\lambda_2(R)} ,$$

and with the fact that $\{f_1\otimes\ldots\otimes f_n\}$ is a complete set in $\otimes^n\lambda_2(R)$, we conclude that $(S_J^{(n)}, \langle\circ,\circ\rangle)$ is an inner product space and that its completion, which is denoted by $\lambda_2(\otimes^n R)$, is isomorphic to $\otimes^n\lambda_2(R)$.

As for the spaces $\Lambda_2(R)$ and $\lambda_2(R)$ we will treat elements of $\Lambda_2(\otimes^n R)$ and $\lambda_2(\otimes^n R)$ as "formal" functions and we will write the inner product in a formal integral form. As before, under some conditions, elements of $\Lambda_2(\otimes^n R)$ and $\lambda_2(\otimes^n R)$ will be representable by functions on T^n in the corresponding sense and in this case we will identify the elements of Λ_2 and λ_2 with the functions (see Theorem 4 and Corollary 6). The important point here is that we have identified the abstract tensor product spaces $\otimes^n\Lambda_2(R)$ and $\otimes^n\lambda_2(R)$ with the (nearly) function spaces $\Lambda_2(\otimes^n R)$ and $\lambda_2(\otimes^n R)$. From now on we will make no distinction between $\otimes^n\Lambda_2(R)$ and $\Lambda_2(\otimes^n R)$, and between $\otimes^n\lambda_2(R)$ and $\lambda_2(\otimes^n R)$.

Let R be of bounded variation on every finite domain of $T\times T$ and let $L_I^{(n)}$ be the set of all measurable functions K on T^n such that the following Lebesgue integrals are finite

$$\iint \cdots \iint |K(t_1,\ldots,t_n)K(s_1,\ldots,s_n)| \; d^2|R|(t_1,s_1) \ldots d^2|R|(t_n,s_n)$$

$$\iint \cdots \iint |K(t_1,\ldots,t_n)| \; 1_{I_1\times\ldots\times I_n}(s_1,\ldots,s_n)d^2|R|(t_1,s_1)\ldots d^2|R|(t_n,s_n)$$

for all bounded half open intervals $I_1,\ldots,I_n \subset T$. The following theorem can be proven like Theorem 1 and thus its proof is omitted.

THEOREM 4. *Let* $R(t,s)$ *be of bounded variation on every finite domain of* $T\times T$. *Then* $L_I^{(n)}$ *is a dense subset of* $\Lambda_2(\otimes^n R)$. *Also if* $K_1,K_2 \in L_I^{(n)}$ *and* $\iint \cdots \iint |K_1(t_1,\ldots,t_n)K_2(t_1,\ldots,t_n)|d^2|R|(t,s)\ldots d^2|R|(t_n,s_n) < \infty$, *then*

$$<K_1,K_2> = \iint \cdots \iint K_1(t_1,\ldots,t_n)K_2(s_1,\ldots,s_n)d^2R(t_1,s_1)\ldots d^2R(t_n,s_n) \; .$$

COROLLARY 5. $L_2(T^p,d^pt) = \otimes^p L_2(T,dt)$.

PROOF. Taking $R(t,s)$ to be the covariance function of a Wiener process, Theorem 4 implies $L_2(T^p,d^pt) \subset \otimes^p L_2(T,dt)$. On the other hand, $\otimes^p L_2(T,dt) \subset L_2(T^p,d^pt)$ since the map $f_1\otimes\ldots\otimes f_p \longmapsto f_1(t_1)\ldots f_p(t_p)$ preserves inner products. Q.E.D.

THEOREM 6. *If* $R(t,s)$ *is continuous on* $T\times T$, *then* $\lambda_2(\otimes^n R) = \Lambda_2(\otimes^n T)$.

PROOF. This follows immediately from the facts that the set $\{f_1\otimes\ldots\otimes f_n,\ f_i\in S_J\}$ is complete in both $\lambda_2(\otimes^n R)$ and $\Lambda_2(\otimes^n \Gamma)$, and that the two inner products are identical on this set. Q.E.D.

Let $L_J^{(n)}$ be the set of all measurable functions K on T^n such that the following Lebesgue integrals are finite

$$\iint \cdots \iint |K(t_1,\ldots,t_n)K(s_1,\ldots,s_n)R(t_1,s_1)\ldots R(t_n,s_n)|\ dt_1ds_1\ldots dt_nds_n$$

$$\iint \cdots \iint |K(t_1,\ldots,t_n)|\ 1_{I_1\times\ldots\times I_n}(s_1,\ldots,s_n)\ |R(t_1,s_1)\ldots R(t_n,s_n)|$$

$$dt_1ds_1\ldots dt_nds_n$$

for all bounded half open intervals $I_1,\ldots,I_n \subset T$. With the usual corresponding convention the following is a corollary of Theorems 4 and 5.

<u>COROLLARY 7.</u> *Let* $R(t,s)$ *be continuous on* $T\times T$. *Then* $L_J^{(n)}$ *is a dense subset of* $\lambda_2(\otimes^n R)$. *Also if* $K_1,K_2 \in L_J^{(n)}$ *and*
$$\iint \cdots \iint |K_1(t_1,\ldots,t_n)K_2(s_1,\ldots,s_n)R(t_1,s_1)\ldots R(t_n,s_n)|dt_1ds_1\ldots dt_nds_n<\infty,$$
then

$$<K_1,K_2> = \iint \cdots \iint K_1(t_1,\ldots,t_n)K_2(s_1,\ldots,s_n)R(t_1,s_1)\ldots R(t_n,s_n)$$

$$dt_1ds_1\ldots dt_nds_n\ .$$

Finally let us consider the symmetric tensor products $\hat{\otimes}^n\Lambda_2(R)$ and $\hat{\otimes}^n\lambda_2(R)$.

For $K \in \Lambda_2(\otimes^n R)$ define

$$\tilde{K}(t_1,\ldots,t_n) = \frac{1}{n!}\sum_{\pi} K(t_{\pi_1},\ldots,t_{\pi_n})$$

where the sum is over all permutations $\pi = (\pi_1,\ldots,\pi_n)$ of $(1,\ldots,n)$, and $\tilde{K}$ is called the <u>*symmetric version of* K</u>. $\tilde{K}$ is well-defined since K is a "function". If $K=\tilde{K}$ then K is said to be a <u>*symmetric*</u> function. Let $\Lambda_2(\hat{\otimes}^n R)$ be the subspace of all symmetric functions in $\Lambda_2(\otimes^n R)$. Then it is easy to show that $\Lambda_2(\hat{\otimes}^n R)$ is a Hilbert space and $\hat{\otimes}^n\Lambda_2(R) \cong \Lambda_2(\hat{\otimes}^n R)$ under the correspondence

$f_1\hat{\otimes}\cdots\hat{\otimes}f_n \longleftrightarrow (f_1(t_1)\cdots f_n(t_n))^{\sim}$. Similarly, let $\lambda_2(\hat{\otimes}^n R)$ be the subspace of all symmetric functions in $\lambda_2(\otimes^n R)$. Then we can show $\hat{\otimes}^n\lambda_2(R) \cong \lambda_2(\hat{\otimes}^n R)$ (under the natural correspondence). As before, we shall hereon identify $\hat{\otimes}^n\Lambda_2(R)$ with $\Lambda_2(\hat{\otimes}^n R)$ and $\hat{\otimes}^n\lambda_2(R)$ with $\lambda_2(\hat{\otimes}^n R)$.

3. Fourier Transform on $\lambda_2(\otimes^n R)$

Consider a continuous stationary covariance function $R(t,s)$, $t,s \in \mathbb{R}$. By Bochner's theorem we have

$$R(t,s) = \int e^{i(t-s)\lambda}\, d\mu(\lambda)$$

with μ a finite measure on the Borel sets of $\mathbb{R}$. For $f \in L_1(\mathbb{R}^n)$ let

$$\hat{f}(\lambda_1,\ldots,\lambda_n) = \int\cdots\int e^{i(t_1\lambda_1+\cdots+t_n\lambda_n)} f(t_1,\ldots,t_n)dt_1\cdots dt_n$$

be its Fourier transform and recall that $\hat{f} \in C_0(\mathbb{R}^n)$, the set of all continuous functions vanishing at infinity. We shall define the Fourier transform of every $f \in \lambda_2(\otimes^n R)$ and for this it is convenient to complexify the space $\lambda_2(\otimes^n R)$.

LEMMA 8. (i) $L_1(\mathbb{R}^n)$ *is a dense subspace of* $\lambda_2(\otimes^n R)$.
(ii) *If* $f,g \in L_1(\mathbb{R}^n)$ *then* $\hat{f},\hat{g} \in L_2(\mathbb{R}^n,\mu^n)$ *and*

$$\langle f,g\rangle_{\lambda_2(\otimes^n R)} = \langle \hat{f},\hat{g}\rangle_{L_2(\mathbb{R}^n,\mu^n)} .$$

PROOF. (i) Let $f \in L_1(\mathbb{R}^n)$. Then

$$\iint\cdots\iint |f(t_1,\ldots,t_n)f(s_1,\ldots,s_n)R(t_1,s_1)\cdots R(t_n,s_n)|dt_1ds_1\cdots dt_nds_n$$
$$\le \mu(\mathbb{R})^n\, ||f||^2_{L_1(\mathbb{R}^n)} < \infty .$$

Similarly the second condition in the definition of $L_J^{(n)}$ is verified and thus $f \in L_J^{(n)}$. By Corollary 7, $f \in \lambda_2(\otimes^n R)$. Since R is continuous, by Theorem 6, $\lambda_2(\otimes^n R) = \Lambda_2(\otimes^n \Gamma)$. Since the set of all step functions is dense in $\lambda_2(\otimes^n \Gamma)$ it follows that $L_1(\mathbb{R}^n)$ is dense in $\lambda_2(\otimes^n R)$.

(ii) Let $f,g \in L_1(\mathbb{R}^n)$. Then $|\hat{f}| \leq ||f||_{L_1(\mathbb{R}^n)}$ implies

$$\int \cdots \int |\hat{f}(\lambda_1,\ldots,\lambda_n)|^2 \, d\mu(\lambda_1)\ldots d\mu(\lambda_n) \leq ||f||^2_{L_1(\mathbb{R}^n)} \mu(\mathbb{R})^n < \infty ,$$

and thus $\hat{f} \in L_2(\mathbb{R}^n,\mu^n)$. Since by (i), $f,g \in L_J^{(n)}$ and

$$\iint \cdots \iint |f(t_1,\ldots,t_n)g(s_1,\ldots,s_n)R(t_1,s_1)\ldots R(t_n,s_n)| dt_1 ds_1 \ldots dt_n ds_n$$

$$\leq \mu(\mathbb{R})^n \, ||f||_{L_1(\mathbb{R}^n)} \, ||g||_{L_1(\mathbb{R}^n)} < \infty ,$$

it follows by Corollary 7 that

$$\langle f,g\rangle = \iint \cdots \iint f(t_1,\ldots,t_n)g(s_1,\ldots,s_n)R(t_1,s_1)\ldots R(t_n,s_n)dt_1 ds_1 \ldots dt_n ds_n .$$

Substituting R and interchanging the order of integration by Fubini's theorem we obtain $\langle f,g\rangle_{\lambda_2(\otimes^n R)} = \langle \hat{f},\hat{g}\rangle_{L_2(\mathbb{R}^n,\mu^n)}$. Q.E.D.

<u>THEOREM 9.</u> *The map* $F: L_1(\mathbb{R}^n) \to L_2(\mathbb{R}^n,\mu^n): f \mapsto \hat{f}$ *can be extended to an isomorphism from* $\lambda_2(\otimes^n R)$ *onto* $L_2(\mathbb{R}^n,\mu^n)$.

The extended map is denoted again by F *and is called the* *<u>Fourier transform</u> on* $\lambda_2(\otimes^n R)$. *We also write* $\hat{f}$ *for* $F(f)$, $f \in \lambda_2(\otimes^n R)$.

<u>PROOF.</u> By Lemma 8, F preserves inner products and thus it can be extended to an isomorphism F on $\lambda_2(\otimes^n R)$ to a closed subspace of $L_2(\mathbb{R}^n,\mu^n)$. We now show that F is onto $L_2(\mathbb{R}^n,\mu^n)$.

Consider $A = \{\hat{f},\ f \in L_2(\mathbb{R}^n)\}$ which is a subalgebra of $C_0(\mathbb{R}^n)$. It is well known that A is self-adjoint, separates points and vanishes at no point. Thus by the Stone-Weinerstrass theorem any $f \in C_0(\mathbb{R}^n)$ is the uniform limit of functions in A. Clearly $C_0(\mathbb{R}^n)$ is dense in $L_2(\mathbb{R}^n,\mu^n)$ (since μ is a finite measure) and hence A is dense in $L_2(\mathbb{R}^n,\mu^n)$. This implies F is onto. Q.E.D.

Heuristically, if we let $R(t,s)$ be the covariance function of the Gaussian white noise $\mathring{W}$ (the derivative of Wiener process W) then $R(t-s)$ is a delta function $\delta_0(t-s)$ and μ is the Lebesgue measure. In this case

$$\lambda_2(R) = \Lambda_2(\Gamma) = L_2(dt)$$

(where $\Gamma(t,s) = EW_tW_s$) and

$$\lambda_2(\otimes^n R) \cong \otimes^n \lambda_2(R) \cong \otimes^n \Lambda_2(\Gamma) \cong \otimes^n L_2(dt) \cong L_2(d^n t) .$$

and thus F becomes the ordinary Fourier transform on $L_2(d^n t)$.

Denote the translation (by $\tau=(\tau_1,\ldots,\tau_n)$) of a function $f(t_1,\ldots,t_n)$ by $f_\tau(t_1,\ldots,t_n) = f(t_1+\tau_1,\ldots,t_n+\tau_n)$. Then we have

<u>THEOREM 10.</u> $\lambda_2(\otimes^n R)$ *is invariant under translations. And for* $f,g \in \lambda_2(\otimes^n R)$, $\langle f_\tau,g_\sigma\rangle = \langle f,g\rangle$; $\hat{f}_\tau(\lambda_1,\ldots,\lambda_n) = e^{-i(\lambda_1\tau_1+\ldots+\lambda_n\tau_n)} \cdot \hat{f}(\lambda_1,\ldots,\lambda_n)$, $\tau=(\tau_1,\ldots,\tau_n)$.

<u>PROOF.</u> Since R is stationary, all these assertions hold for $f \in L_1(\mathbb{R}^n)$. Hence they hold for all $f \in \lambda_2(\otimes^n R)$ by continuity (Lemma 8).

III. NONLINEAR SPACES AND MULTIPLE WIENER INTEGRALS FOR SPHERICALLY INVARIANT PROCESSES

The main effort of this chapter is to obtain an orthogonal decomposition of the nonlinear space of a *spherically invariant process* (SIP) (Theorem 12) and an integral representation for its L_2-functionals (Theorem 21). The integral representation is expressed in terms of *multiple Wiener integrals* (MWI's) which are defined in Section 3. The general properties of SIP's are studied in Section 1.

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 a SIP if all r.v.'s in $H(X-m)$ having the same variance have the same distribution. A SIP 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 $X_{t_1},\dots,X_{t_k}$ $(t_1,\dots,t_k\in T)$ is given by

$$(1)\quad \phi(u_1,\dots,u_k) = \int e^{-\frac{\alpha}{2}\ \Sigma\ R(t_i,t_j)\left(u_i-m(t_i)\right)\left(u_j-m(t_j)\right)}\,dF(\alpha)$$

(Vershik [1964] and Nagornyĭ [1974]). Such a SIP determined by $m(t)$, $R(t,s)$ and $F(\alpha)$ will be denoted, in short, by $SIP(m,R;F)$. A probability measure P 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 $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), P)$ is a SIP. A SIM induced by a $SIP(m,R;F)$ will be denoted by $SIM(m,R;F)$.

THEOREM 1. *Let* P *be a probability measure on* $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ *such that the characteristic functions of its finite dimensional distributions are given by (1), and for each* $\alpha \geq 0$ *let* P_α *be a Gaussian measure on* $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ *with mean* $m(t)$ *and covariance function* $\alpha R(t,s)$. *Then*

(i) $P(E) = \int P_\alpha(E)dF(\alpha) \quad \forall\, E \in \mathcal{B}(\mathbb{R}^T)$.

Furthermore, for any measurable function θ *on* $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), P)$ *that is nonnegative or integrable we have*

(ii) $\mathcal{E}\,\theta = \int \mathcal{E}_\alpha \theta dF(\alpha)$

where $\mathcal{E}\,\theta = \int \theta dP$ *and* $\mathcal{E}_\alpha \theta = \int \theta dP_\alpha$.

PROOF. (i) For each fixed cylinder set E, $P_\alpha(E)$ is a measurable function of α; and the family of sets E such that $P_\alpha(E)$ is α-measurable is a σ-field. Therefore $P_\alpha(E)$ is α-measurable for every $E \in \mathcal{B}(\mathbb{R}^T)$. Thus $\int P_\alpha(E)dF(\alpha)$ is well-defined and is easily checked to be a probability measure. The characteristic functions of its finite dimensional distributions is of course given by (1). Since the characteristic functions of finite dimensional distributions uniquely determine a probability measure on $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$, we have $P(E) = \int P_\alpha(E)dF(\alpha)$ $\forall\, E \in \mathcal{B}(\mathbb{R}^T)$.

(ii). This holds for $\theta = 1_E$ $(E \in \mathcal{B}(\mathbb{R}^T))$ by (i). Hence it holds for θ simple function, and thus for θ nonnegative by considering a sequence of simple functions increasing to θ. For θ integrable consider the positive and negative part separately. Q.E.D.

REMARK. Theorem 1 (i) was proven by Gualtierotti [1974] for P a SIM on a separable Hilbert space.

COROLLARY 2. P *is a* $\mathrm{SIM}(m,R;F)$ *on* $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ *if and only if*

$$P(E) = \int P_\alpha(E)\,dF(\alpha) \quad \forall\, E \in \mathcal{B}(\mathbb{R}^T)$$

where each P_α *is a Gaussian measure with mean* m *and covariance function* αR, *and*

$$\alpha_1 = \int \alpha\, dF(\alpha) < \infty .$$

PROOF. Let $X=(X_t,\ t\in T)$ be the coordinate process on $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), P)$ such that the characteristic functions of its finite dimensional distributions are given by (1). Then X is of second order if and only if $\alpha_1<\infty$, since by Theorem 1 (ii)

$$E\, X_t^2 = \int E_\alpha X_t^2\, dF(\alpha) = \int [\alpha R(t,t)+m^2(t)]\,dF(\alpha) .$$

The result now follows from Theorem 1 (i). Q.E.D.

It follows from Corollary 2 that if X is a coordinate $\mathrm{SIP}(m,R;F)$ then under each P_α, X is a Gaussian process with mean $m(t)$ and covariance function $\alpha R(t,s)$. The following lemma partially generalizes this fact and it is central to the study of the nonlinear spaces of SIP's.

LEMMA 3. *Suppose* X *is a zero mean coordinate* SIP *and suppose* $\{\xi_n\}$ *is a sequence in* $H(X)$. *Then we may take particular versions of* *each* ξ_n *such that under each* P_α *the r.v.'s* (ξ_n) *are jointly Gaussian with zero mean and covariance*

$$\text{(i)}\quad E_\alpha \xi_i \xi_j = \frac{\alpha}{\alpha_1} E\xi_i\xi_j = \alpha E_1 \xi_i \xi_j \ ,$$

and, moreover, for any measurable function θ *on* $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ *we have*

$$\text{(ii)}\quad E_\alpha \theta(\xi_1, \xi_2, \ldots) = E_1 \theta(\sqrt{\alpha}\ \xi_1, \sqrt{\alpha}\ \xi_2, \ldots)$$

whenever one of these expectations exists.

PROOF. For each r.v. $\xi_n(x)$, $x \in \mathbb{R}^T$, there is a sequence $\ell_i^{(n)}(x)$ of finite linear combinations of x such that $\xi_n(x) = \lim \ell_i^{(n)}(x)$ a.e. [P]. Let $C_n = \{x \in \mathbb{R}^T,\ \ell_i^{(n)}(x) \to \circ \text{ as } i \to \infty\}$ and let $C = \cap C_n$. C is clearly a measurable linear space with probability 1. Now take each ξ_n to be $\lim \ell_i^{(n)}(x)$ for $x \in C$ and 0 for $x \notin C$. Since by Theorem 1, $1 = P(C) = \int P_\alpha(C)\,dF(\alpha)$, there is $\alpha_0 \geq 0$ such that $P_{\alpha_0}(C) = 1$. We may assume $\alpha_0 > 0$, because otherwise $dF(\alpha)$ will concentrate at $\alpha = 0$ and in this case the lemma obviously holds by taking each ξ_n to be identically 0. Bring in a zero mean Gaussian process Y on some probability space $(\Omega_0, \mathcal{B}_0, Q)$ with covariance R. Then $\forall\ \alpha \geq 0$

$$P_\alpha(C) = Q(\sqrt{\alpha}\ Y \in C) = Q(\sqrt{\alpha_0}\ Y \in C) = P_{\alpha_0}(C) = 1.$$

(Note that C is a linear space.) Thus $\xi_n = \lim \ell_i^{(n)}$ a.e. $[P_\alpha]$ $\forall\ \alpha \geq 0$. But under each P_α, $\{\ell_i^{(n)},\ n \geq 1,\ i \geq 1\}$ is a Gaussian family. It follows $\{\xi_n\}$ is jointly Gaussian under each P_α.

Since (i) is implied by (ii), we need to prove (ii) only. (ii) holds for $\theta = 1_E$, $E \in \mathcal{B}(\mathbb{R}^\infty)$, since as before $P_\alpha[(\xi_1, \xi_2, \ldots) \in E] = P_1[(\sqrt{\alpha}\ \xi_1, \sqrt{\alpha}\ \xi_2, \ldots) \in E]$ $\forall\ \alpha \geq 0$. Consider now a nonnegative θ. Then θ is the increasing limit of a sequence of simple functions ϕ_n and by the Monotone Convergence Theorem we have

$$
\begin{aligned}
E_\alpha \theta(\xi_1, \xi_2, \ldots) &= \lim E_\alpha \phi_n(\xi_1, \xi_2, \ldots) \\
&= \lim E_1 \phi_n(\sqrt{\alpha}\ \xi_1,\ \sqrt{\alpha}\ \xi_2,\ \ldots) \\
&= E_1 \theta(\sqrt{\alpha}\ \xi_1,\ \sqrt{\alpha}\ \xi_2,\ \ldots) \ .
\end{aligned}
$$

(ii) is now evident. Q.E.D.

We now introduce a r.v. A which will play an important role in studying the nonlinear space of a SIP as well as in defining the multiple Wiener integrals for it. In the following discussion we assume that X is a coordinate $\text{SIP}(0,R;F)$ and X is *nondegenerate*, i.e., $H(X)$ does not have finite dimension.

Pick an orthogonal sequence $\{\xi_n\}$ from $H(X)$ with $E\xi_n^2 = \alpha_1$. Then by Lemma 3 we may assume that under each P_α $\{\xi_n\}$ is a sequence of independent zero mean Gaussian r.v.'s with $E_\alpha \xi_n^2 = \alpha$. Define

$$
A_n = \frac{1}{n} \sum_1^n \xi_i \ .
$$

By the Law of Large Numbers, we have

$$
\lim A_n = \alpha \quad \text{a.e. } [P_\alpha] \ .
$$

Let $C^* = \{x \in \mathbb{R}^T,\ A_n(x) \to \cdot\}$. Then

$$
P(C^*) = \int P_\alpha(C^*)\, dF(\alpha) = \int dF(\alpha) = 1 \ .
$$

Thus A_n converges a.e. $[P]$ and its limit is denoted by A. We remark that $A = \alpha$ a.e. $[P_\alpha]$, and A has distribution function F since

$$
\begin{aligned}
P(A \le a) &= \int P_\alpha(A \le a)\, dF(\alpha) \\
&= \int 1_{[0,a]}(\alpha)\, dF(\alpha) = F(a) \ .
\end{aligned}
$$

If we put $\Omega_\alpha = \{x \in \mathbb{R}^T, A(x)=\alpha\}$ then $P_\alpha(\Omega_\alpha)=1$, and thus the probability measures P_α, $\alpha \geq 0$, are mutually singular.

In the following we will introduce an alternative way of defining the r.v. A and we will give an interesting example of a sample continuous martingale whose family of σ-fields is not continuous.

Consider the same sequence $\{\xi_n\}$ as above. Let

$$W_t = \sum \xi_n \int_0^t e_n(\tau)d\tau \ , \quad 0 \leq t \leq 1 \ ,$$

where $\{e_n\}$ is a CONS in $L_2[0,1]$. This sum, under each P_α, converges a.e. and in mean square and $W = (W_t, 0 \leq t \leq 1)$ is a Wiener process with variance parameter α (e.g. see Shepp [1966]). Define

$$A_n(t) = \sum_{j=1}^{2^n-1} \left[W\left(\frac{j+1}{2^n} t\right) - W\left(\frac{j}{2^n} t\right) \right]^2 .$$

By a theorem of Lévy [Doob 1953]

$$\lim A_n(t) = \alpha t \quad \text{a.e. } [P_\alpha] \ .$$

Applying the same argument as before we can show that for each t $A_n(t)$ converges a.e. $[P]$ and its limit is the quadratic variation of $\{W_s, 0 \leq s \leq t\}$ denoted by $A(t)$. The r.v. A is defined to be $A(1)$. Since $A=\alpha$ a.e. $[P_\alpha]$, we have $A(t)=At$.

Using the fact that W is a Wiener process with variance parameter α under each P_α, it is easy to see that W is a sample continuous martingale under P. Let $\mathcal{B}_t = \mathcal{B}(W_s, 0 \leq s \leq t)$, $0 \leq t \leq 1$, we will show that $\mathcal{B}_t$ is not continuous at $t=0$.

By definition $A(t)=At$ is $\mathcal{B}_t$-measurable and hence A is $\mathcal{B}_t$-measurable for every $t>0$. Suppose A is not a constant (i.e., $dF(\alpha)$ does not concentrate at one point). Then $\bigcap_{t>0} \mathcal{B}_t$ is nontrivial. But

$\mathcal{B}_0$ is trivial because $W_0=0$ a.e. [P]. Therefore $\mathcal{B}_t$ is not continuous at $t=0$.

2. Nonlinear Spaces

Consider a second order process $X = \{X_t, t\in T\}$. We shall briefly recapitulate some results on the nonlinear space $L_2(X)$ which will be useful to us later. We may assume that X has zero mean.

Suppose that every $\xi\in H(X)$ has all moments finite. Let $\mathcal{P}$ be the linear space of all polynomials in elements of $H(X)$ and let $\mathcal{P}_p$ $(p\geq 0)$ be the linear space of all polynomials in $\mathcal{P}$ of degree at most p; hence $\mathcal{P}_0$ is the set of all constants. Let $Q_0=P_0$ and for $p\geq 1$ let Q_p be the set of all polynomials in $\mathcal{P}_p$ orthogonal to $\mathcal{P}_{p-1}$. Denote by $\overline{Q}_p$ the closure of Q_p in $L_2(X)$. Q_p is called the p-th *polynomial chaos* and $\overline{Q}_p$ is called the p-th *homogeneous chaos*.

The following theorem is well known.

THEOREM 4. [Neveu 1968] *If* $e^{\xi} \in L_2(X)$ *for every* $\xi \in H(X)$, *then* $\{e^{\xi}, \xi\in H(X)\}$ *is a complete set in* $L_2(X)$. *Furthermore if* $e^{|\xi|} \in L_2(X)$ *for every* $\xi \in H(X)$, *then* $\mathcal{P}$ *is a dense subspace of* $L_2(X)$ *and hence*

$$L_2(X) = \bigoplus_{p\geq 0} \overline{Q}_p \ .$$

PROOF. Let $\eta \in L_2(X)$ such that $E\eta e^{\xi} = 0 \ \forall \ \xi \in H(X)$. We will show that $\eta=0$ a.e. There exists a sequence $\{\xi_n\}$ in $H(X)$ such that η is $\{\xi_n\}$-measurable [Doob 1953]. Since η is integrable we have by the standard Martingale Convergence Theorem $\eta = \lim E(\eta|\xi_1,\ldots,\xi_n)$ a.e. $E(\eta|\xi_1,\ldots,\xi_n)$ is a Borel measurable function of $\xi_1,\ldots,\xi_n$ and belongs to $L_2(X)$; denote it by $g(\xi_1,\ldots,\xi_n)$. Then

$$Eg(\xi_1,\ldots,\xi_n)e^{t_1\xi_1+\ldots+t_n\xi_n} = E\eta e^{t_1\xi_1+\ldots+t_n\xi_n} = 0$$

$\forall\ t_1,\ldots,t_n \in \mathbb{R}$ which implies $g(\xi_1,\ldots,\xi_n)=0$ a.e. since $\{e^{t_1\xi_1+\ldots+t_n\xi_n},\ t_1,\ldots,t_n \in \mathbb{R}\}$ is complete in $L_2(\xi_1,\ldots,\xi_n)$ (by an application of the Stone-Weierstrass Theorem). Since $E(\eta|\xi_1,\ldots,\xi_n)=0$ a.e. for all n it follows that $\eta=0$ a.e. Thus we conclude that $\{e^{\xi},\ \xi\in H(X)\}$ is complete.

Suppose $e^{|\xi|} \in L_2(X)\ \ \forall\ \xi \in H(X)$. Then we can apply the Bounded Convergence Theorem to $\text{l.i.m.} \sum_{p=1}^{n} \frac{\xi^p}{p!} = e^{\xi}$ and conclude that $e^{\xi} \in \overline{P}$. Thus $L_2(X) = \overline{P}$. Q.E.D.

We now proceed to study the nonlinear space $L_2(X)$ of a nondegenerate SIP(0,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.

$$\int e^{\alpha t}\, dF(\alpha) < \infty \quad \forall\ t \in \mathbb{R}\ .$$

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

$$\begin{aligned} E\ \xi^p &= \int E_\alpha \xi^p\, dF(\alpha) \\ &= \int \alpha^{\frac{p}{2}} E_1\xi^p\, dF(\alpha) \qquad\qquad \text{(by Lemma 3)} \\ &= \alpha_{\frac{p}{2}}\ E_1\xi^p < \infty \end{aligned}$$

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

<u>COROLLARY 5.</u> *If* X *is a* SIP(0,R;F) *satisfying (M) then*

$$L_2(X) = \bigoplus_{p\geq 0} \overline{Q}_p\ .$$

PROOF. It suffices to show that $e^{|\xi|}$ is integrable for every $\xi \in H(X)$. We may assume that X is a coordinate process. Also we may assume by Lemma 3 that under each P_α ξ is a zero mean Gaussian variable with variance $\frac{\alpha}{\alpha_1} E\xi^2$. Then

$$E_\alpha e^{|\xi|} = \frac{2}{\sqrt{2\pi\sigma^2}} \int_0^\infty e^{x - \frac{x^2}{2\sigma^2}} dx \qquad (\sigma^2 = \frac{\alpha}{\alpha_1} E\xi^2)$$

$$= \frac{2e^{\sigma^2}}{\sqrt{2\pi\sigma^2}} \int_{-\sigma^2}^\infty e^{\frac{x^2}{2\sigma^2}} dx$$

$$\leq 2e^{\sigma^2} .$$

Thus

$$E\, e^{|\xi|} = \int E_\alpha e^{|\xi|} \, dF(\alpha)$$

$$< 2 \int e^{\frac{\alpha}{\alpha_1} E\xi^2} dF(\alpha) < \infty .$$

Q.E.D.

REMARK. The use of Lemma 3 in the proof of Corollary 5 is typical. We will hereafter use Lemma 3 without comment.

COROLLARY 6. [Wiener 1968] *If* X *is a zero mean Gaussian process then*

$$L_2(X) = \bigoplus_{p \geq 0} \overline{Q}_p .$$

PROOF. Note that a Gaussian process is a SIP with $dF(\alpha)$ concentrated at $\alpha=1$.

<u>LEMMA 7.</u> *Suppose* H *is a Hilbert space. Then* $\{\xi^{\hat{\otimes} p};\ \xi\in H\}$ *is complete in* $H^{\hat{\otimes} p}$.

<u>PROOF.</u> Let $\{\xi_\gamma,\ \gamma\in\Gamma\}$ be a CONS in H. Then $\{\xi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\cdots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes} p_k}:\ \gamma_1,\dots,\gamma_k\in\Gamma,\ p_1+\dots+p_k=p\}$ is a CONS in $H^{\hat{\otimes} p}$. Let $\eta\in H^{\hat{\otimes} p}$ and $\langle\eta,\xi^{\hat{\otimes} p}\rangle=0\ \ \forall\ \xi\in H$. Write $\eta=\sum\sum a_{\gamma_1,\dots,\gamma_k}^{p_1,\dots,p_k}\,\xi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\cdots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes} p_k}$.

Consider the inner product of η and $s_1\xi_{\gamma_1}+\dots+s_k\xi_{\gamma_k}$, $s_1,\dots,s_k\in\mathbb{R}$; we have

$$0=\langle\eta,\ (s_1\xi_{\gamma_1}+\dots+s_k\xi_{\gamma_k})^{\hat{\otimes} p}\rangle=\sum_{p_1+\dots+p_k=p} a_{\gamma_1\cdots\gamma_k}^{p_1\cdots p_k}\, s_1^{p_1}\cdots s_k^{p_k}$$

$$\forall\ s_1,\dots,s_k\in\mathbb{R}\ .$$

Since this equation has infinite many solutions, it follows $a_{\gamma_1\cdots\gamma_k}^{p_1\cdots p_k}=0$ and thus $\eta=0$. Therefore $\{\xi^{\hat{\otimes} p},\ \xi\in H\}$ is complete. Q.E.D.

Our first step in studying the nonlinear space of a SIP is to investigate to what extent it is determined by the structure of the linear space.

<u>THEOREM 8.</u> *If* X *is a nondegenerate coordinate* SIP(0,R;F) *satisfying (M) then there exists a unique isomorphism* Φ *from* $\bigoplus_{p\geq 0} H^{\hat{\otimes} p}(X)$ *onto a subspace of* $L_2(X)$ *such that*

$$\Phi(e^{\hat{\otimes}\xi})=e^{\xi-\frac{A}{2\alpha_1}E\xi^2}$$

where $e^{\hat{\otimes}\xi}=\sum\left(\frac{\alpha_p}{p!\,\alpha_1^p}\right)^{\frac{1}{2}}\xi^{\hat{\otimes} p}$, $\xi\in H(X)$.

Moreover, Φ *is an isomorphism between* $\bigoplus_{p\geq 0} H^{\hat{\otimes} p}(X)$ *and* $L_2(X)$ *if and only if* dF *is concentrated at one point.*

PROOF. The convergence of the series defining $e^{\hat{\otimes}\xi}$ results from the following:

$$||\xi||^2 = E\xi^2 = \alpha_1 E_1 \xi^2 ,$$

and

$$\begin{aligned}
\sum \frac{\alpha_p}{p!\alpha_1^p} ||\xi^{\hat{\otimes} p}||^2 &= \sum \frac{\alpha_p}{p!\alpha_1^p} ||\xi||^{2p} \\
&= \sum \int \frac{1}{p!} (\alpha E_1 \xi^2)^p \, dF(\alpha) \\
&= \int \Big[\sum \frac{1}{p!} (\alpha E_1 \xi^2)^p \Big] \, dF(\alpha) \\
&\qquad\qquad \text{(by Fubini's Theorem)} \\
&= \int e^{\alpha E_1 \xi^2} \, dF(\alpha) \\
&< \infty \quad \text{(by (M))}.
\end{aligned}$$

Thus $e^{\hat{\otimes}\xi}$ is well-defined. Next we show that the set $\{e^{\hat{\otimes}\xi}, \xi \in H(X)\}$ is complete in $\oplus H^{\hat{\otimes}p}(X)$. Let $\eta \in \underset{p\geq 0}{\oplus} H^{\hat{\otimes}p}(X)$ be orthogonal to this set. We need show that $\eta=0$. Write $\eta = \sum \eta_p$, $\eta_p \in H^{\hat{\otimes}p}(X)$. Then

$$0 = \langle e^{\hat{\otimes}u\xi}, \eta \rangle_{\underset{p\geq 0}{\oplus} H^{\hat{\otimes}p}(X)} = \sum \left(\frac{\alpha_p}{p!\alpha_1^p}\right)^{\frac{1}{2}} u^p \langle \xi^{\hat{\otimes}p}, \eta_p \rangle_{H^{\hat{\otimes}p}(X)}$$

$\forall\, u \in \mathbb{R}$, $\forall\, \xi \in H(X)$, which implies $\eta_p \perp \xi^{\hat{\otimes}p}$ $\forall\, \xi \in H(X)$. By Lemma 7 $\eta_p = 0$ and thus $\eta=0$.

For $\xi, \eta \in H(X)$ we have

$$\begin{aligned}
\langle e^{\hat{\otimes}\xi}, e^{\hat{\otimes}\eta} \rangle_{\underset{p\geq 0}{\oplus} H^{\hat{\otimes}p}(X)} &= \sum \frac{\alpha_p}{p!\alpha_1^p} \langle \xi, \eta \rangle_{H(X)}^p \\
&= \int \sum \frac{1}{p!} \left(\frac{\alpha}{\alpha_1} E\xi\eta\right)^p dF(\alpha) \\
&= \int e^{\frac{\alpha}{\alpha_1} E\xi\eta} \, dF(\alpha) \ ;
\end{aligned}$$

on the other hand we have

$$< e^{\xi - \frac{A}{2\alpha_1} E\xi^2}, e^{\eta - \frac{A}{2\alpha_1} E\eta^2} > = \int E_\alpha\left(e^{\xi - \frac{1}{2} E_\alpha \xi^2} \circ e^{\eta - \frac{1}{2} E_\alpha \eta^2}\right) dF(\alpha)$$

$$= \int e^{E_\alpha \xi\eta} dF(\alpha)$$

$$= \int e^{\frac{\alpha}{\alpha_1} E\xi\eta} dF(\alpha) .$$

Thus Φ preserves the inner products and therefore it can be extended uniquely to the closed subspace spanned by $\{e^{\hat{\otimes}\xi}, \xi \epsilon H(X)\}$ which is indeed $\underset{p\geq 0}{\oplus} H^{\hat{\otimes} p}(X)$. The first statement of the theorem is now proved.

Suppose $dF(\alpha)$ is concentrated at $\alpha=a$. Then $A=a$ a.e. and in this case $\{\Phi(e^{\hat{\otimes}\xi}) = e^{\xi - \frac{1}{2} E\xi^2}, \xi \epsilon H(X)\}$ is a complete set in $L_2(X)$ by Theorem 4. Thus Φ is an onto map and hence an isomorphism. Conversely, suppose dF is not concentrated at one point then there exists a nontrivial r.v. $f(A)$ with zero mean. Note that

$$E\ f(A) e^{\xi - \frac{A}{2\alpha_1} E\xi^2} = \int f(\alpha) E_\alpha e^{\xi - \frac{\alpha}{2\alpha_1} E\xi^2} dF(\alpha)$$

$$= \int f(\alpha) dF(\alpha) = 0.$$

It then follows that Φ is not onto. Q.E.D.

<u>COROLLARY 9.</u> [Neveu 1968] *If* X *is a zero mean Gaussian process then there exists a unique isomorphism from* $\underset{p\geq 0}{\oplus} H^{\hat{\otimes} p}(X)$ *onto* $L_2(X)$ *such that*

$$\Phi(e^{\hat{\otimes}\xi}) = e^{\xi - \frac{1}{2} E\xi^2}$$

where $e^{\hat{\otimes}\xi} = \sum (\frac{1}{p!})^{\frac{1}{2}} \xi^{\hat{\otimes} p}, \quad \xi \epsilon H(X)$.

PROOF. The omission of the assumption that X is a nondegenerate coordinate process needs explanation. Note that in the proof of Theorem 8 this assumption is only used to guarantee the existence of the r.v. A. But in the Gaussian case $dF(\alpha)$ concentrate at $\alpha=1$, and thus the r.v. A always exists and equals to 1. Therefore this assumption can be dropped. The assertion now follows from Theorem 8. Q.E.D.

By the isomorphism Φ in Theorem 8 we shall hereafter identify $\bigoplus_{p\geq 0} H^{\hat{\otimes} p}(X)$ with a subspace of $L_2(X)$.

THEOREM 10. *Suppose* X *is a nondegenerate coordinate* SIP(0,R;F) *satisfying (M), then for* $\xi\in H(X)$

$$\xi^{\hat{\otimes} p} = \left(\frac{\alpha_1^p}{p!\alpha_p}\right)^{\frac{1}{2}} H_{p,\frac{A}{\alpha_1}||\xi||^2}(\xi) \ .$$

More generally, for a finite orthogonal sequence $\xi_1,\ldots,\xi_k$ *in* H(X)

$$\xi_1^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_k^{\hat{\otimes} p_k} = \left(\frac{\alpha_1^p}{p!\alpha_p}\right)^{\frac{1}{2}} H_{p_1,\frac{A}{\alpha_1}||\xi_1||^2}(\xi_1) \ldots H_{p_k,\frac{A}{\alpha_1}||\xi_k||^2}(\xi_k) \ ,$$

$p=p_1+\ldots+p_k$; *in particular,*

$$\xi_1\hat{\otimes}\ldots\hat{\otimes}\xi_k = \left(\frac{\alpha_1^k}{k!\alpha_k}\right)^{\frac{1}{2}} \xi_1\ldots\xi_k \ .$$

PROOF. Since $\{\xi_1,\ldots,\xi_k\}$ is an orthogonal set in H(X), $\{\xi_1^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_k^{\hat{\otimes} p_k},\ p_1,\ldots,p_k\geq 0\}$ is an orthogonal set in $\bigoplus_{p\geq 0} H^{\hat{\otimes} p}(X) \subset L_2(X)$ and the following series converges in $L_2(X)$

$$e^{\hat{\otimes}\Sigma_1^k u_i\xi_i} = \sum_{p\geq 0}\left(\frac{\alpha_p}{p!\alpha_1^p}\right)^{\frac{1}{2}}\left(\sum_1^k u_i\xi_i\right)^{\hat{\otimes} p}$$

$$= \sum_{p\geq 0}\ \sum_{p_1+\ldots+p_k=p}\left(\frac{\alpha_p}{p!\alpha_1^p}\right)^{\frac{1}{2}} \frac{p!}{p_1!\ldots p_k!} u_1^{p_1}\ldots u_k^{p_k}\ \xi_1^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_k^{\hat{\otimes} p_k} \ .$$

On the other hand

$$e^{\hat{\otimes}\Sigma_1^k u_i \xi_i} = e^{\Sigma_1^k (u_i\xi_i - \frac{A}{2\alpha_1}u_i^2||\xi_i||^2)}$$

$$= \prod_1^k e^{u_i\xi_i - \frac{A}{2\alpha_1}u_i^2||\xi_i||^2}$$

$$= \prod_1^k \sum_{p_i\geq 0} \frac{u_i^{p_i}}{p_i!} H_{p_i, \frac{A}{\alpha_1}||\xi_i||^2}(\xi_i) .$$

The theorem now follows by comparing the coefficients in both developments. Q.E.D.

<u>COROLLARY 11.</u> [Neveu 1968] *Suppose* X *is a zero mean Gaussian process. Then for* $\xi \in H(X)$

$$\xi^{\hat{\otimes} p} = (\frac{1}{p!})^{\frac{1}{2}} H_{p,||\xi||^2}(\xi) .$$

More generally, for a finite orthogonal sequence $\xi_1,\ldots,\xi_k$ *in* $H(X)$

$$\xi_1^{\hat{\otimes} p_1} \hat{\otimes} \ldots \hat{\otimes} \xi_k^{\otimes p_k} = (\frac{1}{p!})^{\frac{1}{2}} H_{p_1,||\xi||^2}(\xi) \ldots H_{p_k,||\xi||^2}(\xi) ,$$

$p=p_1+\ldots+p_k$; *in particular,*

$$\xi_1 \hat{\otimes} \ldots \hat{\otimes} \xi_k = (\frac{1}{k!})^{\frac{1}{2}} \xi_1 \ldots \xi_k .$$

Before we prove the main theorem of this section we introduce a "continuity" condition on F:

(C_0) $F(\alpha)$ is continuous at $\alpha=0$.

$L_2(A)$ denotes the nonlinear space of the r.v. A which is a subspace of $L_2(X)$. An element in $L_2(A)$ is of the form $f(A)$, $f \in L_2(dF)$.

THEOREM 12. *If* X *is a nondegenerate coordinate* SIP(0,R;F) *satisfying (M) and* (C_0) *then there exists a unique isomorphism* Ψ *from* $L_2(A)\otimes\left(\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)\right)$ *onto* $L_2(X)$ *such that*

$$\Psi\left(f(A)\otimes e^{\hat{\otimes}\xi}\right) = f(A)\circ e^{\left(\frac{\alpha_1}{A}\right)^{\frac{1}{2}}\xi-\frac{1}{2}E\xi^2}$$

where $e^{\hat{\otimes}\xi} = \sum\frac{1}{\sqrt{p!}}\xi^{\hat{\otimes}p}$ *and* $f\in L_2(dF)$, $\xi\in H(X)$.

REMARK. We have defined $e^{\hat{\otimes}\xi}$ in two different ways (cf. Theorem 8). However, the latter always appears in the form $f(A)\otimes e^{\hat{\otimes}\xi}$ and hence there will be no confusion.

PROOF. The proof is similar to that of Theorem 8. The convergence of the series defining $e^{\hat{\otimes}\xi}$ results from $||\xi^{\hat{\otimes}p}||^2_{H^{\hat{\otimes}p}(X)} = ||\xi||^{2p}_{H(X)}$. Thus $e^{\hat{\otimes}\xi}$ is well-defined, and $\{e^{\hat{\otimes}\xi},\ \xi\in H(X)\}$ is complete in $\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)$ as shown in the proof of Theorem 8. Note that $\{f(A)\otimes e^{\hat{\otimes}\xi}\colon f(A)\in L_2(A),\ \xi\in H(X)\}$ is a complete set in $L_2(A)\otimes\left(\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)\right)$. Let $f,g\in L_2\left(dF(\alpha)\right)$ and $\xi,\eta\in H(X)$. Then we have

$$\begin{aligned}
&\langle f(A)\otimes e^{\hat{\otimes}\xi},\ g(A)\otimes e^{\hat{\otimes}\eta}\rangle_{L_2(A)\otimes\left(\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)\right)} \\
&\qquad = \langle f(A),g(A)\rangle_{L_2(A)}\ \langle e^{\hat{\otimes}\xi},e^{\hat{\otimes}\eta}\rangle_{\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)} \\
&\qquad = E\ f(A)g(A)\circ\sum\frac{1}{p!}\langle\xi,\eta\rangle^p_{H(X)} \\
&\qquad = \int f(\alpha)g(\alpha)dF(\alpha)\circ e^{E\xi\eta}\ ;
\end{aligned}$$

on the other hand, $e^{\left(\frac{\alpha_1}{A}\right)^{\frac{1}{2}}\xi-\frac{1}{2}E\xi^2}$ is well-defined because of (C_0) and

$$<f(A)e^{(\frac{\alpha_1}{A})^{\frac{1}{2}}\xi-\frac{1}{2}E\xi^2}, g(A)e^{(\frac{\alpha_1}{A})^{\frac{1}{2}}\eta-\frac{1}{2}E\eta^2}>_{L_2(X)}$$

$$= \int f(\alpha)g(\alpha) \circ E_\alpha\left(e^{(\frac{\alpha_1}{\alpha})^{\frac{1}{2}}\xi-\frac{1}{2}E\xi^2} \circ e^{\frac{\alpha_1}{\alpha}\eta-\frac{1}{2}E\eta^2}\right) dF(\alpha)$$

$$= \int f(\alpha)g(\alpha)e^{E\xi\eta}\, dF(\alpha)$$

where we have used the fact that under each P_α $(\alpha>0)$ $(\frac{\alpha_1}{\alpha})^{\frac{1}{2}}\xi$ is Gaussian with variance $E\xi^2$. We may now conclude that the map Ψ: $f(A)\otimes e^{\hat{\otimes}\xi} \longmapsto f(A)e^{(\frac{\alpha_1}{A})^{\frac{1}{2}}\xi-E\xi^2}$ can be extended uniquely to an isomorphism from $L_2(A)\otimes\left(\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)\right)$ into $L_2(X)$. We claim that Ψ is onto; the proof is given after Theorem 13. Thus the theorem is proved. Q.E.D.

As before, by the isomorphism Ψ in Theorem 12, we shall hereafter identify $L_2(A)\otimes\left(\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)\right)$ with $L_2(X)$.

<u>THEOREM 13.</u> *Suppose* X *is a nondegenerate coordinate* SIP(0,R;F) *satisfying (M) and* (C_0)*. Then for* $\xi \in H(X)$, $f(A) \in L_2(A)$

$$f(A)\otimes\xi^{\hat{\otimes}p} = f(A) \circ (\frac{1}{p!})^{\frac{1}{2}} H_{p,||\xi||^2}\left((\frac{\alpha_1}{A})^{\frac{1}{2}}\xi\right)$$

$$\left(= f(A)(\frac{1}{p!})^{\frac{1}{2}}(\frac{\alpha_1}{A})^{\frac{p}{2}} H_{p,\frac{A}{\alpha_1}||\xi||^2}(\xi)\ \right).$$

More generally, for a finite orthogonal set $\{\xi_1,\ldots,\xi_k\}$ *in* H(X) *and* $f(A) \in L_2(A)$

$$f(A)\otimes(\xi_1^{\hat{\otimes}p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_k^{\hat{\otimes}p_k})$$

$$= f(A) \circ (\frac{1}{p!})^{\frac{1}{2}} H_{p_1,||\xi||^2}\left((\frac{\alpha}{A})^{\frac{1}{2}}\xi_1\right) \ldots H_{p_k,||\xi_k||^2}\left((\frac{\alpha}{A})^{\frac{1}{2}}\xi_k\right),$$

$p=p_1+\ldots+p_k$; *in particular*

$$1\otimes(\xi_1\hat{\otimes}\ldots\hat{\otimes}\xi_k) = (\frac{1}{k!})^{\frac{1}{2}}(\frac{\alpha_1}{A})^{\frac{k}{2}}\xi_1\ldots\xi_k\ .$$

PROOF. The proof is essentially the same as that of Theorem 10, hence it is omitted. Q.E.D.

REMARK. It is easy to deduce from Theorem 13 that $f(A)\otimes\theta = f(A) \circ 1\otimes\theta$ for $f(A) \in L_2(A)$, $\theta \in \bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)$; and

$$\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X) = \left(\frac{A^p}{\alpha_p}\right)^{\frac{1}{2}} \otimes \Big(\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)\Big).$$

PROOF OF A CLAIM in the proof of Theorem 12.

Since $\{e^{\xi}, \xi\in H(X)\}$ is complete in $L_2(X)$ it is sufficient to show that the e^{ξ} belongs to the range of Ψ. Consider the following series in $L_2(A) \otimes \Big(\bigoplus_{p\geq 0} H^{\hat{\otimes}p}(X)\Big)$

$$\text{(i)} \quad \sum \frac{1}{\sqrt{p!}}\, e^{\frac{A}{2\alpha_1}E\xi^2} \left(\frac{A}{\alpha_1}\right)^{\frac{p}{2}} \otimes \xi^{\hat{\otimes}p}, \quad \xi \in H(X) .$$

The series converges since

$$\sum \left\|\frac{1}{\sqrt{p!}}\, e^{\frac{A}{2\alpha_1}E\xi^2} \left(\frac{A}{\alpha_1}\right)^{\frac{p}{2}} \otimes \xi^{\hat{\otimes}p}\right\|^2_{L_2(A)\otimes\left(\oplus H^{\hat{\otimes}p}(X)\right)}$$

$$= \sum \frac{1}{p!} \int e^{\frac{\alpha}{\alpha_1}E\xi^2} \left(\frac{\alpha}{\alpha_1}\right)^p dF(\alpha) \circ (E\xi^2)^p$$

$$= \int e^{\frac{\alpha}{\alpha_1}E\xi^2} \left[\sum \frac{1}{p!} \left(\frac{\alpha}{\alpha_1}E\xi^2\right)^p\right] dF(\alpha)$$

(by Fubini's Theorem)

$$= \int e^{\frac{2\alpha}{\alpha_1}E\xi^2}\, dF(\alpha)$$

$< \infty$ (by (M)).

From Theorem 13

$$\Psi\Big(\sum \frac{1}{\sqrt{p!}}\, e^{\frac{A}{2\alpha_1}E\xi^2} \big(\frac{A}{\alpha_1}\big)^{\frac{p}{2}} \otimes \xi^{\hat{\otimes} p}\Big) = \sum \frac{1}{\sqrt{p!}} e^{\frac{A}{2\alpha_1}E\xi^2} \big(\frac{A}{\alpha_1}\big)^{\frac{p}{2}} \frac{1}{\sqrt{p!}} H_{p,E\xi^2}\Big(\big(\frac{\alpha_1}{A}\big)^{\frac{1}{2}}\xi\Big)$$

$$= e^{\frac{A}{2\alpha_1}} \sum \frac{1}{p!} H_{p,\frac{A}{\alpha_1}E\xi^2}(\xi)$$

$$= e^{\frac{A}{2\alpha_1}} \cdot e^{\xi - \frac{A}{2\alpha_1}E\xi^2}$$

$$= e^{\xi}$$

where the series converges a.e. [P] and in $L_2(X)$. Thus $e^{\xi} \in \Psi\{L_2(A)\otimes\big(\underset{p\geq 0}{\oplus} H^{\hat{\otimes} p}(X)\big)\}$ and the proof is complete. Q.E.D.

In the following, it is convenient to assume that the index set Γ is linearly ordered. This is no restriction since every set can be linearly ordered.

<u>THEOREM 14.</u> *If* X *is a nondegenerate coordinate* SIP(0,R;F) *satisfying (M) and* (C_0), *and if* $\{\xi_\gamma,\ \gamma\in\Gamma\}$ *and* $\{e_n(A),\ n=1,\ldots,N\}$ *(* N *may be infinite) are* CONS*'s in* $H(X)$ *and* $L_2(A)$ *respectively, then the family*

$$M = \Big\{\big(\frac{p!}{p_1!\ldots p_k!}\big)^{\frac{1}{2}}\, e_n(A)\otimes\big(\xi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes} p_k}\big) : n=1,\ldots,N,\ k\geq 1$$

$$\gamma_1<\ldots<\gamma_k,\ p_1+\ldots+p_k=p.\ p\geq 0\Big\}$$

is a CONS *in* $L_2(A)\otimes\big(\underset{p\geq 0}{\oplus} H^{\hat{\otimes} p}(X)\big)$; *or equivalently the family*

$$\Big\{e_n(A)\cdot \Pi\big(\frac{1}{p_\gamma!}\big)^{\frac{1}{2}} H_{p_\gamma}\Big(\big(\frac{\alpha_1}{A}\big)^{\frac{1}{2}}\xi_\gamma\Big) : n=1,\ldots,N,\ \gamma\in\Gamma,\ p_\gamma\geq 0,\ \sum p_\gamma < \infty\Big\}$$

is a CONS *in* $L_2(X)$.

Hence every $\theta \in L_2(X)$ *has a unique orthogonal development as follows*

$$\theta = \sum_{n\geq 1}\sum_{p\geq 0}\sum_{\substack{p_1+\dots+p_k=p\\ k\geq 0}}\sum_{\gamma_1<\dots<\gamma_k} n^{a^{p_1\dots p_k}_{\gamma_1\dots\gamma_k}}\, e_n(A)\, H_{p_1}\Big(\big(\tfrac{\alpha_1}{A}\big)^{\frac{1}{2}}\xi_{\gamma_1}\Big)\dots H_{p_k}\Big(\big(\tfrac{\alpha_1}{A}\big)^{\frac{1}{2}}\xi_{\gamma_k}\Big) .$$

<u>PROOF.</u> It is easy to check that $L_2(A)\otimes\big(\oplus_{p\geq 0} H^{\hat{\otimes}p}(X)\big) = \oplus_{p\geq 0}\big(L_2(A)\otimes H^{\hat{\otimes}p}(X)\big)$. It is clear that $M\cap\big(L_2(A)\otimes H^{\hat{\otimes}p}(X)\big)$ is a complete set in $L_2(A)\otimes H^{\hat{\otimes}p}(X)$. Thus the set M is complete in $\oplus_{p\geq 0}\big(L_2(A)\otimes H^{\hat{\otimes}p}(X)\big) = L_2(A)\otimes\big(\oplus_{p\geq 0} H^{\hat{\otimes}p}(X)\big)$. It is also clear that M is an orthogonal set.

To show that every element in M has norm 1, it suffices to show

$$\text{(i)}\quad \Big|\Big|\xi_{\gamma_1}^{\hat{\otimes}p_1}\hat{\otimes}\dots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes}p_k}\Big|\Big|^2_{H^{\hat{\otimes}p}(X)} = \frac{p_1!\dots p_k!}{p!} .$$

But by definition

$$\xi_{\gamma_1}^{\hat{\otimes}p_1}\hat{\otimes}\dots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes}p_k} = \frac{1}{p!}\sum_{\nu}\xi_{\nu_1}\otimes\dots\otimes\xi_{\nu_p}$$

where $\nu=(\nu_1,\dots,\nu_p)$ is a permutation of $(\gamma_1,\dots,\gamma_1;\dots;\gamma_k,\dots,\gamma_k)$ and the sum is over all such permutations. Thus the norm in (i) equals

$$\text{(ii)}\quad \Big(\frac{1}{p!}\Big)^2\sum_{\lambda}\sum_{\nu}\langle\xi_{\nu_1}\otimes\dots\otimes\xi_{\nu_p},\ \xi_{\lambda_1}\otimes\dots\otimes\xi_{\lambda_p}\rangle_{H^{\otimes p}(X)}$$

and each summand is either 0 or 1. Consider the number of nonzero summands. The first term in the inner product can vary in $p!$ ways, and for each such term the second term can vary only in $p_1!\dots p_k!$ ways (for a nonzero inner product). Thus (ii) equals to $\dfrac{p_1!\dots p_k!}{p!}$ and the proof for the first assertion is complete.

The second assertion is now a direct consequence of Theorems 12 and 13. Q.E.D.

Theorem 14 for Gaussian processes is the celebrated theorem of Cameron and Martin which was a milestone in the study of nonlinear spaces of Gaussian processes.

COROLLARY 15. [Cameron and Martin 1947] *If* X *is a zero mean Gaussian process and if* $\{\xi_\gamma,\ \gamma\in\Gamma\}$ *is a* CONS *in* H(X), *then*

$$\{\Pi(\frac{1}{p_\gamma})^{\frac{1}{2}} H_{p_\gamma}(\xi):\ \gamma\in\Gamma,\ p_\gamma\geq 0,\ \sum p_\gamma<\infty\}$$

is a CONS *in* $L_2(X)$.

A final remark before closing this section is the following: We have assumed while studying the nonlinear space $L_2(X)$ that X is a coordinate process. This is a convenient assumption rather than a restriction. For if X is defined on an arbitrary probability space $(\Omega,\mathcal{B},P)$ with $\mathcal{B}$ the σ-field generated by X, then $L_2(X) = L_2(\Omega,\mathcal{B},P)$ is isomorphic to $L_2(\mathbb{R}^T,\mathcal{B}(\mathbb{R}^T),PX^{-1})$ under the correspondence $\theta(X(\omega)) \longleftrightarrow \theta(x)$. (Note that every $\mathcal{B}$-measurable function $\rho(\omega)$ is of the form $\theta(X(\omega))$ for some $\mathcal{B}(\mathbb{R}^T)$-measurable function $\theta(x)$ and the r.v.'s $\rho(\omega)$ and $\theta(x)$ have the same probability distribution.)

3. Multiple Wiener Integrals

We shall define the multiple Wiener integrals (MWI's) of the following types

$$I_p(f) = \int\cdots\int f(t_1,\ldots,t_p)\ dX_{t_1}\ldots dX_{t_p}\ ,$$

$$J_p(f) = \int\cdots\int f(t_1,\ldots,t_p)\ X_{t_1}\ldots X_{t_p}\ dt_1\ldots dt_p$$

where $X=(X_t,\ t\in T)$ is a SIP(0,R;F).

Introduce the following "integrability" conditions on X:

(I) X is zero mean, and vanishes at some point $t_0 \in T$ (i.e. $X_{t_0}=0$ a.e.) with T an interval,

(J) X is zero mean and mean square continuous with T an interval.

We will always assume (I) while dealing with integrals I_p, $p\geq 0$; and (J) while dealing with integrals J_p, $p\geq 0$.

The following lemma is well known and is the key in Wiener's definition of MWI's for a Wiener process.

<u>LEMMA 16.</u> *If* $\{\xi_1,\ldots,\xi_n\}$ *is a Gaussian family with zero means then*

$$E\,\xi_1\ldots\xi_n = \sum{}^*\prod{}^*\, E\,\xi_i\xi_j$$

where the sum is over all ways of dividing n *terms into pairs and the product is over all pairs in this way of dividing. (Empty sum is defined to be 0.)*

<u>PROOF.</u> Consider the moment generating function

$$\phi(t_1,\ldots,t_n) = E\,e^{t_1\xi_1+\ldots+t_n\xi_n} = e^{\frac{1}{2}\sum t_i t_j \sigma_{ij}}$$

$$= \sum \frac{1}{p!}\,(\tfrac{1}{2})^p\,(\sum t_i t_j \sigma_{ij})^p$$

where $\sigma_{ij} = E\,\xi_i\xi_j$. Since

$$E\,\xi_1\ldots\xi_n = \left(\frac{\partial^n\phi}{\partial t_1\ldots\partial t_n}\right)_{t_1=\ldots=t_n=0}$$

is exactly the coefficient of the term $t_1\ldots t_n$ in ϕ, we can conclude this lemma after a little algebra. Q.E.D.

<u>LEMMA 17.</u> *Let* X *be a zero mean* SIP *with* $\alpha_{\frac{n}{2}}<\infty$ *and let* $\xi_1,\ldots,\xi_n \in H(X)$. *Then*

$$E\ \xi_1 \cdots \xi_n = \frac{\alpha_p}{\alpha_1^p} \sum{}^* \Pi^* \ E\ \xi_i \xi_j \ , \quad p=\frac{n}{2} \ .$$

<u>PROOF.</u> We may assume X is a coordinate process. Then

$$\begin{aligned} E\ \xi_1 \cdots \xi_n &= \int E_\alpha\ \xi_1 \cdots \xi_n \ dF(\alpha) \\ &= \int \sum{}^* \Pi^* \ \alpha\ E_1 \xi_i \xi_j \ dF(\alpha) \qquad \text{(by Lemma 14)} \\ &= \alpha_p \sum{}^* \Pi^* \frac{1}{\alpha_1} E\ \xi_i \xi_j \\ &= \frac{\alpha_p}{\alpha_1^p} \sum{}^* \Pi^* \ E\ \xi_i \xi_j \ . \end{aligned}$$

Q.E.D.

The following theorem is essential to our approach in defining MWI's for SIP's, and it has its own independent interest.

<u>THEOREM 18.</u> *Let* H *be an infinite dimensional Hilbert space, and let*

$$S = \left\{\phi_1 \otimes \cdots \otimes \phi_p : \{\phi_i,\ 1 \le i \le p\} \text{ an orthogonal set in } H\right\},$$

$$\tilde{S} = \left\{\phi_1 \hat{\otimes} \cdots \hat{\otimes} \phi_p : \{\phi_i,\ 1 \le i \le p\} \text{ an orthogonal set in } H\right\}.$$

Then S *and* $\tilde{S}$ *are complete in* $H^{\otimes p}$ *and* $H^{\hat{\otimes} p}$ *respectively.*

<u>PROOF.</u> Pick a CONS $\{e_\gamma,\ \gamma \in \Gamma\}$ in H, then $\{e_{\gamma_1} \otimes \cdots \otimes e_{\gamma_p},\ \gamma_i \in \Gamma\}$ is complete in $H^{\otimes p}$. To show that S is complete, it suffices to show that each $e_{\gamma_1} \otimes \cdots \otimes e_{\gamma_p}$ can be approximated by linear combinations of elements in S. For $\gamma_1, \ldots, \gamma_p$ distinct we have $e_{\gamma_1} \otimes \cdots \otimes e_{\gamma_p} \in S$, and thus only the case that not all γ's distinct needs consideration. Suppose that $\{\gamma_1, \ldots, \gamma_p\}$ has k distinct elements $\gamma_1, \ldots, \gamma_k$, each repeated $p_1, \ldots, p_k$ times with $p_1 + \cdots + p_k = p$ and $k \le n$. We may choose

an orthonormal set $\{f_1,\dots,f_n;\dots;f_{(k-1)n+1},\dots,f_{kn}\}$ in H such that $e_{\gamma_i} = \frac{1}{n^{\frac{1}{2}}}(f_{(i-1)n+1}+\dots+f_{in})$, $1\le i\le k$. We will show only the existence of $f_1,\dots,f_n$; the existence of the others can be shown similarly.

Take $n-1$ elements $e_{\beta_2},\dots,e_{\beta_n}$ from $\{e_\gamma, \gamma\in\Gamma\} \setminus \{e_{\gamma_1},\dots,e_{\gamma_k}\}$. Let Φ be an isomorphism on the subspace spanned by $e_{\gamma_1}, e_{\beta_1},\dots,e_{\beta_n}$ such that $\Phi\, e_{\gamma_1} = \frac{1}{n^{\frac{1}{2}}}(e_{\gamma_1}+e_{\beta_2}+\dots+e_{\beta_n})$. Thus $e_{\gamma_1} = \frac{1}{n^{\frac{1}{2}}}(\Phi^{-1}e_{\gamma_1}+\Phi^{-1}e_{\beta_2}+\dots+\Phi^{-1}e_{\beta_n})$ and $\Phi^{-1}e_{\gamma_1}, \Phi^{-1}e_{\beta_2}, \dots, \Phi^{-1}e_{\beta_n}$ are our choice of $f_1,\dots,f_n$.

Now assume for convenience $e_{\gamma_1}\otimes\dots\otimes e_{\gamma_p}$ to be

$$e_{\gamma_1}^{\otimes p_1}\otimes\dots\otimes e_{\gamma_k}^{\otimes p_k} = n^{-\frac{p}{2}}\sum f_{\gamma_{11}}\otimes\dots\otimes f_{\gamma_{1n}}\otimes\dots\otimes f_{\gamma_{k1}}\otimes\dots\otimes f_{\gamma_{kn}}$$

where the sum is over all $\gamma_{ij} = ki+1,\dots,k(i+1)$, $1\le i\le k$, $1\le j\le n$. Divide this sum into two parts $\sum'$ and $\sum''$ where $\sum'$ is the sum over all distinct γ_{ij}'s and $\sum''$ is the sum of the remaining terms. $\sum'$ is an element in the span of S and

$$||\textstyle\sum'||^2 = \frac{1}{n^p}\frac{n!}{(n-p_1)!}\cdots\frac{n!}{(n-p_k)!} \to 1 \quad \text{as} \quad n\to\infty.$$

So by choosing n sufficiently large, $e_{\gamma_1}\otimes\dots\otimes e_{\gamma_p}$ can be made arbitrarily close to the span of S. Thus the first assertion is proved.

The second assertion is now obvious. Q.E.D.

We can now proceed to define the MWI's I_p, $p\ge 0$. Suppose X is a nondegenerate $SIP(0,R;F)$ with covariance function $r(t,s)$. Note $r(t,s) = \alpha_1 R(t,s)$. I_0 is defined to be the identity map on $\mathbb{R}$, and I_1 is defined to be the isomorphism

$$I_1 : \Lambda_2(r) \to H(X) : f \mapsto \int f(t)\,dX_t\ .$$

For $p \geq 2$, let $\tilde{S} = \left\{\phi_1 \hat{\otimes} \ldots \hat{\otimes} \phi_p : \{\phi_i\} \text{ is an orthogonal set in } \Lambda_2(r)\right\}$ and consider the map

$$I_p : \tilde{S} \to L_2(X) : \phi_1 \hat{\otimes} \ldots \hat{\otimes} \phi_p \mapsto \int \phi_1 dX \ldots \int \phi_p dX .$$

Then for $\phi_1 \hat{\otimes} \ldots \hat{\otimes} \phi_p$, $\psi_1 \hat{\otimes} \ldots \hat{\otimes} \psi_p \in \tilde{S}$

$$\begin{aligned} E[I_p(\phi_1 \hat{\otimes} \ldots \hat{\otimes} \phi_p) I_p(\psi_1 \hat{\otimes} \ldots \hat{\otimes} \psi_p)] & \\ &= E[\int \phi_1 dX \ldots \int \phi_p dX \circ \int \psi_1 dX \circ \ldots \circ \int \psi_p dX] \\ &= \frac{\alpha_p}{\alpha_1^p} \sum_\pi \prod_{i=1}^p E(\int \phi_1 dX \circ \int \psi_{\pi i} dX) \\ &\qquad\qquad (\pi \text{ over all permutations}) \\ &= \frac{\alpha_p}{\alpha_1^p} \sum_\pi \prod_{i=1}^p \langle \phi_i, \psi_{\pi i} \rangle_{\Lambda_2(r)} , \end{aligned}$$

where the second equality follows from Lemma 17 and the fact that

$$\begin{aligned} E(\int \phi_i dX \int \phi_j dX) &= E(\int \psi_i dX \int \psi_j dX) \\ &= 0 \quad \text{for } i \neq j . \end{aligned}$$

On the other hand,

$$\begin{aligned} \langle \phi_1 \hat{\otimes} \ldots \hat{\otimes} \phi_p, \psi_1 \hat{\otimes} \ldots \hat{\otimes} \psi_p \rangle_{\Lambda_2(\hat{\otimes}^p r)} &= \left(\frac{1}{p!}\right)^2 \sum_\pi \sum_\sigma \langle \phi_{\pi 1} \otimes \ldots \otimes \phi_{\pi p}, \psi_{\sigma 1} \otimes \ldots \otimes \psi_{\sigma p} \rangle_{\Lambda_2(\hat{\otimes}^p r)} \\ &= \left(\frac{1}{p!}\right)^2 \sum_\pi \sum_\sigma \prod_{i=1}^p \langle \phi_{\sigma i}, \psi_{\pi i} \rangle_{\Lambda_2(r)} \\ &= \frac{1}{p!} \sum_\pi \prod_{i=1}^p \langle \phi_i, \psi_{\pi i} \rangle_{\Lambda_2(r)} \end{aligned}$$

where π, σ are over all permutations. Thus

$$\langle I_p(\phi_i \hat{\otimes} \ldots \hat{\otimes} \phi_p), I_p(\psi_1 \hat{\otimes} \ldots \hat{\otimes} \psi_p) \rangle_{L_2(X)} = \frac{p! \alpha_p}{\alpha_1^p} \langle \phi_1 \hat{\otimes} \ldots \hat{\otimes} \phi_p, \psi_1 \hat{\otimes} \ldots \hat{\otimes} \psi_p \rangle_{\Lambda_2(\hat{\otimes}^p r)}$$

and hence $\left(\frac{\alpha_1^p}{p! \alpha_p}\right)^{\frac{1}{2}} I_p$ is a map on $\tilde{S}$ which preserves inner products.

Theorem 18 implies $\left(\frac{\alpha_1^p}{p!\alpha_p}\right)^{\frac{1}{2}} I_p$ can be extended uniquely to an isomorphism on $\Lambda_2(\hat{\otimes}^p r)$. Furthermore define $I_p(f) = I_p(\tilde{f})$ for $f \in \Lambda_2(\otimes^p r)$ where $\tilde{f}$ is the symmetric version of f.

We now summarize our results as follows.

THEOREM 19. *If* X *is a nondegenerate* SIP(0,R;F) *satisfying (I) and* $\alpha_p<\infty$, *then the* MWI I_p *can be defined uniquely on* $\Lambda_2(\otimes^p r)$ $(r=\alpha_1 R)$ *such that* $I_p\colon \Lambda_2(\otimes^p r) \to L_2(X)$ *is a bounded linear operator with* $I_p(\phi_1\otimes\dots\otimes\phi_p) = \int \phi_1 dX \dots \int \phi_p dX$ *for* $\{\phi_i\}$ *an orthogonal set in* $\Lambda_2(r)$.

Moreover $\left(\frac{\alpha_1^p}{p!\alpha_p}\right)^{\frac{1}{2}} I_p$ *restricted to the subspace* $\Lambda_2(\hat{\otimes}^p r)$ *is an isomorphism onto* $\hat{\otimes}^p H(X)$.

The MWI J_p $(p\geq 0)$ cna be defined in exactly the same fashion, and we have the following

THEOREM 19'. *If* X *is a nondegenerate* SIP(0,R;F) *satisfying (J) and* $\alpha_p<\infty$, *then the* MWI J_p *can be defined uniquely on* $\Lambda_2(\otimes^p r)$ $(r=\alpha_1 R)$ *such that* $J_p\colon \Lambda_2(\otimes^p r) \to L_2(X)$ *is a bounded linear operator with* $J_p(\phi_1\otimes\dots\otimes\phi_p) = \int \phi_1 dX \dots \int \phi_p dX$ *for* $\{\phi_i\}$ *an orthogonal set in* $\Lambda_2(r)$.

Moreover $\left(\frac{\alpha_1^p}{p!\alpha_p}\right)^{\frac{1}{2}} I_p$ *restricted to the subspace* $\Lambda_2(\hat{\otimes}^p r)$ *is an isomorphism onto* $\hat{\otimes}^p H(X)$.

We now give another less intuitive definition for I_p. Since $\Lambda_2(r)$ is isomorphic to $H(X)$ under the isomorphism $I\colon f \mapsto \int f dX$, $\Lambda_2(\hat{\otimes}^p r)$ is isomorphic to $H^{\hat{\otimes} p}(X)$. Denote this isomorphism by $I^{\hat{\otimes} p}$. For $\phi_1,\dots,\phi_p$ orthogonal in $\Lambda_2(r)$ we have

$$I^{\hat{\otimes} p}(\phi_1 \hat{\otimes} \dots \hat{\otimes} \phi_p) = \int \phi_1 dX \hat{\otimes} \dots \hat{\otimes} \int \phi_p dX$$
$$= \left(\frac{\alpha_1^p}{p!\alpha_p}\right)^{\frac{1}{2}} \int \phi_1 dX \dots \int \phi_p dX \quad \text{(by Theorem 10)},$$

which suggests to define on $\Lambda_2(\hat{\otimes}^p r)$

$$I_p = \left(\frac{p!\alpha_p}{\alpha_I^p}\right) I^{\hat{\otimes} p} .$$

This definition for I_p is consistent with the previous one since they coincide on $\tilde{S}$ which is complete in $\Lambda_2(\hat{\otimes}^p r)$.

The following results are immediate consequences of Theorems 10 and 12, and the fact that $I_p = \left(\frac{p!\alpha_p}{\alpha_1^p}\right)^{\frac{1}{2}} I^{\hat{\otimes} p}$ is a constant multiple of an isomorphism.

THEOREM 20. *Let* X *be a nondegenerate coordinate* SIP(0,R;F) *satisfying (I), (M). Then the* MMI's I_p, $p \geq 0$, *have the following properties:*

$$I_p(af+bg) = aI_p(f) + bI_p(g) \ , \quad a,b \in \mathbb{R} \ ,$$
$$I_p(f) = I_p(\tilde{f}) \ ,$$
$$\langle I_p(f), I_p(g)\rangle = \frac{p!\alpha_p}{\alpha_1^p} \langle \tilde{f}, \tilde{g}\rangle_{\Lambda_2(\hat{\otimes}^p r)} \ ,$$
$$\langle I_p(f), I_q(g)\rangle = 0 \quad \text{if} \quad p \neq q \ ,$$
$$I_p(\phi_1^{\hat{\otimes} p_1} \hat{\otimes} \dots \hat{\otimes} \phi_k^{\hat{\otimes} p_k}) = H_{p_1, \frac{A}{\alpha_1}||\phi_1||^2}\left(\int \phi_1 dX\right) \dots H_{p_k, \frac{A}{\alpha_1}||\phi_k||^2}\left(\int \phi_k dX\right)$$

where $\{\phi_1, \dots, \phi_k\}$ *is an orthogonal set in* $\Lambda_2(r)$.

THEOREM 21. *Let* X *be a nondegenerate coordinate* SIP(0,R;F) *satisfying (I), (M) and* (C_0). *Let* $\{e_n, n=1,\dots,N\}$ *(* N *may be infinite)* *be a* CONS *in* $L_2(dF)$. *Then every* $\theta \in L_2(X)$ *has an orthogonal development as follows*

$$\theta = \sum_{n\geq 1}\sum_{p\geq 0} \frac{e_n(A)}{A^{\frac{p}{2}}} I_p\left(f_p^{(n)}\right), \quad f_p^{(n)} \in \Lambda_2(\otimes^p r) .$$

If $\sum\sum \frac{e_n(A)}{A^{\frac{p}{2}}}\left(f_p^{(n)}\right) = \sum\sum \frac{e_n(A)}{A^{\frac{p}{2}}}\left(g_p^{(n)}\right)$ *then* $\tilde{f}_p^{(n)} = \tilde{g}_p^{(n)}$ *for all* n *and* p.

PROOF. Pick a CONS $\{\phi_\gamma, \gamma\in\Gamma\}$ in $\Lambda_2(r)$. Let $\xi_\gamma = \int \phi_\gamma dX$. Then $\{\xi_\gamma, \gamma\in\Gamma\}$ is a CONS in $H(X)$. By Theorem 14 we can write

$$\theta = \sum_{n\geq 1}\sum_{p\geq 0}\sum_{p_1+\dots+p_k=p}\sum_{\gamma_1<\dots<\gamma_k} {}_n a^{p_1\dots p_k}_{\gamma_1\dots\gamma_k}\, e_n(A)\otimes\xi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\dots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes} p_k}$$

$$= \sum\sum\sum\sum {}_n a^{p_1\dots p_k}_{\gamma_1\dots\gamma_k}\left(\frac{\alpha_1^p}{p!A^p}\right)^{\frac{1}{2}} e_n(A)\; I_p\left(\phi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\dots\hat{\otimes}\phi_{\gamma_k}^{\hat{\otimes} p_k}\right) .$$

Putting $f_p^{(n)} = \sum_{p_1+\dots+p_k=p}\sum_{\gamma_1<\dots<\gamma_k} {}_n a^{p_1\dots p_k}_{\gamma_1\dots\gamma_k}\left(\frac{\alpha_1^p}{p!}\right)^{\frac{1}{2}} \phi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\dots\hat{\otimes}\phi_{\gamma_k}^{\hat{\otimes} p_k}$ (in $\Lambda_2(\hat{\otimes}^p r)$) we get $\theta = \sum_{n\geq 1}\sum_{p\geq 0}\frac{e_n(A)}{A^{\frac{p}{2}}} I_p\left(f_p^{(n)}\right)$. And the uniqueness of $f_p^{(n)}$ is now clear. Q.E.D.

The corresponding results for J_p can be shown similarly.

THEOREM 20'. *Let* X *be a nondegenerate coordinate* SIP(0,R;F) *satisfying (J), (M) and* (C_0). *Then the* MWI J_p *can be defined uniquely on* $\lambda_2(\otimes^p r)$ *such that the following hold*

$$J_p(af+bg) = aJ_p(f) + bJ_p(g) \ , \quad a,b\in \mathbb{R},$$

$$J_p(f) = J_p(\tilde{f}) \ ,$$

$$<J_p(f),J_p(g)> = \frac{p!\alpha_p}{\alpha_1^p} <\tilde{f},\tilde{g}>_{\lambda_2(\hat{\otimes}^p r)} ,$$

$$<J_p(f),J_p(g)> = 0 \quad \text{if} \quad p \neq q$$

$$J_p(\phi_1^{\hat{\otimes} p_1} \hat{\otimes} \dots \hat{\otimes} \phi_k^{\hat{\otimes} p_k}) = H_{p_1, \frac{A}{\alpha_1}||\phi_1||^2}\left(\int \phi(t)X_t dt\right) \dots H_{p_k, \frac{A}{\alpha_1}||\phi_k||^2}\left(\int \phi(t)X_t dt\right)$$

where $\{\phi_1,\dots,\phi_k\}$ *is an orthogonal set in* $\lambda_2(r)$.

THEOREM 21'. *Let* X *be a nondegenerate coordinate* SIP(0,R;F) *satisfying (J), (M) and* (C_0). *Let* $\{e_n,\ n=1,2,\dots,N\}$ *(N may be infinite) be a* CONS *in* $L_2(dF)$. *Then every* $\theta \in L_2(X)$ *has an orthogonal development as follows*

$$\theta = \sum_{n=1}^{N} \sum_{p\geq 0} \frac{e_n(A)}{A^{\frac{p}{2}}} J_p\left(f_p^{(n)}\right) , \quad f_p^{(n)} \in \lambda_2(\otimes^p r) .$$

If $\sum\sum \frac{e_n(A)}{A^{\frac{p}{2}}} J_p\left(f_p^{(n)}\right) = \sum\sum \frac{e_n(A)}{A^{\frac{p}{2}}} J_p\left(g_p^{(n)}\right)$ *then* $\tilde{f}_p^{(n)} = \tilde{g}_p^{(n)}$ *for all* n *and* p.

COROLLARY 22. *If* X *is a Gaussian process with covariance* R *and satisfies (I) then the* MWI's I_p, $p\geq 0$ *have the following properties:*

$$I_p(af+bg) = aI_p(f)+bI_p(g) , \quad a,b \in \mathbb{R} ,$$

$$I_p(f) = I_p(\tilde{f}) ,$$

$$<I_p(f),I_p(g)> = p! <\tilde{f},\tilde{g}>_{\Lambda_2(\hat{\otimes}^p R)} ,$$

$$<I_p(f),I_q(g)> = 0 \quad \text{if} \quad p\neq q ,$$

$$I_p(\phi_1^{\hat{\otimes} p_1} \hat{\otimes} \dots \hat{\otimes}\ {}_k^{\otimes p_k}) = H_{p_1,||\phi_1||^2}\left(\int \phi_1 dX\right) \dots H_{p_k,||\phi_k||^2}\left(\int \phi_k dX\right)$$

where $\phi_1,\dots,\phi_k$ *are orthogonal in* $\Lambda_2(R)$.

COROLLARY 23. *If* X *is a Gaussian process with covariance* R *and satisfies (I) then every* L_2*-functional* θ *of* X *has an orthogonal development*

$$\theta = \sum I_p(f_p) \ , \quad f_p \in \Lambda_2(\otimes^p R) \ .$$

If $\sum I_p(f_p) = \sum I_p(g_p)$ *then* $\tilde{f}_p = \tilde{g}_p \quad \forall\ p\geq 0$.

COROLLARY 22'. *If* X *is a Gaussian process with covariance* R *and satisfies (J) then the* MWI*'s* I_p, $p\geq 0$, *have the following properties:*

$$J_p(af+bg) = aJ_p(f) + bJ_p(g) \ , \quad a,b \in \mathbb{R} \ ,$$

$$J_p(f) = J_p(\tilde{f}) \ ,$$

$$\langle J_p(f), J_p(g)\rangle = p!\ \langle \tilde{f},\tilde{g}\rangle_{\lambda_2(\hat{\otimes}^p R)} \ ,$$

$$\langle J_p(f), J_p(g)\rangle = 0 \quad \text{if} \quad p\neq q \ ,$$

$$J_p(\phi_1^{\hat{\otimes} p_1}\hat{\otimes}\cdots\hat{\otimes}\phi_k^{\hat{\otimes} p_k}) = H_{p_1,\|\phi_1\|^2}\Big(\int \phi_1(t)X_t dt\Big)\cdots H_{p_k,\|\phi_k\|^2}\Big(\int \phi_k(t)X_t dt\Big)$$

where $\phi_1,\ldots,\phi_k$ *are orthogonal in* $\lambda_2(R)$.

COROLLARY 23'. *If* X *is a Gaussian process with covariance* R *and satisfies (J) then every* L_2*-functional* θ *of* X *has an orthogonal development*

$$\theta = \sum_{p\geq 0} J_p(f_p) \ , \quad f_p \in \lambda_2(\otimes^p R) \ .$$

If $\sum J_p(f_p) = \sum J_p(g_p)$ *then* $\tilde{f}_p = \tilde{g}_p \quad \forall\ p\geq 0$.

IV. EQUIVALENCE AND A 0-1 LAW FOR SPHERICALLY INVARIANT PROCESSES

In Section 1, we combine the representation theorem for SIP's and the dichotomy theorem for Gaussian processes to derive some results on the equivalence of SIP'S (Theorems 5, 7 and 8). (Two processes are said to be *equivalent* if the induced measures are equivalent.) In Section 2, we derive a 0-1 law for SIP'S (Corollary 12) which partially extends the celebrated 0-1 law for Gaussian processes.

1. Equivalence of Spherically Invariant Processes

We first state the general theorem concerning the equivalence $(\sim)$ and the singularity $(\perp)$ of two Gaussian processes.

THEOREM 1. [Neveu 1968] *Let* $X=(X_t,\ t\in T)$ *be a zero mean Gaussian process on the probability space* $(\Omega,\mathcal{B},P)$. *Let* R *be the covariance function of* X. *Let* Q *be a second probability on* $(\Omega,\mathcal{B})$ *under which* X *is a Gaussian process with mean* m *and covariance function* S.

Then either $P\sim Q$ *or* $P\perp Q$. *In order that* $P\sim Q$, *it is necessary and sufficient that there exist* $y\in H(X)$ *and* $U\in H(X)\hat{\otimes}H(X)$ *such that*

$$m(t) = \langle X_t,Y\rangle_{H(X)} + \langle X_t\hat{\otimes}Y;U\rangle_{H(X)\hat{\otimes}H(X)}, \tag{1}$$

$$S(t,s)-R(t,s) = \langle X_t\hat{\otimes}X_s,U\rangle_{H(X)\hat{\otimes}H(X)} \tag{2}$$

and the Hilbert-Schmidt operator U *on* $H(X)$ *does not admit* -1 *as its eigenvalue.*

When $P\sim Q$, *the Radon Nikodym derivative of* Q w.r.t. P *is given by*

$$\frac{dQ}{dP}(X) = [\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}}\ e^{-\frac{1}{2}\sum\frac{\langle Y,\xi_i\rangle^2_{H(X)}}{1-\lambda_i}}\circ e^{\frac{1}{\sqrt{2}}Z+Y} \tag{3}$$

where $Z = \sum\lambda_i\xi_i\hat{\otimes}\xi_i \in H(X)\hat{\otimes}H(X)$, $\{\xi_i\}$ *is an orthonormal sequence in* $H(X)$, $U = \sum u_i\xi_i\hat{\otimes}\xi_i$, $(1-\lambda_i)(1+u_i) = 1$ *and* $Y = \sum\langle Y,\xi_i\rangle_{H(X)}\xi_i$.

REMARK: (i) An element in $H(X)\hat{\otimes}H(X)$ can be viewed as an element in $L_2(X)$ or as a Hilbert-Schmidt operator on $H(X)$ (cf. Appendix).

(ii) X in (3) represents the sample function $X(\omega)$.

Since conditions (1), (2) are very difficult to verify and the expression (3) is very difficult to evaluate, it may be advantageous to restate Theorem 1 as follows.

THEOREM 2. *Under the assumptions in Theorem 1 and the additional assumption that* X *satisfies (I) (resp. (J)), either* $P\sim Q$ *or* $P\perp Q$. *In order that* $P\sim Q$ *it is necessary and sufficient that there exist* $h\in\Lambda_2(R)$ *(resp.* $\lambda_2(R)$ *) and* $K\in\Lambda_2(R)\hat{\otimes}\Lambda_2(R)$ *(resp.* $\lambda_2(R)\hat{\otimes}\lambda_2(R)$ *) such that*

$$(1') \qquad m(t) = \langle 1_t, h\rangle_{\Lambda_2(R)} + \langle 1_t\hat{\otimes}h, K\rangle_{\Lambda_2(R)\hat{\otimes}\Lambda_2(R)}$$

$$\left(\text{resp. } \langle \delta_t, h\rangle_{\lambda_2(R)} + \langle \delta_t\hat{\otimes}h, K\rangle_{\lambda_2(R)\hat{\otimes}\lambda_2(R)}\right)$$

$$(2') \qquad S(t,s)-R(t,s) = \langle 1_t\hat{\otimes}1_s, K\rangle_{\Lambda_2(R)\hat{\otimes}\Lambda_2(R)}$$

$$\left(\text{resp. } \langle \delta_t\hat{\otimes}\delta_s, K\rangle_{\lambda_2(R)\hat{\otimes}\lambda_2(R)}\right)$$

and the Hilbert-Schmidt operator K *on* $\Lambda_2(R)$ *(resp.* $\lambda_2(R)$ *) does not admit* -1 *as its eigenvalue.*

When $P\sim Q$ *then*

$$(3') \qquad \frac{dQ}{dP}(X) = [\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}}\, e^{-\frac{1}{2}\sum\frac{\langle h,e_i\rangle^2_{\Lambda_2(R)}}{1-\lambda_i}} \circ e^{\frac{1}{2}\iint H(t,s)dX_t dX_s + \int h(t)dX_t}$$

$$\left(\text{resp. } [\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}}\, e^{-\frac{1}{2}\sum\frac{\langle h,e_i\rangle^2_{\lambda_2(R)}}{1-\lambda_i}} \circ e^{\frac{1}{2}\iint H(t,s)X_t X_s dtds + \int h(t)X_t dt}\right)$$

where $H=\sum\lambda_i e_i\hat{\otimes}e_i \in \Lambda_2(R)\hat{\otimes}\Lambda_2(R)$ *(resp.* $\lambda_2(R)\hat{\otimes}\lambda_2(R)$ *),* $\{e_i\}$ *is an orthonormal sequence in* $\Lambda_2(R)$ *(resp.* $\lambda_2(R)$ *),* $U=\sum u_i e_i\hat{\otimes}e_i$, $(1-\lambda_i)(1+u_i)=1$, *and* $h=\sum\langle h,e_i\rangle_{\Lambda_2(R)}e_i$ *(resp.* $\sum\langle h,e_i\rangle_{\lambda_2(R)}e_i$ *).*

PROOF. By the isomorphism $\Lambda_2(R)\to H(X)$: $f \mapsto \int f dX$ we can write $X_t=\int 1_t dX$, $Y=\int h dX$ and $\xi_i=\int e_i dX$ $i\geq 1$. Then it is easily seen that (1') and (2') are equivalent to (1) and (2). (3') is equivalent to (3) since

$$\tfrac{1}{2}\iint H(t,s)dX_t dX_s = \tfrac{1}{2}\sum \lambda_i \iint e_i(t)\hat{\otimes}e_i(s)\ dX_t dX_s$$
$$= \tfrac{1}{2}\sum \lambda_i(\sqrt{2}\int e_i dX \hat{\otimes} \int e_i dX)$$
$$= \frac{1}{\sqrt{2}}\sum \lambda_i\ \xi_i\hat{\otimes}\xi_i = \frac{1}{\sqrt{2}}Z.$$

Q.E.D.

If we take X to be Wiener process under P, we get the following well known theorem due to Shepp [1966].

COROLLARY 3. *Let* X *be a Wiener process with* $T=[0,T]$ *or* $[0,\infty)$ *in Theorem 2. Then* $P\sim Q$ *if and only if there exist* $m'\in L_2(T)$ *for which*

$$(1'')\qquad m(t)=\int_0^t m'(u)du$$

and a symmetric function $K\in L_2(T\times T)$ *for which*

$$(2'')\qquad S(t,s)-\min(t,s)=\int_0^s\int_0^t K(u,v)dudv$$

and the Hilbert-Schmidt operator K *on* $L_2(T)$ *does not admit* -1 *as its eigenvalue.*

When $P\sim Q$,

$$(3'')\qquad \frac{dQ}{dP}(X+m) = [\delta(-1)e^{-\mathrm{tr}HK}]^{-\frac{1}{2}}\ e^{\frac{1}{2}\iint H(t,s)dX_t dX_s+\int m'(t)dX_t+\frac{1}{2}\int m'(t)^2dt}$$

where $\delta(\lambda)$ *is the Carleman-Fredholm determinant of* K *and* H *is the Fredholm resolvent of* K *at* -1.

PROOF. Under the present assumption we have $\Lambda_2(R)=L_2(T)$ and $\Lambda_2(R)\hat{\otimes}\Lambda_2(R) = \tilde{L}_2(T\times T)$ (see Appendix). Adopt the notation in Theorem 2 and set $m'=(I+K)h$ (I denotes the identity operator) then it is immediate that (1") and (2") are equivalent to (1') and (2'). Now assume $p\sim Q$ and we will show that (3') implies (3"). Since H is the Fredholm resolvent of K at -1, it is determined by the operator equation

$$\text{(i)} \qquad (I-H)(I+K) = I.$$

Suppose as in Theorem 2 $K = \sum u_i \, e_i \otimes e_i$. Then (i) implies that $H = \sum \lambda_i \, e_i \otimes e_i$, $(1-\lambda_i)(1+u_i)=1$; and $h=(I-H)m'$. (3') now becomes

$$\frac{dQ}{dP}(X) = [\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}} \, e^{-\frac{1}{2}\sum \frac{\langle h,e_k\rangle^2_{L_2(T)}}{1-\lambda_i}} \cdot e^{\frac{1}{2}\sum\lambda_i[(\int e_i dX)^2-1]+\int(I-H)m' dX}.$$

It may be shown that $\int fd(X+m) = \int fdX + \int f(t)m'(t)dt$ for all $f\in L_2(T)$. Thus we have

$$\begin{aligned}
&\tfrac{1}{2}\sum \lambda_i[(\int e_i d(X+m))^2-1] \\
&\quad = \tfrac{1}{2}\sum \lambda_i[(\int e_i dX)^2-1] + \sum \lambda_i \langle m',e_i\rangle_{L_2(T)} \int e_i dX + \tfrac{1}{2}\sum \lambda_i \langle m',e_i\rangle^2_{L_2(T)} \\
&\quad = \tfrac{1}{2}\iint H(t,s)\,dX_t dX_s + \int Hm' dX + \tfrac{1}{2}\langle Hm',m'\rangle_{L_2(T)},
\end{aligned}$$

$$\int (I-H)m' d(X+m) = \int (I-H)m' dX + \langle (I-H)m',m'\rangle_{L_2(T)},$$

and

$$\frac{1}{2} \sum \frac{\langle h, e_i \rangle^2_{L_2(T)}}{1-\lambda_i} = \frac{1}{2} \sum (1+u_i) \langle h, e_i \rangle^2_{L_2(T)}$$

$$= \frac{1}{2} \langle (I+K)h, h \rangle_{L_2(T)}$$

$$= \frac{1}{2} \langle m', (I-H)m' \rangle_{L_2(T)}$$

which imply

$$\frac{dQ}{dP}(X+m) = [\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}} \; e^{\frac{1}{2}\int\int H(t,s)dX_t dX_s + \int m' dX + \frac{1}{2}\int m'(t)^2 dt} \; .$$

(3'') follows by noting that from the definitions of $\delta(\lambda)$ and $\mathrm{tr}HK$

$$\delta(-1)e^{-\mathrm{tr}HK} = [\Pi(1+u_i)e^{-u_i}]e^{-\sum \lambda_i u_i}$$

$$= [\Pi(1-\lambda_i)e^{\lambda_i}]^{-1} \quad \big((1-\lambda_i)(1+u_i)=1\big) \; .$$

Q.E.D.

When P and Q are equivalent the expression (3') serves as a test statistic for the Neyman-Pearson test of the hypothesis

H_0: An observed sample path of X is realized under P

against the alternative

H_1: An observed sample path of X is realized under Q.

Thus one desires to be able to calculate $\frac{dQ}{dP}$ for each sample function of X. Note however that the expression (3') involves the integrals $\iint H(t,s)dX_t dX_s$ and $\int h(t)dX_t$ (respectively $\iint H(t,s)X_t X_s dtds$ and $\int h(t)X_t dt$) which are not defined for each individual sample path but in the mean square sense. It is therefore of interest to see if these integrals can be determined through sample paths. The answer is affirmative and we shall show this for $\iint H(t,s)dX_t dX_s$ only.

Suppose that X is continuous in probability. Let $\{t_1, t_2, \ldots\}$ be a dense subset of the interval T. Then it is known that $\overline{\mathcal{B}}(X) = \overline{\mathcal{B}}(X_{t_k}, k \geq 1)$ and thus by the martingale convergence theorem we have

$$E\left(\iint H(t,s)dX_t dX_s \mid X_{t_1}, \ldots, X_{t_n} \right) \to \iint H(t,s)dX_t dX_s, \quad \text{a.e. } [P] .$$

Orthonormalize $X_{t_i}, \ldots, X_{t_n}$ to obtain $\xi_1, \ldots, \xi_n$,

$$\xi_i = \sum_{j=1}^{i} b_{ij} X_{t_j}, \quad i \leq n.$$

Write

$$\xi_i = \int e_i(t) dX_t$$

where each e_i is a step function.

We may adjoint $e_{n+1}^{(n)}, e_{n+2}^{(n)}, \ldots$ to $e_1, \ldots, e_n$ such that the resulting sequence is orthonormal in $\Lambda_2(R)$ and satisfies

$$H = \sum \langle H, e_i \otimes e_j \rangle e_i \otimes e_j = \sum a_{ij} \, e_i \otimes e_j .$$

Then

$$\iint H(t,s)dX_t dX_s = \sum a_{ij} \, (\xi_i \xi_j - \delta_{ij}).$$

Now

$$(4) \quad E\left(\iint H(t,s)dX_t dX_s \mid X_{t_1}, \ldots, X_{t_n} \right) = E\left(\iint H(t,s)dX_t dX_s \mid \xi_1, \ldots, \xi_n \right)$$
$$= \sum a_{ij} \, E(\xi_i \xi_j - \delta_{ij} \mid \xi_1, \ldots, \xi_n)$$
$$= \sum_{i,j=1}^{n} a_{ij} \, (\xi_i \xi_j - \delta_{ij}).$$

Thus $\iint H(t,s)dX_t dX_s$ can be determined as the a.e. limits of (4) which is obtainable through each sample path.

From now on $X=(X_t, t\epsilon T)$ will be a coordinate SIP(0,R;F) on $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T), P)$. Recall that R is a covariance function and F is a probability distribution on $\mathbb{R}_+$ with finite first moment α_1. P is a mixture of Gaussian measures, specifically,

$$(5) \qquad P(E) = \int P_\alpha(E)dF(\alpha)$$

where each P_α $(\alpha\geq 0)$ is a zero mean Gaussian measure with covariance function αR.

Suppose X is nondegenerate and let A be the r.v. associated with X as defined in Chapter III. Let H(X) and $H_\alpha(X)$ be the linear space of X under P_α and R respectively. We will use the following

LEMMA 4. *If $\{\xi_i\}$ is a sequence in H(X) then we may take particular versions of each ξ, so that $\{\xi_i\}$ is a sequence in $H_\alpha(X)$ for every $\alpha\geq 0$. Conversely, if $\{\xi_i\}$ is a sequence in $H_{\alpha_0}(X)$ for a fixed α_0 then we may take particular versions of each ξ_i so that $\{\xi_i\}$ is a sequence in H(X) and $H_\alpha(X)$ for every $\alpha\geq 0$.*

PROOF. This is implicit in the proof of Lemma 3 in Chapter III. Q.E.D.

Now consider a second probability measure Q on $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ under which X is a SIP(m,S;G). Then

$$(6) \qquad Q(E) = \int Q_\alpha(E)dG(\alpha)$$

where each Q_α $(\alpha\geq 0)$ is a Guassian measure with mean m and covariance αS.

We are interested in the equivalence and mutual singularity of the measures P and Q.

THEOREM 5. *Suppose that* $P_1 \sim Q_1$. *Then*

(i) if $F \sim G$, $P \sim Q$

(ii) if $F \perp G$, $P \perp Q$

(iii) if neither $F \sim G$ *nor* $F \perp G$, *neither* $P \sim Q$ *nor* $P \perp Q$.

PROOF. (i) follows from (5), (6) and the fact that $P_\alpha \sim Q_\alpha$ for $\alpha \geq 0$ if $P_1 \sim Q_1$. Suppose $F \perp G$. Then there exists $E \in \mathcal{B}(\mathbb{R})$ such that $F(E) = G(E') = 1$. Since $A = \alpha$ a.e. $[P_\alpha]$ for $\alpha \geq 0$, we have

$$P(A \in E) = \int P_\alpha(A \in E)\, dF(\alpha) = F(E) = 1$$

and

$$\begin{aligned} Q(A \in E') &= \int Q_\alpha(A \in E')\, dG(\alpha) \\ &= \int P_\alpha(A \in E')\, dG(\alpha) \quad (P_\alpha \sim Q_\alpha) \\ &= G(E') = 1 \end{aligned}$$

which implies (ii).

The proof of (iii) is analogous to that of (ii). Q.E.D.

REMARK. This theorem was first stated in Gualtierotti [1974] for P and Q SIM's on a separable Hilbert space.

If $P \sim Q$, it is of interest to find an expression of $\frac{dQ}{dP}$ in terms of $\rho_\alpha = \frac{dQ_\alpha}{dP_\alpha}$ a.e. $[P_\alpha]$ and $\frac{dG}{dF}$. To this end we prepare the following

LEMMA 6. *Suppose* $P_1 \sim Q_1$ *and* $F \sim G$. *If* ρ_A *is a measurable function, then*

$$\frac{dQ}{dP}(x) = \rho_{A(x)}(x)\, \frac{dG}{dF}\big(A(x)\big) \quad \text{a.e. } [P]. \tag{7}$$

PROOF. For $E \in \mathcal{B}(\mathbb{R}^T)$,

$$\int_E \rho_A \circ \frac{dG}{dF}(A)\,dP = \int dF(\alpha) \circ \int_E \rho_A \circ \frac{dG}{dF}(A)\,dP_\alpha$$
$$= \int dF(\alpha) \int_E \rho_\alpha \circ \frac{dG}{dF}(\alpha)\,dP_\alpha$$
$$= \int Q_\alpha(E)\,dG(\alpha) = Q(E).$$

Thus $\frac{dQ}{dP} = \rho_A \circ \frac{dG}{dF}(A)$ a.e. $[P]$. Q.E.D.

REMARK. The measurability of ρ_A is not automatic since ρ_α can be arbitrarily changed on a set of P_α-measure 0.

In the following we assume that F satisfy (C_0).

THEOREM 7. *Suppose* $R=S$ *and* $F \sim G$. *Then either* $P \sim Q$ *or* $P \perp Q$. *In order that* $P \sim Q$, *it is necessary and sufficient that* m *belongs to the reproducing kernel Hilbert space of* R $\big(m \in \mathcal{R}(R)\big)$, *i.e., there exists* $Y \in H(X)$ *for which*

$$m(t) = \langle X_t, Y \rangle_{H(X)}. \tag{8}$$

When $P \sim Q$,

$$\frac{dQ}{dP} = e^{\frac{\alpha_1}{A}(Y - \frac{1}{2}EY^2)} \cdot \frac{dG}{dF}(A). \tag{9}$$

PROOF. Suppose $m \notin \mathcal{R}(R)$. Then by a well known theorem of Fortet [1973] we have $P \perp Q$. Conversely, assume (8). We may take a version of Y such that $Y \in H_\alpha(X)$ $\forall\ \alpha \geq 0$, (Lemma 4). Then

$$\langle X_t, \frac{\alpha_1}{\alpha} Y \rangle_{H_\alpha(X)} = \frac{\alpha_1}{\alpha} E_\alpha X_t Y$$
$$= E\, X_t Y = m(t).$$

Thus by Theorem 1 $P_\alpha \sim Q_\alpha$ and

$$P_\alpha = \frac{dQ_\alpha}{dP_\alpha} = e^{\frac{\alpha_1}{\alpha}Y - \frac{1}{2}E_\alpha(\frac{\alpha_1}{\alpha}Y)^2}$$

$$= e^{\frac{\alpha_1}{\alpha}(Y - \frac{1}{2}EY^2)} .$$

Now $P \sim Q$ follows from Theorem 5(i), and (9) follows from Lemma 6.

REMARK. Sytaya [1969] derived the expression (9) for P and Q. SIM's on a separable Hilbert space and $F=G$.

Two covariance functions R and S are said to be *equivalent* ($R \sim S$) if the corresponding zero mean Gaussian measures are equivalent. We now state the main result of this section.

THEOREM 8. *Suppose* $R \sim S$ *and* $F \sim G$. *Then either* $P \sim Q$ *or* $P \perp Q$; *and* $P \sim Q$ *if and only if* $m \in R(R)$.

When $P \sim Q$ *there exist* $Y \in H(X)$ *and* $U \in H(X) \hat{\otimes} H(X)$ *such that*

$$m(t) = \langle X_t, Y\rangle_{H(X)} + \langle X_t \hat{\otimes} Y, U\rangle_{H(X)\hat{\otimes}H(X)} , \tag{10}$$

$$S(t,s) - R(t,s) = \frac{1}{\alpha_1}\langle X_t \hat{\otimes} X_s, U\rangle_{H(X)\hat{\otimes}H(X)} \tag{11}$$

and the Hilbert-Schmidt operator U *on* $H(X)$ *does not admit* -1 *as its eigenvalue, and*

$$\frac{dQ}{dP} = [\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}} \, e^{-\frac{1}{2}\frac{\alpha_1}{A}\sum \frac{\langle Y,\xi_i\rangle^2_{H(X)}}{1-\lambda_i}} \circ e^{\frac{\alpha_1}{A}\left(\frac{1}{\sqrt{2}}\frac{\alpha_2^{\frac{1}{2}}}{A}Z+Y\right)} \circ \frac{dG}{dF}(A) \tag{12}$$

where $Z = \sum \lambda_i \, \xi_i \hat{\otimes} \xi_i \in H(X)\hat{\otimes}H(X)$, $\{\xi_i\}$ *is an orthonormal sequence in* $H(X)$, $U = \sum u_i \, \xi_i \hat{\otimes} \xi_i$, $(1-\lambda_i)(1+u_i) = 1$, *and* $Y = \sum \langle Y, \xi_i\rangle_{H(X)} \xi_i$.

PROOF. Bring in the probability measure Q' defined by $Q'(E) = Q(E+m)$, $E \in \mathcal{B}(\mathbb{R}^T)$. It is clear that under Q' X is a $SIP(0,S;G)$, thus $P\sim Q'$ because $R\sim S$ and $F\sim G$. From Theorem 7, either $Q\sim Q'$ or $Q\perp Q'$, and $Q\sim Q'$ if and only if $m\in\mathcal{R}(S)$. The first part of the theorem now follows by noting that $\mathcal{R}(S)=\mathcal{R}(R)$ since $R\sim S$.

Suppose $P\sim Q$. Then $Q'\sim Q$, hence by Theorem 7 $m\in\mathcal{R}(S)=\mathcal{R}(R)$ which implies $P_1\sim Q_1$. Thus it follows from Theorem 7 that there exist $Y_1\in H_1(X)$ and $U_1\in H_1(X)\hat{\otimes}H_1(X)$ such that

$$\text{(i)}\quad m(t) = \langle X_t, Y_1\rangle_{H_1(X)} + \langle X_t\hat{\otimes}Y_1, U_1\rangle_{H_1(X)\hat{\otimes}H_1(X)}$$

$$\text{(ii)}\qquad S(t,s)-R(t,s) = \langle X_t\hat{\otimes}X_s, U_1\rangle_{H_1(X)\hat{\otimes}H_1(X)},$$

U_1 does not admit -1 as its eigenvalue.

Let $\{\sqrt{\alpha_1}\,\xi_i\}$ be an orthonormal sequence in $H_1(X)$ such that

$$\text{(iii)}\quad U_1 = \sum u_i \sqrt{\alpha_1}\,\xi_i \,\hat{\otimes}\, \sqrt{\alpha_1}\,\xi_i = \alpha_1 \sum u_i\, \xi_i\hat{\otimes}\xi_i,$$

$$\text{(iv)}\quad Y_1 = \sum \langle Y_1, \sqrt{\alpha_1}\,\xi_i\rangle_{H_1(X)} \sqrt{\alpha_1}\,\xi_i = \alpha_1 \sum \langle Y_1,\xi_i\rangle_{H_1(X)}\xi_i.$$

We may assume by Lemma 4 that Y_1 and ξ_i $(i\geq 1)$ are elements in $H(X)$ and $H_\alpha(X)$ $\forall\ \alpha\geq 0$. Note that $E\,\xi_i^2 = \alpha_1\, E_1\,\xi_i^2 = 1$. Define

$$Y = \frac{1}{\alpha_1} Y_1 \in H(X)$$

$$U = \frac{\alpha_1}{(2\alpha_2)^{\frac{1}{2}}} \sum u_i\left(\xi_i^2 - \frac{A}{\alpha_1}\right)$$

$$= \sum u_i\ \xi_i\hat{\otimes}\xi_i \in H(X)\hat{\otimes}H(X)$$

$$Z = \frac{\alpha_1}{(2\alpha_2)^{\frac{1}{2}}} \sum \lambda_i\left(\xi_i^2 - \frac{A}{\alpha_1}\right)$$

$$= \sum \lambda_i\ \xi_i\hat{\otimes}\xi_i \in H(X)\hat{\otimes}H(X)$$

where $(1-\lambda_i)(1+u_i)=1\ \ \forall\ \ i\geq 1$.

Note that

$$\text{(v)}\quad \sum \langle Y,\xi_i\rangle_{H(X)}\xi_i = \alpha_1 \sum \langle Y,\xi_i\rangle_{H_1(X)}\xi_i = \sum \langle Y_1,\xi_i\rangle_{H_1(X)}\xi_i$$

$$= \frac{Y_1}{\alpha_1} = Y \qquad \text{(by (iv))}.$$

It is important to keep in mind that under each P_α

$$Y \in H_\alpha(X),$$

$$\text{(vi)}\quad U = \frac{\alpha_1}{\sqrt{2\alpha_2}} \sum u_i\left(\xi_i^2 - \frac{\alpha}{\alpha_1}\right) = \frac{\alpha_1}{\sqrt{\alpha_2}} \sum u_i\ \xi_i\hat{\otimes}\xi_i \in H_\alpha(X)\hat{\otimes}H_\alpha(X),$$

and

$$\text{(vii)}\quad Z = \frac{\alpha_1}{\sqrt{2\alpha_2}} \sum \lambda_i\left(\xi_i^2 - \frac{\alpha}{\alpha_1}\right) = \frac{\alpha_1}{\sqrt{\alpha_2}} \sum \lambda_i\ \xi_i\hat{\otimes}\xi_i \in H_\alpha(X)\hat{\otimes}H_\alpha(X).$$

In the following computations we will repeatedly use the fact that if $\xi,\eta\in H(X)$ and $H_\alpha(X)$ then

$$\langle\xi,\eta\rangle_{H(X)} = \frac{\alpha_1}{\alpha} \langle\xi,\eta\rangle_{H_\alpha(X)}.$$

We now show (10) and (11):

$$\langle X_t,Y\rangle_{H(X)} + \langle X_t\hat{\otimes}Y,U\rangle_{H(X)\hat{\otimes}H(X)} = \frac{1}{\alpha_1}\langle X_t,Y_1\rangle_{H(X)} + \langle X_t\hat{\otimes}Y, \sum u_i\ \xi_i\hat{\otimes}\xi_i\rangle_{H(X)\hat{\otimes}H(X)}$$

$$= \langle X_t,Y_1\rangle_{H_1(X)} + \sum u_i \langle X_t,\xi_i\rangle_{H(X)}\langle Y,\xi_i\rangle_{H(X)}$$

$$= \langle X_t,Y_1\rangle_{H_1(X)}$$

$$+ \alpha_1 \sum u_i\langle X_t,\xi_i\rangle_{H_1(X)} \langle Y_1,\xi_i\rangle_{H_1(X)}$$

$$= \langle X_t,Y_1\rangle_{H_1(X)}$$

$$+ \langle X_t\hat{\otimes}Y_1;\ \alpha_1 \sum u_i\ \xi_i\hat{\otimes}\xi_i\rangle_{H_1(X)\hat{\otimes}H_1(X)}$$

(cont.)

(cont.)

$$\langle X_t, Y\rangle_{H(X)} + \langle X_t \hat{\otimes} Y, U\rangle_{H(X)\hat{\otimes}H(X)} = m(t)$$

(by (i) and (iii);

$$\frac{1}{\alpha_1}\langle X_t \hat{\otimes} X_s, U\rangle_{H(X)\hat{\otimes}H(X)} = \frac{1}{\alpha_1}\sum u_i \langle X_t, \xi_i\rangle_{H(X)}\langle X_s, \xi_i\rangle_{H(X)}$$

$$= \alpha_1 \sum u_i \langle X_t, \xi_i\rangle_{H_1(X)}\langle X_s, \xi_i\rangle_{H_1(X)}$$

$$= \langle X_t \hat{\otimes} X_s, \alpha_1 \sum u_i \xi_i \hat{\otimes} \xi_i\rangle_{H_1(X)\hat{\otimes}H_1(X)}$$

$$= S(t,s) - R(t,s) \qquad \text{(by (ii), (iii)).}$$

In order to show (12), we will first find an expression for $\rho_\alpha = \frac{dQ_\alpha}{dP_\alpha}$. According to Theorem 1, it is necessary to determine the corresponding Y_α, U_α and Z_α for ρ_α. We claim that $Y_\alpha = \frac{\alpha_1}{\alpha} Y \in H_\alpha(X)$, $U_\alpha = \frac{\alpha_2^{1/2}}{\alpha} U = \frac{\alpha_1}{\alpha}\sum u_i \xi_i \hat{\otimes} \xi_i \in H_\alpha(X) \otimes H_\alpha(X)$ (by (iv)); for the same calculations with those carried out for $\alpha = 1$ give

$$\langle X_t, Y_\alpha\rangle_{H_\alpha(X)} + \langle X_t \hat{\otimes} Y_\alpha, U_\alpha\rangle_{H_\alpha(X)\hat{\otimes}H_\alpha(X)} = m(t).$$

and

$$\langle X_t \hat{\otimes} X_s, U_\alpha\rangle_{H_\alpha(X)\hat{\otimes}H_\alpha(X)} = \alpha S(t,s) - \alpha R(t,s).$$

Since under P_α, $U_\alpha = \sum u_i \left(\frac{\alpha_1}{\alpha}\right)^{1/2}\xi_i \hat{\otimes} \left(\frac{\alpha_1}{\alpha}\right)^{1/2}\xi_i$ and $\{\left(\frac{\alpha_1}{\alpha}\right)^{1/2}\xi_i\}$ is an orthonormal sequence in $H_\alpha(X)$; and since

$$\sum \langle Y_\alpha, \left(\frac{\alpha_1}{\alpha}\right)^{1/2}\xi_i\rangle_{H_\alpha(X)}\left(\frac{\alpha_1}{\alpha}\right)^{1/2}\xi_i = \frac{\alpha_1}{\alpha}\sum \langle Y, \xi_i\rangle_{H(X)}\xi_i$$

$$= \frac{\alpha_1}{\alpha} Y \qquad \text{(by (v))}$$

$$= Y_\alpha.$$

we have, by Theorem 1, under P_α

$$Z_\alpha = \sum \lambda_i \left(\frac{\alpha_1}{\alpha}\right)^{\frac{1}{2}} \xi_i \hat{\otimes} \left(\frac{\alpha_1}{\alpha}\right)^{\frac{1}{2}} \xi_i$$

$$= \frac{\alpha_1 \alpha_2^{\frac{1}{2}}}{\alpha^2} \qquad \text{(from (v)).}$$

Thus, it follows from Theorem 1 that

$$\rho_\alpha = \frac{dQ_\alpha}{dP_\alpha} = [\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}} \, e^{-\frac{1}{2}\sum \frac{\langle Y_\alpha, \left(\frac{\alpha_1}{\alpha}\right)^{\frac{1}{2}} \xi_i\rangle^2_{H_\alpha(X)}}{1-\lambda_i}} \circ e^{\frac{1}{\sqrt{2}} Z_\alpha + Y_\alpha}$$

$$= [\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}} \, e^{-\frac{1}{2}\frac{\alpha_1}{\alpha}\sum \frac{\langle Y, \xi_i\rangle^2_{H(X)}}{1-\lambda_i}} \circ e^{\frac{\alpha_1\alpha_2^{\frac{1}{2}}}{\sqrt{2}\alpha^2} Z + \frac{\alpha_1}{\alpha} Y}.$$

We can now conclude (12) from Lemma 6. Thus the theorem is proved.

Q.E.D.

As in the Gaussian case, Theorem 8 can be stated as follows.

THEOREM 9. *Suppose* $R \sim S$ *and* $F \sim G$. *Then either* $P \sim Q$ *or* $P \perp Q$, *and* $P \sim Q$ *if and only if* $m \in \mathcal{R}(R)$.

If $P \sim Q$ *and if, in addition,* $\alpha_2 < \infty$ *and* X *satisfies (I) (resp. (J)), then there exist* $h \in \Lambda_2(\alpha_1 R)$ *(resp.* $\lambda_2(\alpha_1 R)$ *) and* $K \in \Lambda_2(\alpha_1 R) \hat{\otimes} \Lambda_2(\alpha_1 R)$ *(resp.* $\lambda_2(\alpha_1 R) \hat{\otimes} \lambda_2(\alpha_1 R)$*) such that*

$$m(t) = \langle 1_t, h\rangle_{\Lambda_2(\alpha_1 R)} + \langle 1_t \hat{\otimes} h, K\rangle_{\Lambda_2(\alpha_1 R)\hat{\otimes}\Lambda_2(\alpha_1 R)}$$

$$\text{(resp. } \langle \delta_t, h\rangle_{\lambda_2(\alpha_1 R)} + \langle \delta_t \hat{\otimes} h, K\rangle_{\lambda_2(\alpha_1 R)\hat{\otimes}\lambda_2(\alpha_1 R)} \text{),}$$

$$S(t,s) - R(t,s) = \frac{1}{\alpha_1} \langle 1_t \hat{\otimes} 1_s, K\rangle_{\Lambda_2(\alpha_1 R)\hat{\otimes}\Lambda_2(\alpha_1 R)}$$

$$\text{(resp. } \frac{1}{\alpha_1} \langle \delta_t \hat{\otimes} \delta_s, K\rangle_{\lambda_2(\alpha_1 R)\hat{\otimes}\lambda_2(\alpha_1 R)} \text{)}$$

and the Hilbert-Schmidt operator K *does not admit* -1 *as its eigenvalue and*

$$\frac{dQ}{dP} = [\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}} \; e^{-\frac{1}{2}\frac{\alpha_1}{A}\sum \frac{\langle h,e_i\rangle^2_{\lambda_2(\alpha_1 R)}}{1-\lambda_i}} \circ e^{\frac{\alpha_1}{A}\left(\frac{\alpha_1}{2A}\int\int H(t,s)\,dX_t\,dX_s + \int h(t)\,dX_t\right)}$$

(resp. $[\Pi(1-\lambda_i)e^{\lambda_i}]^{\frac{1}{2}} \; e^{-\frac{1}{2}\frac{\alpha_1}{A}\frac{\langle h,e_i\rangle^2_{\lambda_2(\alpha_1 R)}}{1-\lambda_i}} \circ e^{\frac{\alpha_1}{A}\left(\frac{\alpha_1}{2A}\int\int H(t,s)X_t X_s\,dt\,ds + \int h(t)X_t\,dt\right)}$ *)*

where $H = \sum \lambda_i \, e_i \hat{\otimes} e_i \in \Lambda_2(\alpha_1 R)\hat{\otimes}\Lambda_2(\alpha_1 R)$ *(resp.* $\lambda_2(\alpha_1 R)\hat{\otimes}\lambda_2(\alpha_1 R)$ *)*, $\{e_i\}$ *is an orthonormal sequence in* $\Lambda_2(\alpha_1 R)$ *(resp.* $\lambda_2(\alpha_1 R)$ *)*, $K = \sum u_i \, e_i \hat{\otimes} e_i$, $(1-\lambda_i)(1+u_i)=1$, *and* $h = \sum \langle h,e_i\rangle_{\Lambda_2(\alpha_1 R)} e_i$ *(resp.* $\sum \langle h,e_i\rangle_{\lambda_2(\alpha_1 R)} e_i$ *)*.

PROOF. The proof is analogous to that of Theorem 2. Q.E.D.

2. A 0-1 Law

The following celebrated 0-1 law for Gaussian measures is due to Kallianpur (see Cambanis and Rajput [1973] or Le Page [1973]).

THEOREM 10. *Let* P *be a Gaussian measure on* $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ *and let* G *be a* $\overline{\mathcal{B}}(\mathbb{R}^T)$*-measurable (additive) subgroup of* $\mathbb{R}^T$*. Then* $P(G)=0$ *or* 1.

In the application of Theorem 10 it happens very often that P has zero mean and G is a linear subspace. Under these additional assumptions we can extend Theorem 10 to the following

THEOREM 11. *Let* P *on* $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ *be a* SIM(0,R;F) *with* $F(\{0\})=\varepsilon$, *and let* L *be a* $\overline{\mathcal{B}}(\mathbb{R}^T)$*-measurable linear subspace of* $\mathbb{R}^T$. *Then* $P(L)=\varepsilon$ *or* 1.

Furthermore, $P(L)=1$ *if and only if there exists* $\alpha>0$ *such that* L *is* $\overline{\mathcal{B}}^\alpha$*-measurable and* $P_\alpha(L)=1$. *(* $\overline{\mathcal{B}}^\alpha$, $\alpha\geq 0$, *denotes the completion of* $\mathcal{B}(\mathbb{R}^T)$ w.r.t. P_α *.)*

PROOF. L is $\overline{\mathcal{B}}(\mathbb{R}^T)$-measurable implies that there exist L^*, $N\in\mathcal{B}(\mathbb{R}^T)$ and $N^*\subset N$ such that $L=L^*\cup N^*$ and $P(N)=0$. Let E be the set of α's with $P_\alpha(N)=0$. Then $F(E)=1$ and $N^*\in\overline{\mathcal{B}}^\alpha(\mathbb{R}^T)$ $\forall\ \alpha\in E$. Thus $L=L^*\cup N^*\in\overline{\mathcal{B}}^\alpha(\mathbb{R}^T)$ and $P_\alpha(L)=P_\alpha(L^*)$ $\forall\ \alpha\in E$.

Now it follows that

(i) $P(L)=P(L^*) = \int P_\alpha(L^*)dF(\alpha) = P_0(L)\varepsilon + \int_{(0,\infty)\cap E} P_\alpha(L)dF(\alpha)$

We may assume that $\varepsilon<1$; otherwise, the theorem clearly holds. Fix $\alpha_0\in E\setminus\{0\}$. Since L is a linear space, we have

(ii) $P_\alpha(L) = P_{\alpha_0}\left(\left(\frac{\alpha_0}{\alpha}\right)^{\frac{1}{2}} L\right) = P_{\alpha_0}(L)$ $\forall\ \alpha\in E\setminus\{0\}$,

(iii) $P_0(L)=1$.

(i), (ii) and (iii) yield $P(L) = \varepsilon + P_{\alpha_0}(L)(1-\varepsilon)$. Thus $P(L)=\varepsilon$ or 1 since by Theorem 10 $P_{\alpha_0}(L)=0$ or 1.

The second assertion is now obvious. Q.E.D.

COROLLARY 12. *Let* P *on* $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ *be a* SIM(0,R;F) *satisfying* (C_0), *and let* L *be a* $\overline{\mathcal{B}}(\mathbb{R}^T)$*-measurable linear subspace of* $\mathbb{R}^T$. *Then* $P(L)=0$ *or* 1.

Furthermore, $P(L)=1$ *if and only if there exists* $\alpha>0$ *such that* L *is* $\overline{\mathcal{B}}^\alpha$*-measurable and* $P_\alpha(L)=1$.

By applying Corollary 12 we can obtain a number of 0-1 laws for path properties of a zero mean SIP. For example, we can show that if

$X=(X_t,\ t\epsilon T)$ is a separable zero mean SIP satisfying (C_0) where T is an interval on the real line then with probability 0 or 1 the paths of X are

(i) continuous on T

(ii) differentiable on T

(iii) absolutely continuous on every compact subinterval of T

(iv) bounded on T

(v) free of oscillatory discontinuities on T.

(Cambanis and Rajput [1973] proved these results for a separable Gaussian processes, but their proof is also available for the present case.)

Indeed most known 0-1 laws for Gaussian processes extend to zero mean SIP's since their properties correspond to linear subspaces. Also, by Corollary 12, necessary and sufficient conditions for each alternative (whenever known) remain the same for zero mean SIP's. Some sample path properties for which necessary and sufficient conditions are known are integrability, continuity (stationary case), absolute continuity, analyticity. For example, a necessary and sufficient condition for the p-th order sample integrability $(1\leq p<\infty)$ of a measurable Gaussian process (and hence also, SIP) with covariance function R is $\int_T R^{\frac{p}{2}}(t,t)dt < \infty$ [Rajput 1972]; a necessary and sufficient condition for the sample continuity of a separable stationary Gaussian (and hence also, SI) process is $\sum_{n\geq 1} \frac{1}{q^n} \sqrt{H(\frac{1}{q^n})} < \infty$ for some $q>1$ where H is the metric entropy for this process [Fernique 1974]. (H is defined at the end of this section.)

In order to illustrate the techniques employed in proving 0-1 laws, we sketch, in the following, the proof of the 0-1 law and of a sufficient condition for sample continuity of a SIP.

Let $X=(X_t,\ t\in[a,b])$, defined on $(\Omega,\overline{\mathcal{B}}(X),P)$, be a separable (w.r.t. closed sets) SIP $(0,R;F)$ satisfying (C_0). Let μ be the induced SIM on the sample space $(\mathbb{R}^{[a,b]},\mathcal{B}(\mathbb{R}^{[a,b]}))$. Then $\mu(E) = \int \mu_\alpha(E)dF(\alpha)\quad \forall\ E\in\mathcal{B}(\mathbb{R}^{[a,b]})$ where each μ_α is a Gaussian measure with zero mean and covariance αR.

Let $S\subset[a,b]$ be the separating set and let $\Omega_0\subset\Omega$ be the exceptional set in the definition of separability of X. Consider the following sets

$$M = \{\omega\in\Omega:\ X_t(\omega)\ \text{ is continuous}\}$$
$$M^* = \{\omega\in\Omega:\ X_t(\omega)\ \text{ is uniformly continuous on } S\}$$
$$L^* = \{x\in\mathbb{R}^{[a,b]}:\ x(\circ)\ \text{ is uniformly continuous on } S\}.$$

Then L^* is easily shown to be a $\mathcal{B}(\mathbb{R}^{[a,b]})$- measurable linear subspace of $\mathbb{R}^{[a,b]}$. Thus $M^* = X^{-1}(L^*)\in\mathcal{B}(X)$. By separability it follows that $M\in\overline{\mathcal{B}}(X)$, and $P(M)=P(M^*)=\mu(L^*)$, which is either 0 or 1 by Corollary 12. Therefore with probability 0 or 1 the paths of X are continuous.

To show that the best known sufficient condition for the sample continuity of Gaussian processes remains the same for SIP's, we need a lemma which is of independent interest.

<u>LEMMA 13.</u> *There exist probability measures* P_α, $\alpha\geq 0$, *on* $(\Omega,\mathcal{B}(X))$ *such that* $\mu_\alpha=P_\alpha\circ X^{-1}$. *Hence under each* P_α, X *is a zero mean Gaussian process with covariance function* $\alpha R(t,s)$.

<u>PROOF.</u> Let $C(\mathbb{R}^{[a,b]})\subset\mathcal{B}(\mathbb{R}^{[a,b]})$ be the class of all cylinder sets, which forms a field. Let $C(X) = X^{-1}\circ C(\mathbb{R}^{[a,b]})\subset\mathcal{B}(X)$. For each $D\in C(X)$, define $P_\alpha(D)=\mu_\alpha(E)$ if $D=X^{-1}(E)$. We claim that P_α is well-defined. It is not hard to see (by looking at their characteristic functions) that when restricted to a finite dimensional space $\mathbb{R}^n$ μ_α is

absolutely continuous w.r.t. μ. Thus if $D=X^{-1}(E)=X^{-1}(E_1)$, $E,E_1 \in C(\mathbb{R}^{[a,b]})$, then $\mu(E)=P(D)=\mu(E_1)$ and hence $\mu_\alpha(E)=\mu_\alpha(E_1)$. Consequently, P_α is a probability measure on the field $C(X)$ and hence it can be extended uniquely to a probability measure on $\mathcal{B}(X)$. Now it follows that $\mu_\alpha = P_\alpha \circ X^{-1}$ since they coincide on $C(\mathbb{R}^{[a,b]})$. Q.E.D

Suppose $(\xi_t, t\in[a,b])$ is a second order process. Define $d(t,s) = [E(\xi_t-\xi_t)^2]^{\frac{1}{2}}$. Then d is a pseudo-metric on $[a,b]$. Let $N(\varepsilon)$ be the smallest number of (closed) ε-balls (w.r.t. pseudo-metric d) which covers $[a,b]$. The <u>*metric entropy*</u> H of ξ is defined by $H(\varepsilon)=\log N(\varepsilon)$. The following result is well-known: If ξ is a separable measurable mean square continuous Gaussian process and if

$$(10) \qquad \exists\ \delta>0, \quad \int_0^\delta H(\varepsilon)d\varepsilon < \infty$$

then with probability 1 the sample paths of ξ are continuous.

Assume that X is also measurable and mean square continuous. We now show that (10) is a sufficient condition for the sample continuity of X. By Corollary 12 it suffices to show that (10) implies

$$(11) \qquad \exists\ \alpha_0>0, \quad \mu_{\alpha_0}(L^*) = 1.$$

Thus assume (10).

We first note that under (dF) almost all P_α X is a separable process since $\Omega_0 \in \overline{\mathcal{B}}^\alpha(X)$ and $P_\alpha(\Omega_0)=0$ for almost all α as before. It can be shown that under almost all α X is measurable. It is clear that X is mean square continuous under each P_α. Now pick $\alpha_0>0$ such that under P_{α_0} X is a separable, measurable, mean square continuous Gaussian process. Let us compare the metric entropies of X under P

and X under P_{α_0}. Note that

$$d^2_{\alpha_0}(t,s) = E_{\alpha_0}(X_t - X_s)^2 = \frac{\alpha_0}{\alpha_1} E\,(X_t - X_s)^2 = \frac{\alpha_0}{\alpha_1} d^2(t,s).$$

Thus $N_{\alpha_0}(\varepsilon) = N\left[\left(\frac{\alpha_1}{\alpha_0}\right)^{\frac{1}{2}} \varepsilon\right]$, i.e. $H_{\alpha_0}(\varepsilon) = H\left[\left(\frac{\alpha_1}{\alpha_0}\right)^{\frac{1}{2}} \varepsilon\right]$.

Now (10) yields that $\exists\ \delta > 0$,

$$\infty > \int_0^\delta H(\varepsilon)d\varepsilon = \left(\frac{\alpha_1}{\alpha_0}\right)^{\frac{1}{2}} \int_0^{\left(\frac{\alpha_0}{\alpha_1}\right)^{\frac{1}{2}}\delta} H\left[\left(\frac{\alpha_1}{\alpha_0}\right)^{\frac{1}{2}} \varepsilon\right] d\varepsilon = \left(\frac{\alpha_1}{\alpha_0}\right)^{\frac{1}{2}} \int_0^{\left(\frac{\alpha_1}{\alpha_0}\right)^{\frac{1}{2}}\delta} H_{\alpha_0}(\varepsilon)d\varepsilon$$

which implies $P_{\alpha_0}(M) = 1$. Thus $\mu_{\alpha_0}(L^*) = P_{\alpha_0}(M^*) = P_{\alpha_0}(M) = 1$ as was to be proved.

A similar result is obtained by Besson [1974] with a different approach, using the fact that a SIP is the product of a Gaussian process and an independent nonnegative r.v.

V. NONLINEAR ESTIMATION AND PREDICTION

Via the tensor product space structure of the nonlinear space, we are able to solve the general *nonlinear estimation problem* for SIP's (in particular for Gaussian processes), in the sense that we can reduce the nonlinear problem to a standard linear estimation problem, the theory of which has been well-developed. This is done in Section 1.

In Section 2 we introduce the concept of *super prediction* for a class of prediction problems and derive a lower bound for the mean square error of the nonlinear prediction.

The notation introduced in Chapter III is used here.

1. Nonlinear Estimation

Let $X = (X_t,\ t \epsilon T)$ be a second order process with 0 mean. We consider the following estimation problem: We observe X_t for $t \epsilon S$, a subset of T, and we want to estimate an L_2-functional θ of X based on the observations. We are interested in finding the best estimate $\hat{\theta}$, an L_2-functional of $(X_t,\ t \epsilon S)$ which minimizes the mean square error of estimation $E(\hat{\theta}-\theta)^2$.

It is well known that $\hat{\theta}$ can be obtained as the conditional expectation of θ given $(X_t,\ t \epsilon S)$

$$\hat{\theta} = E(\theta \mid X_t,\ t \epsilon S) .$$

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 need the following notation. Let $L_2(X;S) = L_2(X_t,\ t\in S)$, $H(X;S) = H(X_t,\ t\in S)$, and $\{\xi_\gamma,\ \gamma\in\Gamma\}$ be a CONS in $H(X)$ with Γ a linearly ordered set. Let $\{e_n,\ n=1,2,\ldots,N\}$ be a CONS in $L_2(dF)$ (where N may be ∞).

<u>THEOREM 1.</u> *Let* $X = (X_t,\ t\in T)$ *be a nondegenerate* SIP(0,R;F) *satisfying (M) and* (C_0). *Let* $\theta \in L_2(X)$ *have the following orthogonal development*

$$\theta = \sum_{n=1}^{N} \sum_{p\geq 0} \sum_{\substack{p_1+\ldots+p_k=p \\ \gamma_1<\ldots<\gamma_k}} {}_n a^{p_1\ldots p_k}_{\gamma_1\ldots\gamma_k}\, e_n(A)\otimes\left(\xi_{\gamma_1}^{\hat{\otimes}p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_{\gamma_2}^{\hat{\otimes}p_k}\right) .$$

Suppose $(X_t,\ t\in S)$ *is nondegenerate. Then*

$$\hat{\theta} = \sum_{n=1}^{N} \sum_{p\geq 0} \sum_{\substack{p_1+\ldots+p_k=p \\ \gamma_1<\ldots<\gamma_k}} {}_n a^{p_1\ldots p_k}_{\gamma_1\ldots\gamma_k}\, e_n(A)\otimes\left(\hat{\xi}_{\gamma_1}^{\hat{\otimes}p_1}\hat{\otimes}\ldots\hat{\otimes}\hat{\xi}_{\gamma_k}^{\hat{\otimes}p_k}\right)$$

where

$$\hat{\xi}_{\gamma_i} = \mathrm{Proj}_{H(X,S)}\ \xi_{\gamma_i} .$$

<u>PROOF.</u> 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

$$\text{(i)} \quad \rho = \sum_{m=1}^{N} \sum_{q\geq 0} \sum_{\substack{q_1+\ldots+q_j=q \\ \beta_1<\ldots<\beta_j}} {}_m b^{q_1\ldots q_j}_{\beta_1\ldots\beta_j}\, e_m(A) \otimes \left(\eta_{\beta_1}^{\hat{\otimes}p_1}\hat{\otimes}\ldots\hat{\otimes}\eta_{\beta_k}^{\hat{\otimes}p_k}\right) .$$

where $\{\eta_\beta,\ \beta\in\mathcal{B}\}$ is a CONS in $H(X;S)$.

We have

$$\hat{\theta} = E(\theta \mid X_t,\ t\in S) = \mathrm{Proj}_{L_2(X;S)}\theta$$

$$= \sum_{n=1}^{N} \sum_{p\geq 0} \sum_{\substack{p_1+\ldots+p_k=p \\ \gamma_1<\ldots<\gamma_k}} {}_n a^{p_1\ldots p_k}_{\gamma_1\ldots\gamma_k}\ \mathrm{Proj}_{L_2(X;S)}\ e_n(A) \otimes\left(\xi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes} p_k}\right) .$$

Thus to show the theorem it suffices to show

$$\text{(ii)}\qquad \mathrm{Proj}_{L_2(X;S)}\ e_n(A)\otimes\left(\xi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes} p_k}\right) = e_n(A)\otimes\left(\hat{\xi}_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\hat{\xi}_{\gamma_k}^{\hat{\otimes} p_k}\right) .$$

If $\xi_1,\ldots,\xi_p\in H(X)$ and $\eta_1,\ldots,\eta_p\in H(X;S)$ then it follows from $\langle\xi_i,\eta_j\rangle_{H(X)} = \langle\hat{\xi}_i,\eta_j\rangle_{H(X)}$ that

$$\text{(iii)}\qquad \langle\xi_1\hat{\otimes}\ldots\hat{\otimes}\xi_p,\eta_1\hat{\otimes}\ldots\hat{\otimes}\eta_p\rangle_{H^{\hat{\otimes} p}(X)} = \langle\hat{\xi}_1\hat{\otimes}\ldots\hat{\otimes}\hat{\xi}_p,\eta_1\hat{\otimes}\ldots\hat{\otimes}\eta_p\rangle_{H^{\hat{\otimes} p}(X)} .$$

Now

$$\langle e_n(A)\otimes\left(\xi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes} p_k}\right),\rho\rangle_{L_2(X)}$$

$$= \sum_{j=1}^{N} \sum_{q\geq 0} \sum_{\substack{q_1+\ldots+q_j=q \\ \gamma_1<\ldots<\gamma_j}} {}_m b^{q_1\ldots q_j}_{\beta_1\ldots\beta_j}\ \langle e_n(A)\otimes\left(\xi_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\xi_{\gamma_k}^{\hat{\otimes} p_k}\right),\ e_m(A)\otimes\left(\eta_{\beta_1}^{\hat{\otimes} q_1}\hat{\otimes}\ldots\hat{\otimes}\eta_{\beta_j}^{\hat{\otimes} q_j}\right)\rangle_{L_2(X)}$$

$$= \sum\sum\sum {}_m b^{q_1\ldots q_j}_{\beta_1\ldots\beta_j}\ \langle e_n(A)\otimes\left(\hat{\xi}_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\hat{\xi}_{\gamma_k}^{\hat{\otimes} p_k}\right),\ e_m(A)\otimes\left(\eta_{\beta_1}^{\hat{\otimes} q_1}\hat{\otimes}\ldots\hat{\otimes}\eta_{\beta_j}^{\hat{\otimes} q_j}\right)\rangle_{L_2(X)} \quad \text{(by (iii))}$$

$$= \langle e_n(A)\otimes\left(\hat{\xi}_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\hat{\xi}_{\gamma_k}^{\hat{\otimes} p_k}\right),\rho\rangle_{L_2(X)} .$$

Since $\rho\in L_2(X;S)$ is arbitrary and $e_n(A)\otimes\left(\hat{\xi}_{\gamma_1}^{\hat{\otimes} p_1}\hat{\otimes}\ldots\hat{\otimes}\hat{\xi}_{\gamma_k}^{\hat{\otimes} p_k}\right) \in L_2(X;S)$, (ii) follows. Q.E.D.

<u>REMARK</u>. If $\theta=\xi\in H(X)$ then it follows $E(\xi \mid X_t,\ t\in S) = \mathrm{Proj}_{H(X;S)}\xi$ which is also a characterizing property for SIP's.

COROLLARY 2. *Let* X *be a zero mean Gaussian process. Let* $\theta \in L_2(X)$ *have the following orthogonal development*

$$\theta = \sum_{p\geq 0} \sum_{\substack{p_1+\ldots+p_k=p \\ \gamma_1<\ldots<\gamma_k}} a^{p_1\ldots p_k}_{\gamma_1\ldots\gamma_k} \xi_{\gamma_1}^{\hat{\otimes}p_1} \hat{\otimes} \ldots \hat{\otimes} \xi_{\gamma_k}^{\hat{\otimes}p_k} .$$

Then

$$\hat{\theta} = \sum_{p\geq 0} \sum_{\substack{p_1+\ldots+p_k=p \\ \gamma_1<\ldots<\gamma_k}} a^{p_1\ldots p_k}_{\gamma_1\ldots\gamma_k} \hat{\xi}_{\gamma_1}^{\hat{\otimes}p_1} \hat{\otimes} \ldots \hat{\otimes} \hat{\xi}_{\gamma_k}^{\hat{\otimes}p_k}$$

where $\hat{\xi}_i = \mathrm{Proj}_{H(X;S)}\xi_i$ $(= E(\xi_i \mid X_t, t\in S))$.

COROLLARY 3. *If* X *is a zero mean Gaussian process then*

$$\begin{aligned}\mathrm{Proj}_{H^{\hat{\otimes}p}(X)} \circ \mathrm{Proj}_{L_2(X;S)} &= \mathrm{Proj}_{L_2(X;S)} \circ \mathrm{Proj}_{H^{\hat{\otimes}p}(X)} \\ &= \mathrm{Proj}_{H^{\hat{\otimes}p}(X;S)} .\end{aligned}$$

EXAMPLES. Suppose X is a SIP(0,R;F) satisfying (M) and (C_0). Consider the nonlinear estimation of $\theta = f(A)H_{p,||\xi||^2}\left(\left(\frac{\alpha_1}{A}\right)^{\frac{1}{2}}\xi\right)$, $\xi\in H(X)$ and $f(A)\in L_2(A)$. Write

$$f(A)H_{p,||\xi||^2}\left(\left(\frac{\alpha_1}{A}\right)^{\frac{1}{2}}\xi\right) = (p!)^{\frac{1}{2}} f(A)\otimes(\xi^{\hat{\otimes}p}) .$$

(Theorem 13, Chapter III)

By Theorem 1,

$$\hat{\theta} = (p!)^{\frac{1}{2}} f(A)\otimes(\hat{\xi}^{\hat{\otimes}p}) = f(A)H_{p,||\xi||^2}\left(\left(\frac{\alpha_1}{A}\right)^{\frac{1}{2}}\hat{\xi}\right) .$$

If $f(A) = \left(\frac{A}{\alpha_1}\right)^{\frac{p}{2}}$ then $\theta = H_{p,\frac{A}{\alpha_1}||\xi||^2}(\xi)$ and

$$\hat{\theta} = H_{p,\frac{A}{\alpha_1}||\hat{\xi}||^2}(\hat{\xi}) .$$

Next consider the nonlinear estimation of an exponential function $\theta = e^{\xi - \frac{A}{2\alpha_1}||\xi||^2}$. We have

$$\hat{\theta} = E(e^{\xi - \frac{A}{2\alpha_1}||\xi||^2} \mid X_t,\ t\epsilon S)$$

$$= \sum \frac{1}{p!} E(H_{p,\frac{A}{\alpha_1}||\xi||^2}(\xi) \mid X_t,\ t\epsilon S)$$

$$= \sum \frac{1}{p!} H_{p,\frac{A}{\alpha_1}||\hat{\xi}||^2}(\hat{\xi})$$

$$= e^{\hat{\xi} - \frac{A}{2\alpha_1}||\hat{\xi}||^2} .$$

In the most important case where X is Gaussian we have

$$E\big(H_{p,||\xi||^2}(\xi) \mid X_t,\ t\epsilon S\big) = H_{p,||\hat{\xi}||^2}(\hat{\xi}) ,$$

$$E(e^{\xi - \frac{1}{2}||\xi||^2} \mid X_t,\ t\epsilon S) = e^{\hat{\xi} - \frac{1}{2}||\hat{\xi}||^2} .$$

<u>COROLLARY 4.</u> *Let* X *be a zero mean Gaussian martingale. Then*

$$(Y_t = H_{p,||X_t||^2}(X_t),\ t\epsilon T) ,$$

$$(Z_t = e^{X_t - \frac{1}{2}||X_t||^2},\ t\epsilon T)$$

are martingales.

<u>PROOF.</u> This follows readily from the above example. Q.E.D.

<u>REMARK.</u> For X a Wiener process it is well known that $(X_t^2 - t,\ t\geq 0)$ and $(e^{X_t - \frac{t}{2}},\ t\geq 0)$ are martingales.

In the following we shall use MWI's to present another solution of the nonlinear estimation problem. For simplicity as well as for practical purposes we will consider the Gaussian case only.

Let $X=(X_t,\ t\in T)$ be a zero mean Gaussian process satisfying (I) or (J). We will then employ the orthogonal development in terms of I_p in the former case and J_p in the latter case.

For f a function on S^p, S a subinterval of T, let $\overline{f}$ denote the obvious extension of f to T^p (i.e. $\overline{f}=f$ on S^p and $\overline{f}=0$ on $T^p\backslash S^p$). We will consider the Λ_2-spaces and λ_2-spaces corresponding to $(X_t,\ t\in T)$ and $(X_t,\ t\in S)$; and we denote them by $\Lambda_2(R)$, $\Lambda_2(R;S)$ and $\lambda_2(R)$, $\lambda_2(R;S)$ respectively. Since

$$
\begin{aligned}
\langle f,g\rangle_{\Lambda_2(R;S)} &= \int_S \int_S f(t)g(s)\ d^2R(t,s) \\
&= \int_T \int_T \overline{f}(t)\overline{g}(s)\ d^2R(t,s) \\
&= \langle \overline{f},\overline{g}\rangle_{\Lambda_2(R)}
\end{aligned}
$$

for all f,g step function on S, we have

$$
\langle f,g\rangle_{\otimes^p\Lambda_2(R;S)} = \langle \overline{f},\overline{g}\rangle_{\otimes^p\Lambda_2(R)}
$$

for all f,g step functions on S^p which form a dense subspace of $\otimes^p\Lambda_2(R;S)$. Therefore we can embed $\otimes^p\Lambda_2(R;S)$ as a closed subspace of $\otimes^p\Lambda_2(R)$ by identifying each step function f on S^p with $\overline{f}$ on T^p. Moreover, we have

$$
\int_S \cdots \int_S f(t_1,\ldots,t_p)dX_{t_1}\cdots dX_{t_p} = \int_T \cdots \int_T \overline{f}(t_1,\ldots,t_p)dX_{t_1}\cdots dX_{t_p}
$$

since this is true for f step functions.

Similarly we can conclude that $\otimes^p\lambda_2(R;S)$ is a closed subspace of

$\otimes^p \lambda_2(R)$ and

$$\int_S \cdots \int_S f(t_1,\ldots,t_p)\, X_{t_1}\cdots X_{t_p}\, dt_1\ldots dt_p$$
$$= \int_T \cdots \int_T \bar{f}(t_1,\ldots,t_p)\, X_{t_1} \cdots X_{t_p}\, dt_1\ldots dt_p$$

for all $f \in \otimes^p \Lambda_2(R;S)$.

<u>THEOREM 5.</u> *Let* $X=(X_t,\ t\in T)$ *be a Gaussian process satisfying (I) and let* $\theta\in L_2(X)$ *have the orthogonal development*

$$\theta = \sum_{p\geq 0} I_p(f_p)\ ,\quad f_p \in \otimes^p \Lambda_2(R)\ .$$

Then

$$E(\theta \mid X_t,\ t\in S) = \sum_{p\geq 0} I_p\ (f_p^*)$$

where S *is a subinterval of* T *and*

$$f_p^* = \mathrm{Proj}_{\otimes^p \Lambda_2(R;S)}\ f_p\ .$$

<u>PROOF.</u> f_p^* is well-defined since $\otimes^p\Lambda_2(R;S)$ is a closed subspace of $\otimes^p\Lambda_2(R)$. Also $I_p(f_p^*) \in L_2(X;S)$.

Consider $\rho\in L_2(X;S)$ with the orthogonal development $\rho = \sum I_p(g_p)$, $g_p \in \otimes^p\Lambda_2(R;S)$. Then

$$\begin{aligned} E\theta\rho &= \left\langle \sum I_p(f_p), \sum I_p(g_p)\right\rangle_{L_2(X)} \\ &= \sum p!\ \langle \tilde{f}_p, \tilde{g}_p\rangle_{\otimes^p\Lambda_2(R)} \\ &= \sum p!\ \langle f_p, \tilde{g}_p\rangle_{\otimes^p\Lambda_2(R)} \\ &= \sum p!\ \langle f_p^*, \tilde{g}_p\rangle_{\otimes^p\Lambda_2(R)} \quad \left(g_p \in \otimes^p\Lambda_2(R;S)\right) \end{aligned}$$

(cont.)

(cont.)

$$E\theta\rho = \sum p! \langle \tilde{f}_p^*, \tilde{g}_p \rangle_{\otimes^p \Lambda_2(R)}$$

$$= \langle \sum I_p(f_p^*), \sum I_p(g_p) \rangle_{L_2(X)}$$

which implies

$$E(\theta \mid X_t,\ t\in S) = \mathrm{Proj}_{L_2(X;S)}(\theta) = \sum I_p(f_p^*) .$$

Q.E.D.

REMARK. $f_p^* = \mathrm{Proj}_{\otimes^p \Lambda_2(R;S)} f_p$ can be obtained as follows. Pick a CONS $\{e_\gamma,\ \gamma\in\Gamma\}$ in $\Lambda_2(R)$ and expand f_p into its orthogonal development $f_p = \sum a_{\gamma_1\ldots\gamma_p} e_{\gamma_1}\otimes\ldots\otimes e_{\gamma_p}$. Then it can be shown as in Theorem 1 that $f_p^* = \sum a_{\gamma_1\ldots\gamma_p} e^*_{\gamma_1}\otimes\ldots\otimes e^*_{\gamma_p}$ where $e^*_{\gamma_i} = \mathrm{Proj}_{\Lambda_2(R;S)} e_{\gamma_i}$.

In general, the explicit form of f_p^* is very difficult to determine. However, when $\Lambda_2(R)$ is an L_2-space (e.g. X is a Wiener process) f_p^* is simply the restriction of f_p to S^p.

COROLLARY 6. *Let* $X=(X_t,\ t\in T)$ *be a Gaussian process satisfying (I) and let* $B_t = B(X_s,\ t\geq s\in T)$. *Then* $(M_t, B_t, t\in T)$ *is a square-integrable martingale if and only if*

$$M_t = \sum_{p\geq 0} I_p(\mathrm{Proj}_{\otimes^p \Lambda_2^t(R)} f_p) ,$$

$f_p \in \otimes^p \Lambda_2(R)$ *and* $\Lambda_2^t(R) = \Lambda_2(R;\ \{s\in T:\ s\leq t\})$.

PROOF. $(M_t, B_t, t\in T)$ is an square-integrable martingale if and only if $M_t = E(\theta \mid B_t)$, $\theta \in L_2(X)$, [Meyer 1966; p. 163]. So the assertion follows from Theorem 5.

Q.E.D.

THEOREM 5'. *Let* $X=(X_t,\ t\in T)$ *be a Gaussian process satisfying (J) and let* $\theta\in L_2(X)$ *have the orthogonal development*

$$\theta = \sum_{p\geq 0} J_p(f_p)\ ,\quad f_p \in \otimes^p\lambda_2(R)\ .$$

Then

$$E(\theta \mid X_t,\ t\in S) = \sum_{p\geq 0} J_p(f_p^*)$$

where S *is a subinterval of* T *and*

$$f_p^* = \mathrm{Proj}_{\otimes^p\lambda_2(R;S)}\ f_p\ .$$

COROLLARY 6'. *Let* $X=(X_t,\ t\in T)$ *be a Gaussian process satisfying (J) and let* $\mathcal{B}_t = \mathcal{B}(X_s,\ t\geq s\in T)$. *Then* $(M_t,\ \mathcal{B}_t,\ t\in T)$ *is a square-integrable martingale if and only if*

$$M_t = \sum_{p\geq 0} J_p(\mathrm{Proj}_{\otimes^p\lambda_2^t(R)}\ f_p)\ ,$$

$f_p \in \otimes^p\lambda_2(R)$ *and* $\lambda_2^t(R) = \lambda_2(R;\ \{s\in T:\ s\leq t\})$.

2. Nonlinear Prediction

Consider the following prediction problem for a certain class of processes: Let $X=(X_t,\ t\in T)$, T an interval, be a second order process and let $Y_t=\theta_t(X_t)$ with θ_t a real function such that $EY_t^2<\infty\ \forall\ t\in T$. Suppose that on the basis of the past values of $Y=(Y_t,\ t\in T)$ up to time t we want to find the best prediction of the future value $Y_{t+\tau}$, $\tau>0$.

Two predictions are of special interest: the optimal linear prediction $\hat{Y}_t^{\ell}(\tau)$ and the optimal nonlinear prediction $\hat{Y}_t^{n\ell}(\tau)$. The optimality is in the sense of minimizing the mean square error within the class of all linear and nonlinear predictions respectively. It is well known that

$$\hat{Y}_t^{n\ell}(\tau) = E(Y_{t+\tau} \mid Y_s,\ s\le t) ,$$

$$\hat{Y}_t^{\ell}(\tau) = \mathrm{Proj}_{H(Y_s;\, s\le t)} Y_{t+\tau} .$$

We denote the corresponding mean square prediction errors by

$$\sigma_{n\ell}^2 = E(\hat{Y}_t^{n\ell}(\tau)-Y_{t+\tau})^2 ,$$

$$\sigma_{\ell}^2 = E(\hat{Y}_t^{\ell}(\tau)-Y_{t+\tau})^2 .$$

Now introduce a *super prediction* $\hat{Y}_t^s(\tau)$ to be the nonlinear prediction of $Y_{t+\tau}$ based on $(X_s,\ s\le t)$, i.e.

$$\hat{Y}_t^s(\tau) = E(Y_{t+\tau} \mid X_s,\ s\le t) ,$$

whose mean square prediction error is denoted by σ_s^2. It is clear that

$$\sigma_s^2 \le \sigma_{n\ell}^2 \le \sigma_\ell^2$$

and thus σ_s^2 is a lower bound for the mean square errors of linear and nonlinear predictions. If X is a SIP then σ_s^2 can be obtained by solving an estimation problem (cf. Section 1). If, in addition, for each t, θ_t happens to be a 1-1 function then the σ-fields generated by X_t and Y_t coincide. In this case

$$\hat{Y}_t^{n\ell}(\tau) = \hat{Y}_t^s(\tau) = E(Y_{t+\tau} \mid X_s,\ s\le t) ,$$

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

In the important particular 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 $\theta_t=\theta$ is a polynomial we can calculate the lower bound σ_s^2 as follows.

Write

$$Y_t = \theta(X_t) = a_1 H_{1,\sigma^2}(X_t)+\ldots+a_n H_{n,\sigma^2}(X_t)$$

(where H_{p,σ^2}, $1\le p\le n$, are the Hermite polynomials and $\sigma^2 = EX_0^2$). Note that

$$\begin{aligned} E\, H_{p,\sigma^2}(X_t)\, H_{p,\sigma^2}(X_s) &= p!\ \langle \hat{X}_t^{\otimes p}, \hat{X}_s^{\otimes p}\rangle_{L_2(X)} \\ &= p!\ \langle X_t^{\otimes p}, X_s^{\otimes p}\rangle_{\otimes^p H(X)} \\ &= p!\ \langle X_t, X_s\rangle_{H(X)}^p \\ &= p!\ R^p(t,s) \end{aligned}$$

and

$$\begin{aligned} E\, H_{p,\sigma^2}(X_t) H_{q,\sigma^2}(X_s) &= \langle \hat{X}_t^{\otimes p}, \hat{X}_s^{\otimes q}\rangle_{L_2(X)} \\ &= 0 \quad \text{if} \quad p\ne q\ . \end{aligned}$$

Thus the covariance function of Y is

$$\begin{aligned} E\, Y_t Y_s &= \sum_{p,q=1}^{n} a_p a_q\ E\big(H_{p,\sigma^2}(X_t) H_{q,\sigma^2}(X_t)\big) \\ &= \sum_{p=1}^{n} p!\ a_p^2\ R^p(t-s)\ . \end{aligned}$$

Let $\hat{X}_t(\tau) = E(X_{t+\tau} \mid X_s,\ s\le t)$ be the optimal nonlinear prediction of $X_{t+\tau}$ (which is also the optimal linear prediction since X is Gaussian), and let σ_0^2 be the mean square prediction error. Then

$$\begin{aligned} \hat{Y}_t^s(\tau) &= E(Y_{t+\tau} \mid X_s,\ s\le t) \\ &= \sum_{p=1}^{n} a_p\ H_{p,||\hat{X}_t(\tau)||^2}\big(\hat{X}_t(\tau)\big) \end{aligned}$$

(see Example in Section 1)

and

$$\sigma_s^2 = E\left(Y_{t+\tau}-\hat{Y}_t^s(\tau)\right)^2$$
$$= E\, Y_{t+\tau}^2 - E\left(\hat{Y}_t^s(\tau)\right)^2$$
$$= \sum_1^n p!\ a_p^2 \sigma^{2p} - \sum_1^n p!\ a_p^2\ (\sigma^2-\sigma_0^2)^p$$
$$= \sum_1^n p!\ a_p^2 [\sigma^{2p} - (\sigma^2-\sigma_0^2)^p]\ .$$

It is well known from the general theory of stationary process [Doob 1953; Gikhman and Skorokhod 1969] that σ_0^2 can be obtained analytically (if not explicitly) through the Wiener-Paley Factorization Theorem if X is *regular* (i.e. $\bigcap_{t\in\mathbb{R}} H(X_t;\ s\le t) = \{0\}$). When X is regular we now show that so is Y, and hence σ_ℓ^2 can be obtained analytically.

Y is clearly stationary. To prove that Y is regular we invoke the one-side moving average representation for X, i.e.,

$$X_t = \int_{-\infty}^{t} C(t-u)\,dW_u$$

where $C(u)$ is a square integrable function vanishing on positive real line and W is a Wiener process. Let $\mathcal{B}_t = \mathcal{B}(W_u,\ u\le t)$ and $\mathcal{B}_{-\infty} = \bigcap_t \mathcal{B}_t$. Then $\mathcal{B}_{-\infty}$ is trivial [Doob 1953]. X_t is $\mathcal{B}_t$-measurable and hence Y_t is $\mathcal{B}_t$-measurable. Suppose $\xi \in \bigcap_t H(Y_s,\ s\le t)$. Then ξ is $\mathcal{B}_t$-measurable for all t, therefore it is $\mathcal{B}_{-\infty}$-measurable which implies that ξ is a constant. But $E\xi=0$, so $\xi=0$. Thus Y is regular. (The regularity of Y also follows from a general result which states that for a Gaussian process the regularity condition is equivalent to the triviality of the remote past.)

Summing up these results, we now state the following

THEOREM 7. *Let* $X=(X_t, t\in\mathbb{R})$ *be a zero mean stationary Gaussian process with covariance function* $R(t-s)$ *and* $R(0)=\sigma^2$. *Let* $Y_t = \sum_1^n a_p H_{p,\sigma^2}(X_t)$, $t\in\mathbb{R}$. *Then*

(i) $Y=(Y_t, t\in\mathbb{R})$ *is a zero mean stationary process with covariance function* $\sum_1^n p!\ a_p^2\ R^p(t-s)$; *moreover,* Y *is regular if* X *is regular.*

(ii) We have

$$\sigma_s^2 \le \sigma_{n\ell}^2 \le \sigma_\ell^2$$

where $\sigma_s^2 = \sum_1^n p!\ a_p^2\ [\sigma^{2p}-(\sigma^2-\sigma_0^2)^p]$.

REMARK: The corresponding theorem for a stationary Gaussian sequence also holds.

Jaglom [1970] has considered the problem of comparing the performance of optimal linear and nonlinear predictions for polynomial functionals of certain stationary Markov processes. Donelson III and Maltz [1972] studied this problem in detail for polynomial functionals of the Ornstein-Uhlenbeck process. The inequality (ii) in Theorem 7 plays a central role in such studies.

As an example, consider $(X_t, t\in\mathbb{R})$ the Ornstein-Uhlenbeck process, i.e. a zero mean Gaussian process with covariance function $R(t,s)=e^{-|t-s|}$. By the Markov property we have

$$\hat{X}_t(\tau) = E\ (X_{t+\tau}\ |\ X_s,\ s\le t) = e^{-\tau}X_t\ .$$

Thus

$$E\ \left(H_p(X_{t+\tau})\ |\ X_s,\ s\le t\right) = H_{p,||\hat{X}_t(\tau)||^2}\left(\hat{X}_t(\tau)\right)$$

(cont.)

$$E\left(H_p(X_{t+\tau}) \mid X_s,\ s\le t\right) = H_{p,e^{-2\tau}}(e^{-\tau}X_t)$$
$$= e^{-p\tau} H_p(X_t)\ ,$$

and hence

$$\hat{Y}_t^s(\tau) = \sum_1^n a_p\ e^{-p\tau}\ H_p(X_t)\ .$$

So

$$\sigma_s^2 = E\ Y_{t+\tau}^2 - E\left(\hat{Y}_t^s(\tau)\right)^2 = \sum_1^n p!\ a_p^2\ (1-e^{-2p\tau})$$

which provides a lower bound (sometimes the greatest lower bound, e.g., $Y_t=H_p(X_t)$) to the mean square prediction error of $Y_{t+\tau}$.

This result has been obtained by Donelson III and Maltz using a different approach, who also compared σ_s^2 to σ_ℓ^2 and found that they are frequently very closed to each other.

VI. NONLINEAR NOISE

A (strictly) stationary process $Y=(Y_t, -\infty<t<\infty)$ with $EY_t=0$ and $EY_t^2<\infty$ is called *noise*. Wiener [1958] liked to think of such a noise as the output of a "black box" θ: We put in a white noise $\mathring{W}=(\mathring{W}_t, -\infty<t<\infty)$ (formally the derivative of a Wiener process) and we get $Y_0=\theta(\mathring{W}_t, -\infty<t<\infty)$ out; the noise $(Y_t, -\infty<t<\infty)$ is produced by shifting the input by the flow of the white noise $\mathring{W}(\circ) \to \mathring{W}(\circ+t)$. In order for Y to be a noise we require that $E\theta=0$ and $E\theta^2<\infty$.

Since θ has the orthogonal development (cf. III. 2.)

$$(1) \qquad \theta = \sum_{p\geq 1} \int \cdots \int f_p(t_1,\ldots,t_p)\, \mathring{W}_{t_1}\cdots\mathring{W}_{t_p}\, dt_1\ldots dt_p$$

$$\left(= \sum_{p\geq 1} \int \cdots \int f_p(t_1,\ldots,t_p)\, dW_{t_1}\ldots dW_{t_p} \right),$$

where $f_p \in L_2(\mathbb{R}^p)$, the noise Y obtained by shifting the incoming white noise through the nonlinear device θ can be expressed as

$$(2) \quad Y_t = \sum_{p\geq 1} \int \cdots \int f_p(t_1-t,\ldots,t_p-t)\, \mathring{W}_{t_1}\cdots\mathring{W}_{t_p}\, dt_1\ldots dt_p \;,$$

and the covariance function of Y is readily seen to be

$$\begin{aligned} E\, Y_tY_s &= E\, Y_0Y_\tau \quad (\tau=t-s) \\ &= \sum_{p\geq 1} \int \cdots \int f_p(t_1,\ldots,t_p)\, f_p(t_1-\tau,\ldots,t_p-\tau)\, dt_1\ldots dt_p \;. \end{aligned}$$

Wiener's theory of nonlinear noise starts from this idea. He also proved a profound theorem which was clarified by Nisio [1960] and which states that every *ergodic* noise can be approximated in law by noises of the form (2). Note that not every ergodic noise has the representation (2), and a necessary condition is <u>*strong mixing*</u> [McKean, 1973]. (For a discussion of the ergodicity of stationary processes see Doob [1953].)

More generally, instead of sending white noise through a nonlinear device θ, we may send an arbitrary mean square continuous Gaussian noise $X=(X_t, -\infty<t<\infty)$ with covariance R. Then the noise Y obtained by shifting the incoming Gaussian noise X can be expressed as

$$(3)\quad Y_t = \sum_{p\geq 1} \int \cdots \int f_p(t_1-t,\ldots,t_p-t)\; X_{t_1}\cdots X_{t_p}\; dt_1\cdots dt_p$$

where $f_p \in \lambda_2(\otimes^p R)$, and the covariance function of Y is again readily seen to be

$$\begin{aligned}(4)\quad E\,Y_tY_s &= E\,Y_0Y_\tau \quad (\tau=t-s)\\ &= \sum_{p\geq 1} \iint \cdots \iint f_p(t_1,\ldots,t_p) f_p(s_1-\tau,\ldots,s_p-\tau) R(t_1,s_1)\cdots \\ &\qquad\qquad R(t_p,s_p)\; dt_1 ds_1\cdots dt_p ds_p \\ &= \sum_{p\geq 1} p!\, \langle f_p(\circ,\ldots,\circ),\, f_p(\circ-\tau,\ldots,\circ-\tau)\rangle_{\Lambda_2(\otimes^p R)}\end{aligned}$$

(cf. II. 3.).

Although (2) and (3) are intuitively clear, they require proof. The proof of (3) follows from the following property (and (2) is shown similarly).

<u>LEMMA 1.</u> *If* X *is a Gaussian noise with covariance* R *then, for* $f_p \in \lambda_2(\otimes^p R)$,

$$\int \cdots \int f_p(t_1,\ldots,t_p)\; X_{t_1+t}\cdots X_{t_p+t}\; dt_1\cdots dt_p = \int \cdots \int f_p(t_1-t,\ldots,t_p-t)\; X_{t_1}\cdots X_{t_p}\; dt_1\cdots dt_p\,.$$

PROOF. Both integrals are well-defined since X is stationary. Pick a CONS $\{\phi_\gamma, \gamma\epsilon\Gamma\}$ in $\lambda_2(R)$. Since $\{\phi_{\gamma_1}^{\hat{\otimes}p_1}\hat{\otimes}\cdots\hat{\otimes}\phi_{\gamma_k}^{\hat{\otimes}p_k}: \gamma_1,\ldots,\gamma_k\epsilon\Gamma, p_1+\ldots+p_k=p, k\geq 0\}$ is complete in $\lambda_2(\hat{\otimes}^p R)$ and $J_p(f_p) = J_p(\tilde{f}_p)$, it suffices to prove this assertion for $f_p = \phi_{\gamma_1}^{\hat{\otimes}p_1}\hat{\otimes}\cdots\hat{\otimes}\phi_{\gamma_k}^{\hat{\otimes}p_k}$. But for such f_p, the assertion becomes

$$H_{p_1,||\phi_1||^2}\left(\int\phi_1(t_1)X_{t_1+t}dt_1\right)\cdots H_{p_k,||\phi_k||^2}\left(\int\phi_k(t_k)X_{t_k+t}dt_k\right)$$

$$= H_{p_1,||\phi_1||^2}\left(\int\phi_1(t_1-t)X_{t_1}dt_1\right)\cdots H_{p_k,||\phi_k||^2}\left(\int_k(t_k-t)X_{t_k}dt_k\right)$$

and thus we only need to show that

$$\int \phi(u)X_{u+t}du = \int \phi(u-t)X_u du \ .$$

This is true for $\phi\epsilon S_I$ and hence for $\phi \epsilon \lambda_2(R)$. The proof is now complete. Q.E.D.

When Y has representation (3), we say that Y is *X-presentable*. Note that X is always X-presentable since $X_t = \int \delta_0(u-t)X_u du$. If X is not strongly mixing then X is not white noise-presentable, however, it is X-presentable.

In the following we shall prove the analogue of Wiener-Nisio's theorem using Nisio's approach as simplified (for convergence in law) by McKean [1973].

Let X be a sample continuous ergodic Gaussian noise, and let $x=(x(t), -\infty<t<\infty)$ be the corresponding coordinate process. We assume that X satisfies the following condition (S):

(S) $$\Pr(X_t>0,\ 0\le t\le n) > 0 \quad \forall\ n\ge 1 .$$

THEOREM 2. *Let* Y *be a measurable ergodic noise (defined on any probability space). Then* Y *can be approximated in law by a sequence of* X-*presentable noises.*

The proof of Theorem 2 requiring several auxillary results will be given after the following theorem which shows that not every ergodic noise is X-presentable.

THEOREM 3. *If* X *is a Gaussian noise with absolutely continuous spectral distribution then every* X-*presentable noise* Y *is strongly mixing.*

PROOF. Having introduced the Fourier transform on $\Lambda_2(\otimes^p R)$ in II. 3., the proof is similar to Mckean's proof for X white noise.

We need to show that for $A,B \in \mathcal{B}(\mathbb{R}^{\mathbb{R}})$

$$\lim_{\tau\to\infty} \Pr(Y\in A,\ Y^\tau\in B) = \Pr(Y\in A) \circ \Pr(Y\in B)$$

where Y^τ denotes the shift of Y by τ, i.e. $Y^\tau_t = Y_{t+\tau}$. In fact we will show that for $\theta,\rho \in L_2(Y)$

$$\lim_{\tau\to\infty} E\,\theta\rho^\tau = E\theta \circ E\rho$$

where ρ^τ denotes the functional of shifted paths, i.e. $\rho^\tau(Y) = \rho(Y^\tau)$. Then the strong mixing property is self-evident. By the presentability of Y, we can expand θ and ρ as follows.

$$\theta = E\theta + \sum_{p\ge 1} \int \cdots \int f_p(t_1,\dots,t_p)\, X_{t_1}\cdots X_{t_p}\, dt_1\cdots dt_p$$

$$\rho = E\rho + \sum_{p\geq 1} \int \cdots \int g_p(t_1,\ldots,t_p)\, X_{t_1}\cdots X_{t_p}\, dt_1\cdots dt_p \;,$$

$f_p, g_p \in \lambda_2(\hat{\otimes}^p R)$, and we have

$$E\,\theta\rho^\tau = E\theta \circ E\rho + \sum_{p\geq 1} p!\, \langle f_p, g_p^\tau\rangle_{\lambda_2(\otimes^p R)}$$

where $g_p^\tau(\cdot,\ldots,\cdot) = g_p(\cdot-\tau,\ldots,\cdot-\tau)$. By Theorem 10 in Chapter II,

$$\langle f_p, g_p^\tau\rangle_{\lambda(\otimes^p R)} = \langle \hat{f}_p, \widehat{g_p^\tau}\rangle_{L_2(d^p F)}$$

$$= \int \cdots \int e^{i\tau(\lambda_1+\cdots+\lambda_n)}\, \hat{f}_p(\lambda_1,\ldots,\lambda_p)\widehat{g_p^\tau}(\lambda_1,\ldots,\lambda_p)$$

$$F'(\lambda_1)\cdots F'(\lambda_p)\,d\lambda_1\cdots d\lambda_p$$

(where F is the spectral distribution of X) which approaches to 0 as τ approaches to ∞ by the Riemann-Lebesgue theorem. Also

$$\sum_{p\geq 1} p!\, \left|\langle f_p, g_p^\tau\rangle_{\lambda(\otimes^p R)}\right| \leq \sum_{p\geq 1} p!\, \|f_p\|_{\lambda(\otimes^p R)} \circ \|g_p\|_{\lambda(\otimes^p R)}$$

$$\leq \tfrac{1}{2} \sum_{p\geq 1} p!\left(\|f_p\|^2_{\lambda(\otimes^p R)} + \|g_p\|^2_{\lambda(\otimes^p R)}\right)$$

$$= \tfrac{1}{2}\,(\mathrm{Var}\ \theta + \mathrm{Var}\ \rho) < \infty\,.$$

It then follows that $\lim_{\tau\to\infty} E\,\theta\rho^\tau = E\theta \circ E\rho$. Q.E.D.

For the proof of Theorem 2 we shall at first introduce a sequence of r.v.'s $\{a_n(x)\}$ which will play a fundamental role in the proof. Let

$$S(x) = \{t:\ x(t) > 0\}\ ,\quad x \in \mathbb{R}^{\mathbb{R}}\,.$$

Because of the continuity of the paths of X, $S(x)$ is an open set with probability 1 and can therefore be expressed as a denumerable disjoint sum of open intervals $I_i(x)$, $i\geq 1$. Put

$$S_n(x) = \bigcup_{i\geq 1} \{I_i(x): |I_i(x)|>n,\ I_i(x)\subset(-n,\infty)\} .$$

Denote by x_s^+ the shifted path of x by s, i.e. $x_s^+(t)=x(t+s)$.

LEMMA 4. $S_n(x)$ *is nonempty for every* $n\geq 1$ *and for almost all* x.

PROOF. Let $f(x)$ be an integrable function. Then by the ergodicity of X we have

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_a^{a+\tau} f(x_t^+)dt = \mathcal{E}f \quad \text{a.e.}$$

Now put

$$f(x) = \begin{cases} 1 & \text{if } x(t)>0,\ 0\leq t<n \\ 0 & \text{if otherwise} \end{cases}$$

and $a=-n$. Then it follows that $S_n(x)$ is nonempty if $\mathcal{E}f>0$. But indeed $\mathcal{E}f = \Pr(X_t>0,\ 0\leq t<n)>0$ by the assumption (S). Thus the proof is complete. Q.E.D.

Now we shall define $a_n(x)$ by

$$a_n(x) = n + \inf S_n(x) .$$

By Lemma 4 $a_n(x)$ is finite with probability 1.

Next we shall determine the probability law of $a_n(x)$.

LEMMA 5. *The probability distribution of* $a_n(x)$ *is absolutely continuous and its density function, say* p_n, *is flat on* $[0,n]$ *and decreasing on* (n,∞).

PROOF. Note that "$a_n(x)=t$" is equivalent to the condition "$t-n\notin S(x)$, $(t-n,t]\subset S(x)$ if $0\leq t\leq n$; $t-n\notin S(x)$, $(t-n,t)\subset S(x)$, $(-n,t-n)\cap S_n(x)=\phi$ if $t\geq n$"; we remark that if $0\leq t\leq n$, then

$t-n \notin S(x)$ implies $(-n,t-n) \cap S_n(x) = \phi$.

For $0<t<t'$, $s>0$, $t+s < t'+s < n$ we have

$$
\begin{aligned}
\Pr\left(a_n(x) \in (t+s,t'+s)\right) &= \Pr\left(\exists\, h\in(t+s,t+s');\ h-n\notin S(x),\ (h-n,h)\subset S(x)\right)\\
&= \Pr\left(\exists\, h\in(t,t');\ h-n\notin S(x_s^+),\ (h-n,h)\subset S(x_s^+)\right)\\
&= \Pr\left(\exists\, h\in(t,t');\ h-n\notin S(x),\ (h-n,h)\subset S(x)\right)\\
&\qquad\qquad \text{(by the stationarity of } X\text{)}\\
&= \Pr\left(a_n(x)\in(t,t')\right)\ ;
\end{aligned}
$$

while for $t'+s\geq n$, $t>0$, $s>0$ we get

$$
\begin{aligned}
\Pr\left(a_n(x)\in(t+s,t'+s)\right) &= \Pr\left(\exists\, h\in(t+s,t'+s);\ h-n\notin S(x),\ (h-n,h)\subset S(x),\right.\\
&\qquad\qquad \left.(-n,h-n)\cap S_n(x)=\phi\right)\\
&= \Pr\left(\exists\, h\in(t,t');\ h-n\notin S(x_s^+),\ (h-n,h)\subset S(x_s^+),\right.\\
&\qquad\qquad \left.(-n-s,h-n)\cap S_n(x_s^+)=\phi\right)\\
&\leq \Pr\left(\exists\, h\in(t,t');\ h-n\notin S(x_s^+),\ (h-n,h)\subset S(x_s^+),\right.\\
&\qquad\qquad \left.(-n,h-n)\cap S_n(x_s^+)=\phi\right)\\
&= \Pr\left(a_n(x)\in(t,t')\right)\ .
\end{aligned}
$$

Now the assertion follows easily from these two equations. Q.E.D

<u>PROOF OF THEOREM 2.</u> Since $p_n(z)$ is nonnegative, decreasing and integrable, we have $\lim_{z\to\infty} p_n(z)=0$ and $\lim_{z\to\infty} zp_n(z)=0$. Therefore

$$p_n(t) = \int_t^\infty -dp_n(z) = \int_0^\infty C_z(t)\,z\left(-dp_n(z)\right)$$

where $C_z(t) = \begin{cases} \frac{1}{z} & 0\leq t\leq z \\ 0 & t>z \end{cases}$, and

$$\int_0^\infty z\left(-dp_n(z)\right) = -z\ p_n(z)\ \Big|_0^\infty + \int_0^\infty p_n(z)\,dz = 1$$

which implies that $p_n(t)$ is a convex combination of $C_z(t)$ with the weight $d\sigma_n(z) = \left(-z dp_n(z)\right)$.

Pick a "typical" sample path y of Y to be specified further below. Define $\theta_n(x) = y\left(-a_n(x)\right)$. Then

$$\begin{aligned} E\,\theta_n^2 &= \int_0^\infty y^2(-t)p_n(t)\,dt \\ &= \int_0^\infty y^2(-t) \circ \int C_z(t)\,d\sigma_n(z) \circ dt \\ &= \int_0^\infty \int_0^z \frac{1}{z}\,y^2(t)dt\,d\sigma_n(z) < \infty\,, \end{aligned}$$

since $\lim\limits_{z\to\infty} \frac{1}{z}\int_0^z y^2(-t)dt = E\,Y_0^2 < \infty$ by the ergodicity of Y. Thus $\theta_n(X) \in L_2(X)$.

Now we can define a sequence of X-presentable noises by

$$X_t^{(n)}(\omega) = \theta_n\left(X_t^+(\omega)\right)$$

which will approximate Y in law.

Note that $a_n(x_s^+) = a_n(x)-s$ if $a_n(x)\geq s$. So the probability of the cylinder set

$$X^{(n)} \in B = \{x\in \mathbb{R}^{\mathbb{R}}\colon\ a_i \leq x(t_i) < b_i,\ 1\leq i\leq m\}$$

is the same as

$$\begin{aligned} \Pr\left[\bigcap_1^m \left(a_i\leq y\left(-a_n(X_{t_i}^+)\right)<b_i\right)\right] &= \Pr\left[\bigcap_1^m \left(a_i\leq y\left(-a_n(X)+t_i\right)<b_i\right); a_n(X)\geq \max_i t_i\right] \\ &\quad + \Pr\left[\bigcap_1^m \left(a_i\leq y\left(-a_n(X_{t_i}^+)<b_i\right); a_n(X)<\max_i t_i\right)\right] \end{aligned}$$

and as $n\uparrow\infty$ the second term goes to 0 since $\Pr\left(a_n(X) < \max t_i\right) = O(\frac{1}{n})$; the first term becomes

$$\lim \int_0^\infty 1_{\cap\left(a_i \le y(t_i - t) < b_i\right)}(t)\; p_n(t)\; dt$$

$$= \lim \int_n^\infty \int_0^z \frac{1}{z}\, 1_{\cap\left(a_i \le y(t_i - t) < b_i\right)}(t)\; dt\; d\sigma_n(z)$$

$$= \Pr(Y \in B)\ ,$$

since

$$\frac{1}{z} \int_0^z 1_{\cap\left(a_i \le y(t_i - t) < b_i\right)}(t)dt \to \Pr(Y \in B) \quad \text{as} \quad z \to \infty$$

by the ergodicity of Y.

Thus we have shown that

$$\Pr(X^{(n)} \in B) \to \Pr(Y \in B) \quad \text{as} \quad n \to \infty\ ,$$

i.e., $X^{(n)}$ converges to Y in law. Q.E.D.

We will conclude this chapter after a discussion of the assumption (S): $\Pr(X_t > 0,\ 0 \le t \le n) > 0,\ \forall\ n \ge 0$.

As before we assume that X is a sample continuous ergodic Gaussian noise with covariance function $R(t,s) = R(t-s)$. We believe that (S) always holds, yet we are not able to prove it. Instead, we find two sufficient conditions for (S) which indicate that (S) is a mild assumption (if it is a restriction at all). These two sufficient conditions are the following:

(S_1) $\qquad \forall\ n \ge 0 \quad \exists\ f_n \in R(R) \ni f_n > 0 \quad \text{on} \quad [0,n]$;

(S_2) $\qquad R(\tau) \ge 0 \quad \forall\ \tau \in \mathbb{R}$.

Now we prove that (S_1) implies (S). (I owe this proof to L. Pitt.) First note that X is mean square continuous since it is a sample continuous Gaussian process. Thus $R(t,s)$ is continuous and every $f\in R(R)$ is continuous. Assume (S_1). Then $cf_n\in R(R)$ $\forall\ c\in\mathbb{R}$ and thus $P^{(cf_n)} \sim P$ (cf. IV. 1. and recall that $P^{(cf_n)}$ is the translated measure of $\mathbb{P}$ by cf_n). By the sample continuity of X and the continuity of $f_n>0$, there exists $c>0$ such that

$$P^{(cf_n)}(X_t>0,\ 0\le t\le n) = P(X_t+cf_n(t)>0,\ 0\le t\le n)$$
$$> 0$$

(since $\{\omega\in\Omega:\ X_t(\omega)+cf_n(t)>0,\ 0\le t\le n\} \uparrow \Omega$ as $c\uparrow\infty$). (S) now follows from the equivalence of P and $P^{(cf_n)}$.

Next we prove that (S_2) implies (S). The key to the proof is a theorem of Slepian [1962; Theorem 1, p. 469] which implies that if $R(\tau)\ge 0$ and $P(X_t>0,\ 0\le t\le n) = 0$ then $P(X_t>0,\ 0\le t\le\delta) = 0,\ 0\le\delta\le n$. Now assume (S_2). If (S) does not hold then $P(X_t>0,\ 0\le t\le\delta) = 0\ \ \forall\ \delta\le n$, which implies $P(X_0>0) = 0$ since, by the sample continuity of X, $\{\omega\in\Omega:\ X_t(\omega)>0,\ 0\le t\le\delta\}\uparrow \{\omega\in\Omega:\ X_t(\omega)>0\}$ as $\delta\downarrow 0$. This is obviously a contradiction since X_0 is a nontrivial Gaussian r.v. Thus (S) holds.

(S_1) is satisfied by all stationary processes with rational spectral densities. For it is well known that for a stationary process with rational spectral density $R(R) = W_2^m$ (with $2m$ = degree of denominator - degree of nominator), the set of all functions possessing on every finite interval absolutely continuous derivatives up to order $m-1$ and square integrable m-th derivative. (S_2) is satisfied by the processes with triangular covariance functions.

VII. APPENDIX

1. Tensor Product of Hilbert Spaces

Let H_1 and H_2 be two Hilbert spaces with inner products $<\circ,\circ>_1$ and $<\circ,\circ>_2$. For each $h_1 \in H_1$, $h_2 \in H_2$ define the map $h_1 \otimes h_2 : H_1 \times H_2 \to \mathbb{R}$ by

$$(h_1 \otimes h_2)(g_1, g_2) = <h_1, g_1>_1 <h_2, g_2>_2 \quad \forall\ g_1 \in H_1,\ g_2 \in H_2 .$$

We then have the obvious relation:

$$\left(\sum_1^N a_n h_1^{(n)}\right) \otimes h_2 = \sum_1^N a_n (h_1^{(n)} \otimes h_2)$$
$$h_1 \otimes \left(\sum_1^N a_n h_2^{(n)}\right) = \sum_1^N a_n (h_1 \otimes h_2^{(n)}) .$$

Let

$$\mathcal{H} = \left\{\sum_1^N h_1^{(n)} \otimes h_2^{(n)} : h_1^{(n)} \in H_1,\ h_2^{(n)} \in H_2,\ N \geq 1\right\} .$$

Then

<u>LEMMA 1.</u> *$\mathcal{H}$ is an inner product space under the inner product*

$$<A,B> = \sum_{n=1}^{N} \sum_{m=1}^{M} <h_1^{(n)}, g_1^{(m)}>_1 <h_2^{(n)}, g_2^{(m)}>_2$$

where $A = \sum_1^N h_1^{(n)} \otimes h_2^{(n)}$, $B = \sum_1^M g_1^{(m)} \otimes g_2^{(m)}$.

The *tensor product* $H_1 \otimes H_2$ *of* H_1 *and* H_2 is the completion of $\mathcal{H}$ w.r.t. its inner product.

The following properties are useful to us.

<u>THEOREM 2.</u> *If* $\{f_\alpha, \alpha\in A\}$ *is complete in* H_1 *and* $\{g_\beta, \beta\in B\}$ *is complete in* H_2 *then* $\{f_\alpha\otimes g_\beta, \alpha\in A, \beta\in B\}$ *is complete in* $H_1\otimes H_2$. *Moreover if* $\{f_\alpha, \alpha\in A\}$ *and* $\{g_\beta, \beta\in B\}$ *are* CONS'*s then* $\{f_\alpha\otimes g_\beta, \alpha\in A, \beta\in B\}$ *is a* CONS *in* $H_1\otimes H_2$.

<u>EXAMPLE.</u> Let (X_1,S_1,μ_1) and (X_2,S_2,μ_2) be two measure spaces. Then

$$L_2(X_1,S_1,\mu_1) \otimes L_2(X_2,S_2,\mu_2) \cong L_2(X_1\times X_2,\ S_1\times S_2,\ \mu_1\times\mu_2)$$

with corresponding elements $f_1\otimes f_2$ and $f_1(x_1)f_2(x_2)$.

Similarly we can define the tensor product $H_1\otimes\ldots\otimes H_p$ for p Hilbert spaces $H_1,\ldots,H_p$. The inner product is such that

$$\langle f_1\otimes\ldots\otimes f_p,\ g_1\otimes\ldots\otimes g_p\rangle = \langle f_1,g_1\rangle_1 \cdots \langle f_p,g_p\rangle_p$$

and all properties carry over from the case $p=2$ to the case of general p in an obvious manner. If $H_1=\ldots=H_p=H$, we write $\otimes^p H$ or $H^{\otimes p}$ for $H\otimes\ldots\otimes H$ and $\otimes^p f$ or $f^{\otimes p}$ for $f\otimes\ldots\otimes f$.

We will define the symmetric tensor product $\hat{\otimes}^p H$ as consisting of the symmetric elements (tensors). Let Π_p be the set of all permutations of $(1,2,\ldots,p)$. Let $\pi\in\Pi_p$. Then $\pi=(\pi_1,\ldots,\pi_p)$ is a permutation of $(1,2,\ldots,p)$.

<u>LEMMA 3.</u> For each $\pi\in\Pi_p$ there is a unique unitary operator U_π on $\otimes^p H$ such that for all $f_1,\ldots,f_p\in H$

$$U_\pi(f_1\otimes\ldots\otimes f_p) = f_{\pi_1}\otimes\ldots\otimes f_{\pi_p} .$$

An element f in $\otimes^p H$ is called *symmetric* if $U_\pi f = f \quad \forall\ \pi \in \pi_p$. The *symmetric tensor product space* $\hat{\otimes}^p H$ is the set of all symmetric tensors in $\otimes^p H$.

THEOREM 4. *$\hat{\otimes}^p H$ is a closed subspace of $\otimes^p H$ and the operator $\frac{1}{p!} \sum_\pi U_\pi$ is the projection operator onto $\hat{\otimes}^p H$. Hence*

$$\hat{\otimes}^p H = \text{sp. } \{\frac{1}{p!} \sum_\pi f_{\pi_1} \otimes \ldots \otimes f_{\pi_p},\ f_1, \ldots, f_p \in H\} .$$

EXAMPLE. Let (X, S, μ) be an arbitrary measure space and a function on X^p is called symmetric if

$$f(x_1, \ldots, x_p) = f(x_{\pi_1}, \ldots, x_{\pi_p}) \quad \forall\ \pi \in \Pi_p .$$

Denote $\tilde{L}_2(X^p, S^p, \mu^p)$ the set of all symmetric functions in $L_2(X^n, S^n, \mu^n)$. Then

$$\hat{\otimes}^p L_2(X, S, \mu) \cong \tilde{L}_2(X^p, S^p, \mu^p) .$$

THEOREM 5. *If $\{e_\gamma, \gamma \in \Gamma\}$ is a complete set in H, then $\{e_{\gamma_1} \otimes \ldots \otimes e_{\gamma_p} : \gamma_1, \ldots, \gamma_p \in \Gamma\}$ and $\{e_{\gamma_1} \hat{\otimes} \ldots \hat{\otimes} e_{\gamma_p} : \gamma_1, \ldots, \gamma_p \in \Gamma\}$ are complete sets in $\otimes^p H$ and $\hat{\otimes}^p H$ respectively.*

THEOREM 6. *Let H be a Hilbert space. For each element $U \in H \otimes H$ there corresponds a bounded linear operator on H, say $\tilde{U}$, such that*

$$\langle \tilde{U} h_1, h_2 \rangle = \langle h_1 \otimes h_2, U \rangle_{H \otimes H} , \quad \forall\ h_1, h_2 \in H .$$

This operator is such that

$$\sum ||\tilde{U} h_\gamma||^2 = ||U||^2 < \infty$$

for any CONS $\{h_\gamma, \gamma\in\Gamma\}$.

Conversely every bounded linear operator V *on* H *with* $\sum ||Vh_\gamma||^2 < \infty$ *for a* CONS $\{h_\gamma\}$ *of* H, *is the operator* $\tilde{U}$ *associated with a unique element in* $H\otimes H$.

Moreover $U\in H\hat{\otimes}H$ *if and only if* $\tilde{U}$ *is self-adjoint.*

The operators $\tilde{U}$ on H associated with an element $U\in H\otimes H$ as above are called *Hilbert-Schmidt*.

In virtue of Theorem 6, we shall make no distinction between U and $\tilde{U}$.

THEOREM 7. *Every* $U\in H\hat{\otimes}H$ *can be written in the form*

$$U = \sum \lambda_i f_i \otimes f_i$$

where $\{f_i, i\geq 1\}$ *is an orthonormal sequence, and* $\sum |\lambda_i|^2 < \infty$. *Also we have*

$$Uf_i = \lambda_i f_i \; , \quad i\geq 1 \; .$$

2. Hermite Polynomials

Let X be a Gaussian variable with zero mean and variance t. Consider the sequence of random variables in $L_2(X)$

$$1, X, X^2, X^3, \ldots \; .$$

Applying the Gram-Schmidt procedure to orthogonalize this sequence, we obtain the orthogonal sequence

$$H_{0,t}(X), \; H_{1,t}(X), \; H_{2,t}(X), \; \ldots$$

$H_{p,t}$ $(p=0,1,2,\ldots)$ is called the *Hermite polynomial of degree p* (with parameter t). $H_{p,t}(X)$ is a polynomial in both variables t and X. The first few Hermite polynomials are

$$H_{0,t}(X) = 1$$
$$H_{1,t}(X) = X$$
$$H_{2,t}(X) = X^2-t$$
$$H_{3,t}(X) = X^3-3tX \ .$$

The Hermite polynomials $H_{p,t}(X)$, $p\geq 0$, satisfy the following:

$$E\, H_{p,t}(X)\, H_{q,t}(X) = p!\ \delta_{pq}\ t^p$$

$$H_{p,t}(X) = XH_{p-1,t}(X) - (p-1)tH_{p-2,t}(X) \ , \quad p\geq 2$$

$$e^{uX-\frac{t}{2}u^2} = \sum_{p\geq 0} H_{p,t}(X)\ \frac{1}{p!}\ u^p \qquad \forall\ u\in \mathbb{R}$$

$$H_{p,t}(\sigma X) = \sigma^p\, H_{p,\frac{t}{\sigma^2}}(X) \ , \quad \sigma>0 \ .$$

BIBLIOGRAPHY

Besson, J. L. [1974] Continuité des fonctions aléatoires spheriquement invariantes, Publication du Département de Mathématiques, Lyon, pp. 59-69.

Cambanis, S. [1974] *Stochastic Processes*, (Lecture Notes), University of North Carolina at Chapel Hill.

Cambanis, S. and Rajput, B. S. [1973] Some zero-one laws for Gaussian processes, *Annals of Probability, 1:* pp. 304-312.

Cameron, R. H. and Martin, W. T. [1947], The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals, *Annals of Mathematics, 48:* pp. 385-392.

Cramér, H. [1951] A contribution to the theory of stochastic processes, *Proc. Berkeley Symposium on Math. Statistics and Probability*, pp. 329-339, Berkeley: University of California Press.

Doob, J. L. [1953] *Stochastic Processes*, New York: Wiley.

Donelson, J. III and Maltz, F. [1972] A comparison of linear versus nonlinear prediction for polynomial functions of the Ornstein-Uhlenbeck process, *Journal of Applied Probability, 9:* pp. 725-744.

Fernique, X. [1974] Regularité des trajectoires des fonctions aléatoires Gaussiennes, École d'Été, Calcul des Probabilitíes, St. Flour.

Fortet, R. [1973] Espaces à noyau reproduisant et lois de probabilités des fonctions aléatoires, *Annales de l'Institut Henri Poincaré Section B, IX:* pp. 41-58.

Gikhman, I. I. and Skorokhod, A. V. [1964] *Introduction to the Theory of Random Processes*, Philadelphia: Saunders.

Gualtierotti, A. F. [1974] Some remarks on spherically invariant distributions, *Journal of Multivariate Analysis, 4:* pp. 347-349.

Itô, K. [1951] Multiple Wiener integrals, *Journal of Mathematical Society of Japan, 13:* pp. 157-169.

Itô, K. [1956] Spectral type of the shift transformation of differential processes with stationary increments, *Transactions of American Mathematical Society, 81:* pp. 253-263.

Jaglom, A. M. [1970] Examples of optimal nonlinear extrapolation of stationary processes, *Selected Translations of Mathematical Statistics and Probability, 9:* American Mathematical Society, pp. 273-298.

Kallianpur, G. [1970] The role of reproducing kernel Hilbert spaces in the study of Gaussian processes. In Ney, P. (Editor). *Advances in Probability and Related Topics*. New York: Marcel Dekker.

Le Page, R. [1973] Subgroups of paths and reproducing kernels, *Annals of Probability, 1:* pp. 345-347.

Loève, M. [1955] *Probability Theory*. New York: Van Nostrand.

McKean, H. P. [1973] Wiener's theory of nonlinear noise. In *Stochastic Differential Equations, SIAM-AMS Proceedings, Vol. VI,* pp. 191-209.

Meyer, P. A. [1966] *Probability and Potentials*. Waltham, Mass.: Blaisdell.

Nagornyi, V. N. [1974] Random processes for which optimal prediction is linear, *Theory of Probability and Mathematical Statistics, 1:* pp. 147-154.

Neveu, J. [1968] *Processus Aléatoires Gaussiens*. Montréal: Les Presses de L'université de Montréal.

Nisio, M. [1960] On polynomial approximation for strictly stationary processes, *Journal of Mathematical Society of Japan, 12:* pp. 207-226.

Rajput, B. S. [1972] Gaussian measures on L_p spaces, $1 \le p < \infty$, *Journal of Multivariate Analysis, 2:* pp. 382-403.

Sytaya, G. N. [1969] On admissible displacements of suspended Gaussian measures, *Theory of Probability and its Applications, 14:* pp. 506-509.

Shepp, L. A. [1966] Radon-Nikodym derivatives of Gaussian measures, *Annals of Mathematical Statistics, 27:* pp. 321-354.

Slepian, D. [1962] The one-sided barrier problem for Gaussian noise, *Bell System Technical Journal, 41:* pp. 463-501.

Vershik, A. M. [1964] Some characteristic properties of Gaussian stochastic processes, *Theory of Probability and its Applications, 9:* pp. 353-356.

Wiener, N. [1938] The homogeneous chaos, *American Journal of Mathematics, 60:* pp. 897-936.

Wiener, N. [1958] *Nonlinear Problems in Random Theory*. New York: Wiley.