

* Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514. This research was supported by the National Science Foundation under Grant GU-2059 and by the Air Force Office of Scientific Research under Grant AFOSR-68-1415.

REPRESENTATION OF STOCHASTIC PROCESSES
OF SECOND ORDER AND LINEAR OPERATIONS

by

Stamatis Cambanis*
Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 735

February, 1971

REPRESENTATION OF STOCHASTIC PROCESSES
OF SECOND ORDER AND LINEAR OPERATIONS

Stamatis Cambanis*

ABSTRACT

Series representations are obtained for the entire class of measurable, second order stochastic processes defined on any interval of the real line. They include as particular cases all earlier representations; they suggest a notion of "smoothness" that generalizes well known continuity notions; and they decompose the stochastic process into two orthogonal parts, the smooth part and a strongly discontinuous part. Also linear operations on measurable, second order processes are studied; it is shown that all "smooth" linear operations on a process, and in particular all linear operations on a "smooth" process, can be approximated arbitrarily closely by linear operations on the sample paths of the process.

* Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514. This research was supported by the National Science Foundation under Grant GU-2059 and by the Air Force Office of Scientific Research under Grant AFOSR-68-1415.

1. INTRODUCTION

Series and integral representations of stochastic processes provide a powerful tool in the study of structural properties of the processes as well as in the analysis of a number of problems in statistical communication theory and stochastic systems. The well known Karhunen-Loève representation [6, p.478] applies to mean square continuous processes defined on compact intervals of the real line. Representations valid over the entire real line have been obtained for subclasses of mean square continuous processes, the wide sense stationary processes [7] and the harmonizable processes [2]. In an attempt to relax both kinds of restrictive assumptions, series representations valid on any interval are obtained in [3] for the class of weakly continuous processes. In Section 2, two series representations are given in Theorems 1 and 3 for the entire class of measurable, second order stochastic processes, valid on any interval of the real line. Even though these representations are not unique, they have some interesting invariant properties. For instance, it is shown in Theorem 2 that they provide a unique decomposition of every second order process into two orthogonal parts, one of which can be considered as being the "smooth" part of the process. This smooth part, as expressed in the representation of Theorem 1, can be obtained explicitly in a straightforward way, while in the representation of Theorem 3 it is expressed in terms of eigenvalues and eigenfunctions that may be difficult to find explicitly. As a corollary to these representations, a general representation is given in Theorem 5 for symmetric, nonnegative definite kernels on any square of the plane, which generalizes the well known Mercer's theorem.

In Section 3, two kinds of linear operations on a second order stochastic process are considered, the first defined in the stochastic mean, and the second defined on the sample paths of the process. It is shown in Theorem 9 that all

"smooth" linear operations of the first kind can be approximated arbitrarily closely in the stochastic mean by linear operations of the second kind.

These basic results presented in Sections 2 and 3 can be used in obtaining a general, explicit solution to linear, mean square "signal" or "system" estimation problems (including additive and/or multiplicative noise problems, stochastic system identification problems, etc.) and also in providing some new insight into the problem of discrimination between two stochastic processes, both in the Gaussian and the non-Gaussian case. These applications will be given in a subsequent paper.

The results presented in this paper are stated and discussed in Sections 2 and 3 and their proofs are given in Section 4.

2. REPRESENTATION OF STOCHASTIC PROCESSES OF SECOND ORDER

The following notation and assumptions will be used throughout this paper.

(I). $\{x(t, \omega), t \in T\}$ is a measurable, second order stochastic process defined on the probability space (Ω, \mathcal{F}, P) , with T any interval of the real line, open or closed, bounded or unbounded. $r_x(t, s)$ is the autocorrelation function of $x(t, \omega)$ and $H(x)$ denotes the subspace of $L_2(\Omega, \mathcal{F}, P)$ spanned by the random variables $\{x(t, \omega), t \in T\}$.

(II). μ is any measure on $(T, \mathcal{B}(T))$ ($\mathcal{B}(T)$ is the σ -algebra of Lebesgue measurable subsets of T) which is equivalent to the Lebesgue measure (i.e., mutually absolutely continuous) and satisfies

$$(2.1) \quad \int_T r_x(t, t) d\mu(t) < +\infty.$$

That such measures μ exist follows from the following particular choice [3]:

Define μ_0 on $(T, \mathcal{B}(T))$ by $[\frac{d\mu_0}{d\text{Leb}}](t) = f(t)g(t)$, where $g(t) > 0$ a.e.[Leb] on

T , $g \in L_1(T, \mathcal{B}(T), \text{Leb.})$ and $f(t) = 1$ for $0 \leq r_x(t,t) \leq 1$ and $f(t) = r_x^{-1}(t,t)$ for $1 < r_x(t,t)$. It is clear that μ_0 satisfies (2.1), $\mu_0 \sim \text{Leb.}$ and μ_0 is finite. $\{f_k(t)\}_{k=1}^\infty$ is any complete set in $L_2(\mu) = L_2(T, \mathcal{B}(T), \mu)$. A way to find explicitly such complete sets is presented in [1], from which it is clear that the $f_k(t)$'s can be chosen to be continuous functions on T .

(III). The random variables $\{\eta_k(\omega)\}_{k=1}^\infty$ are defined by

$$(2.2) \quad \eta_k(\omega) = \int_T x(t, \omega) f_k^*(t) d\mu(t)$$

almost surely and are of second order as it is easily seen from (2.1).

$H(x, \{f_k\}_k, \mu)$ denotes the subspace of $L_2(\Omega, \mathcal{F}, P)$ spanned by the random variables $\{\eta_k(\omega)\}_{k=1}^\infty$. $\{\zeta_k(\omega)\}_{k=1}^\infty$ are the random variables derived from $\{\eta_k(\omega)\}_{k=1}^\infty$ by the Gram-Schmidt orthonormalization procedure: $\zeta_k(\omega) = \sum_{j=1}^k c_{kj} \eta_j(\omega)$. We then have

$$(2.3) \quad \zeta_k(\omega) = \int_T x(t, \omega) g_k^*(t) d\mu(t) \quad \text{a.s.}$$

where $g_k(t) = \sum_{j=1}^k c_{kj}^* f_j(t)$.

We always have $H(x, \{f_k\}_k, \mu) \subseteq H(x)$ [3, Theorem 3] and if $x(t, \omega)$ is weakly continuous equality holds [3, Theorem 4]. The following two theorems establish the properties of $H(x, \{f_k\}_k, \mu)$ as a subspace of $H(x)$ and its consequences in representing $x(t, \omega)$.

THEOREM 1. Every measurable, second order stochastic process $\{x(t, \omega), t \in T\}$ admits the representation

$$(2.4) \quad x(t, \omega) = \sum_{k=1}^{\infty} a_k(t) \zeta_k(\omega) + w(t, \omega)$$

for all $t \in T$ in the stochastic mean, where the convergence of the series is in the stochastic mean. The orthonormal random variables $\{\zeta_k(\omega)\}_{k=1}^\infty$ are given by (2.3) and the time functions $\{a_k(t)\}_{k=1}^\infty$ by

$$(2.5) \quad a_k(t) = \int_T r_x(t,s) g_k(s) d\mu(s)$$

for all $t \in T$. $r_x(t,s)$ admits the representation

$$(2.6) \quad r_x(t,s) = \sum_{k=1}^{\infty} a_k(t) a_k^*(s) + r_w(t,s)$$

for all $t, s \in T$, where the convergence of the series is absolute in t and s on $T \times T$, and $r_w(t,s)$ is the autocorrelation function of $w(t, \omega)$. The stochastic process $\{w(t, \omega), t \in T\}$ has the following properties:

(1) If $H(w)$ is the subspace of $L_2(\Omega, \mathcal{F}, P)$ spanned by the random variables $\{w(t, \omega), t \in T\}$ then

$$(2.7) \quad H(x) = H(x, \{f_k\}_k, \mu) \oplus H(w).$$

$$(11) \quad E[|w(t, \omega)|^2] = r_w(t, t) = 0 \text{ a.e. [Leb] on } T.$$

THEOREM 2. $H(x, \{f_k\}_k, \mu)$ is independent of the measure μ satisfying (II) and of the complete set $\{f_k\}_k$ in $L_2(\mu)$. Hence, denote $H(x, \{f_k\}_k, \mu)$ by $H(x, \text{smooth})$.

It is clear from Theorem 2 that the decomposition of $x(t, \omega)$ into two orthogonal terms given in (2.4) is unique. However, the representation of the first term in this decomposition by the series $\sum_{k=1}^{\infty} a_k(t) \zeta_k(\omega)$ depends clearly on the choice of μ and $\{f_k\}_k$. For a particular choice of a complete set $\{f_k\}_k$ in $L_2(\mu)$ we obtain the following theorem.

THEOREM 3. Let $\{\phi_k(t)\}_{k=1}^{\infty}$ and $\{\lambda_k\}_{k=1}^{\infty}$ be the corresponding eigenfunctions and nonzero eigenvalues of the integral type operator on $L_2(\mu)$ with kernel $r_x(t,s)$. Then

$$(2.8) \quad H(x, \{\phi_k\}_k, \mu) = H(x, \text{smooth})$$

i.e., the random variables $\{\xi_k(\omega)\}_{k=1}^{\infty}$ defined by

$$(2.9) \quad \xi_k(\omega) = \int_T x(t, \omega) \phi_k^*(t) d\mu(t) \quad \text{a.s.}$$

are complete in $H(x, \text{smooth})$. Also, if $\{\phi_k(t)\}_{k=1}^{\infty}$ are the versions of the eigenfunctions which are defined for all $t \in T$ by

$$(2.10) \quad \phi_k(t) = \frac{1}{\lambda_k} \int_T r_x(t, s) \phi_k(s) d\mu(s),$$

then $x(t, \omega)$ admits the representation

$$(2.11) \quad x(t, \omega) = \sum_{k=1}^{\infty} \phi_k(t) \xi_k(\omega) + w(t, \omega)$$

for all $t \in T$ in the stochastic mean, where the convergence in the series is in the stochastic mean, and $w(t, \omega)$ is as in Theorem 1. Also

$$(2.12) \quad r_x(t, s) = \sum_{k=1}^{\infty} \lambda_k \phi_k(t) \phi_k^*(s) + r_w(t, s)$$

for all $t, s \in T$, where the convergence of the series is absolute in t and s on $T \times T$.

Because of Theorem 2, (2.7) is written as

$$(2.13) \quad H(x) = H(x, \text{smooth}) \oplus H(w).$$

It is clear from (ii) of Theorem 1 that $H(w)$ depends on the random variables of the process $x(t, \omega)$ on a zero Lebesgue measure subset of T . On the other hand, $H(x, \text{smooth})$ represents the "smooth" part of the process $x(t, \omega)$ in a sense that is made precise in the next section. In general, both subspaces $H(x, \text{smooth})$ and $H(w)$ appear in (2.13). It is, therefore, interesting to characterize the stochastic processes for which we have $H(x) = H(x, \text{smooth})$ or $H(x) = H(w)$. This is done in the following theorem. Let us denote by $\text{RKHS}(r_x)$ the reproducing kernel Hilbert space of $r_x(t, s)$. Every function in $\text{RKHS}(r_x)$ is of the form $f(t) = E[\xi x^*(t)]$ for all $t \in T$ and some $\xi \in H(x)$, and the correspondence $\xi \leftrightarrow f$ is an isomorphism between $H(x)$ and $\text{RKHS}(r_x)$ [8, Theorem 5D].

THEOREM 4. (1) The following are equivalent:

(1.1) $H(x) = H(x, \text{smooth})$.

(1.2) If $f \in \text{RKHS}(r_x)$ and $f(t) = 0$ a.e. [Leb] on T , then $f(t) = 0$ for all $t \in T$.

(1.3) $r_x(t,s) = \sum_{k=1}^{\infty} \lambda_k \phi_k(t) \phi_k^*(s)$ for all $t,s \in T$, where the ϕ_k 's are given by (2.10).

(1.4) $r_x(t,s) = \sum_{k=1}^{\infty} a_k(t) a_k^*(s)$ for all $t,s \in T$, where the a_k 's are given by (2.5).

(2) The following are equivalent:

(2.1) $H(x) = H(w)$.

(2.2) There is a set $T_0 \in \mathcal{B}(T)$ with $\text{Leb}(T_0) = 0$ such that if $f \in \text{RKHS}(r_x)$ then $f(t) = 0$ for all $t \in T \sim T_0$.

(2.3) There is a set $T_0 \in \mathcal{B}(T)$ with $\text{Leb}(T_0) = 0$ such that $r_x(t,t) = 0$ for all $t \in T \sim T_0$.

The process $x(t,\omega)$ will be called "smooth" if it satisfies any of the conditions in Theorem 4.(1). Note that the weak continuity of $x(t,\omega)$ is equivalent to the continuity of all functions in $\text{RKHS}(r_x)$ [8, Theorem 5E]. Since condition (1.2) requires only those functions in $\text{RKHS}(r_x)$ which are equal to zero almost everywhere to be continuous, it is clear that the smoothness of $x(t,\omega)$ in the sense of Theorem 4.(1) is a weaker condition than the weak, and therefore the mean square, continuity of $x(t,\omega)$. It is also clear by Theorem 4.(2) that $H(w)$ contains a strongly discontinuous part of $x(t,\omega)$.

Let us note that if $x(t,\omega)$ is smooth in the sense of Theorem 4.(1) then, as it is seen from Theorem 2, (2.3), (2.5) and from (2.8), (2.9), (2.10), the sets $\{a_k^*(t)\}_{k=1}^{\infty}$ and $\{\lambda_k^{\frac{1}{2}} \phi_k^*(t)\}_{k=1}^{\infty}$ determined by (2.5) and (2.10) are orthonormal and complete in $\text{RKHS}(r_x)$.

A useful criterion for $x(t, \omega)$ to be Gaussian is obtained as a straightforward consequence of Theorem 2 and the closure property of Gaussian families of random variables with respect to limits in the stochastic mean.

COROLLARY 1. If $x(t, \omega)$ is smooth in the sense that it satisfies any of the conditions in Theorem 4.(1), then it is Gaussian if and only if for some complete set $\{f_k(t)\}_{k=1}^{\infty}$ in $L_2(\mu)$ the family of random variables $\{\eta_k(\omega)\}_{k=1}^{\infty}$ defined by (2.2) is Gaussian.

We lastly remark that the stochastic process representations obtained in Theorems 1 and 3 imply the following general result about kernel representation.

THEOREM 5. Every symmetric, nonnegative definite kernel $r(t, s)$ on $T \times T$ admits the representations (2.6) and (2.12). If $r(x, t)$ is weakly continuous, or continuous, on $T \times T$, then the representations (1.3) and (1.4) of Theorem 4 hold, where the convergence is absolute for all $t, s \in T$ and uniform on $T' \times T'$ for any compact subset T' of T .

Necessary and sufficient conditions for $r(t, s)$ to be weakly continuous are given in [8, Theorem 2E]. The uniformity of the convergence of (1.3) and (1.4) of Theorem 4 on compact subsets T' of T is shown easily by using the continuity of the ϕ_k 's as defined by (2.10) for all $t \in T$, which in turn follows from the weak continuity of $r(t, s)$. The last statement in Theorem 5 includes as particular case Mercer's theorem, for which it provides an alternate proof. In this sense, Theorem 5 can be considered as a generalization of Mercer's theorem to arbitrary symmetric nonnegative kernels on arbitrary intervals. It should be emphasized that the representations (2.6) and (2.12) are not trivial. As it is obvious from $r(t, s) \in L_2(\mu \times \mu)$, (1.3) and (1.4) of Theorem 4 hold a.e. [Leb \times Leb] on $T \times T$; the significance of (2.6) and (2.12) is in that they imply that (1.3) and (1.4) of Theorem 4 hold on $(T \sim T_0) \times (T \sim T_0)$, with $\text{Leb}(T_0) = 0$.

3. LINEAR OPERATIONS ON STOCHASTIC PROCESSES OF SECOND ORDER

A linear operation on $\{x(t, \omega), t \in T\}$ is a transformation which maps $\{x(t, \omega), t \in T\}$ into $\{y(s, \omega), s \in S\}$, where $y(s, \omega) \in H(x)$ for all $s \in S$, and S is an arbitrary index set. This definition generalizes in a natural way the one usually given for wide sense stationary processes and defines a linear operation in the stochastic mean sense.

On the other hand, it follows by (2.1) and the measurability of $x(t, \omega)$, that $x(\cdot, \omega) \in L_2(\mu)$ a.s. Hence bounded linear operations on the realizations of the stochastic process $x(t, \omega)$ can be considered, and they are of the form

$$(3.1) \quad \xi(\omega) = \int_T x(t, \omega) f^*(t) d\mu(t) \quad \text{a.s.}$$

with $f \in L_2(\mu)$. Clearly, by taking in (3.1) instead of $f \in L_2(\mu)$, $f(s, t)$ with $f(s, \cdot) \in L_2(\mu)$ for all $s \in S$, we obtain linear operations resulting in $\{y(s, \omega), s \in S\}$.

Thus we have introduced two kinds of linear operations on $\{x(t, \omega), t \in T\}$: (1) in the stochastic mean sense, and (2) on the sample paths of the process. Linear operations of the second kind can be realized in a straightforward manner by means of the weighting of filter functions $f(t)$. On the other hand, linear operation of the first kind are very important for linear mean square estimation problems, where the estimate of a second order random variable (or stochastic process) based on $\{x(t, \omega), t \in T\}$ is its projection onto $H(x)$. It is therefore important to study the relationship between the two classes of linear operations. It is shown in Theorem 6 that linear operations of the second kind form a subset of linear operations of the first kind.

THEOREM 6. If $\xi(\omega)$ is defined by (3.1) a.s. with $f \in L_2(\mu)$, then $\xi \in H(x, \text{smooth})$.

The inclusion relationship between linear operations of the two kinds is proper as it is seen from the fact that the random variables $x(t, \omega)$ cannot be obtained by a linear operation of the second kind (3.1) with $f \in L_2(\mu)$. The question which naturally arises at this point is whether it is possible to approximate linear operations of the first kind by linear operations of the second kind. It is shown in Theorem 9 that the answer is yes for the subset of linear operations of the first kind that result in random variables in $H(x, \text{smooth})$; and hence, if x is smooth in the sense of Theorem 4.(1) the answer is unqualified yes. Two interesting properties leading to Theorem 9 are given in Theorems 7 and 8.

From now on, we assume for simplicity that μ is a finite measure, $\mu(T) < +\infty$. That finite measures μ satisfying (II) exist is shown by the construction of a particular finite measure in (II). (Note that, as it is clear from (3.3), it suffices to choose μ so that $\mu \sim \text{Leb}$ and $\int_T [r_X(t, t) + \sqrt{r_X(t, t)}] d\mu(t) < +\infty$. In this case X is defined by (3.2) for all $B \in \mathcal{B}(T)$ with $\mu(B) < +\infty$ and, because of $|r_X|(T \times T) < +\infty$, it can be extended to $\mathcal{B}(T)$.) For every $B \in \mathcal{B}(T)$ define

$$(3.2) \quad X(B, \omega) = \int_B x(t, \omega) d\mu(t) \quad \text{a.s.}$$

By applying Theorem 6 to $f(t) = I_B(t)$, the indicator function of the set B , we have $X(B, \omega) \in H(x, \text{smooth})$ for all $B \in \mathcal{B}(T)$. It is also clear from (3.2) that X is a random measure on $(T, \mathcal{B}(T))$, i.e., a countably additive function defined on $\mathcal{B}(T)$ to $H(x, \text{smooth}) \subset L_2(\Omega, \mathcal{F}, P)$. Its corresponding complex measure r_X defined on $\mathcal{B}(T \times T)$ by $r_X(B_1 \times B_2) = E[X(B_1)X^*(B_2)]$ for all $B_1, B_2 \in \mathcal{B}(T)$, clearly satisfies $[\frac{dr_X}{d(\mu \times \mu)}](t, s) = r_X(t, s)$ and is also of bounded variation since

$$(3.3) \quad |r_X|(T \times T) \leq 2 \int_T \int_T |r_X(t, s)| d\mu(t) d\mu(s) \\ \leq 2 \left[\int_T \sqrt{r_X(t, t)} d\mu(t) \right]^2 \leq 2\mu(T) \int_T r_X(t, t) d\mu(t) < +\infty.$$

Let $H(X)$ be the subspace of $L_2(\Omega, \mathcal{F}, P)$ spanned by the random variables $\{X(B, \omega), B \in \mathcal{B}(T)\}$ and let $\Lambda_2(r_X)$ be the set of all complex valued, $\mathcal{B}(T)$ -measurable functions f on T such that $\int_{T \times T} f(t) f^*(s) r_X(dt, ds)$ is finite. Then upon considering two functions f and g in $\Lambda_2(r_X)$ as identical if and only if

$$\int_{T \times T} [f(t) - g(t)] [f^*(s) - g^*(s)] r_X(dt, ds) = 0,$$

$\Lambda_2(r_X)$ becomes a Hilbert space with inner product

$$\langle f, g \rangle_{\Lambda_2(r_X)} = \int_{T \times T} f(t) g^*(s) r_X(dt, ds)$$

and $\Lambda_2(r_X) = \text{sp}\{I_B(t), B \in \mathcal{B}(T)\}$ [4]. The Hilbert spaces $H(X)$ and $\Lambda_2(r_X)$ are isomorphic with corresponding elements $X(B, \omega)$ and $I_B(t)$, $B \in \mathcal{B}(T)$, and integration of functions in $\Lambda_2(r_X)$ with respect to X is defined by $\xi(\omega) = \int_T f(t) X(dt, \omega)$, where ξ and f are corresponding elements under the isomorphism. As it is shown in [2, p.196], for every $f \in \Lambda_2(r_X)$ we have

$$(3.4) \quad \xi(\omega) = \int_T f(t) X(dt, \omega) = \int_T f(t) x(t, \omega) d\mu(t)$$

in the stochastic mean.

THEOREM 7. $H(x, \text{smooth}) = H(X)$.

It follows from Theorem 7 that every random variable ξ in $H(x, \text{smooth})$ has the representation (3.4) for some $f \in \Lambda_2(r_X)$; and hence so does $x(t, \omega)$ if it is smooth in the sense of Theorem 4.(1).

It is shown in the proof of Theorem 7 that $L_2(\mu)$ is a subset of $\Lambda_2(r_X)$. It should be clear that $\Lambda_2(r_X)$ is a much bigger function space than $L_2(\mu)$. In fact, $\Lambda_2(r_X)$ contains functions with properties similar to those of delta functions. This is seen as follows: Assume that $x(t, \omega)$ satisfies any of the conditions in Theorem 4.(1) so that $H(x) = H(x, \text{smooth})$ (we may assume that

$x(t, \omega)$ is mean square continuous). Then, for all $t \in T$, since $x(t, \omega) \in H(x) = H(x, \text{smooth}) = H(X)$, there exists $f(t, \cdot) \in \Lambda_2(r_X)$ such that

$$x(t, \omega) = \int_T f(t, u) X(du, \omega) = \int_T f(t, u) x(u, \omega) d\mu(u)$$

in the stochastic mean. It follows that for all $t, s \in T$

$$r_X(t, s) = \int_T \int_T f(t, u) f^*(s, v) r_X(u, v) d\mu(u) d\mu(v)$$

which is a delta function type property for $f(t, u)$. The following theorem describes the properties of $L_2(\mu)$ as a subset of $\Lambda_2(r_X)$.

THEOREM 8. $L_2(\mu)$ is a dense subset of $\Lambda_2(r_X)$. Specifically:

- (i) if $\{f_k(t)\}_{k=1}^{\infty}$ is a complete set in $L_2(\mu)$, then the set $\{f_k^*(t)\}_{k=1}^{\infty}$ is complete in $\Lambda_2(r_X)$; and
- (ii) if $\{\phi_k(t)\}_{k=1}^{\infty}$ are the eigenfunctions of $r_X(t, s)$ corresponding to nonzero eigenvalues, then the set $\{\lambda_k^{-1/2} \phi_k^*(t)\}_{k=1}^{\infty}$ is orthonormal and complete in $\Lambda_2(r_X)$.

The question raised after Theorem 6 can be answered now.

THEOREM 9. Given any random variable $\xi(\omega)$ in $H(x, \text{smooth})$ and any $\epsilon > 0$ there exists $f_\epsilon(t) \in L_2(\mu)$ such that $E[|\xi - \xi_\epsilon|^2] < \epsilon^2$, where

$$(3.5) \quad \xi_\epsilon(\omega) = \int_T x(t, \omega) f_\epsilon^*(t) d\mu(t) \text{ a.s.}$$

Clearly, if $x(t, \omega)$ is smooth in the sense of Theorem 4.(1) then every linear operation on x of the first kind can be approximated by a linear operation on x of the second kind. The proof of Theorem 9 given in Section 4 makes use of Theorems 7 and 8, which are interesting in their own right. By using only Theorem 2, we can obtain another proof of Theorem 9 and also an explicit expression for $f_\epsilon(t)$ as follows. If $\xi \in H(x, \text{smooth})$ and $\rho(t) = E[x(t)\xi^*]$ then,

as in the proof of Theorem 2 $\rho \in L_2(\mu)$ and $\xi(\omega) = \sum_{n=1}^{\infty} \langle g_n, \rho \rangle_{L_2(\mu)} \zeta_n(\omega)$ in the stochastic mean. Hence, given any $\varepsilon > 0$ there exists $N(\varepsilon)$ such that

$$E[|\xi - \xi_\varepsilon|^2] < \varepsilon^2, \text{ where } \xi_\varepsilon \text{ is given by (3.5) and}$$

$$f_\varepsilon(t) = \sum_{n=1}^{N(\varepsilon)} \langle \rho, g_n \rangle_{L_2(\mu)} g_n(t) \text{ a.e. [Leb] on } T.$$

It is reasonable to assume that values of realizations of $x(t, \omega)$ at fixed $t \in T$ cannot be observed; this would require systems with zero inertia. Thus observations $\xi(\omega)$ of $x(t, \omega)$ obtained by means of linear systems with nonzero inertia are of the form $\xi(\omega) = \int_T x(t, \omega) h(t) dt$ a.s. If ξ has finite second moment, it is shown at the end of Section 4 (proof of a claim) that

$\xi \in H(x, \text{smooth})$. As a consequence, among the observations or measurements of the realizations of the process $x(t, \omega)$ by means of nonzero inertia linear systems, those that have finite second moments are in $H(x, \text{smooth})$. It follows that the linear mean square estimate of a second order random variable (or stochastic process) based on observations of the realizations of $x(t, \omega)$, is the projection of the random variable onto $H(x, \text{smooth})$.

4. PROOFS

PROOF OF THEOREM 1. For every fixed $t \in T$ denote by $y(t, \omega)$ the projection of $x(t, \omega)$ onto $H(x, \{f_k\}_k, \mu)$, and set $w(t, \omega) = x(t, \omega) - y(t, \omega)$. Since the set $\{\zeta_k(\omega)\}_{k=1}^{\infty}$ is orthonormal and complete in $H(x, \{f_k\}_k, \mu)$, we have

$$y(t, \omega) = \sum_{k=1}^{\infty} a_k(t) \zeta_k(\omega)$$

for every $t \in T$, where the convergence is in $L_2(\Omega, F, P)$, and where for every k and $t \in T$,

$$\begin{aligned} a_k(t) &= E[y(t) \zeta_k^*] = E[x(t) \zeta_k^*] \\ &= \int_T r_x(t, s) g_k(s) d\mu(s). \end{aligned}$$

Thus (2.4) and (2.5) are shown. (2.6) follows from (2.4) and the definition of $w(t, \omega)$, which implies that $w(t, \omega) \perp H(x, \{f_k\}_k, \mu)$ for every $t \in T$. The absolute convergence of the series $\sum_k a_k(t) a_k^*(s)$ on $T \times T$ follows from $\sum_k |a_k(t)|^2 = E[|y(t)|^2] < +\infty$ for all $t \in T$.

(i) is shown as follows. If we denote by H' the orthogonal complement of $H(x, \{f_k\}_k, \mu)$ in $H(x)$, it suffices to show that $H(w) = H'$. By the definition of $w(t, \omega)$ we have that $w(t, \omega) \in H'$ for all $t \in T$, and thus $H(w) \subseteq H'$. Hence it suffices to prove that $H' \subseteq H(w)$ or, equivalently, that $\xi \in H'$ and $\xi \perp H(w)$ imply $\xi = 0$. Assume that $\xi \in H'$ and $\xi \perp H(w)$. It follows that $\xi \perp \zeta_k$, $k = 1, 2, \dots$, and $\xi \perp w(t)$ for all $t \in T$. Hence, by (2.4), $\xi \perp x(t)$ for all $t \in T$ and $\xi \perp H(x)$. Now $\xi \in H' \subseteq H(x)$ and $\xi \perp H(x)$ imply $\xi = 0$.

The simplest way of proving (ii) is by making use of Theorem 3. Note that Theorems 2 and 3 are proven without employing (ii) and hence they can be used to provide a proof of (ii). We first note that the measurability of $r_w(t, s)$ and $r_w(t, t)$ follows from (2.6) and the measurability of $\sum_k a_k(t) a_k^*(s)$ and $\sum_k |a_k(t)|^2$. It follows from (2.12) and the monotone convergence theorem that

$$(4.1) \quad \int_T r_x(t, t) d\mu(t) = \sum_{k=1}^{\infty} \lambda_k + \int_T r_w(t, t) d\mu(t).$$

Also, by using Tonelli's theorem, Parseval's relationship and the monotone convergence theorem we have

$$(4.2) \quad \begin{aligned} \int_T r_x(t, t) d\mu(t) &= E \left[\int_T |x(t, \omega)|^2 d\mu(t) \right] \\ &= E \left[\sum_{k=1}^{\infty} |\langle x(\cdot, \omega), e_k(\cdot) \rangle_{L_2(\mu)}|^2 \right] \\ &= \sum_{k=1}^{\infty} E [|\langle x(\cdot, \omega), e_k(\cdot) \rangle_{L_2(\mu)}|^2] \\ &= \sum_{k=1}^{\infty} \langle r e_k, e_k \rangle_{L_2(\mu)} = \text{tr}(r) = \sum_{k=1}^{\infty} \lambda_k \end{aligned}$$

where $\{e_k(t)\}_{k=1}^{\infty}$ is any complete orthonormal set in $L_2(\mu)$, r is the integral type operator on $L_2(\mu)$ with kernel $r_x(t,s)$ and $\text{tr}(r)$ is its trace. It follows from (4.1) and (4.2) that

$$\int_T r_w(t,t) d\mu(t) = 0$$

and hence $r_w(t,t) = 0$ a.e. [Leb] on T .

PROOF OF THEOREM 2. Let μ and μ' be any measures on $(T, \mathcal{B}(T))$ satisfying (II) and let $\{f_k(t)\}_{k=1}^{\infty}$ and $\{f'_k(t)\}_{k=1}^{\infty}$ be any complete sets in $L_2(\mu)$ and $L_2(\mu')$ respectively. We first show that $H(x, \{f_k\}_k, \mu) \subseteq H(x, \{f'_k\}_k, \mu')$. It suffices to show that $\xi \in H(x, \{f_k\}_k, \mu)$ and $\xi \perp H(x, \{f'_k\}_k, \mu')$ imply $\xi = 0$. Assume that $\xi \in H(x, \{f_k\}_k, \mu)$ and $\xi \perp H(x, \{f'_k\}_k, \mu')$. It follows by $\xi \perp H(x, \{f'_k\}_k, \mu')$ that for all k

$$(4.3) \quad 0 = E[\xi \eta_k'^*] = \int_T E[\xi x^*(t)] f_k'(t) d\mu'(t).$$

The function $f^*(t) = E[\xi x^*(t)]$, $t \in T$, belongs to $L_2(\mu')$, since $|f^*(t)|^2 \leq E[|\xi|^2] r_x(t,t)$ and $r_x(t,t) \in L_1(\mu')$ by (2.1). Since by (4.3) $f(t)$ is orthogonal to the complete set $\{f'_k(t)\}_{k=1}^{\infty}$ in $L_2(\mu')$, it follows that $f = 0$ in $L_2(\mu')$, hence $f(t) = 0$ a.e. $[\mu']$ on T and, by $\mu' \sim \text{Leb}$, $f(t) = 0$ a.e. [Leb] on T . It follows that

$$(4.4) \quad E[\xi x^*(t)] = 0 \quad \text{a.e. [Leb] on } T.$$

Now $\xi \in H(x, \{f_k\}_k, \mu)$ implies that

$$(4.5) \quad \xi(\omega) = \sum_{k=1}^{\infty} a_k \zeta_k(\omega)$$

in $L_2(\Omega, \mathcal{F}, P)$, where for every k we obtain from (4.4) and (4.5) that

$$a_k = E[\xi \zeta_k^*] = \int_T E[\xi x^*(t)] g_k(t) d\mu(t) = 0.$$

It follows that $\xi = 0$ in $L_2(\Omega, \mathcal{F}, P)$. In the same way it is shown that $H(x, \{f'_k\}_k, \mu') \subseteq H(x, \{f_k\}_k, \mu)$ (these two statements are symmetric) and thus the theorem is proven.

PROOF OF THEOREM 3. It follows by (2.1) that $r_x(t, s) \in L_2(T \times T, \mathcal{B}(T \times T), \mu \times \mu) = L_2(\mu \times \mu)$ and thus the integral type operator r on $L_2(\mu)$ with kernel $r_x(t, s)$ is Hilbert-Schmidt. Let $\{\phi_k(t)\}_{k=1}^{\infty}$ and $\{\lambda_k\}_{k=1}^{\infty}$ be its corresponding eigenfunctions and nonzero, hence positive, eigenvalues. Let $\{\psi_j(t)\}_{j=1}^{\infty}$ be any complete set in the orthogonal complement of the subspace of $L_2(\mu)$ spanned by $\{\phi_k(t)\}_{k=1}^{\infty}$. Then Theorem 2 implies

$$(4.6) \quad H(x, \{\phi_k\}_k \cup \{\psi_j\}_j, \mu) = H(x, \text{smooth})$$

i.e., the random variables $\{\xi_k(\omega)\}_{k=1}^{\infty} \cup \{\eta_j(\omega)\}_{j=1}^{\infty}$ defined by (2.9) and by

$$\eta_j(\omega) = \int_T x(t, \omega) \psi_j^*(t) d\mu(t) \quad \text{a.s.}$$

are complete in $H(x, \text{smooth})$. But we have for all j that

$$E[\eta_j^2] = \langle r\psi_j, \psi_j \rangle_{L_2(\mu)} = \langle 0, \psi_j \rangle_{L_2(\mu)} = 0$$

since $r\psi_j = 0$ for all j . Thus $\eta_j = 0$, $j = 1, 2, \dots$, in $L_2(\Omega, \mathcal{F}, P)$ and (2.8) follows from (4.6).

It follows now from (2.8) that for every $t \in T$

$$(4.7) \quad x(t, \omega) = \sum_{k=1}^{\infty} a_k(t) \xi_k(\omega) + w(t, \omega)$$

in the stochastic mean, where $w(t, \omega)$ is as in Theorem 1. Since by (2.9),

$E[\xi_k \xi_j^*] = \langle r\phi_k, \phi_j \rangle_{L_2(\mu)} = \lambda_k \delta_{kj}$ it follows by (4.7) that for all $t \in T$ and k we have

$$\lambda_k a_k(t) = E[x(t) \xi_k^*] = \int_T r_x(t, s) \phi_k^*(s) d\mu(s)$$

and thus, if we consider the versions of the eigenfunctions defined pointwise everywhere on T by (2.10), we obtain $a_k(t) = \phi_k(t)$ for all $t \in T$ and k . Now (2.12) follows from (2.11).

PROOF OF THEOREM 4. (1). We first show that (1.1) implies (1.2). Let $f \in \text{RKHS}(r_x)$ with $f(t) = 0$ a.e. [Leb] on T . Then $f(t) = E[\xi x^*(t)]$ for all $t \in T$ for some $\xi \in H(x)$. Since $H(x) = H(x, \text{smooth}) = H(x, \{f_k\}_k, \mu)$ we have $\xi(\omega) = \sum_{k=1}^{\infty} a_k \zeta_k(\omega)$ in the stochastic mean, where for all k , $a_k = E[\xi \zeta_k^*] = \int_T f(t) g_k(t) d\mu(t) = 0$. Hence $\xi = 0$ in $L_2(\Omega, \mathcal{F}, P)$ and $f(t) = E[\xi x^*(t)] = 0$ for all $t \in T$.

We now show that (1.2) implies (1.1). It suffices to show that $\xi \in H(x)$ and $\xi \perp H(x, \text{smooth})$ imply $\xi = 0$. As in the proof of Theorem 2, $\xi \perp H(x, \text{smooth})$ implies $f(t) = E[\xi x^*(t)] = 0$ a.e. [Leb] on T . Since $f \in \text{RKHS}(r_x)$, (1.2) implies that $f(t) = E[\xi x^*(t)] = 0$ for all $t \in T$. Hence $\xi \perp H(x)$ and since $\xi \in H(x)$, it follows that $\xi = 0$.

The equivalence of (1.1) with (1.3) and (1.4) follows from (2.6) and (2.12) respectively and the fact that $H(x) = H(x, \text{smooth})$ if and only if $w(t, \omega) = 0$ a.s. for all $t \in T$, i.e., if and only if $r_w(t, s) = 0$ for all $t, s \in T$.

(2). It is shown in Theorem 1.(ii) that $r_w(t, t) = 0$ a.e. [Leb] on T . Let $T_0 = \{t \in T: r_w(t, t) \neq 0\}$. Then $T_0 \in \mathcal{B}(T)$ and $\text{Leb}(T_0) = 0$.

We first show that (2.1) implies (2.2). Let $f \in \text{RKHS}(r_x)$. Then $f(t) = E[\xi x^*(t)]$ for all $t \in T$ and some $\xi \in H(x)$. Since $H(x) = H(w)$, $x(t, \omega) = w(t, \omega)$ a.s. for all $t \in T$, and by (ii) of Theorem 1, we have $f(t) = E[\xi w^*(t)] = 0$ for all $t \in T \setminus T_0$.

We now show that conversely, (2.2) implies (2.1). It suffices to show that $\xi \in H(x, \text{smooth})$ implies $\xi = 0$. Let $\xi \in H(x, \text{smooth}) = H(x, \{f_k\}_k, \mu)$. Then $\xi(\omega) = \sum_{k=1}^{\infty} a_k \zeta_k(\omega)$ in the stochastic mean, where for all k , $a_k = E[\xi \zeta_k^*] = \int_T E[\xi x^*(t)] g_k(t) d\mu(t)$. But $f(t) = E[\xi x^*(t)] \in \text{RKHS}(r_x)$ and by (2.2),

$f(t) = 0$ a.e. [Leb] on T . Hence $a_k = 0$ for all k , and $\xi = 0$.

For the equivalence between (2.1) and (2.3) we first show that (2.1) implies (2.3). It follows by (2.1) that $H(x, \text{smooth}) = \{0\}$, hence $x(t, \omega) = w(t, \omega)$ a.s. for all $t \in T$, and $r_x(t, t) = r_w(t, t)$ for all $t \in T$. Now (2.3) follows by Theorem 1.(ii).

Conversely, if $r_x(t, t) = 0$ a.e. [Leb] on T , then $r_x(t, s) = 0$ a.e. [Leb \times Leb] on $T \times T$, since $|r_x(t, s)|^2 \leq r_x(t, t)r_x(s, s)$ for all $t, s \in T$. Then $r = 0$ and $E[|\eta_k|^2] = \langle r f_k, f_k \rangle_{L_2(\mu)} = 0$ for all k . Hence $\eta_k = 0$ for all k , $H(x, \text{smooth}) = H(x, \{f_k\}_k, \mu) = \{0\}$ and $H(x) = H(w)$.

PROOF OF THEOREM 6. Let $y(t, \omega)$ be the projection of $x(t, \omega)$ on $H(x, \text{smooth})$ for every $t \in T$, as in the proof of Theorem 1. Then $x(t, \omega) = y(t, \omega) + w(t, \omega)$, $r_x(t, s) = r_y(t, s) + r_w(t, s)$ and by Theorem 1.(ii), $r_x(t, s) = r_y(t, s)$ a.e. [Leb \times Leb] on $T \times T$. The integral

$$(4.8) \quad \eta(\omega) = \int_T y(t, \omega) f^*(t) d\mu(t)$$

exists in the stochastic mean sense as an element in $H(y)$ if and only if

$$(4.9) \quad \sigma^2 = \int_T \int_T r_y(t, s) f^*(t) f(s) d\mu(t) d\mu(s) < +\infty$$

[5, p.33]. But (4.9) is satisfied since $f \in L_2(\mu)$ and since $r_y \in L_2(\mu \times \mu)$, because $r_x(t, s) = r_y(t, s)$ a.e. [Leb \times Leb] on $T \times T$ and $r_x \in L_2(\mu \times \mu)$. Hence the integral $\eta(\omega)$ exists and $\eta \in H(x, \text{smooth})$ since the definition of $y(t, \omega)$ implies $H(y) \subseteq H(x, \text{smooth})$ (in fact it can be easily shown that $H(y) = H(x, \text{smooth})$). It follows from (3.1), (4.9) and $r_x = r_y$ a.e. on $T \times T$ that

$$(4.10) \quad E[|\xi|^2] = \sigma^2.$$

By (4.8) and a property of the stochastic mean integral [5, p.30] we have

$$(4.11) \quad E[|\eta|^2] = \sigma^2.$$

Also by a property of the stochastic mean integral, we get

$$\begin{aligned}
 (4.12) \quad E[\xi\eta^*] &= \int_T E[\xi y^*(s)] f(s) d\mu(s) \\
 &= \int_T \left\{ \int_T E[x(t)y^*(s)] f^*(t) d\mu(t) \right\} f(s) d\mu(s) \\
 &= \sigma^2 = E[\eta\xi^*].
 \end{aligned}$$

It follows from (4.10), (4.11) and (4.12) that

$$E[|\xi - \eta|^2] = E[|\xi|^2] - E[\xi\eta^*] - E[\eta\xi^*] + E[|\eta|^2] = 0.$$

Hence $\xi = \eta \in H(x, \text{smooth})$.

PROOF OF THEOREM 7. It follows from (3.2) and Theorem 6 that $X(B, \omega) \in H(x, \text{smooth})$ for all $B \in \mathcal{B}(T)$. Hence $H(X) \subseteq H(x, \text{smooth})$. We now show that $H(x, \text{smooth}) \subseteq H(X)$. Since by Theorem 2, $H(x, \text{smooth}) = H(x, \{f_k\}_k, \mu)$, it suffices to prove that $\eta_k \in H(X)$ for all $k = 1, 2, \dots$, where the η_k 's are defined by (2.2). It can be shown, as in the proof of Theorem 6, that the integral in (2.2) defined almost surely or in the stochastic mean gives the same random variable in $L_2(\Omega, \mathcal{F}, P)$. Note that if $f \in L_2(\mu)$ then $f \in \Lambda_2(r_X)$ since

$$\begin{aligned}
 \|f\|_{\Lambda_2(r_X)}^2 &\leq \int_{T \times T} |f(t)| |f^*(s)| |r_X| (dt, ds) \\
 &\leq 2 \int_T \int_T |f(t)| |f^*(s)| |r_X(t, s)| d\mu(t) d\mu(s) \\
 &\leq 2 \left(\int_T |f(t)| \sqrt{r_X(t, t)} d\mu(t) \right)^2 \\
 &\leq 2 \|f\|_{L_2(\mu)}^2 \cdot \int_T r_X(t, t) d\mu(t) < +\infty.
 \end{aligned}$$

Hence, for all k , $f_k \in \Lambda_2(r_X)$ and by (3.4), $\eta_k(\omega) = \int_T f_k^*(t) X(dt, \omega) \in H(X)$.

PROOF OF THEOREM 8. (i) Let $\{f_k(t)\}_{k=1}^\infty$ be any complete set in $L_2(\mu)$. Then the random variables $\{\eta_k(\omega)\}_{k=1}^\infty$ defined by (2.2) are complete in $H(x, \text{smooth}) = H(X)$. Also, it is shown in the proof of Theorem 7 that $\eta_k(\omega) = \int_T f_k^*(t) X(dt, \omega)$. Thus, it follows from the isomorphism between $H(X)$ and $\Lambda_2(r_X)$ that the set $\{f_k^*(t)\}_{k=1}^\infty$ is complete in $\Lambda_2(r_X)$. Since $\{f_k^*(t)\}_{k=1}^\infty$ is also complete in $L_2(\mu)$, $L_2(\mu)$ is dense in $\Lambda_2(r_X)$.

(ii) Since the random variables $\xi_k(\omega) = \int_T x(t, \omega) \phi_k^*(t) d\mu(t)$ are shown in Theorem 3 to be complete in $H(x, \text{smooth})$, it follows as in (i) that the set $\{\phi_k^*(t)\}_{k=1}^\infty$ is complete in $\Lambda_2(r_X)$. The set $\{\lambda_k^{-1/2} \phi_k^*(t)\}_{k=1}^\infty$ is orthonormal since

$$\begin{aligned} \langle \lambda_k^{-1/2} \phi_k^*, \lambda_j^{-1/2} \phi_j^* \rangle_{\Lambda_2(r_X)} &= (\lambda_k \lambda_j)^{-1/2} \int_{T \times T} \phi_k^*(t) \phi_j^*(s) r_X(dt, ds) \\ &= (\lambda_k \lambda_j)^{-1/2} \langle r \phi_k, \phi_j \rangle_{L_2(\mu)} = \delta_{kj}. \end{aligned}$$

PROOF OF THEOREM 9. If $\xi \in H(x, \text{smooth}) = H(X)$, there exists $f^* \in \Lambda_2(r_X)$ such that $\xi(\omega) = \int_T f^*(t) X(dt, \omega)$. Since $L_2(\mu)$ is dense in $\Lambda_2(r_X)$ by Theorem 8, given any $\varepsilon > 0$ there exists $f_\varepsilon^* \in L_2(\mu)$ such that $\|f^* - f_\varepsilon^*\|_{\Lambda_2(r_X)} < \varepsilon$. Hence, if $\xi_\varepsilon(\omega) = \int_T f_\varepsilon^*(t) X(dt, \omega)$, we have $E[|\xi - \xi_\varepsilon|^2] < \varepsilon^2$. It follows by (3.4) that $\xi_\varepsilon(\omega)$ is given by (3.5) with the integral defined in the stochastic mean, and this completes the proof because, as it is shown in the proof of Theorem 6, the integral in (3.5) defined almost surely and in the stochastic mean gives the same random variable in $L_2(\Omega, \mathcal{F}, P)$.

PROOF OF A CLAIM (made in last paragraph of Section 3). If $\xi(\omega) = \int_T x(t, \omega) h(t) dt$ a.s. has finite second moment, then

$$(4.13) \quad E[|\xi|^2] = \int_T \int_T r_x(t, s) h(t) h^*(s) dt ds < +\infty$$

Let μ be any measure as in (II) of Section 2 and $F(t) = [\frac{d\mu}{d\text{Leb}}](t)$. Then $F(t) \neq 0$ a.e. [Leb] on T and by (4.13), $f(t) = \frac{h(t)}{F(t)} \in \Lambda_2(r_X)$. It now follows from $\xi(\omega) = \int_T x(t, \omega) f(t) d\mu(t)$ a.s., the argument used in the proof of Theorem 6, and (3.3) that $\xi(\omega) = \int_T f(t) X(dt, \omega)$. Hence $\xi \in H(X) = H(x, \text{smooth})$, by Theorem 7.

REFERENCES

- [1] S. Cambanis, Bases in L_2 spaces with applications to stochastic processes with orthogonal increments, Proc. Amer. Math. Soc., to appear.
- [2] S. Cambanis and B. Liu, On harmonizable stochastic processes, Information and Control 17(1970), pp. 183-202.
- [3] S. Cambanis and E. Masry, On the representation of weakly continuous processes, Information Sci., to appear. Also, in a shortened version, Proc. Eighth Annual Allerton Conference on Circuit and System Theory, October 7-9, 1970, pp. 610-618.
- [4] H. Cramèr, A contribution to the theory of stochastic processes, Proc. Second Berkeley Symp. Math. Stat. Probability, Univ. of California Press, Berkeley, 1951, pp. 329-339.
- [5] K. Karhunen, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, Ann. Acad. Sci. Fen. Ser. AI, No. 37(1947), pp. 1-79.
- [6] M. Loève, Probability Theory, Van Nostrand, Princeton, New Jersey, 1963.
- [7] E. Masry, B. Liu and K. Steiglitz, Series expansion of wide-sense stationary random processes, IEEE Trans. Information Theory IT-14(1968), pp. 792-796.
- [8] E. Parzen, Statistical inference on time series by Hilbert space methods, in Time Series Analysis Papers, Holden-Day, San Francisco, California, 1967, pp. 251-382.