* *The major part of this work was done while the author was at the Electrical Engineering Department, Princeton University. His work was supported in part by the National Science Foundation under Grant GK-1439 with Princeton University and by the National Science Foundation under Grant GU-2059 with the University of North Carolina at Chapel Hill.*

** *This author's work was supported by the National Science Foundation under Grant GK-13192.*

ON THE REPRESENTATION OF WEAKLY CONTINUOUS STOCHASTIC PROCESSES

by

STAMATIS CAMBANIS*
Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

ELIAS MASRY**
Department of Applied Physics and Information Science
University of California at San Diego

Institute of Statistics Mimeo Series No. 674

March, 1970.

ON THE REPRESENTATION OF WEAKLY CONTINUOUS STOCHASTIC PROCESSES

by

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

Elias Masry**
Department of Applied Physics
and Information Science
University of California at San Diego

ABSTRACT

A novel approach to obtaining series representations in the stochastic mean for weakly, and therefore mean square, continuous stochastic processes is presented. Two distinct orthogonal series representations are derived for the entire class of weakly continuous stochastic processes on any Lebesgue set of the real line, and a constructive procedure to obtain them explicitly is given. They include as particular cases all earlier representations. Also they are shown to converge almost surely in the norm of an L_2 space. Two general results, on which the development of this paper is based, are also presented.

* The major part of this work was done while the author was at the Electrical Engineering Department, Princeton University. His work was supported in part by the National Science Foundation under Grant GK-1439 with Princeton University and by the National Science Foundation under Grant GU-2059 with the University of North Carolina at Chapel Hill.

** This author's work was supported by the National Science Foundation under Grant GK-13192.

1. Introduction

Series representations of stochastic processes are of considerable use in several problems of statistical communication theory. If $\{x(t,\omega), t\epsilon T\}$ is a stochastic process of second order defined on the Lebesgue measurable set T of the real line, a general series representation is given by

$$x(t,\omega) = \sum_k a_k(t)\ \eta_k(\omega) \tag{1}$$

where the convergence is usually in the stochastic mean. The significance of such a representation is in the fact that it decomposes the stochastic process into the random, time independent, part $\{\eta_k(\omega)\}$ and the time dependent deterministic part $\{a_k(t)\}$, thus providing some insight into the structure of the stochastic process as well as a tool for the study of many particular problems.

Various series representations have been obtained under certain assumptions. Before mentioning them, let us introduce some notation. Let R^1 be the real line, $\mathcal{B}^1$ the σ-algebra of Lebesgue measurable sets on R^1, $\mathcal{B}_T^1$ the restriction of $\mathcal{B}^1$ to $T\epsilon\mathcal{B}^1$, R^2 the plane, $\mathcal{B}^2$ the σ-algebra of Lebesgue sets on R^2 and m the Lebesgue measure on the real line.

A second order, mean square continuous stochastic process $\{x(t,\omega), t\epsilon I\}$ on a compact interval of the real line, has the series representation [6,8]

$$x(t,\omega) = \sum_k f_k(t)\ \xi_k(\omega) \tag{2}$$

where the convergence is in the stochastic mean uniformly on I. $\{f_k(t)\}_k$

and $\{\lambda_k\}_k$ are the corresponding eigenfunctions and nonzero eigenvalues of the operator on $L_2(I, B_I^1, m)$ with kernel $R(t,s)$, where $R(t,s)$ is the autocorrelation function of $x(t,\omega)$. $\{\xi_k(\omega)\}_k$ is a set of random variables defined by

$$\xi_k(\omega) = \int_I x(t,\omega)\, f_k^*(t)\, m(dt) \tag{3}$$

in the stochastic mean and satisfying $E[\xi_k \xi_\ell^*] = \lambda_k \delta_{k\ell}$. Representation (2) is known as the Karhunen-Loève expansion and its great advantage is that both the time functions and the random variables are orthogonal.

The Karhunen-Loève representation makes the minimum of assumptions on the process, but in general it is not valid over the entire real line or over non-compact intervals of the real line. A series representation similar to (2) and valid over a possibly infinite interval $I = (a,b)$, $-\infty \leq a < b \leq +\infty$, of the real line is given in [6] for the subclass of second order, mean square continuous stochastic processes $\{x(t,\omega), t\epsilon I\}$ whose autocorrelation functions satisfy

$$\int_I R(t,t)\, m(dt) < \infty . \tag{4}$$

Equations (2) and (3) are valid in this case, the only difference being that the convergence in (2) is not in general uniform on I. If I is a compact interval, then (4) is satisfied and the convergence in (2) is uniform.

However, for any Lebesgue measurable set T of the real line, it is known that a second order stochastic process $\{x(t,\omega), t\epsilon T\}$ has a series representation of the form (1) if and only if the Hilbert space $H(x,T)$ spanned in the mean square sense by the random variables $\{x(t,\omega), t\epsilon T\}$

is separable [7, p.27]. On the other hand, if $\{x(t,\omega),\ t\epsilon T\}$ is weakly continuous on T, then $H(x,T)$ is separable [10, Thm. 2D]. Hence every weakly continuous stochastic process, and a fortiori every mean square continuous stochastic process, has a series representation of the form (1) on any Lebesgue set T. What remains to be established then, is an explicit way of determining the time functions $\{a_k(t)\}$ and the random variables $\{\eta_k(\omega)\}$. Let us note that if the set $\{\eta_k(\omega)\}$ is orthonormal, then the time functions are given by

$$a_k(t) = E[x(t)\ \eta_k^*]. \tag{5}$$

The time functions and the random variables in the representation (1), valid over the entire real line, have been obtained in [9] for the class of mean square continuous, wide sense stationary stochastic processes, and in [2] for the class of harmonizable stochastic processes.

A general way to obtain complete sets of random variables in the span $H(x,T)$ of any weakly continuous stochastic process $\{x(t,\omega),\ t\epsilon T\}$, and therefore series representations of the form (1), is presented in Theorem 4. The random variables are defined by linear operations on the stochastic process. The significance of this novel approach in obtaining series representations for weakly continuous stochastic processes is illustrated by the representations derived in subsequent theorems as immediate consequences of Theorem 4. Theorems 5 and 6 provide two series representations for any weakly continuous stochastic process; that is, they provide two distinct ways of determining the time functions and the random variables in (1). The time functions in the representation of Theorem 5 can be obtained explicitly in a straightforward way, while in order to obtain the time functions in the representation of Theorem 6 the computation of the eigenfunctions of an integral equation is required. The novelty of the representations presented in Theorems 5 and 6 is in the fact that they hold for all weakly, and therefore mean square, continuous stochastic processes over any Lebesgue set of the real line, finite or infinite, compact or non-compact. In contrast, the representations of mean square continuous processes given in [6,8] hold either only on compact subsets of the real line or on any Lebesgue set of the real line with the additional condition (4) and they are included in Theorem 6 as particular cases. Both representations given in Theorems 5 and 6 can be written in the form of the representations given in [9] for mean square continuous, wide sense stationary stochastic processes, and in [2] for harmonizable stochastic processes. They also provide two absolutely convergent series representations for $R(t,s)$.

For all measurable second order stochastic processes, a series representation of the form (1) is obtained in Section 4, where the convergence is not in the sotchastic mean anymore but in the norm of an appropriate L_2 space almost surely. It is also shown that both representations of Theorems 5 and 6 converge also almost surely in an L_2 space sense.

All the results presented in Sections 3 and 4 rely on two general results presented as Theorems 1 and 2, in Section 2. They essentially say that for every measurable second order stochastic process defined on any Lebesgue measurable set T of the real line, there exists an appropriate L_2 space over T with the properties that (i) almost all sample functions of the process belong to it, and (ii) the autocorrelation function of the process, considered as kernel, defines a completely continuous operator on it of the trace class. It is believed that the significance of these two results is beyond their implications presented in this paper.

2. Some General Results

This section presents two general results on measurable, second order stochastic processes defined on any Lebesgue set T of the real line, that are used in subsequent sections.

Theorem 1.

For every measurable, second order stochastic process $\{x(t,\omega),\ t\in T\}$ there exists a finite, non-negative measure μ on $(T,\ \mathcal{B}_T^1)$ such that

$$\int_T R(t,t)\ \mu(dt) < \infty \tag{6}$$

where $R(t,s)$ is the autocorrelation function of $x(t,\omega)$.

Proof: Define a function f on T by

$$f(t) = \begin{cases} 1 & \text{for} \quad R(t,t) \le 1 \\ \dfrac{1}{R(t,t)} & \text{for} \quad R(t,t) > 1. \end{cases} \tag{7}$$

Then f is measurable and $0 \le R(t,t)\, f(t) \le 1$ on T. If ν is any finite, non-negative measure on $(T,\ \mathcal{B}_T^1)$ and if the measure μ on $(T,\ \mathcal{B}_T^1)$ is defined by $[\frac{d\mu}{d\nu}] = f$, then (6) is satisfied. Furthermore, μ is non-negative and finite since $0 \le f(t) \le 1$ on T.

Q.E.D.

Let us point out that if $R(t,s)$ is uniformly bounded on $T\times T$, then μ can be chosen to be any finite, non-negative measure on $(T,\ \mathcal{B}_T^1)$. This includes the cases where x is mean square continuous, wide sense stationary or harmonizable. If $R(t,t)$ is integrable over T, then μ can be chosen to be the Lebesgue measure m on T.

If the measure ν is chosen to be absolutely continuous with respect to the Lebesgue measure m with Radon-Nikodym derivative ϕ, then μ will have Radon-Nikodym derivative with respect to m:

$$[\frac{d\mu}{dm}](t) = F(t) = f(t)\phi(t). \tag{8}$$

Since $x(t,\omega)$ is of second order, it may be assumed that the set $T_0 = \{t\epsilon T;\ E[|x(t,\omega)|^2] = R(t,t) = 0\}$ has Lebesgue measure zero; otherwise one may consider $x(t,\omega)$ defined on $T\sim T_0$. Hence $f(t)$ may be chosen positive a.e. [m], as for instance in (7). If ϕ is also chosen nonzero on T a.e. [m], then the function F defined on T by (8) will be nonzero a.e. [m], non-negative and Lebesgue integrable over T. Hence we have the following.

COROLLARY:

The measure μ of Theorem 1 can always be chosen to be absolutely continuous with respect to the Lebesgue measure m such that $[\frac{d\mu}{dm}](t) = F(t) \neq 0$ a.e. [m] on T.

It should be pointed out that the measure μ of Theorem 1 is not uniquely determined. In the construction of μ given in the proof of Theorem 1, the choice of f is clearly not unique and ν is an arbitrary finite measure; neither is this construction the only way to obtain a measure μ with the properties stated in Theorem 1. However, it should be emphasized that the considerable freedom in the choice of the measure μ with the properties stated in Theorem 1, far from being a drawback, represents an advantage to be appropriately exploited in every particular case.

This point is best illustrated by an example. Let $T = [0, +\infty)$ and $R(t,s) = \min(t,s)$. As it will become clear in subsequent sections, the construction of a complete set in $L_2(T,\mathcal{B}_T^1,\mu)$ will frequently be needed. For this purpose, an appropriate choice for f and ϕ gives $F(t) = e^{-t}$ and enables one to choose as complete set in $L_2(T,\mathcal{B}_T^1,\mu)$ the Laguerre polynomials or just the powers of t: $\{t^k, k = 0,1,2,\ldots\}$.

Theorem 2.

Let $\{x(t,\omega), t\epsilon T\}$ be a measurable, second order stochastic process and let μ be the measure introduced in Theorem 1. Then $x(t,\omega) \epsilon L_2(T,\mathcal{B}_T^1,\mu) = L_2(\mu)$ almost surely.

Proof: The measurability of x and Fubini's Theorem imply

$$\int_T R(t,t)\mu(dt) = \int_T E[|x(t,\omega)|^2]\mu(dt) = E\left[\int_T |x(t,\omega)|^2\mu(dt)\right]. \tag{9}$$

Hence, by (9), $x(t,\omega)\epsilon L_2(\mu)$ a.s.

Q.E.D.

It should be pointed out that the applicability of the property proven in Theorem 2, i.e., of the fact that almost all sample functions of any measurable, second order process belong to some L_2 space, is far beyond its use in Section 4. For instance, this property can be used in generalizing a number of results established in the literature in the particular case of stochastic processes whose sample functions are almost surely square integrable with respect to the Lebesgue measure.

3. Derivation of Series Representations

Throughout this and the next section, the following notation is used:

(i) $\{x(t,\omega),\ t\in T\}$ is a measurable, second order stochastic process defined on any Lebesgue measurable set T of the real line, and

(ii) μ is a non-negative measure on $(T,\mathcal{B}_T^1)$ such that (6) is satisfied and $[\frac{d\mu}{dm}](t) \neq 0$ a.e. [m] on T. As shown in Section 2 (Theorem 1 and its corollary), such a measure always exists and can be found explicitely if $R(t,t)$ is known.

Notice that in order to include in the subsequent study the case considered in [6], i.e., $\mu = m$, the assumption is not made that μ is finite.

Our goal, as made clear in the introduction, is to find complete sets of random variables in the span $H(x,T)$ of a second order stochastic process. Such a set of random variables can be used in obtaining series representations of the process $x(t,\omega)$ and of linear operations on $x(t,\omega)$.

A convenient set of random variables $\{\eta_k(\omega)\}_k$ is defined by

$$\eta_k(\omega) = \int_T x(t,\omega)\ \phi_k^*(t)\ \mu(dt) \tag{10}$$

almost surely, where $\{\phi_k(t)\}_k$ is an arbitrary complete set of functions in $L_2(\mu)$.

Theorem 3.

The random variables $\{\eta_k(\omega)\}_k$ defined by (10) are of second order. If $H(\eta)$ is the Hilbert space spanned in the mean square sense by the random variables $\{\eta_k(\omega)\}_k$, then

$$H(\eta) \subseteq H(x,T). \tag{11}$$

Proof: Eq. (6) implies $R(t,s)\epsilon L_2(T\times T,\ B^2_{T\times T},\ \mu\times\mu) = L_2(\mu\times\mu)$. Hence $\int_T\int_T R(t,s)\phi^*_k(t)\phi_k(s)\mu(dt)\mu(ds)$ is finite and the random variables $\{\eta_k(\omega)\}_k$ are of second order.

To show (11), it suffices to show that $\eta_k\epsilon H(x,T)$ for all k, or that if η_k is orthogonal to $x(t)$ for all $t\epsilon T$ then $E[|\eta_k|^2] = 0$. Assume that η_k is orthogonal to $x(t)$ for all $T\epsilon T$. Then

$$0 = E[x(t)\eta^*_k] = \int_T R(t,s)\phi_k(s)\mu(ds) = (R\phi_k)(t) \tag{12}$$

for all $t\epsilon T$, and

$$\begin{aligned} E[|\eta_k|^2] &= \int_T \int_T R(t,s)\phi^*_k(t)\ \phi_k(s)\ \mu(dt)\ \mu(ds) \\ &= (R\phi_k,\ \phi_k)_{L_2(\mu)} = 0. \end{aligned} \tag{13}$$

Q.E.D.

A natural question to ask is under what conditions equality holds in (11). Before pursuing this question further let us make the following remark. If $T'\subset T$ is such that $m(T\sim T') = 0$, then $H(\eta)\subseteq H(x,T')$. This is shown as Theorem 3; the only difference is that in this case $(R\phi_k)(t) = 0$ for $t\epsilon T'$, i.e., a.e. [m] on T, but this suffices to imply $R\phi_k = 0$ in $L_2(\mu)$ and hence $E[|\eta_k|^2] = 0$. This remark suggests that $H(\eta)$ may not contain the random variables $x(t)$ corresponding to points of discontinuity of x of some kind, for instance discontinuity in the mean square sense; it also suggests that $H(\eta)$ contains only the "smooth" part of x and therefore in order to have $H(\eta) = H(x,T)$ the process x has to be "smooth" in some sense. A continuity ("smoothness") condition on x which implies $H(x,T) = H(\eta)$ is given in the next theorem. Let us note that a stochastic process $\{x(t,\omega), t\in T\}$ is called weakly continuous on T if for every $t\in T$ and every $\xi\in H(x,T)$, $\lim_{\tau\to t} E[x(\tau)\xi^*] = E[x(t)\xi^*]$.

THEOREM 4.

If $\{x(t,\omega), t\in T\}$ is weakly continuous on T, then

$$H(\eta) = H(x,T). \tag{14}$$

Proof: In view of (11), it suffices to show $H(x,T)\subseteq H(\eta)$, or equivalently that if $\zeta\in H(x,T)$ is orthogonal to η_k for all k then $E[|\zeta|^2] = 0$. Assume that $\zeta\in H(x,T)$ is orthogonal to η_k for all k. Then

$$0 = E[\zeta\eta_k^*] = \int_T E[\zeta x^*(t)]\, \phi_k(t)\, \mu(dt) \tag{15}$$

for all k. The function $f(t) = E[x(t)\zeta^*]$ belongs to $L_2(\mu)$ since

$$\int_T |f(t)|^2\, \mu(dt) \le E[|\zeta|^2] \int_T R(t,t)\, \mu(dt) < \infty \tag{16}$$

and is orthogonal to the complete set $\{\phi_k(t)\}_k$ by Eq. (15). Hence $|f(t)|^2 = 0$ on T a.e. $[\mu]$, or by (8) $|f(t)|^2 F(t) = 0$ on T a.e. $[m]$, and since $F(t) \neq 0$ on T a.e. $[m]$, it follows that $f(t) = 0$ on T a.e. $[m]$. We further note that $f(t)$ is a continuous function since x is weakly continuous. Consequently, $f(t) = 0$ for all $t\in T$ so that $\zeta(\omega)$ is orthogonal to the set $\{x(t,\omega), t\in T\}$, i.e., to $H(x,T)$. Thus $E[|\zeta|^2] = 0$.

Q.E.D.

It should be pointed out that Theorem 4 and all subsequent results of this section, are true a fortiori for mean square continuous processes, since mean square continuity implies weak continuity.

It would be interesting to find a necessary and sufficient condition for $H(x,T) = H(\eta)$. The weak continuity of x is shown here to be only a sufficient condition.

Theorem 4 implies that every weakly continuous process x admits the representation

$$x(t,\omega) = \sum_k a_k(t)\, \eta_k(\omega) \; ; \quad t \in T \tag{17}$$

in the stochastic mean on T. However, the time functions $\{a_k(t)\}_k$ in (17) cannot be explicitly obtained because the set $\{\eta_k(\omega)\}_k$ is not orthogonal. An explicit expression for the time functions and the random variables in the representation (1) is given in Theorem 5, which follows directly from Theorem 4.

THEOREM 5.

Every stochastic process $\{x(t,\omega), t\epsilon T\}$ weakly continuous on T admits the representation

$$x(t,\omega) = \sum_k b_k(t)\, \xi_k(\omega) \; ; \quad t \in T \tag{18}$$

where the convergence is in the stochastic mean on T. The random variables $\{\xi_k(\omega)\}_k$ are derived from $\{\eta_k(\omega)\}$ by the Gram-Schmidt orthonormalization procedure: $\xi_k(\omega) = \sum_{\ell=1}^{k} c_{k\ell}\eta_\ell(\omega)$. They are given by

$$\xi_k(\omega) = \int_T x(t,\omega)\, g_k^*(t)\, \mu(dt) \tag{19}$$

almost surely, where $g_k(t) = \sum_{\ell=1}^{k} c_{k\ell}^* \phi_\ell(t)$, satisfy $E[\xi_k \xi_\ell^*] = \delta_{k\ell}$ and constitute a complete set in $H(x,T)$. The time functions $\{b_k(t)\}_k$ are given by

$$b_k(t) = \int_T R(t,s)\, g_k(s)\, \mu(ds). \tag{20}$$

Also $R(t,s)$ admits the representation

$$R(t,s) = \sum_k b_k(t)\, b_k^*(s) \;; \quad t,s \in T \tag{21}$$

where the convergence is absolute in t and s on $T \times T$.

Let us note that (20) implies $b_k(t) \in L_2(\mu)$ for all k, and by the orthonormality of the set $\{\xi_k(\omega)\}_k$, (19) and (20) we have

$$\int_T g_k^*(t)\, b_\ell(t)\, \mu(dt) = \delta_{k\ell} \tag{22}$$

i.e., the set $\{b_k(t)\}_k$ is biorthogonal to the set $\{g_k(t)\}_k$.

In the case where μ is a finite measure, as a complete set of functions in $L_2(T, \mathcal{B}_T^1, \mu)$ one can use the general orthonormal and complete set of functions $\{\phi_k(t)\}_k$ given in [9, Thm.2] by

$$\left\{ \phi_k(t) = \frac{1}{\sqrt{\mu(T)}} \exp\left[\frac{ik2\pi}{\mu(T)}\, \mu\{(-\infty,t) \cap T\}\right]; \quad k = 0, \pm 1, \pm 2, \ldots \right\}. \tag{23}$$

If μ is the Lebesgue measure and T is the entire real line or the half line then as a complete set $\{\phi_k(t)\}_k$ one can choose the Tchebysheff-Hermite or the Tchebysheff-Laguerre functions respectively. On the other hand, if T is a finite interval then the Legendre polynomials or the trigonometric functions are applicable.

Let us illustrate how the representation (18) can be explicitly obtained, i.e., how the functions $\{b_k(t)\}_k$ and $\{g_k(t)\}_k$ can be explicitly found, in the following

<u>Example 1</u>: Let $x(t,\omega)$ be a real stochastic process defined on $T = [0, +\infty)$ with $R(t,s) = \min(t,s)$. The measure μ can be chosen as follows. Choose f as in (7): $f(t) = 1$ for $0 \le t \le 1$, and $= \frac{1}{t}$ for $1 < t < +\infty$. Choose ϕ in (8) by: $\phi(t) = e^{-t}$ for $0 \le t \le 1$, and $= te^{-t}$ for $1 < t < +\infty$. Then by (8),

$$[\frac{d\mu}{dm}](t) = F(t) = e^{-t} . \tag{24}$$

As a complete set $\{\phi_k(t)\}_k$ in $L_2(T,\mathcal{B}_T^1,\mu)$ one can choose

$$\phi_k(t) = t^k; \quad k = 0,1,2,\dots \tag{25}$$

Then the functions $\{b_k(t)\}_k$ and $\{g_k(t)\}_k$ in (18) and (19) are given by

$$g_k(t) = \sum_{\ell=0}^{k} c_{k\ell}\, t^{\ell}; \quad k = 0,1,2,\dots \tag{26}$$

$$b_k(t) = \sum_{\ell=0}^{k} c_{k\ell}[\gamma(\ell+2,t) + t\Gamma(\ell+1,t)]; \quad k = 0,1,2,\dots \tag{27}$$

where γ and Γ are the incomplete gamma functions [4], and the coefficients in the Gram-Schmidt orthonormalization procedure are easily expressed in terms of the values of the integrals

$$I_{n,m} = E[\eta_n \eta_m] = \int_0^{+\infty}\int_0^{+\infty} \min(t,s)t^n s^m e^{-t} e^{-s}\, dt\, ds$$

$$= n!(m+1)! + J_{n+1,m} - J_{n,m+1} \tag{28}$$

where $J_{n,m} = (n!m!/2^{n+1})\sum_{k=0}^{m}(1/2^k)\binom{n+k}{k}$.

Upon making a particular choice for the complete set of functions $\{\phi_k(t)\}_k$ used in (10), a representation very similar to the usual Karhunen-Loève representation is obtained in Theorem 6 and is valid over any Lebesgue set of the real line. The results are similar to those in [6]. However, the representation derived in Theorem 6 applies to all weakly continuous stochastic processes, all mean square continuous processes included, while the representation of [6] applies to the class of mean square continuous stochastic processes satisfying (4), which clearly does not contain the entire class of mean square continuous stochastic processes.

THEOREM 6.

Every stochastic process $\{x(t,\omega),\ t\epsilon T\}$ weakly continuous on T admits the representation

$$x(t,\omega) = \sum_k f_k(t)\ \xi_k(\omega)\ ;\quad t \in T \tag{29}$$

where the convergence is in the stochastic mean on T. $\{f_k(t)\}_k$ and $\{\lambda_k\}_k$ are the corresponding eigenfunctions and nonzero eigenvalues of the operator on $L_2(\mu)$ with kernel $R(t,s)$. The random variables $\{\xi_k(\omega)\}_k$ are defined by

$$\xi_k(\omega) = \int_T x(t,\omega)\ f_k^*(t)\ \mu(dt) \tag{30}$$

almost surely, satisfy $E[\xi_k\xi_\ell^*] = \lambda_k\delta_{k\ell}$ and constitute a complete set in $H(x,T)$. Also $R(t,s)$ admits the representation

$$R(t,s) = \sum_k \lambda_k\ f_k(t)\ f_k^*(s);\quad t,s \in T \tag{31}$$

where the convergence is absolute in t and s on $T\times T$.

<u>**Proof:**</u> It follows from $R(t,s)\epsilon L_2(\mu\times\mu)$ that $R(t,s)$ is the kernel of a Hilbert-Schmidt operator R from $L_2(\mu)$ to $L_2(\mu)$. Hence R is a self-adjoint, non-negative definite, completely continuous operator [1, pp.54-56, 58-59]; it is also of the trace class. Its nonzero eigenvalues $\{\lambda_k\}_k$ and the corresponding eigenfunctions $\{f_k(t)\}_k$ satisfy the integral equation $Rf_k = \lambda_k f_k$ [1, pp.124-129]. The set $\{f_k(t)\}_k$ is orthonormal in $L_2(\mu)$ and complete in the range of R. Let $\{h_\ell(t)\}_\ell$ be an orthonormal basis in the orthogonal complement of the range of R in $L_2(\mu)$. By applying Theorem 4, we obtain

$$x(t,\omega) = \sum_k a_k(t)\ \xi_k(\omega) + \sum_\ell d_\ell(t)\ \zeta_\ell(\omega); \quad t \in T \tag{32}$$

in the stochastic mean, where the ξ_k's are given by (30) and $\zeta_\ell(\omega) = \int_T x(t,\omega)h_\ell^*(t)\mu(dt)$ almost surely. Since $Rh_\ell = 0$ in $L_2(\mu)$ for all ℓ, it follows that $E[|\zeta_\ell|^2] = (Rh_\ell,h_\ell)_{L_2(\mu)} = 0$ for all ℓ. Also $\lambda_k a_k(t) = E[x(t)\xi_k^*] = (Rf_k)(t) = \lambda_k f_k(t)$ for all k and $t\epsilon T$. Hence (29) follows from (32).

Q.E.D.

An example of a representation (29) valid over an infinite interval of the real line is now given.

<u>Example 2</u>: Consider the same stochastic process as in example 1. In order to obtain the representation (29) explicitly it suffices to find the f_k's, i.e., the eigenfunctions of the integral equation

$$\lambda f(t) = \int_0^{+\infty} \min(t,s)\ f(s)\ \mu(ds); \quad t \in [0,+\infty). \tag{33}$$

For reasons that will become clear in the sequel we choose, in a way similar to that of Example 1, μ such that

$$[\frac{d\mu}{dm}](t) = F(t) = e^{-2t} . \tag{34}$$

Then (33) can be reduced to the differential equation

$$\lambda f''(t) + e^{-2t} f(t) = 0 ; \quad t \in [0,+\infty) \tag{35}$$

with boundary conditions $f(o) = 0$ and $\lim_{t\to+\infty} f(t)$: finite. It follows from the solution of the differential equation (35) [5, Section 5.12], that the nonzero eigenvalues $\{\lambda_k\}_k$ and the corresponding eigenfunctions $\{\phi_k(t)\}_k$ of the integral equation (33) are given by

$$\left\{\lambda_k = \frac{1}{\mu_k^2} , \quad f_k(t) = \frac{\sqrt{2}}{J_1(\mu_k)} J_0(\mu_k e^{-t}); \quad k = 1,2,\ldots\right\} \tag{36}$$

where J_p is the Bessel function of the first kind of order p, and $\{\mu_k\}_k$ are the positive zeros of J_0.

In connection with the representations in the stochastic mean sense given in Theorem 5 and 6 for any weakly continuous stochastic process, and a fortiori for any mean square continuous stochastic process, the following remarks should be made.

Remark 1: In comparing the two representations given in Theorems 5 and 6, the following should be noticed:

(i) In both representations, the random variables are orthogonal.

(ii) The time functions in the representation of Theorem 6 are orthogonal, while in the representation of Theorem 5 they are not.

(iii) The representation of Theorem 6 requires the computation of the eigenfunctions of an integral equation, while the representation of Theorem 5 can be obtained explicitly in a straightforward way.

Remark 2: For series of independent random variables, it is known [8, p.251] that convergence in the mean square sense implies almost sure convergence. It follows that, if the weakly continuous stochastic process x is Gaussian, then both representations (18) and (29) of Theorems 5 and 6 converge almost surely for every $t \in T$.

Remark 3: It follows from the representations of $R(t,s)$ given in Theorems 5 and 6, (21) and (31), and the bounded convergence theorem that

$$\sum_k \int_T |b_k(t)|^2 \mu(dt) = \int_T R(t,t) \mu(dt) = \sum_k \lambda_k < \infty . \quad (37)$$

The measurability of x and (37) imply that if $x_n(t,\omega)$ is the sum of the first n terms in the representation (18) of Theorem 5 or in the representation (29) of Theorem 6, then

$$\lim_{n \to \infty} E\left[\int_T |x(t,\omega) - x_n(t,\omega)|^2 \mu(dt)\right] = \lim_{n \to \infty} \int_T E[|x(t,\omega) - x_n(t,\omega)|^2] \mu(dt)$$

$$= 0. \quad (38)$$

Remark 4: By using the properties of the autocorrelation function of a weakly continuous stochastic process [10, Thm. 2E], it can be shown that every function $f(t)$ of the form $f = Rg$, where $g \in L_2(\mu)$, is continuous in t on T. In particular, the functions $b_k(t)$ in the representation (18) of Theorem 5, given by (20), as well as the eigenfunctions $f_k(t)$ of $R(t,s)$ used in the representation (29) of Theorem 6, are continuous functions of t on T.

Remark 5: If x is mean square continuous on T, it can be shown, by using Dini's theorem [11, p.136], that the convergence in both representations of x given in Theorems 5 and 6 is uniform on compact subsets of T. The same is true for the representations of $R(t,s)$, (21) and (31). If in particular T is compact, then an alternative proof to the Karhunen-Loève representation is thus obtained, a proof which does not rely upon Mercer's theorem. Moreover, an alternative proof to Mercer's theorem is obtained.

Remark 6: As a brief comment on the relationship of the main earlier representations [6,8,9,2] to those derived in Theorems 5 and 6 the following should be noted. Karhunen's representation [6] of mean square continuous stochastic processes satisfying (4) is clearly a particular case of Theorem 6. As noted in Remark 5, the so-called Karhunen-Loève representation is also a particular case of Theorem 6. Also, it is easily seen that both representations of Theorems 5 and 6 can be written in the form of the representations derived in [9] and in [2] over the entire real line for the class of mean square continuous, wide sense stationary stochastic processes and for the class of harmonizable stochastic processes respectively.

Remark 7: It should be noted that for the results of this section the measurability of the process $x(t,\omega)$ is not essential, provided $R(t,s)$ is measurable and "smooth" in the sense specified in the sequel. If the measurability of $R(t,s)$ is assumed, but not that of $x(t,\omega)$, then the random variables $\{\eta_k(\omega)\}_k$ cannot be defined by (10) almost surely anymore. Let us note that the functions $\{\phi_k(t)\}_k$ can always be chosen continuous on T, as for instance in (23). Then, if $R(t,s)$ is such that the integrals $\int_T\int_T R(t,s)\phi_k^*(t)\phi_k(s)\mu(dt)\mu(ds)$, which always exist and are finite in the Lebesgue sense, exist also in the Riemann sense, the integrals (10) can be defined in the usual stochastic mean sense [8]. This is clearly possible if x is mean square continuous on T. It should be noted in conclusion that whenever both integrals exist, their values are almost surely equal and that the almost sure integrals are more convenient from a practical point of view, since they can be calculated from realizations of the process.

4. Sample Function Representation

In this section, the process x and the measure μ are defined as in the beginning of Section 3.

The fact that almost all sample functions of the process x belong to $L_2(\mu)$, shown in Theorem 2 of Section 2, implies in a straightforward way the following.

Theorem 7.

If $\{\phi_k(t)\}_k$ is an arbitrary complete set of orthonormal functions in $L_2(\mu)$, then x admits the representation

$$x(t,\omega) = \sum_k \phi_k(t)\, \eta_k(\omega) \tag{39}$$

where the convergence is in $L_2(\mu)$ almost surely, and the random variables $\{\eta_k(\omega)\}_k$ are given almost surely by

$$\eta_k(\omega) = \int_T x(t,\omega)\, \phi_k^*(t)\, \mu(dt). \tag{40}$$

The remark which follows Theorem 5 on how the functions $\{\phi_k(t)\}_k$ can be found in all cases under consideration, applies here also.

It follows from (39), Parceval's relationship and the monotone convergence theorem that

$$\int_T R(t,t)\, \mu(dt) = \sum_k \rho_{kk} \tag{41}$$

where $\rho_{kk} = E[|\eta_k|^2] = (R\phi_k,\phi_k)_{L_2(\mu)}$. If $x_n(t,\omega)$ is the sum of the first n terms in the representation (39) of Theorem 7, then (41) implies that

$$\lim_{n \to \infty} E\left[\int_T |x(t,\omega) - x_n(t,\omega)|^2 \mu(dt)\right] = 0. \quad (42)$$

The representation of Theorem 7 clearly applies to weakly continuous processes x. For this class of processes, it is now shown that both representations given in Theorems 5 and 6 are also sample function representations.

THEOREM 8.

For every weakly continuous stochastic process x the representations (18) and (29) given in Theorems 5 and 6 converge also in $L_2(\mu)$ almost surely.

Proof: (i) For the representation (18) of Theorem 5. Consider the functions $\{b_k(t)\}_k$ given by (20): $b_k = Rg_k$, and let $\{d_\ell(t)\}_\ell$ be an orthonormal basis in the orthogonal complement of the span of the set $\{b_k(t)\}_k$ in $L_2(\mu)$. Then we have

$$x(t,\omega) = \sum_k b_k(t)\, \beta_k(\omega) + \sum_\ell d_\ell(t)\, \delta_\ell(\omega) \quad (43)$$

in $L_2(\mu)$ almost surely (a.s.). Eq. (22) implies that $\beta_k(\omega) = \xi_k(\omega)$ a.s., where the ξ_k's are given by (19). Also $\delta_\ell(\omega) = \int_T x(t,\omega) d_\ell^*(t)\mu(dt)$ a.s. It follows that $E[\delta_\ell \xi_k^*] = (Rg_k, d_\ell)_{L_2(\mu)} = (b_k, d_\ell)_{L_2(\mu)} = 0$ for all k and ℓ, and hence δ_ℓ is orthogonal to $H(x,T)$ for all ℓ. Since Theorem 3 implies $\delta_\ell \epsilon H(x,T)$ for all ℓ, we have $E[|\delta_\ell|^2] = 0$ for all ℓ. It follows now from (43) that (18) converges in $L_2(\mu)$ a.s.

(ii) For the representation (29) of Theorem 6. Apply Theorem 7 for the complete orthonormal set $\{f_k(t)\}_k \cup \{h_\ell(t)\}_\ell$ in $L_2(\mu)$ introduced in the proof of Theorem 6 and proceed as in (i).

Q.E.D.

It should be noted in conclusion that implicit in Theorem 8 is the fact that almost all sample functions of the weakly continuous process x belong to the span of the eigenfunctions $\{f_k(t)\}_k$, which is the orthogonal complement in $L_2(\mu)$ of the null space of the operator R.

REFERENCES

1. N.I. Akhiezer and I.M. Glazman, Theory of Linear Operators in Hilbert Space, Vol. 1, Ungar Publishing Co., New York, 1961.

2. S. Cambanis and B. Liu, On Harmonizable Stochastic Processes, submitted to Information and Control.

3. J.L. Doob, Stochastic Processes, Wiley, New York, 1953.

4. A. Erdélyi, Higher Transcendental Functions, Vol. II. McGraw-Hill, New York, 1953.

5. F.B. Hildebrand, Advanced Calculus for Applications, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

6. K. Karhunen, Zur Spectraltheorie stochastischer Prozesse, Ann. Acad. Scient. Fennicae, Ser. AI, No. 34, (1946), pp. 1-7.

7. K. Karhunen, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, Ann. Acad. Scient. Fennicae, Ser. AI, No. 37 (1947), pp. 1-79.

8. M. Loève, Probability Theory, Van Nostrand, Princeton, New Jersey, 1963.

9. E. Masry, B. Liu and K. Steiglitz, Series Expansion of Wide-Sense Stationary Random Processes, IEEE Trans. Information Theory, Vol. IT-14 (1968), pp.792-796.

10. E. Parzen, Statistical Inference on Time Series by Hilbert Space Methods, in Time Series Analysis Papers, E. Parzen, Holden-Day, San Francisco, 1967, pp. 251-382.

11. W. Rudin, Principles of Mathematical Analysis, McGraw-Hill, New York, 1964.