

* Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514. This research was supported by the Air Force Office of Scientific Research under Grant AFOSR-68-1415 and by the Office of Naval Research under Grant N00014-67A-0321-0006.

THE EQUIVALENCE OR SINGULARITY OF STOCHASTIC PROCESSES
AND OF THE MEASURES THEY INDUCE ON L_2

by

Stamatis Cambanis*

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

Institute of Statistics Mimeo Series No. 759

July, 1971

THE EQUIVALENCE OR SINGULARITY OF STOCHASTIC PROCESSES
AND OF THE MEASURES THEY INDUCE ON L_2

Stamatis Cambanis*

ABSTRACT

It is shown that the proposition: "Two stochastic processes are equivalent or singular if and only if the measures they induce on an appropriate L_2 space are equivalent or singular respectively", is not true in general, and sufficient conditions are given for its validity. For a wide class of square integrable martingales it is shown that this proposition is valid and a number of results are obtained which generalize known results for the Wiener process.

* Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514. This research was supported by the Air Force Office of Scientific Research under Grant AFOSR-68-1415 and by the Office of Naval Research under Grant N00014-67A-0321-0006.

1. INTRODUCTION

Problems in detection theory and in classification theory or pattern recognition reduce to studying the equivalence or singularity of the probability measures induced on the probability space by the stochastic processes under consideration; and in computing the Radon-Nikodym derivative in the former case. These processes induce probability measures on appropriate L_2 spaces, and the original detection or discrimination problem is often treated on this Hilbert space where a powerful structure is available (see for instance [16, 17, 1]).

It is the purpose of this paper to study the validity of the proposition stated in the abstract. It is shown that this proposition is not true in general (Proposition 1 and Theorem 4); specifically, two stochastic processes may be singular and yet their induced measures on L_2 be equivalent (Theorem 4). Under appropriate continuity conditions it is shown that the proposition is valid (Theorem 2). The second order stochastic processes for which the proposition is shown to be valid are those that are continuous in probability or smooth, a notion introduced in [4]. Weakly, as well as mean square, continuous processes are smooth and a further class of smooth processes is given in Section 3.

It is shown there that all square integrable martingales satisfying certain conditions are smooth stochastic processes (Theorem 5). Also their linear span, their reproducing kernel Hilbert space, and the range of the square root of their covariance operator are characterized (Theorem 6); these spaces play a significant role, especially in the Gaussian case. These results generalize well known facts for the Wiener process and, in the case of the reproducing kernel Hilbert space, for mean square continuous processes with orthogonal increments.

Finally, the relationship between the reproducing kernel Hilbert space of a stochastic process and the range of the square root of its covariance operator is established in the most general case (Theorem 3), generalizing a known result for mean square continuous processes defined on compact intervals. This relationship is of primary importance in the Gaussian case.

2. THE RELATIONSHIP BETWEEN THE EQUIVALENCE OR SINGULARITY OF STOCHASTIC PROCESSES AND OF THE MEASURES THEY INDUCE ON L_2 .

Throughout this section the following notation is used. P_0 and P_1 are two probability measures on (Ω, \mathcal{F}) with respect to which the measurable stochastic process $\{x(t, \omega), t \in T\}$ is second order with mean, autocorrelation and covariance functions $m_0(t)=0$, $r_0(t, s)$, $R_0(t, s)$ and $m_1(t)=m(t)$, $r_1(t, s)$, $R_1(t, s)$ respectively; \mathcal{F} is the σ -algebra of subsets of Ω generated by the random variables $\{x(t, \omega), t \in T\}$, and T is any interval on the real line R , open or closed, bounded or unbounded.

DEFINITION 1. N is the class of measures ν on $(T, \mathcal{B}(T))$ ($\mathcal{B}(T)$ is the σ -algebra of Lebesgue measurable subsets of T) which are equivalent to the Lebesgue measure, $\nu \sim \text{Leb}$, (i.e. mutually absolutely continuous) and satisfy

$$\int_T r_i(t, t) d\nu(t) < +\infty, \quad i=0,1. \quad (1)$$

N is a nonempty class of measures, as it is demonstrated by the following construction. Define ν on $(T, \mathcal{B}(T))$ by $[d\nu/d \text{Leb}](t) = f(t)g(t)$, where $g \in L_1(T, \mathcal{B}(T), \text{Leb})$, $g(t) > 0$ a.e. $[\text{Leb}]$ on T , and

$$f(t) = \begin{cases} 1 & \text{if } 0 \leq \max\{r_0(t,t), r_1(t,t)\} \leq 1 \\ (\max\{r_0(t,t), r_1(t,t)\})^{-1} & \text{if } 1 < \max\{r_0(t,t), r_1(t,t)\} \end{cases}$$

Clearly $v \in \mathbb{N}$ and also v is finite.

For every $v \in \mathbb{N}$, the measurability of $x(t, \omega)$ and (1) imply that $x(\cdot, \omega) \in L_2(T, \mathcal{B}(T), \nu) = L_2(\nu)$ a.s. $[P_i]$, $i=0,1$. Also the map $T: (\Omega, \mathcal{F}) \rightarrow (H^v = L_2(\nu), \mathcal{B}(H^v))$, where $\mathcal{B}(H^v)$ is the σ -algebra generated by the open subsets of H^v , defined by

$$T \omega = x(\cdot, \omega) \quad (2)$$

is measurable [15], and thus the probability measures P_i , $i=0,1$, on (Ω, \mathcal{F}) induce probability measures μ_i^v on $(H^v, \mathcal{B}(H^v))$ defined by

$$\mu_i^v(B) = P_i\{\omega \in \Omega: x(\cdot, \omega) \in B\} \quad (3)$$

for all $B \in \mathcal{B}(H^v)$. It follows by (1) and (3) that

$$\begin{aligned} E_i[|||u|||_{H^v}^2] &= \int_{H^v} |||u|||_{H^v}^2 d\mu_i^v(u) = \int_{\Omega} |||x(\cdot, \omega)|||_{H^v}^2 dP_i(\omega) \\ &= \int_T r_i(t,t) d\nu(t) < +\infty. \end{aligned}$$

Hence [13] the mean elements and the covariance operators of the probability measures μ_i^v , $i=0,1$, are defined as the elements $u_i^v \in H^v$ and as the bounded, linear, nonnegative, self-adjoint, trace class operators S_i^v on H^v which satisfy

$$\begin{aligned} E_i[<u, v>_{H^v}] &= <u_i^v, v>_{H^v} \\ E_i[<u-u_i^v, v>_{H^v} <u-u_i^v, w>_{H^v}] &= <S_i^v v, w>_{H^v} \end{aligned}$$

for all $v, w \in H^v$. It is easily seen that $u_i^v = m_i$ and that S_i^v is an integral operator with kernel $R_i(t,s)$ [15]. It is also shown in [15] that if the stochastic process $\{x(t,\omega), t \in T\}$ is Gaussian with respect to P_i then the induced measure μ_i^v is Gaussian.

A remark should be made about the need to consider the family of measures N . What is really needed is that the probability measures P_i induce probability measures μ_i on some space of square integrable functions on the interval T , and that μ_i have covariance operators. A sufficient condition for this induction to be possible is that $v \in N$, in which case the space of square integrable functions on T is H^v . In particular, it is shown in [15], that if $x(t,\omega)$ is Gaussian with respect to P_i , then $v \in N$ is also a necessary condition. If the Lebesgue measure satisfies (1), then the natural choice of $v \in N$ is clearly $v = \text{Leb}$. This is the case if, for instance, $x(t,\omega)$ is mean square continuous with respect to both P_0 and P_1 and T is a compact interval. However the Lebesgue measure is not always in N . For example, if $x(t,\omega)$ is wide sense stationary with respect to P_0 and P_1 and T is an unbounded interval, then the Lebesgue measure is not in N , but every finite measure equivalent to the Lebesgue measure belongs to N .

Before we consider the relationship between the pairs of measures (P_0, P_1) and (μ_0^v, μ_1^v) , we prove that the equivalence (\sim) or singularity (\perp) of the measures μ_0^v and μ_1^v does not depend on $v \in N$.

THEOREM 1. If any one of the following relations is satisfied for some $v \in N$:

$$\mu_1^v \ll \mu_0^v, \quad \mu_1^v \sim \mu_0^v, \quad \mu_1^v \perp \mu_0^v$$

then it is satisfied for all $v \in N$.

For the proof of Theorem 1 we need the following

LEMMA 1 [4, Theorem 2]. Let $v \in N$, $\{f_k^v(t)\}_{k=1}^\infty$ be a complete set of functions in $L_2(v)$, and the random variables $\{\eta_k^v(\omega)\}_{k=1}^\infty$ be defined by

$$\eta_k^v(\omega) = \int_T x(t, \omega) f_k^v(t) dv(t) \quad \text{a.e. } [P_i]. \quad (4)$$

Then the subspace $H_i(x, \{f_k^v\}_k, v)$ of $L_2(\Omega, F, P_i)$, which is spanned by the random variables $\{\eta_k^v(\omega)\}_k$, does not depend on $v \in N$ and on the complete set $\{f_k^v\}_k$ in $L_2(v)$; it is denoted by $H(x, \text{smooth})$.

PROOF OF THEOREM 1. For $v \in N$ and $H^v = L_2(v)$ define

$$G^v = T^{-1}[B(H^v)] = \{T^{-1}(B), B \in B(H^v)\}.$$

Then G^v is a σ -algebra of subsets of Ω and because of the measurability of T (defined by (2)), $G^v \subseteq F$. It is clear from (3) that the measures μ_i^v depend only on the values of the measures P_i on G^v . We will denote by G_i^v, F_i the completed σ -algebras with respect to the measure $P_i, i=0,1$, and prove that for all $v, \lambda \in N$

$$G_i^v = G_i^\lambda, \quad i=0,1. \quad (5)$$

If $\{f_k^v\}_k$ and $\{\eta_k^v\}_k$ are as in Lemma 1, then $B(H^v)$ is the smallest σ -algebra of subsets of H^v with respect to which the linear functionals $\{F_k^v(u) = \langle u, f_k^v \rangle_{H^v}\}_{k=1}^\infty$ are measurable, and since $\eta_k^v(\omega) = (F_k^v \circ T)(\omega)$, it follows that $G^v = F^v(\eta)$, where $F^v(\eta)$ is the σ -algebra of subsets of Ω generated by $\{\eta_k^v(\omega)\}_{k=1}^\infty$. Because of the measurability of $x(t, \omega)$ and (4), $F^v(\eta) \subseteq F$. Thus it suffices to prove

$$F_i^v(\eta) = F_i^\lambda(\eta), \quad i=0,1. \quad (6)$$

Because of the symmetry, it is enough to show $F_i^\lambda(\eta) \subseteq F_i^v(\eta)$, and this for

$i=0$. Since by Lemma 1 $H_0(x, \{f_k^\nu\}_k, \nu) = H_0(x, \{f_k^\lambda\}_k, \lambda)$, it follows that for every j , $\eta_j^\lambda \in H_0(x, \{f_k^\nu\}_k, \nu)$ and thus for some constants $\{a_{jk}\}_k$,

$$\eta_j^\lambda(\omega) = \sum_{k=1}^{\infty} a_{jk} \eta_k^\nu(\omega)$$

where the convergence is in $L_2(\Omega, \mathcal{F}, P_0)$. Hence along some subsequence

$$\eta_j^\lambda(\omega) = \lim_{n \rightarrow \infty} \sum_{k=1}^n a_{jk} \eta_k^\nu(\omega) \quad \text{a.e. } [P_0]$$

which implies that $\eta_j^\lambda(\omega)$ is $F_0^\nu(\eta)$ - measurable for all j , and thus $F_0^\lambda(\eta) \subseteq F_0^\nu(\eta)$.

Note that (5) implies that every set F in G^λ is of the form $F = A_i \cup C_i$, where $A_i \in G^\nu$ and C_i is a subset of a G^ν set of P_i measure zero. Clearly C_i can be taken disjoint from A_i (if not replace by $C_i \setminus A_i$) and in this case $C_i \in F$. This will be used in the following.

We now complete the proof of the theorem by proving the following properties

$$\mu_1^\nu \ll \mu_0^\nu \quad \text{if and only if} \quad P_{1/G^\nu} \ll P_{0/G^\nu} \quad (7)$$

$$P_{1/G^\nu} \ll P_{0/G^\nu} \quad \text{if and only if} \quad P_{1/G^\lambda} \ll P_{0/G^\lambda} \quad (8)$$

$$\mu_1^\nu \perp \mu_0^\nu \quad \text{if and only if} \quad P_{1/G^\nu} \perp P_{0/G^\nu} \quad (9)$$

$$P_{1/G^\nu} \perp P_{0/G^\nu} \quad \text{if and only if} \quad P_{1/G^\lambda} \perp P_{0/G^\lambda} \quad (10)$$

where P_{i/G^ν} denotes the restriction of P_i to G^ν . Clearly (7) and (8) show that for any $\nu, \lambda \in N$, $\mu_1^\nu \ll \mu_0^\nu$ if and only if $\mu_1^\lambda \ll \mu_0^\lambda$ and hence, $\mu_1^\nu \sim \mu_0^\nu$ if and only if $\mu_1^\lambda \sim \mu_0^\lambda$, and (9) and (10) imply that $\mu_1^\nu \perp \mu_0^\nu$ if and only if $\mu_1^\lambda \perp \mu_0^\lambda$. For notational convenience we will put $P_{i/G^\nu} = P_i^\nu$.

Proof of (7). Assume $\mu_1^\nu \ll \mu_0^\nu$ and let $F \in G^\nu$ be such that $P_0(F) = 0$. Since $F \in G^\nu$, $F = T^{-1}(B)$ for some $B \in \mathcal{B}(H^\nu)$. Then $P_0(F) = 0$ and (3) imply

$\mu_0^v(B) = 0$, hence $\mu_1^v(B) = 0$ and by (3), $P_1(F) = 0$. Thus $P_1^v \ll P_0^v$.

Conversely assume $P_1^v \ll P_0^v$ and let $B \in \mathcal{B}(H^v)$ be such that $\mu_0^v(B) = 0$. Then, by (3), $P_0(F) = 0$, where $F = T^{-1}(B) \in G^v$. Hence $P_1(F) = 0$ and by (3), $\mu_1^v(B) = 0$, which implies $\mu_1^v \ll \mu_0^v$.

Proof of (8). Because of the symmetry it suffices to prove that $P_1^v \ll P_0^v$ implies $P_1^\lambda \ll P_0^\lambda$. Assume $P_1^v \ll P_0^v$ and let $F \in G^\lambda$ be such that $P_0(F) = 0$. Since $F \in G^\lambda$, we have $F = A_0 \cup C_0$, where $A_0 \in G^v$, $C_0 \subseteq D_0 \in G^v$ and $P_0(D_0) = 0$, which implies $P_1(D_0) = 0$. Now $P_0(F) = 0$ implies $P_1(A_0) = 0$ and thus $P_1(F) \leq P_1(A_0 \cup D_0) \leq P_1(A_0) + P_1(D_0) = 0$, i.e. $P_1(F) = 0$, which proves that $P_1^\lambda \ll P_0^\lambda$.

Proof of (9). Assume $\mu_1^v \perp \mu_0^v$. Then there exists $B \in \mathcal{B}(H^v)$ such that $\mu_1^v(B) = 1 = \mu_0^v(B^c)$. If $F = T^{-1}(B)$, then $F \in G^v$ and by (3), $P_1(F) = 1 = P_0(F^c)$ which proves that $P_1^v \ll P_0^v$.

Conversely assume $P_1^v \perp P_0^v$. Then there exists $F \in G^v$ such that $P_1(F) = 1 = P_0(F^c)$. Since $F \in G^v$, we have $F = T^{-1}(B_1)$, $F^c = T^{-1}(B_0)$ for some $B_0, B_1 \in \mathcal{B}(H^v)$. It follows by (3) that $\mu_1^v(B_1) = 1 = \mu_0^v(B_0)$. But $T^{-1}(B_1 \cap B_0) = T^{-1}(B_1) \cap T^{-1}(B_0) = F \cap F^c = \emptyset$ and by (3), $\mu_1^v(B_1 \cap B_0) = 0$. Hence $\mu_1^v(B_1 \setminus B_0) = 1 = \mu_0^v(B_0 \setminus B_1)$, which implies $\mu_1^v \perp \mu_0^v$.

Proof of (10). Because of the symmetry it suffices to show that $P_1^v \perp P_0^v$ implies $P_1^\lambda \perp P_0^\lambda$. Assume $P_1^v \perp P_0^v$. Then there exists $F \in G^v$ such that $P_1(F) = 1 = P_0(F^c)$. Since $F, F^c \in G^v$ we have $F = A_1 \cup C_1$, $F^c = A_0 \cup C_0$, where $A_i \in G^\lambda$, $C_i \in F$ and $P_i(C_i) = 0$, $i=0,1$. It follows that $P_1(A_1) = 1 = P_0(A_0)$ and since $A_1 \cap A_0 = \emptyset$, $P_1^\lambda \perp P_0^\lambda$. \square

We now consider the relationship between the equivalence or singularity of the pairs of measures (P_0, P_1) and (μ_0^v, μ_1^v) , $v \in N$. In view of Theorem 1 it suffices to do this for a fixed $v \in N$. Thus in the following the superscript indicating the dependence of H , μ_i and S_i on $v \in N$ will be omitted.

The proof of the following two propositions is contained in the proof of Theorem 1.

PROPOSITION 1.

- (1) $P_1 \ll P_0$ implies $\mu_1 \ll \mu_0$
- (2) $P_1 \sim P_0$ implies $\mu_1 \sim \mu_0$
- (3) $\mu_1 \perp \mu_0$ implies $P_1 \perp P_0$

In general the inverses of (1), (2) and (3) of Proposition 1 are not true. This is not surprising because of the following. Let G be the σ -algebra of subsets of Ω which is the inverse image of $B(H)$ under T , defined by (2). Since T is measurable, $G \subset F$ and it is clear from (3) that the measures μ_i depend only on the values of the measures P_i on G . Thus if F is larger than G it may happen that $P_1 \perp P_0$ on F and yet $P_1 \sim P_0$ on G . That this is not an hypothetical situation, it is proven in Theorem 4. Thus Proposition 1 cannot be improved unless restrictive assumptions are made. An appropriate assumption would be that F is not essentially larger than G , specifically that $G_i = F_i$, where the subscript i denotes completion with respect to the measure P_i , $i=0,1$. The implications of this assumption are stated in Proposition 2 and it is shown in Theorem 2 that a large class of processes satisfies this assumption.

PROPOSITION 2. If $G_i = F_i$, $i=0,1$, then

- (1) $P_1 \ll P_0$ if and only if $\mu_1 \ll \mu_0$
- (2) $P_1 \sim P_0$ if and only if $\mu_1 \sim \mu_0$
- (3) $P_1 \perp P_0$ if and only if $\mu_1 \perp \mu_0$

DEFINITION 2. Define the following classes of real, measurable stochastic processes:

S_1 is the class of continuous in probability stochastic processes $\{x(t, \omega), t \in T\}$ for which there exists a σ -finite measure ν on $(T, \mathcal{B}(T))$ such that $\text{Leb} \ll \nu$ and $x(\cdot, \omega) \in L_2(\nu)$ a.s.

S_2 is the class of continuous in probability, second order stochastic processes $\{x(t, \omega), t \in T\}$.

S_3 is the class of smooth, second order stochastic processes $\{x(t, \omega), t \in T\}$, defined in [4, Theorem 4].

S is the union of S_1, S_2 and S_3 .

It is shown in [4] that the weakly continuous, and therefore the mean square continuous processes are smooth. Further classes of smooth second order processes are given in Section 3.

THEOREM 2. If with respect to both probabilities P_0 and P_1 , $\{x(t, \omega), t \in T\}$ belongs to the class S , then (1), (2) and (3) of Proposition 2 are valid.

PROOF. For $x(t, \omega)$ in S_1 a measure ν on $(T, \mathcal{B}(T))$ is given. For $x(t, \omega)$ in $S_2 \cup S_3$ take ν to be any measure in N . Let $\{f_k(t)\}_{k=1}^{\infty}$ be any complete set in $L_2(\nu)$ and the random variables $\{\eta_k(\omega)\}_{k=1}^{\infty}$ be defined as in (4). Denote by $F(\eta)$ the σ -algebra of subsets of Ω generated by $\{\eta_k(\omega)\}_{k=1}^{\infty}$. Then

$$F_i(\eta) = F_i \quad (11)$$

This is shown in [3, 11] for the class S_1 ; in [11] for the class S_2 ; and for the class S_3 in the same way as (6) or as in [11, Theorem 3]. It is shown in the proof of Theorem 1 (between (5) and (6)) that $G = F(\eta)$. It follows by (11) that $G_i = F_i$ and thus Proposition 2 applies. \square

Before we proceed we need the result stated in Theorem 3. Let $H_0(x)$ be the subspace of $L_2(\Omega, \mathcal{F}, P_0)$ spanned by the random variables $\{x(t, \omega), t \in T\}$ and let $\text{RKHS}(R_0)$ be the reproducing kernel Hilbert space of the covariance $R_0(t, s)$. It is well known [14] that $H_0(x)$ and $\text{RKHS}(R_0)$ are isomorphic with corresponding elements $x(t, \omega)$ and $R_0(\cdot, t)$ respectively, and that $\text{RKHS}(R_0)$ consists of all real valued functions f on T of the form $f(t) = E_0[\xi x(t)]$ for all $t \in T$ and some $\xi \in H_0(x)$. Let also A be the integral type operator from $H_0(x)$ to $L_2(v)$ with kernel $x(t, \omega)$. Since by (1),

$$\int_{\Omega} \int_T x^2(t, \omega) dv(t) dP_0(\omega) = \int_T r_0(t, t) dv(t) < +\infty$$

A is a Hilbert-Schmidt operator.

THEOREM 3. (1) If $f \in \text{RKHS}(R_0)$ then $f \in \text{range}(S_0^{\frac{1}{2}})$. Conversely, if $f \in \text{range}(S_0^{\frac{1}{2}})$ then f is equal a.e.[Leb] on T to a function in $\text{RKHS}(R_0)$; i.e. every equivalence class in $\text{range}(S_0^{\frac{1}{2}})$ contains a function in $\text{RKHS}(R_0)$.

$$(2) \text{ range}(A) = \text{range}(S_0^{\frac{1}{2}}).$$

The relationship between the reproducing kernel Hilbert space of a stochastic process and the range of the square root of its covariance operator, established in Theorem 3.1, plays a significant role for Gaussian processes as it is demonstrated in Theorem 4. Theorem 3 generalizes a well known result for zero mean, mean square continuous stochastic processes defined on a closed and bounded interval (see for example [8]). Note that it follows from Theorem 3 that the linear manifold $\text{range}(S_0^{\nu \frac{1}{2}})$ is invariant of $\nu \in \mathbb{N}$, in the sense explained in the theorem.

For the proof of Theorem 3 we need the following

LEMMA 2[4, Theorem 3]. Let $\{\lambda_k\}_{k=1}^{\infty}$ and $\{\phi_k\}_{k=1}^{\infty}$ be the nonzero eigenvalues and the corresponding eigenfunctions of S_0 , let $\{\phi_k(t)\}_{k=1}^{\infty}$ be the versions of the eigenfunctions which are defined for all $t \in T$ by $\lambda_k \phi_k(t) = \int_T R_0(t,s) \phi_k(s) dv(s)$, and let the random variables $\{\xi_k(\omega)\}_{k=1}^{\infty}$ be defined by $\xi_k(\omega) = \int_T x(t,\omega) \phi_k(t) dv(t)$ a.e. $[P_0]$. Then $\xi_k \in H_0(x)$, $E_0[\xi_k \xi_j] = \lambda_k \delta_{kj}$, and for all $t \in T$

$$x(t,\omega) = \sum_{k=1}^{\infty} \phi_k(t) \xi_k(\omega) + w(t,\omega) \quad (12)$$

where the equality as well as the convergence of the series are in $L_2(\Omega, F, P_0)$, and $E_0[w^2(t)] = 0$ for all $t \in T \setminus T_0$ with $\text{Leb}(T_0) = 0$.

PROOF OF THEOREM 3. Note that

$$\text{range}(S_0^{\frac{1}{2}}) = \left\{ f = \sum_{k=1}^{\infty} a_k \phi_k \text{ in } L_2(v), \sum_{k=1}^{\infty} \frac{a_k^2}{\lambda_k} < +\infty \right\} \quad (13)$$

(1.1) Let $f \in \text{RKHS}(R_0)$. Then for some $\xi \in H_0(x)$ and all $t \in T$, we obtain by (12),

$$f(t) = E_0[\xi x(t)] = \sum_{k=1}^{\infty} E_0[\xi \xi_k] \phi_k(t) + E_0[\xi w(t)]$$

and if $a_k = E_0[\xi \xi_k]$,

$$f(t) = \sum_{k=1}^{\infty} a_k \phi_k(t) \quad \text{for all } t \in T \setminus T_0 \quad (14)$$

$$\sum_{k=1}^{\infty} \frac{a_k^2}{\lambda_k} = \sum_{k=1}^{\infty} E_0\left[\xi \frac{\xi_k}{\lambda_k^{\frac{1}{2}}}\right]^2 \leq E_0[\xi^2] < +\infty \quad (15)$$

Hence $\sum_{k=1}^{\infty} a_k^2 < +\infty$ and $\sum_{k=1}^{\infty} a_k \phi_k(t)$ converges in $L_2(v)$ to a function which, because of (14), is equal a.e. $[\text{Leb}]$ on T to $f(t)$. Hence

$f = \sum_{k=1}^{\infty} a_k \phi_k$ in $L_2(\nu)$ and by (13) and (15), $f \in \text{range}(S_0^{\frac{1}{2}})$.

(1.ii) Conversely, let $f \in \text{range}(S_0^{\frac{1}{2}})$. Then $f = \sum_{k=1}^{\infty} a_k \phi_k$ in $L_2(\nu)$ with $\sum_{k=1}^{\infty} \lambda_k^{-1} a_k^2 < +\infty$. The series $\sum_{k=1}^{\infty} \lambda_k^{-1} a_k \xi_k(\omega)$ converges in $L_2(\Omega, F, P_0)$ since $\sum_{k=1}^{\infty} \lambda_k^{-2} a_k^2 E_0[\xi_k^2] = \sum_{k=1}^{\infty} \lambda_k^{-1} a_k^2 < +\infty$. Let $\xi = \sum_{k=1}^{\infty} \lambda_k^{-1} a_k \xi_k$ in $L_2(\Omega, F, P_0)$. Then $\xi \in H_0(x)$ and $E_0[\xi \xi_k] = a_k$. It follows from $f = \sum_{k=1}^{\infty} E_0[\xi \xi_k] \phi_k$ in $L_2(\nu)$, that there exists a subsequence $\{N_k\}_{k=1}^{\infty}$ such that

$$\begin{aligned} f(t) &= \lim_{k \rightarrow \infty} \sum_{n=1}^{N_k} E_0[\xi \xi_n] \phi_n(t) \quad \text{a.e. [Leb] on } T \\ &= \lim_{k \rightarrow \infty} E_0[\xi \sum_{n=1}^{N_k} \phi_n(t) \xi_n] \\ &= E_0[\xi \{x(t) - w(t)\}] = E_0[\xi x(t)] \quad \text{a.e. [Leb] on } T. \end{aligned}$$

and thus f equals a.e. [Leb] on T a function in $\text{RKHS}(R_0)$, namely $E_0[\xi x(t)]$.

(2.i) Let $f \in \text{range}(A)$. Then there exists $\xi \in H_0(x)$ such that $f(t) = (A\xi)(t) = E_0[x(t)\xi]$ a.e. [ν] on T , hence a.e. [Leb] on T . It follows as in (1.i) that $f \in \text{range}(S_0^{\frac{1}{2}})$.

(2.ii) Conversely, let $f \in \text{range}(S_0^{\frac{1}{2}})$. Then $f = \sum_{k=1}^{\infty} a_k \phi_k$ in $L_2(\nu)$ with $\sum_{k=1}^{\infty} \lambda_k^{-1} a_k^2 < +\infty$ and as in (1.ii) we obtain $f(t) = E[\xi x(t)]$ a.e. [Leb] on T , where $\xi = \sum_{k=1}^{\infty} \lambda_k^{-1} a_k \xi_k$ in $L_2(\Omega, F, P_0)$. It follows that $\xi \in H_0(x)$ and $f(t) = E_0[\xi x(t)]$ a.e. [ν] on T , i.e. $f = A\xi$ and $f \in \text{range}(A)$. \square

We now consider the case where $\{x(t, \omega), t \in T\}$ is Gaussian with respect to both probability measures P_0 and P_1 . Then the probability measures μ_0 and μ_1 are also Gaussian [15]. It is well known that both pairs of Gaussian

measures (P_0, P_1) and (μ_0, μ_1) are either equivalent or singular [9].

Necessary and sufficient conditions for $P_1 \sim P_0$ are given in [18] and for $\mu_1 \sim \mu_0$ are given in [16]. In the particular case where $R_0(t,s) = R_1(t,s) = R(t,s)$, and thus $S_0 = S_1 = S$, these conditions are:

$$P_1 \sim P_0 \quad \text{if and only if} \quad m \in \text{RKHS}(R) \quad (16)$$

$$\mu_1 \sim \mu_0 \quad \text{if and only if} \quad m \in \text{range}(S^{\frac{1}{2}}) \quad (17)$$

The relationship between the equivalence or singularity of the pairs of measures (P_0, P_1) and (μ_0, μ_1) is given in Theorem 4. This theorem demonstrates that a detection or discrimination problem, which is defined on the probability space Ω , can be treated on the Hilbert space of square integrable functions $H = L_2(\nu)$, where powerful analytic tools are available, only when case (iii) can be excluded. This is the case, for instance, if the process belongs to the class S , as it is shown in Theorem 2. However, in general, it may be that the two processes are singular and yet their induced measures on $H = L_2(\nu)$ are equivalent.

THEOREM 4. If $\{x(t, \omega), t \in T\}$ is Gaussian with respect to both probabilities P_0 and P_1 , then one of the following will always be satisfied

$$(i) \quad P_1 \sim P_0 \quad \text{and} \quad \mu_1 \sim \mu_0$$

$$(ii) \quad P_1 \perp P_0 \quad \text{and} \quad \mu_1 \perp \mu_0$$

$$(iii) \quad P_1 \perp P_0 \quad \text{and} \quad \mu_1 \sim \mu_0$$

PROOF. In view of Proposition 1 and the fact that both pairs of Gaussian measures (P_0, P_1) and (μ_0, μ_1) are either equivalent or singular, it suffices to prove that case (iii) is possible. This is shown by the following example.

Take $T = [a, b]$, $-\infty < a < b < +\infty$. We can find a probability space (Ω, \mathcal{A}, P) and independent, zero mean and unit variance, Gaussian random variables $\xi_1(\omega)$, $\xi_2(\omega)$, $\xi_3(\omega)$ on it such that the completion of $F(\xi_1, \xi_3)$ be properly included in the completion of $F(\xi_1, \xi_2, \xi_3)$, where $F(\xi_j, j \in J)$ denotes the sub- σ -algebra of \mathcal{A} generated by the random variables $\{\xi_j, j \in J\}$. Define the stochastic processes $\{x_i(t, \omega), t \in T\}$, $i=0,1$, by

$$x_0(t, \omega) = \begin{cases} d(t)\xi_1(\omega) & \text{for } a \leq t < c \\ d(c)\xi_2(\omega) & \text{for } t = c \\ d(t)\xi_3(\omega) & \text{for } c < t \leq b \end{cases} \quad \text{and all } \omega \in \Omega$$

$$x_1(t, \omega) = m(t) + x_0(t, \omega) \quad \text{for all } t \in T, \omega \in \Omega$$

where $d(t)$ and $m(t)$ are defined and finite valued for all $t \in T$, $d(c) \neq 0$, and $d(t) \neq 0$ a.e.[Leb] on T . Then clearly x_0 and x_1 are measurable, Gaussian processes. Also $F = F(x_0(t, \omega), t \in T) = F(x_1(t, \omega), t \in T)$ and x_0, x_1 induce Gaussian probabilities P_0, P_1 on (Ω, F) respectively. Take $\nu \in \mathcal{N}$ (note that if an example for a given $\nu \sim \text{Leb}$ is to be constructed it suffices to take $d(t) \in L_2(\nu)$). Then Gaussian measures μ_0, μ_1 are induced on $(H = L_2(\nu), \mathcal{B}(H))$. It is easily seen that $G_i = F_i(\xi_1, \xi_3)$ and thus G_i is properly contained in F_i (so that Proposition 2 does not apply). Note that $R_0(t, s) = R_1(t, s)$ and thus (16) and (17) apply. It is shown in Lemma 3 that given $R_0(t, s)$, i.e. given $d(t)$, one can always find a function $m(t)$, $t \in T$, such that $m \in \text{range}(S_0^{\frac{1}{2}})$ and $m \notin \text{RKHS}(R_0)$. For this choice of $d(t)$ and $m(t)$, it follows from (16) and (17) that we have $P_1 \perp P_0$ and $\mu_1 \sim \mu_0$. \square

LEMMA 3. There always exist real valued functions $f(t)$ defined for all $t \in T$ such that $f \in \text{range}(S_0^{\frac{1}{2}})$ and $f \notin \text{RKHS}(R_0)$.

PROOF. Let $g \in \text{RKHS}(R_0)$. Then, by Theorem 3.1, $g \in \text{range}(S_0^{\frac{1}{2}})$. Also there exists some $\xi \in H_0(x)$ such that $g(t) = E_0[\xi x(t)]$ for all $t \in T$, and by (12)

$$g(t) = \sum_{k=1}^{\infty} a_k \phi_k(t) + E_0[\xi w(t)] \quad (18)$$

for all $t \in T$, where

$$a_k = E_0[\xi \xi_k] = \int_T E_0[\xi x(t)] \phi_k(t) dv(t) = \langle g, \phi_k \rangle_{L_2(v)}$$

Note that for every function $f(t)$, $t \in T$, such that $f \in \text{RKHS}(R_0)$ and $f(t) = g(t)$ a.e.[Leb] on T , we have

$$f(t) = \sum_{k=1}^{\infty} a_k \phi_k(t) + E_0[\eta w(t)] \quad (19)$$

for all $t \in T$ and some $\eta \in H_0(x)$. Since by Lemma 2, $E_0[w^2(t)] = 0$ for all $t \in T \sim T_0$ and $\text{Leb}(T_0) = 0$, it follows by (18) and (19) that $f(t) = g(t)$ for all $t \in T \sim T_0$. Hence, if we choose $f(t)$, $t \in T$, equal almost everywhere [Leb] on T to $g(t)$ but not equal to $g(t)$ on $T \sim T_0$ (for instance $f(t) = g(t)$ for $t \in T \sim \{t_0\}$ and $f(t_0) \neq g(t_0)$, where $t_0 \in T \sim T_0$) we have $f \in \text{range}(S_0^{\frac{1}{2}})$ and $f \notin \text{RKHS}(R_0)$. \square

It follows from (16), (17) and Theorem 3 that, in the case where $R_0(t,s) = R_1(t,s)$, necessary and sufficient conditions for (i) to (iii) of Theorem 4 are respectively

- (i)' $m \in \text{RKHS}(R_0)$
- (ii)' $m \notin \text{range}(S_0^{\frac{1}{2}})$
- (iii)' $m \notin \text{RKHS}(R_0)$ and $m \in \text{range}(S_0^{\frac{1}{2}})$.

Necessary and sufficient conditions for the general case $R_0(t,s) \neq R_1(t,s)$ can be obtained in a similar way; they are not given here for space considerations.

It is shown in Theorem 3.1 that $f \in \text{range}(S_0^{\frac{1}{2}})$ implies that f is equal a.e. [Leb] on T to a function in $\text{RKHS}(R_0)$. By combining (17), Theorem 2 and (16) we obtain for the mean function the following

COROLLARY. If with respect to both probabilities P_0 and P_1 , $\{x(t, \omega), t \in T\}$ is Gaussian and belongs to the class S , then $m \in \text{range}(S_0^{\frac{1}{2}})$ implies $m \in \text{RKHS}(R_0)$.

3. A CLASS OF SMOOTH SQUARE INTEGRABLE MARTINGALES

Let $\{x(t, \omega), t \in T\}$ be a real, measurable, second order stochastic process defined on the probability space (Ω, \mathcal{F}, P) with autocorrelation and covariance functions $r(t, s)$ and $R(t, s)$; T is any interval on the real line. Denote by $H(x)$ the subspace of $L_2(\Omega, \mathcal{F}, P)$ spanned by the random variables $\{x(t, \omega), t \in T\}$. The subspace $H(x, \text{smooth})$ defined in Lemma 1 is always included in $H(x)$ (in the present case N is as in Definition 1 for one value of the index i , so that $P_i = P$). The stochastic process $\{x(t, \omega), t \in T\}$ is called *smooth* if and only if $H(x) = H(x, \text{smooth})$ [4].

In the following the class of square integrable martingales will be considered. A stochastic process $\{x(t, \omega), t \in T\}$ is a square integrable martingale if $x(t, \omega) \in L_2(\Omega, \mathcal{F}, P)$ for all $t \in T$, and if $E[x(t, \omega) | \mathcal{F}_s] = x(s, \omega)$ a.s. for all $s \leq t$ in T , where \mathcal{F}_s is the sub- σ -algebra of \mathcal{F} generated by the random variables $\{x(u, \omega), u \in T, u \leq s\}$. Note that all stochastic processes with orthogonal increments are square integrable martingales. If $\{x(t, \omega), t \in T\}$ is a square integrable martingale then we have the following. There exists a monotone nondecreasing function $F(t)$ on T such that

$$E[\{x(t) - x(s)\}^2] = F(t) - F(s) \quad (20)$$

for all $s \leq t$ in T . [12]. Since $F(t)$ is a monotone nondecreasing function, the left and right limits $F(t^-)$ and $F(t^+)$ exist at every point $t \in T$, and the set D of points of discontinuity of $F(t)$ is at most countable. It follows from (20) that the left and right mean square limits $x(t^-, \omega)$ and $x(t^+, \omega)$ exist at every point $t \in T$ and that D is the set of points of mean square discontinuity of $x(t, \omega)$. It is also easily seen that $x(t, \omega)$ is not weakly continuous at the points of D . Indeed let $t \in D$ and let $F(t) \neq F(t^-)$. Then it follows by (20) that $x(t, \omega) \neq x(t^-, \omega)$ in $L_2(\Omega, \mathcal{F}, P)$. Hence if $t_n \uparrow t$ and $\xi \in H(x)$ is not orthogonal to $x(t) - x(t^-)$ we have

$$\lim_n E[x(t_n)\xi] = E[x(t^-)\xi] \neq E[x(t)\xi]$$

which shows that $x(t, \omega)$ is not weakly continuous at $t \in D$. The following condition will be considered:

(C1) At every point of mean square discontinuity ($t \in D$) the square integrable martingale equals either its left or its right mean square limits

or equivalently,

(C2) At every point of discontinuity ($t \in D$), F equals either its left or its right limit.

THEOREM 5. If a square integrable martingale satisfies (C1) then it is smooth.

PROOF. By [4, Theorem 4.1] it suffices to prove that if $f \in \text{RKHS}(r)$ and $f(t) = 0$ a.e. [Leb] on T , then $f(t) = 0$ for all $t \in T$. Indeed assume that $f(t) = E[\xi x(t)]$ for all $t \in T$, where $\xi \in H(x)$, and that $f(t) = 0$ for all $t \in T \sim T'$, where $\text{Leb}(T') = 0$. Now let $t \in T'$. Since $\text{Leb}(T') = 0$, there exists a sequence $\{t_n\}_n$ of points in $T \sim T'$ con-

verging to t ; in particular there always exist increasing (decreasing) such sequences. We clearly have $F(t_n) = 0$. If $t \notin D$, it follows that $f(t) = \lim_n f(t_n) = 0$. If $t \in D$, by (C1), either $F(t) = F(t^-)$ or $F(t) = F(t^+)$, and we obtain $f(t) = 0$ as when $t \notin D$ by choosing an increasing or decreasing sequence $\{t_n\}_n$ respectively. Thus $f(t) = 0$ for all $t \in T$ and the theorem is proven. \square

REMARK. The proof of Theorem 5 applies to all second order stochastic processes $\{x(t, \omega), t \in T\}$ satisfying the condition

(C2) For all $t \in T$ the mean square limits $x(t^-, \omega)$ and $x(t^+, \omega)$ exist and $x(t, \omega)$ equals either $x(t^-, \omega)$ or $x(t^+, \omega)$

which are thus smooth. This follows from the fact that if (C2) is satisfied then the set of mean square discontinuity points of $x(t, \omega)$ is at most countable [6, Lemma 1].

We now proceed to characterize the reproducing kernel Hilbert space and the range of the square root of the covariance operator of a square integrable martingale. These spaces play an important role in the Gaussian case, as it is seen from Theorem 4 and the conditions (i)', (ii)' and (iii)'. It will be assumed that the following simplifying condition is satisfied:

(C3) T is of the form $[0, a]$, $a < +\infty$, or $[0, a)$, $a \leq +\infty$, and $x(0, \omega) = 0$ a.s.

A simple stochastic integral with respect to a square integrable martingale will also be used here [5,7]. We define F' by

$$F'(t) = F(t^+) \quad \text{for all } t \in T \quad (21).$$

and we denote by m or dF' the measure induced on $(T, \mathcal{B}(T))$ by F' in the usual way, $m\{(s, t]\} = F'(t) - F'(s) = F(t^+) - F(s^+)$ for all $s \leq t$ in T . For all $f \in L_2(dF')$ the stochastic integral $\int_T f(t) dx(t, \omega)$ is a well

defined random variable in $L_2(\Omega, F, P)$ and has the property

$$E\left[\int_T f(t)dx(t, \omega) \cdot \int_T g(t)dx(t, \omega)\right] = \int_T f(t)g(t)dF'(t) \quad (22)$$

THEOREM 6. For a square integrable martingale $\{x(t, \omega), t \in T\}$ satisfying (C1) and (C3) the following are true:

$$(1) \quad H(x) = \{\xi(\omega) = \int_T g(t)dx(t, \omega), \quad g \in L_2(dF')\} \quad (23)$$

(2) If $E[x(t)] = 0$ for all $t \in T$, then

$$\text{RKHS}(R) = \{f(t) = \int_0^t g(u)dF'(u) \text{ for all } t \in T, g \in L_2(dF')\} \quad (24)$$

where the integral in (24) is over $[0, t)$ for $t \in T \setminus D$ and for $t \in D$ with $F(t) = F(t^-)$, and over $[0, t]$ for $t \in D$ with $F(t) = F(t^+)$; and

$$\text{range}(S^{\frac{1}{2}}) = \{f(t) = \int_0^t g(u)dF'(u) \text{ a.e. [Leb] on } T, g \in L_2(dF')\} \quad (25)$$

The results of Theorem 6 are known for the Wiener process, for which $F(t) = t$. The characterization of the reproducing kernel Hilbert space (24) is also known for mean square continuous processes with orthogonal increments [10]. As it is suggested by Theorem 6, a number of results established for the Wiener process can be extended in an appropriate way to all square integrable Gaussian martingales satisfying certain conditions (like (C1) and (C3)). Results in this direction will be included in a forthcoming paper.

PROOF OF THEOREM 6. (1) Let $L(x) = \{\xi(\omega) = \int_T g(t)dx(t, \omega), g \in L_2(dF')\}$.

It follows by (C3) that

$$x(t^-, \omega) = x(t^-, \omega) - x(0, \omega) = \int_T I_{[0, t)}(u)dx(u, \omega) \in L(x) \quad (26)$$

for all $t \in T$. Similarly $x(t^+, \omega) \in L(x)$ and thus, by (C1), $H(x) \subseteq L(x)$. Hence it suffices to show that $L(x) \subseteq H(x)$ or equivalently that $\xi \in L(x)$ and $\xi \perp H(x)$ imply $\xi = 0$. Assume $\xi \in L(x)$ and $\xi \perp H(x)$. Then $\xi(\omega) = \int_T g(u) dx(u, \omega)$ for some $g \in L_2(dF')$ and by (26) and (22) we have for all $t \in T$:

$$0 = E[x(t^-)\xi] = \langle I_{[0,t)}, g \rangle_{L_2(dF')}.$$

Since the set of functions $\{I_{[0,t)}(u), t \in T\}$ is dense in $L_2(dF')$, it follows that $g = 0$ in $L_2(dF')$ and by (22), $E[\xi^2] = \|g\|_{L_2(dF')}^2 = 0$, i.e. $\xi = 0$ in $L_2(\Omega, \mathcal{F}, P)$.

(2) It follows by the definition of reproducing kernel Hilbert space, (23) and (22) that

$$\begin{aligned} \text{RKHS}(R) &= \{f(t) = E[x(t)\xi] \text{ for all } t \in T, \xi \in H(x)\} \\ &= \{f(t) = \int_0^t g(u) dF'(u) \text{ for all } t \in T, g \in L_2(dF')\} \end{aligned}$$

where the integral is over $[0, t)$ or $[0, t]$ as explained in the statement of the theorem. The characterization of $\text{range}(S^{\frac{1}{2}})$ follows directly from Theorem 3.1 and (24). An alternative proof of independent interest is the following. Note that, as it follows from (20) and (C3), $R(t, s) = r(t, s) = F(\min(t, s))$. Let $\nu \in \mathcal{N}$, i.e., $\nu \ll \text{Leb}$ and $\int_T F(t) d\nu(t) < +\infty$. Then it is easily seen that $S = LL^*$, where L is a bounded linear operator from $L_2(dF')$ to $L_2(\nu)$ defined by

$$(Lg)(t) = \int_0^t g(u) dF'(u) \quad \text{a.e.}[\text{Leb}] \text{ on } T, \quad g \in L_2(dF')$$

and has norm $\|L\|^2 = \int_T F(t) d\nu(t) < +\infty$. It follows by a straightforward extension of a result given in [2, Corollary 2.c] that $\text{range}(S^{\frac{1}{2}}) = \text{range}(L)$ and hence (25). \square

REFERENCES

- [1] C. R. Baker, On equivalence of probability measures, *Institute of Statistics Mimeo Series No. 701*, University of North Carolina at Chapel Hill.
- [2] C. R. Baker, On covariance operators, *SIAM J. Math. Anal.*, to appear.
- [3] B. H. Bharucha and T.T. Kadota, On the representation of continuous parameter processes by a sequence of random variables, *IEEE Trans. Information Theory IT-16* (1970), pp. 139-141.
- [4] S. Cambanis, Representation of stochastic processes of second order and linear operations, submitted to *SIAM J. Appl. Math.*
- [5] P. Courrège, Intégrales stochastiques et martingales de carrée intégrable, Séminaire de Théorie du Potentiel, M. Brelot, G. Choquet and J. Deny, ed., *Institut Henri Poincaré*, Paris, 7e année, 1962-63, No. 7.
- [6] H. Cramèr, On the structure of purely non-deterministic stochastic processes, *Ark. Mat.*, 4(1961), pp. 249-266.
- [7] J.L. Doob, *Stochastic Processes*, Wiley, New York, 1953.
- [8] B. Eisenberg, Translating Gaussian processes, *Ann. Math. Stat.*, 41(1970), pp. 888-893.
- [9] J. Feldman, Equivalence and perpendicularity of Gaussian processes, *Pacific J. Math.*, 8(1958), 699-708.
- [10] I.I. Gikhman and A.V. Skorokhod, On the densities of probability measures in function spaces, *Russian Math. Surveys*, 21(1966), pp. 83-156.
- [11] E. Masry and S. Cambanis, The representation of stochastic processes without loss of information, submitted to *Information and Control*.
- [12] P.A. Meyer, *Probability and Potentials*, Blaisdell, Waltham, Mass., 1966.
- [13] E. Mourier, Éléments aléatoires dans un espace de Banach, *Ann. Inst. H. Poincaré Sect. B*, 13(1953), pp. 161-244.

- [14] E. Parzen, Statistical inference on time series by Hilbert space methods, *Time Series Analysis Papers*, Holden-Day, San Francisco, 1967, pp. 251-382.
- [15] B.S. Rajput and S. Cambanis, Gaussian stochastic processes and Gaussian measures, *Institute of Statistics Mimeo Series No. 705*, University of North Carolina at Chapel Hill.
- [16] C.R. Rao and V.S. Varadarajan, Discrimination of Gaussian processes, *Sankhyā Ser. A*, 25(1963), pp. 303-330.
- [17] W.L. Root, Singular Gaussian measures in detection theory, *Proc. of Symp. on Time Series Analysis*, M. Rosenblatt, ed., Wiley, New York, 1962.
- [18] Yu. A. Rosanov, Some remarks on the paper on Gaussian distribution densities and Wiener-Hopf integral equations, *Theor. Prob. Appl.* 11(1966), 483-485.