The Annals of Probability
1973, Vol. 1, No. 1, 104-122

NON-ANTICIPATIVE REPRESENTATIONS OF EQUIVALENT GAUSSIAN PROCESSES[1]

BY G. KALLIANPUR AND H. OODAIRA

University of Minnesota

Given two equivalent Gaussian processes the notion of a *non-anticipative* representation of one of the processes with respect to the other is defined. The main theorem establishes the existence of such a representation under very general conditions. The result is applied to derive such representations explicitly in two important cases where one of the processes is (i) a Wiener process, and (ii) a N-ple Gaussian Markov process. Radon-Nikodym derivatives are also discussed.

1. Introduction. Let $(X(t), P)$ and $(X(t), Q)$, $(0 \leqq t \leqq 1)$ be equivalent Gaussian processes given on some space $(\Omega, \mathscr{A})$ which are quadratic mean continuous, have zero mean functions and covariance functions Γ_P and Γ_Q. The term "equivalent" is here used in the sense that the probability measures P and Q are mutually absolutely continuous with respect to the σ-field $\mathscr{B}$ generated by the random variables $(X(t))$.

By a representation of $(X(t), P)$ in terms of $(X(t), Q)$ we mean a family of random variables $(Y(t))$ $(0 \leqq t \leqq 1)$ on Ω such that

$$(Y(t), Q) \tag{1.1}$$

is a quadratic mean continuous Gaussian process with zero mean and covariance Γ_P, and for each t

$$Y(t) \in L(X; 1) \tag{1.2}$$

where $L(X; 1)$ is the linear space of the process $(X(t), Q)$. (The precise definitions of terms and notation used here are given in Sections 3 and 4.) When $(X(t), Q)$ is a standard Wiener process the following representation for all $(X(t), P)$ equivalent to it was given by Shepp (1966).

$$Y(t) = X(t) - \int_0^t [\int_0^1 M(s, u)\, dX(u)]\, ds \tag{1.3}$$

where $M(s, u)$ is a square integrable kernel related to Γ_P. From the point of view of statistical or engineering applications a drawback of this representation is that in order to find $Y(t)$ from (1.3) we need to know all the values of $X(s)$, $(0 \leqq s \leqq 1)$. It is desirable to obtain a representation which involves only the values $(X(s), 0 \leqq s \leqq t)$, i.e. the "present and past" but not the "future" values $X(s)$ $(s > t)$. Just such a representation has recently been given by Hitsuda (1968). Such a representation will be called non-anticipative (see Section 4).

Received September 13, 1971; revised May 23, 1972.

[1] This work was supported in part by NSF Grant GP-1188-8.

Key words and phrases. Equivalent Gaussian processes, representation, Volterra operators, factorization.

The main purpose of this paper is to establish the existence of non-anticipative representations in the most general case. This is done in Section 4 and is based on the fundamental work of Gohberg and Krein on the factorization of operators of the form $I - T$ where T is Hilbert-Schmidt ([2] Chapter IV). In Sections 5 and 6 we apply the general theorem to derive non-anticipative representations when $(X(t), Q)$ is, respectively, the non-standard Wiener process (thus incidentally obtaining an alternative proof of Hitsuda's result) and a Gaussian N-ple Markov process. The problem of the Radon-Nikodym derivative is considered in Section 7. The possibility of using Gohberg and Krein's ideas in this connection was raised by Kailath (1970) who has formally derived Hitsuda's result using the result of [2]. (See comments at the end of Section 5.)

The following remarks form the starting point of our investigations which, we hope, put our problem in its most general setting. Starting from a given representation $(Y(t), Q)$ and using the necessary and sufficient conditions for equivalence obtained in Kallianpur and Oodaira ([5], 1963), (see Theorem 4.2 below), it is easy to verify that (1.1) and (1.2) imply the existence of a bounded linear operator $\tilde{F}$ on $L(X; 1)$ with the following properties: For every t

$$Y(t) = \tilde{F}X(t)\,, \tag{1.4}$$

$$\tilde{S} = \tilde{F}^*\tilde{F}\,, \tag{1.5}$$

$\tilde{S}$ being the operator in $L(X; 1)$ corresponding to the operator S of Theorem 4.2, so that $\tilde{S} = I - \tilde{T}$ where $\tilde{T}$ is Hilbert-Schmidt and 1 is not a point of the spectrum of $\tilde{T}$. Conversely, if $\tilde{F}$ is an operator on $L(X; 1)$ satisfying (1.5) then $(Y(t), Q)$ where $Y(t)$ is defined by (1.4), is a representation of $(X(t), P)$ in the sense of (1.1) and (1.2). Thus every representation of $(X(t), P)$ uniquely corresponds (the uniqueness is easily seen) to a factorization of S (or $\tilde{S}$) of the type (1.5). What we intend to do is to pick out the particular factorization that corresponds to the non-anticipative representation. That such a factorization exists is not obvious a priori and that is where Gohberg and Krein's theory of special factorization enters into the picture in a natural fashion. Before considering it in Section 4 we discuss some basic ideas and results in the next section.

2. Factorization of self-adjoint positive invertible operators. Let H be a separable Hilbert space. A family of orthoprojectors $\pi = \{P\}$ is called a chain if for any distinct $P_1, P_2 \in \pi$, either $P_1 < P_2$ or $P_2 < P_1$, where $P_1 < P_2$ means $P_1H \subset P_2H$, i.e., $P_1P_2 = P_2P_1 = P_1$. We shall write $P_1 \leqq P_2$ if either $P_1 < P_2$ or $P_1 = P_2$. A chain π is said to be bordered if $\pi \ni 0, I$. The closure of a chain π is the set of all operators which are the strong limits of sequences in π. The closure of a chain is again a chain and if a chain coincides with its closure, it is said to be closed. A pair (P^-, P^+) of orthoprojectors in a closed chain π with $P^- < P^+$ is called a gap of π if for any $P \in \pi$ either $P \leqq P^-$ or $P \geqq P^+$, and the dimension of $P^+ - P^-$, i.e., $\dim[P^+H \ominus P^-H]$, is called the dimension of the gap (P^-, P^+). A chain is said to be maximal if it cannot be enlarged, or, equivalently, if it is bordered, closed and its gaps (if any) are one-dimensional. A chain π is called an eigenchain of a bounded linear operator A on H if $PAP = AP$ for all $P \in \pi$.

Let π be a closed chain. A partition ζ of π is a chain consisting of a finite number of elements $\{P_0 < P_1 < \cdots < P_n\}$ of π such that $P_0 = \min_{P \in \pi} P$ and $P_n = \max_{P \in \pi} P$. Let $F(P)$ be an operator function defined on π and having as its values bounded linear operators on H. For a partition $\zeta = \{P_0 < P_1 < \cdots < P_n\}$ of π, define

$$S(\zeta) = \sum_{j=1}^{n} F(P_{j-1})\Delta P_j , \qquad \Delta P_j = P_j - P_{j-1} .$$

An operator A is called the limit in norm of $S(\zeta)$, denoted by

$$A = (m) \int_\pi F(P)\, dP , \tag{2.1}$$

if for any $\varepsilon > 0$ there exists a partition $\zeta(\varepsilon)$ of π such that, for every partition $\zeta \supset \zeta(\varepsilon)$, $||S(\zeta) - A|| < \varepsilon$. If the limit of $S(\zeta)$ exists, we shall say that the integral (2.1) converges. The integral

$$B = (m) \int_\pi dPF(P)$$

is defined analogously.

The dual $\pi^\perp$ of a chain π is a chain consisting of all orthoprojectors of the form $P^\perp = I - P$, $P \in \pi$. If π is an eigenchain of an operator A, then the dual chain $\pi^\perp$ is an eigenchain of the adjoint operator A^*.

By a special factorization of an operator A along a chain π we mean the representation of A in the form

$$A = (I + X_+)D(I + X_-) , \tag{2.2}$$

where X_+ and X_- are Volterra operators (i.e., completely continuous operators with the one-point spectrum $\lambda = 0$) having π and $\pi^\perp$ as eigenchains respectively, D commutes with all $P \in \pi$, and $D - I$ is completely continuous.

The factors $I + X_+$, $I + X_-$ are invertible, and if A is invertible, so is the factor D. If an invertible operator A admits a special factorization relative to a maximal chain π, then the factorization is unique, and from the uniqueness it follows that if a self-adjoint invertible operator $A = A^*$ has such a factorization, then $X_+^* = X_-$ and $D^* = D$.

The following theorem is a special case of Theorems 6.1 and 6.2, Chapter IV, Gohberg–Krein [2]. We denote by $\mathscr{S}_2$ the class of all Hilbert-Schmidt operators on H.

THEOREM 2.1. *Let π be a maximal chain. Then, for every operator $T \in \mathscr{S}_2$ such that each of the operators $I - PTP$, $P \in \pi$, is invertible, the integrals*

$$\begin{aligned} X_+ &= (m) \int_\pi (I - PTP)^{-1}PT\, dP , \\ X_- &= (m) \int_\pi dPTP(I - PTP)^{-1} \end{aligned} \tag{2.3}$$

converge in norm, and the operator $A = (I - T)^{-1}$ has a special factorization (2.2) along π with $X_+, X_-, D - I \in \mathscr{S}_2$ and

$$D = I + \sum_j (P_j^+ - P_j^-)[(I - P_j^+TP_j^+)^{-1} - I](P_j^+ - P_j^-) , \tag{2.4}$$

where $\{(P_j^-, P_j^+)\}$ is the set of all gaps in the chain π.

For the convenience of the reader it is perhaps worth pointing out that the deduction of this result from the above-mentioned theorems of Gohberg and Krein is based on the fact that if $T \in \mathscr{S}_2$ then the integral $(m) \int_\pi PT\,dP$ converges in uniform norm (in fact, in Hilbert-Schmidt norm) and belongs to $\mathscr{S}_2$. The verification is simple and is a part of the proof of Theorem 10.1, Chapter I of [2].

LEMMA 2.1. *Let $T \in \mathscr{S}_2$. If $I - T$ is self-adjoint, positive and invertible, then, for any orthoprojector P, $I - PTP$ is invertible.*

The proof is immediate as can be seen from the following inequality.

$$\begin{aligned}\langle (I - PTP)f, f\rangle &= \langle (I - P)f, f\rangle + \langle (I - T)Pf, Pf\rangle \\ &= \|(I - P)f\|^2 + \|(I - T)^{\frac{1}{2}}Pf\|^2 \\ &\geqq \|(I - P)f\|^2 + c^2\|Pf\|^2 \geqq c_1^2\|f\|^2 ,\end{aligned}$$

where c is some positive constant, $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ are the inner product and the norm of H, and $c_1^2 = \min(1, c^2)$.

LEMMA 2.2. *If a self-adjoint, positive, invertible operator A has a special factorization* (2.2), *then the factor D is self-adjoint, positive and invertible.*

PROOF. We need only to prove the positive definiteness of D. Since $I + X_+$, $I + X_-$ are invertible,

$$\langle Df, f\rangle = \langle (I + X_+)^{-1}A(I + X_-)^{-1}f, f\rangle .$$

Set $(I + X_-)^{-1}f = g$. Then

$$\begin{aligned}\langle Df, f\rangle &= \langle (I + X_+)^{-1}Ag, (I + X_-)g\rangle \\ &= \langle (I + X_-)^*(I + X_+)^{-1}Ag, g\rangle \\ &= \langle (I + X_+)(I + X_+)^{-1}Ag, g\rangle \\ &= \langle Ag, g\rangle \\ &\geqq 0 .\end{aligned}$$

LEMMA 2.3. *If V is a Volterra operator $\in \mathscr{S}_2$ with π as an eigenchain, then the operator $W = (I + V)^{-1} - I$ is also Volterra $\in \mathscr{S}_2$ and has π as an eigenchain.*

PROOF. That $W \in \mathscr{S}_2$ follows immediately from the relation

$$W + V + VW = 0 . \tag{2.5}$$

Since V and $I + W$ are permutable, we have (cf. [6] page 426)

$$r_{V+VW} \leqq r_V \cdot r_{I+W}$$

where r_A denotes the spectral radius of A, i.e., the radius of the smallest closed disk centered at 0 which contains all the spectrum of A. By assumption $r_V = 0$, so $r_{V+VW} = 0$, i.e., $V + VW$ is Volterra. Hence, from (2.5), W is Volterra. Since $I + W$ is the resolvent at 1 of the operator $-V$,

$$W = \sum_{n=1}^{\infty} (-1)^n V^n ,$$

the right-hand side converging in norm. It is readily verified that, for $P \in \pi$, $PV^nP = V^nP$ for any n. Hence we have $PWP = WP$.

From Theorem 2.1 and Lemmas 2.1–2.3 we have the following theorem.

THEOREM 2.2. *Let $S = I - T$ with $T \in \mathscr{S}_2$ be a self-adjoint positive and invertible operator. Then S and S^{-1} have the factorizations along any maximal chain $\pi = \{P\}$*

$$S = (I + W_-)D^{-1}(I + W_+)$$
$$S^{-1} = (I + X_+)D(I + X_-) ,$$

where

(a) *W_+, W_-, X_+, X_- are Volterra operators $\in \mathscr{S}_2$, X_+, X_- are given by* (2.3), *the integral converging in norm, and $I + W_+ = (I + X_+)^{-1}$, $I + W_- = (I + X_-)^{-1}$,*
(b) *W_+, X_+ have π and W_-, X_- have $\pi^\perp$ as eigenchains,*
(c) *$W_+^* = W_-$ and $X_+^* = X_-$,*
(d) *D is a self-adjoint, positive and invertible operator given by* (2.4), *and*
(e) *$D - I \in \mathscr{S}_2$, $DP = PD$ for all $P \in \pi$.*

3. Chains of orthoprojectors associated with a Gaussian process. Let $\{X(t), t \in [a, b]\}$ be a Gaussian process defined on a probability space $(\Omega, \mathscr{A}, Q)$ with $E_Q X(t) \equiv 0$ and covariance function $\Gamma_Q(s, t)$, where $[a, b]$ is taken to be either a finite closed or an infinite interval. For the sake of simplicity we assume that $[0, 1] \subset [a, b]$. Throughout the paper we make the following assumption:

(3.1) $\{X(t)\}$ is continuous in quadratic mean (q.m.) .

For $0 \leqq t \leqq 1$, let $L(X; t)$ be the closed linear subspace spanned by $X\{(\tau), 0 \leqq \tau \leqq t\}$ of $L^2(\Omega, \mathscr{A}, Q)$, and let $\tilde{P}(t)$ be the orthoprojector defined on $L(X; 1)$ with range $L(X; t)$. We are interested in a maximal chain containing the chain $\{\tilde{P}(t), 0 \leqq t \leqq 1\}$ (or $\{\tilde{P}(t), 0 < t \leqq 1\}$).

Let

$$L(X; t+) = \textstyle\bigcap_{s>t} L(X; s)$$

and

$L(X; t-) =$ the smallest closed linear space containing all $L(X; s)$, $s < t$.

Obviously $L(X; 0) \subset L(X; 0+)$ and $L(X; t-) \subset L(X; t) \subset L(X; t+)$ for $t > 0$. It is also easy to verify that $L(X; t-) = L(X; t)$ for all $0 < t \leqq 1$. Since by the assumption (3.1) the Hilbert space $L(X; 1)$ is separable, the set of discontinuities $D = \{t \in [0, 1] \mid L(X; t) \neq L(X; t+)\}$ is at most countable. Let $\tilde{P}(t_j+)$ be the orthoprojector with range $L(X; t_j+)$ for $t_j \in D$. The closure of the chain $\{\tilde{P}(t), 0 \leqq t \leqq 1\}$ consists of $\{\tilde{P}(t), 0 < t \leqq 1\}$ and $\{\tilde{P}(t_j+), t_j \in D\}$. If $D \neq \varnothing$, it has gaps $(\tilde{P}(t_j), \tilde{P}(t_j+))$, $t_j \in D$. If the dimension of gap $(\tilde{P}(t_j), \tilde{P}(t_j+))$ is $n_j > 1$, we write the space $(\tilde{P}(t_j+) - \tilde{P}(t_j))L(X; 1)$ as the orthogonal sum of one-dimensional subspaces $L(j, i)$:

$$(\tilde{P}(t_j+) - \tilde{P}(t_j))L(X, 1) = \textstyle\sum_{i=1}^{n_j} \oplus L(j, i) .$$

Let $\tilde{Q}(j, k)$ be the orthoprojector with range $\sum_{i=1}^{k} \oplus L(j, i)$. Now consider the family of orthoprojectors $\tilde{\pi}$ consisting of 0, $\{\tilde{P}(t), 0 \leqq t \leqq 1\}$, $\{\tilde{P}(t_j+), t_j \in D\}$ and $\{\tilde{P}(t_j) + \tilde{Q}(j, k), k = 1, \cdots, n_j - 1, t_j \in D\}$. It is clear that $\tilde{\pi}$ is a chain and maximal.

REMARK 1. A maximal chain containing the chain $\{\tilde{P}(t), 0 \leqq t \leqq 1\}$ is obviously not unique, in general. If the dimension of gap $(\tilde{P}(t_j), \tilde{P}(t_j+))$ is > 1, we may take different orthoprojectors $\tilde{Q}(j, k)$.

REMARK 2. The gap $(\tilde{P}(0), \tilde{P}(0+))$, if it exists and $\tilde{P}(0) \neq 0$, is special. Instead of filling in the gap $(\tilde{P}(0), \tilde{P}(0+))$, we may insert any set of orthoprojectors $\{\tilde{Q}(j), j = 0, 1, \cdots, n\}$ such that $\tilde{Q}(0) = 0$, $\tilde{Q}(n) = \tilde{P}(0+)$ and $\dim(\tilde{Q}(j) - \tilde{Q}(j-1)) = 1$. The maximal chain thus obtained will suffice for our purposes. In other words, we need only a maximal chain $\tilde{\pi}$ containing $\{\tilde{P}(t), 0 < t \leqq 1\}$.

As we shall see the space $L(X; 0+)$ is of particular interest. It can be trivial, can be n-dimensional ($1 \leqq n < \infty$), or even infinite dimensional.

EXAMPLE 1. If $\{X(t), 0 \leqq t \leqq 1\}$ is a Wiener process, then $L(X; t+) = L(X; t)$ for all $t \in [0, 1]$, i.e., $D = \emptyset$. If $X(0) = 0$, i.e., if $\{X(t)\}$ is the standard Wiener process, the chain $\tilde{\pi} = \{\tilde{P}(t), 0 \leqq t \leqq 1\}$ is maximal. However, if $X(0) \neq 0$, the chain $\tilde{\pi} = \{0, \tilde{P}(t), 0 \leqq t \leqq 1\}$ is maximal and has a one-dimensional gap $(0, \tilde{P}(0) = \tilde{P}(0+))$.

EXAMPLE 2. Let $\{X(t), a \leqq t \leqq b\}$ be an N-ple Gaussian Markov process, where $a < 0$, $b \geqq 1$ (see Section 6). Then $L(X; t+) = L(X; t)$ for all $t \in (0, t]$ and the space $L(X; 0+)$ is N-dimensional.

Let ψ denote the congruence (isometric isomorphism) from $L(X; 1)$ onto the reproducing kernel Hilbert space (RKHS) $H \equiv H(\Gamma_Q)$ with reproducing kernel $\Gamma_Q(s, t)$, $0 \leqq s, t \leqq 1$, such that $\psi X(t) = \Gamma_Q(\cdot, t)$, $0 \leqq t \leqq 1$. We note the following relation between subspaces of $L(X; 1)$ and of H.

Let

$$F(t) = \{f \in H \mid f(s) = 0, 0 \leqq s \leqq t\}$$

and

$$M(t) = H \ominus F(t), \quad \text{the orthogonal complement of} \quad F(t).$$

LEMMA 3.1. $\psi[L(X; t)] = M(t)$.

PROOF. Let $f \in F(t)$ and $\xi = \psi^{-1}f$. Then, for all $s \in [0, t]$,

$$0 = f(s) = \langle f(\cdot), \Gamma(\cdot, s)\rangle = (\xi, X(s)),$$

where $\langle \cdot, \cdot \rangle$, $(\cdot, \cdot)$ denote respectively the inner products of H and $L(X; 1)$. Hence $\xi \perp L(X; t)$, and thus $f \in F(t)$ if and only if $\psi^{-1}f \in L(X; 1) \ominus L(X; t)$. This is equivalent to the assertion of the lemma.

Let $P(t)$ denote the orthoprojector on H with range $M(t)$. Then $P(t) = \psi\tilde{P}(t)\psi^{-1}$. If $\tilde{\pi} = \{\tilde{P}\}$ is a maximal chain containing $\{\tilde{P}(t), 0 \leqq t \leqq 1\}$, then, obviously, the chain $\pi = \{\psi\tilde{P}\psi^{-1}\}$ is maximal and contains $\{P(t), 0 < t \leqq 1\}$. We shall consistently use the following notation. If A is any linear operator on H, $\tilde{A}$ denotes the operator on $L(X; 1)$ given by $\tilde{A} = \psi^{-1}A\psi$.

4. Non-anticipative representations of equivalent Gaussian processes. Let $\mathscr{B}$ denote the σ-field generated by the random variables $\{X(t), 0 \leqq t \leqq 1\}$. Let P be another probability measure on $(\Omega, \mathscr{B})$ such that $\{X(t), 0 \leqq t \leqq 1, P\}$ is a q.m. continuous Gaussian process with $E_P X(t) = 0$ and covariance function $\Gamma_P(s, t)$. ($\{X(t), 0 \leqq t \leqq 1, Q\}$ is, by assumption, a q.m. continuous Gaussian process with $E_Q X(t) = 0$ and covariance function $\Gamma_Q(s, t)$.) Assume that P and Q are equivalent, i.e., mutually absolutely continuous relative to $\mathscr{B}$.

A non-anticipative representation of a Gaussian process with respect to another is defined as follows.

DEFINITION. The process $\{X(t), 0 \leqq t \leqq 1, P\}$ has a non-anticipative representation with respect to $\{X(t), 0 \leqq t \leqq 1, Q\}$ if there is a Gaussian process $\{Y(t), 0 \leqq t \leqq 1, Q\}$, having zero mean and Γ_P for its covariance, with the following property:

$$Y(t) \in L(X; t) \quad \text{for each} \quad t \in [0, 1]. \tag{4.1}$$

REMARK. (4.1) implies $L(Y; t) \subset L(X; t)$ for $t \in [0, 1]$.

We shall now prove the following main theorem.

THEOREM 4.1. *Every Gaussian process $\{X(t), 0 \leqq t \leqq 1, P\}$ which is equivalent to a given Gaussian process $\{X(t), 0 \leqq t \leqq 1, Q\}$ has a non-anticipative representation with respect to $\{X(t), 0 \leqq t \leqq 1, Q\}$. The processes are assumed to be* q.m. *continuous.*

The proof is based on Theorem 2.2. and the following necessary and sufficient conditions for equivalence of P and Q (cf. [5]).

THEOREM 4.2. *Gaussian measures P and Q are equivalent if and only if Γ_P defines an operator S on the* RKHS *$H(\Gamma_Q)$ with the following properties*:

(a) $\Gamma_P(\cdot, t) = S\Gamma_Q(\cdot, t)$ *for* $0 \leqq t \leqq 1$,
(b) *S is a bounded, self-adjoint, positive operator*,
(c) $T = I - S \in \mathscr{S}_2$,
(d) $1 \notin \sigma(T)$, *the spectrum of T.*

PROOF OF THEOREM 4.1. Consider a maximal chain π in $H(\Gamma_Q)$ described in the preceding section. Applying Theorem 2.2 to the operator S defined in Theorem 4.2, we have

$$S = (I + W_-)\Delta(I + W_+),$$

where $\Delta = D^{-1}$. The operator Δ is self-adjoint and positive. Since D commutes with all $P \in \pi$, we have $\Delta^{\frac{1}{2}}P = P\Delta^{\frac{1}{2}}$ for all $P \in \pi$. If we write

$$F = \Delta^{\frac{1}{2}}(I + W_+)$$

then

$$S = F^*F,$$

because $F^* = (I + W_+^*)\Delta^{\frac{1}{2}} = (I + W_-)\Delta^{\frac{1}{2}}$.

Consider now the operator $\tilde{F}$ on $L(X; 1)$ corresponding to F:

$$\tilde{F} = \tilde{\Delta}^{\frac{1}{2}}(I + \tilde{W}_+) . \tag{4.2}$$

Define

$$Y(t) = \tilde{F}X(t) , \qquad t \in [0, 1] .$$

Since $\tilde{F}$ is a linear operator on $L(X; 1)$, $\{Y(t), 0 \leqq t \leqq 1, Q\}$ is Gaussian. Furthermore, $E_Q Y(t) \equiv 0$ and

$$\begin{aligned} E_Q Y(s)Y(t) &= (\tilde{F}X(s), \tilde{F}X(t)) \\ &= \langle F\Gamma_Q(\cdot, s), F\Gamma_Q(\cdot, t)\rangle \\ &= \langle S\Gamma_Q(\cdot, s), \Gamma_Q(\cdot, t)\rangle \\ &= \langle \Gamma_P(\cdot, s), \Gamma_Q(\cdot, t)\rangle \\ &= \Gamma_P(s, t) . \end{aligned}$$

Thus $\{Y(t), 0 \leqq t \leqq 1, Q\}$ is a Gaussian process with zero mean and covariance function $\Gamma_P(s, t)$. We have also for each $t \in [0, 1]$,

$$\begin{aligned} Y(t) &= \tilde{\Delta}^{\frac{1}{2}}(I + \tilde{W}_+)X(t) \\ &= \tilde{\Delta}^{\frac{1}{2}}(I + \tilde{W}_+)\tilde{P}(t)X(t) \\ &= \tilde{\Delta}^{\frac{1}{2}}\tilde{P}(t)(I + \tilde{W}_+)\tilde{P}(t)X(t) \\ &= \tilde{P}(t)\tilde{\Delta}^{\frac{1}{2}}(I + \tilde{W}_+)\tilde{P}(t)X(t) \\ &= \tilde{P}(t)Y(t) . \end{aligned}$$

This shows that $Y(t) \in L(X; t)$. The proof of the theorem is complete.

We have shown that the non-anticipative representation $Y(t)$ is given by

$$Y(t) = \tilde{\Delta}^{\frac{1}{2}}(I + \tilde{W}_+)X(t) . \tag{4.3}$$

We are indebted to Yu. A. Rozanov for the remark that the representation (4.3) can be cast in the form

$$Y(t) = X(t) + \tilde{G}X(t) \tag{4.4}$$

where $\tilde{G}$ is a Hilbert-Schmidt operator in the space $L(X; 1)$ such that for each t

$$\tilde{G}L(X; t) \subseteqq L(X; t) . \tag{4.5}$$

This is easily deduced from Theorems 2.2 and 4.1 as follows:
Write $\tilde{G} = \tilde{F} - I$. From (4.2) we have

$$\tilde{G} = \tilde{G}_1 + \tilde{G}_2 \tag{4.6}$$

where $\tilde{G}_1 = \tilde{\Delta}^{\frac{1}{2}} - I$ and $\tilde{G}_2 = \tilde{\Delta}^{\frac{1}{2}}\tilde{W}_+$.
Now

$$\tilde{G}_1 = \tilde{D}^{-\frac{1}{2}}(I - \tilde{D}^{\frac{1}{2}}) = \tilde{D}^{-\frac{1}{2}}(I + \tilde{D}^{\frac{1}{2}})^{-1}(I - \tilde{D}) .$$

Hence it follows that $\tilde{G}_1$ is Hilbert-Schmidt (i.e. belongs to $\mathscr{S}_2$), self-adjoint and that $\tilde{G}_1\tilde{P} = \tilde{P}\tilde{G}_1$ for all $\tilde{P} \in \tilde{\pi}$. Since W_+ belongs to $\mathscr{S}_2$ and has $\tilde{\pi}$ as an eigenchain and since $\tilde{\Delta}^{\frac{1}{2}}$ is a self-adjoint operator with $\tilde{\pi}$ as an eigenchain we conclude that $\tilde{G}_2 = \tilde{\Delta}^{\frac{1}{2}}\tilde{W}_+$ belongs to $\mathscr{S}_2$ and has $\tilde{\pi}$ as an eigenchain. This proves the assertions (4.4) and (4.5).

In fact, we can strengthen the above remark and show that (4.6) yields a decomposition of the non-anticipative Hilbert-Schmidt operator $\tilde{G}$ into the sum of two such operators $\tilde{G}_1$ and $\tilde{G}_2$ *where* $\tilde{G}_2$ *is Volterra*. That $\tilde{G}_2$ is Volterra follows from the fact that the Volterra operator $\tilde{W}_+$ and the bounded self-adjoint operator $\Delta^{\frac{1}{2}}$ have the maximal chain $\tilde{\pi}$ for a common eigenchain. For if $(\tilde{P}^-, \tilde{P}^+)$ is any gap of $\tilde{\pi}$ we have

$$(\tilde{P}^+ - \tilde{P}^-)\tilde{\Delta}^{\frac{1}{2}}\tilde{W}_+(\tilde{P}^+ - \tilde{P}^-) = \tilde{\Delta}^{\frac{1}{2}}(\tilde{P}^+ - \tilde{P}^-)\tilde{W}_+(\tilde{P}^+ - \tilde{P}^-) = 0 .$$

From Theorem 5.1 of Chapter I of [2] it follows that $\tilde{\Delta}^{\frac{1}{2}}\tilde{W}_+$ has a triangular representation and is therefore, of course, a Volterra operator. The triangular representation is given by

$$\tilde{\Delta}^{\frac{1}{2}}\tilde{W}_+ = \int_{\tilde{\pi}} \tilde{P}\tilde{H}\,d\tilde{P} \tag{4.7}$$

where

$$\tilde{H} = (\tilde{\Delta}^{\frac{1}{2}}\tilde{W}_+) - (\tilde{\Delta}^{\frac{1}{2}}\tilde{W}_+)^* = \tilde{\Delta}^{\frac{1}{2}}\tilde{W}_+ - \tilde{W}_-\tilde{\Delta}^{\frac{1}{2}} . \tag{4.8}$$

Let us recall that $\tilde{\pi}$ is a maximal chain containing the orthoprojectors $\tilde{P}(t)$ $(0 \leqq t \leqq 1)$ defined in Section 3. Since $\tilde{P}(t)X(t) = X(t)$ it follows from (4.7) that

$$\tilde{G}_2X(t) = \int_{\tilde{\pi}} \tilde{P}\tilde{H}\,d\tilde{P}X(t) = \int_{\tilde{\pi}} \tilde{P}\tilde{H}\,d\tilde{P}\tilde{P}(t)X(t) = \int_{\tilde{\pi}_t} \tilde{P}\tilde{H}\,d\tilde{P}X(t) \tag{4.9}$$

where $\tilde{\pi}_t$ is the chain $\{\tilde{P} \in \tilde{\pi} : \tilde{P} \leqq \tilde{P}(t)\}$. The last step in (4.9) is easily verified from the definition of the operator integral.

We thus arrive at an alternate and perhaps more interesting version of Theorem 4.1.

THEOREM 4.3. *Every Gaussian process* $\{X(t), 0 \leqq t \leqq 1, P\}$ *(satisfying the conditions stated at the beginning of the section) which is equivalent to a given Gaussian process* $\{X(t), 0 \leqq t \leqq 1, Q\}$ *has a non-anticipative representation* $Y(t)$ *given by*

$$Y(t) = X(t) + (\tilde{\Delta}^{\frac{1}{2}} - I)X(t) + \int_{\tilde{\pi}_t} \tilde{P}\tilde{H}\,d\tilde{P}X(t) . \tag{4.10}$$

If we know more about the nature of the $(X(t), Q)$ process it is reasonable to expect a more "concrete" representation for the operator $\tilde{G}$, leading to an explicit evaluation of $\tilde{\Delta}^{\frac{1}{2}}$ and the replacement of the last term in (4.10) by an expression involving conventional stochastic integrals. We investigate this question in the following sections and obtain explicit forms for $Y(t)$ for some important special cases.

5. Non-anticipative representation of a Gaussian process equivalent to a Wiener process. Suppose that $\{X(t), 0 \leqq t \leqq 1, Q\}$ is a Wiener process with $EX(t) = 0$, $E(X(t) - X(0))^2 = t$ and $EX^2(0) = \sigma^2 \geqq 0$. Then

$$\begin{aligned} \Gamma_Q(s, t) &= EX(s)X(t) \\ &= \sigma^2 + \min(s, t) \\ &= \int_0^1 \chi(s, u)\chi(t, u)\mu(d\mu) , \end{aligned} \tag{5.1}$$

where

$$\chi(t, u) = 1 \quad \text{if} \quad 0 \leqq u \leqq t$$
$$= 0 \quad \text{if} \quad t < u \leqq 1$$

and the measure μ assigns point mass σ^2 at $u = 0$ and is Lebesgue measure over $(0, 1]$. (5.1) implies that there is an isometric isomorphism from $H(\Gamma_Q)$ onto $L^2([0, 1], \mu)$ which sends $\Gamma_Q(\cdot, t)$ to $\chi(t, u)$, and any element $f \in H(\Gamma_Q)$ is represented in the form

$$f(t) = \int_0^t \hat{f}(u)\mu(du) = \hat{f}(0) + \int_0^t \hat{f}(u)\, du$$

with $\hat{f} \in L^2([0, 1], \mu)$. Correspondingly, there is an isometric isomorphism θ from $L(X; 1)$ onto $L^2([0, 1], \mu)$ such that $\theta X(t) = \chi(t, u)$ and, for any $\xi \in L(X; 1)$, we have

$$\xi = \hat{f}(0) + \int_0^1 \hat{f}(u)\, dB(u)\,, \tag{5.2}$$

where $B(u) = X(u) - X(0)$ and $\hat{f} = \theta\xi$.

It is easy to see that

$$\theta[L(X; t)] = \{\hat{f} \in L^2([0, 1], \mu) \,|\, \hat{f}(s) = 0 \ \text{ a.e. } \ \mu \ \text{ for } \ t < s \leqq 1\}\,.$$

The chain $\pi = \{0, P(t) = \theta\tilde{P}(t)\theta^{-1}, 0 \leqq t \leqq 1\}$ is maximal and $P(t) \in \pi$ is characterized by

$$P(t)\hat{f}(u) = \hat{f}(u)\chi(t, u)\,. \tag{5.3}$$

We shall denote by A the operator on $L^2([0, 1], \mu)$ corresponding to an operator $\tilde{A}$ on $L(X; 1)$ (or A on $H(\Gamma_Q)$), $L^2([0, 1], \mu)$ being a representation of $H(\Gamma_Q)$.

LEMMA 5.1. *Let $K(u, v)$ be the kernel in $L^2([0, 1] \times [0, 1], \mu \times \mu)$ corresponding to a Hilbert-Schmidt operator K on $L^2([0, 1], \mu)$. If the chain π is an eigenchain for K, then*

$$K(u, v) = 0 \quad \text{a.e.} \quad \mu \times \mu \quad \textit{for} \quad u > v$$

and if, in addition, K is Volterra and $\sigma^2 > 0$,

$$K(0, 0) = 0\,.$$

PROOF. Since, by assumption,

$$KP(t)\hat{f} = P(t)KP(t)\hat{f}\,, \qquad 0 \leqq t \leqq 1\,,$$

for $\hat{f} \in L^2([0, 1], \mu)$, it follows from (5.3) that

$$\int_0^t K(u, v)\hat{f}(v)\mu(dv) = 0 \quad \text{a.e.} \quad \mu \quad \text{for} \quad u > t\,. \tag{5.4}$$

Hence we have

$$\int_0^1\int_0^1 K(u, v)(1 - \chi(t, u))\chi(t, v)\hat{g}(u)\hat{f}(v)\mu(d\mu)\mu(dv) = 0$$

for all $0 \leqq t \leqq 1$ and for any $\hat{f}, \hat{g} \in L^2([0, 1], \mu)$. Let $\hat{f}(u) = \chi((a, b], u) = \chi(b, u) - \chi(a, u)$ and $\hat{g}(v) = \chi((c, d], v) = \chi(d, v) - \chi(c, v)$. Then, for either $b \leqq c$ or $a \leqq d$,

$$\int_0^1\int_0^1 (1 - \chi(u, v))K(u, v)\chi((a, b], u)\chi((c, d], v)\mu(d\mu)\mu(dv) = 0\,.$$

Since the family $\{\chi((a, b], u)\chi((c, d], v)$ with $b \leqq c$ or $a \geqq d\}$ spans $L^2([0, 1] \times [0, 1]$, Lebesgue measure), we have

$$(1 - \chi(u, v))K(u, v) = 0 \quad \text{a.e.} \quad \mu \times \mu \qquad \begin{array}{l} \text{for} \quad u, v \geqq 0 \quad \text{if} \quad \sigma^2 = 0 \\ \text{for} \quad u, v > 0 \quad \text{if} \quad \sigma^2 > 0 , \end{array}$$

i.e.,

$$K(u, v) = 0 \quad \text{a.e.} \quad \mu \qquad \begin{array}{l} \text{for} \quad u > v \geqq 0 \quad \text{if} \quad \sigma^2 = 0 \\ \text{for} \quad u > v > 0 \quad \text{if} \quad \sigma^2 > 0 . \end{array}$$

In the case $\sigma^2 > 0$, setting $t = 0$ in (5.4), we immediately obtain

$$K(u, 0) = 0 \quad \text{a.e.} \quad \mu \qquad \text{for} \quad u > 0 .$$

Finally, if $\sigma^2 > 0$, we have, for $\hat{f}(v) = \chi(0, v)$,

$$K\hat{f} = K(0, 0)\sigma^2\hat{f} ,$$

and hence, if K is Volterra, $K(0, 0) = 0$, for otherwise $K(0, 0)\sigma^2$ would be a nonzero eigenvalue. The proof is complete.

LEMMA 5.2. *Let $T(u, v)$ be the $L^2([0, 1] \times [0, 1], \mu \times \mu)$ kernel corresponding to the Hilbert-Schmidt operator $T = I - S$. Then for $\hat{f} \in L^2([0, 1], \mu)$,*

$$\Delta^{\frac{1}{2}}\hat{f}(u) = [1 - \sigma^2 T(0, 0)]^{\frac{1}{2}}\hat{f}(0) \quad \textit{with} \quad 1 - \sigma^2 T(0, 0) > 0 \quad \textit{for} \quad u = 0 \tag{5.5}$$
$$= \hat{f}(u) \qquad \textit{for} \quad 0 < u \leqq 1 .$$

PROOF. The only possible gap is $(0, P(0) = P(0+))$, and so $D = I + P(0)[(I - P(0)TP(0))^{-1} - I]P(0)$. Direct verification shows that $\Delta = D^{-1} = I - P(0)TP(0)$. From (5.2) it follows that

$$\Delta\hat{f}(u) = [1 - \sigma^2 T(0, 0)]\hat{f}(0) \qquad \text{if} \quad u = 0$$
$$= \hat{f}(u) \qquad \text{if} \quad 0 < u \leqq 1 .$$

Hence we have (5.5). If $\sigma^2 > 0$, then $1 - \sigma^2 T(0, 0) > 0$, because $\Delta^{\frac{1}{2}}$ is positive.

Consider now

$$\theta Y(t) = \Delta^{\frac{1}{2}}(I + W_+)\theta X(t) = \Delta^{\frac{1}{2}}(I + W_+)\chi(t, \cdot) .$$

Applying Lemma 5.1 to the Volterra operator W_+, we have

$$(I + W_+)\chi(t, \cdot) = \chi(t, \cdot) + \int_0^1 W_+(\cdot, v)\chi(t, v)\mu(dv)$$
$$= \chi(t, \cdot) + \int_0^t W_+(\cdot, v)\, dv ,$$

where $W_+(u, v)$ is the Volterra kernel (i.e., $W_+(u, v) = 0$ a.e. μ for $u > v$ and $W_+(0, 0) = 0$ if $\sigma^2 > 0$) in $L^2([0, 1] \times [0, 1], \mu \times \mu)$ corresponding to W_+. Hence, by Lemma 5.2,

$$\theta Y(t)(u) = [1 - \sigma^2 T(0, 0)]^{\frac{1}{2}}\{1 + \int_0^t W_+(0, v)\, dv\} \qquad \text{for} \quad u = 0$$
$$= \chi(t, u) + \int_0^t W_+(u, v)\, dv \qquad \text{for} \quad 0 < u \leqq 1 .$$

Taking $\xi = Y(t)$ in (5.2), we thus obtain the non-anticipative representation

$$
\begin{aligned}
Y(t) &= [1 - \sigma^2 T(0, 0)]^{\frac{1}{2}}\{1 + \textstyle\int_0^t W_+(0, v)\, dv\}X(0) \\
&\quad + \textstyle\int_0^1 \chi(t, u)\, dB(u) + \int_0^1 \int_0^t W_+(u, v)\, dv\, dB(u) \\
&= [1 - \sigma^2 T(0, 0)]^{\frac{1}{2}}\{1 + \textstyle\int_0^t W_+(0, v)\, dv\}X(0) \\
&\quad + B(t) + \textstyle\int_0^t \{\int_0^v W_+(u, v)\, dB(u)\}\, dv\,,
\end{aligned} \tag{5.6}
$$

where $W_+(u, v)$ is a Volterra kernel in $L^2([0, 1] \times [0, 1])$ and $1 - \sigma^2 T(0, 0) > 0$.

The non-anticipative representation of Gaussian process equivalent to the standard Wiener process. If, in particular, $X(0) = 0$, i.e., if $\{X(t), 0 \leqq t \leqq 1, Q\}$ is the standard Wiener process $\{B(t), 0 \leqq t \leqq 1, Q\}$, then (5.6) becomes

$$Y(t) = X(t) + \textstyle\int_0^t \{\int_0^v W_+(u, v)\, dX(u)\}\, dv\,.$$

This formula has been obtained by Hitsuda (1968) using martingale theory and Girsanov's theorem, and also formally by Kailath (1970) using Gohberg–Krein's results.

6. Non-anticipative representation of Gaussian process equivalent to an N-ple Gaussian-Markov process. Suppose that the process $\{X(t), a \leqq t \leqq b, Q\}$ $(a < 0 < 1 \leqq b)$ with $E_Q X(t) \equiv 0$ has the representation of the form

$$X(t) = \textstyle\int_a^t F(t, u)\, dB(u)\,,$$

where $\{B(u), a \leqq u \leqq b\}$ is the standard Wiener process,

$$F(t, u) = \textstyle\sum_{j=1}^N f_j(t) g_j(u)$$

and $\{f_j(t)\}$, $\{g_j(u)\}$ satisfy the following conditions.

(a) $f_j(t) \in C^{N-1}[a, b]$, $j = 1, \cdots, N$, and $\det(f_j(t_i)) \neq 0$ for any choice of N distinct indices $\{t_i\}$,

(b) $g_j(u) \in L^2([a, b] \cap (-\infty, t])$ for all $t < \infty$, $j = 1, \cdots, N$, and $\{g_j\}$ is linearly independent,

(c) $\partial^k/\partial t^k F(t, u)|_{t=u} = 0$, $k = 0, 1, \cdots, N-2$, $\partial^{N-1}/\partial t^{N-1} F(t, u)|_{t=u} \neq 0$ on $[a, b]$.

Then, for $0 \leqq t \leqq 1$, we have

$$X(t) = \textstyle\sum_{j=1}^N f_j(t)\eta_j + \int_0^1 \chi(t, u) F(t, u)\, dB(u)\,. \tag{6.1}$$

where

$$\eta_j = \textstyle\int_a^0 g_j(u)\, dB(u)\,, \qquad j = 1, 2, \cdots, N\,.$$

Let C be the covariance matrix of η_j, $j = 1, 2, \cdots, N$, and let H_N be the N-dimensional space consisting of column vectors α with inner product $\langle \alpha, \beta \rangle = \alpha^* C \beta$, $*$ denoting the transpose. From the assumptions (a)—(c) it follows that the family of functions $\{\chi(\tau, u)F(\tau, u), 0 \leqq \tau \leqq t\}$ spans the space $L^2[0, t]$ for each $0 \leqq t \leqq 1$. Then (6.1) implies that there is an isometric isomorphism θ from $L(X; 1)$ onto the space $H_N \oplus L^2[0, 1]$ such that

$$\theta X(t) = \{(f_j(t)); \chi(t, u)F(t, u)\} \in H_N \oplus L^2[0, 1]$$

and any element $\xi \in L(X; 1)$ has the representation

$$\xi = \sum_{j=1}^{N} \alpha_j \eta_j + \int_0^1 \varphi(u)\, dB(u) , \tag{6.2}$$

where $\alpha = (\alpha_1, \cdots, \alpha_N)^* \in H_N$; $\varphi \in L^2[0, 1]$ and $\theta\xi = \{a; \varphi\}$. Just as in the preceding section, the operator on $H_N \oplus L^2[0, 1]$ corresponding to an operator $\tilde{A}$ on $L(X; 1)$ will be denoted by A.

Since $\theta[L(X; t)] = H_N \oplus L^2[0, t]$ for $0 < t \leqq 1$, we have $L(X; t+) = L(X; t)$ for all $0 < t \leqq 1$, and $\theta[L(X; 0+)] = H_N$. Thus the chain $\tilde{\pi} = \{0, \tilde{P}(0+), \tilde{P}(t), 0 < t \leqq 1\}$ is closed. It has an N-dimensional gap $(0, \tilde{P}(0+))$. Let H_j, $j = 0, 1, \cdots, N$, denote the subspaces of H_N consisting of vectors of the form $(\alpha_1, \cdots, \alpha_j, 0, \cdots, 0)^*$. Define the orthoprojectors Q_j, $j = 0, 1, \cdots, N$, on $H_N \oplus L^2[0, 1]$ with range H_j. Then the chain $\pi = \{0 = Q_0, Q_1, \cdots, Q_{N-1}, Q_N = P(0+), P(t), 0 < t \leqq 1\}$, where $P(t) = \theta\tilde{P}(t)\theta^{-1}$, is maximal, all gaps (Q_{j-1}, Q_j) being one-dimensional. Q_j, $P(t)$ are characterized by

$$\begin{aligned} Q_j[H_N \oplus L^2[0, 1]] &= H_j , \qquad j = 0, 1, \cdots, N , \\ P(t)[H_N \oplus L^2[0, 1]] &= H_N \oplus L^2[0, t] . \end{aligned} \tag{6.3}$$

It is convenient to use the following matrix form for linear operators K on $H_N \oplus L^2[0, 1]$. We shall write

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} ,$$

where K_{11} is an operator on H_N, i.e., an $N \times N$ matrix, K_{12} an operator from $L^2[0, 1]$ into H_N, which may be represented in the form of a column vector $(K(i, v))$, K_{21} an operator from H_N into $L^2[0, 1]$, which may be written as a row vector $(K(u, j))$, and K_{22} an operator on $L^2[0, 1]$. If we write an element $h = \{h_1; h_2\} \in H_N \oplus L^2[0, 1]$, $h_1 \in H_N$, $h_2 \in L^2[0, 1]$, as a column vector, then Kh is obtained by the usual multiplication rule.

LEMMA 6.1. *Let K be a Hilbert-Schmidt Volterra operator on $H_N \oplus L^2[0, 1]$ having π as an eigenchain. Then,*

(i) $K_{11} = (k_{ij})$ *is a Volterra matrix*, i.e. $k_{ij} = 0$ *for* $i \geqq j$,
(ii) $K_{21} = (K(u, j)) = 0$ a.e.
(iii) $K_{22}(u, v) = 0$ a.e. *for* $u > v$,

where $K_{22}(u, v)$ is the $L^2([0, 1] \times [0, 1])$ kernel corresponding to K_{22}.

PROOF. By assumption,

$$Q_n K Q_n = K Q_n , \qquad n = 0, 1, 2, \cdots, N , \tag{6.4}$$

and

$$P(t) K P(t) = K P(t) , \qquad 0 < t \leqq 1 . \tag{6.5}$$

Let $h = \{\alpha; \varphi\} \in H_N \oplus L^2[0, 1]$, where $\alpha = (\alpha_1, \cdots, \alpha_N)^* \in H_N$ and $\varphi \in L^2[0, 1]$. Then, from (6.3) and (6.4),

$$Q_N K Q_N h = \{K_{11}\alpha; 0\} = K Q_N h = \{K_{11}\alpha; K_{21}\alpha\} .$$

Hence $K_{21} = 0$, which is (ii). Also

$$KQ_n h = \{(\sum_{j=1}^n k_{ij}\alpha_j); \sum_{j=1}^n K(u, j)\alpha_j\}$$
$$= \{(\sum_{j=1}^n k_{ij}\alpha_j); 0\}$$

and

$$Q_n K Q_n h = \{(\sum_{j=1}^n k_{1j}\alpha_j, \cdots, \sum_{j=1}^n k_{nj}\alpha_j, 0, \cdots, 0)^*; 0\}.$$

Thus $\sum_{j=1}^n k_{ij}\alpha_j = 0$ for $i > n$. Taking $\alpha = e_j = (\delta_{ij})$, $j = 1, \cdots, N$, we have $k_{ij} = 0$ for $i > j$. We have also $k_{jj} = 0$ for all j, for otherwise e_j would be an eigenvector with non-zero eigenvalue k_{jj}. (iii) follows by the same argument as in the proof of Lemma 5.1, using (6.5) and taking $h = \{0; \varphi\}$.

From the lemma we see that the operator W_+ is of the form

$$W_+ = \begin{bmatrix} W_{11} & W_{12} \\ 0 & W_{22} \end{bmatrix}$$

where W_{11} is a Volterra matrix (w_{jk}) with $w_{jk} = 0$ for $j \geqq k$ and the kernel $W_{22}(u, v)$ of W_{22} is a Volterra kernel.

Now the operator $\Delta^{\frac{1}{2}}$ is given in the following form.

LEMMA 6.2.

$$\Delta^{\frac{1}{2}} = \sum_{j=1}^N (d_j/d_{j-1})^{\frac{1}{2}}(Q_j - Q_{j-1}) + (I - Q_N),$$

where $d_0 = 1$, $d_j > 0$, $j = 1, 2, \cdots, N$, are the principal minors of the matrix $(I - T_{11})$, T_{11} being the $N \times N$ matrix as the component of the operator T.

PROOF.

$$D = I + \sum_{j=1}^N (Q_j - Q_{j-1})[(I - Q_j T Q_j)^{-1} - I](Q_j - Q_{j-1})$$
$$= I - Q_N + \sum_{j=1}^N (Q_j - Q_{j-1})(I - Q_j T Q_j)^{-1}(Q_j - Q_{j-1}).$$

Now

$$I - Q_j T Q_j = I - Q_j + Q_j(I(j) - T(j))Q_j,$$

where $I(j)$ is the $j \times j$ identity matrix and $T(j) = Q_j T Q_j$ regarded as a $j \times j$ matrix. Hence we see that

$$(I - Q_j T Q_j)^{-1} = I - Q_j + Q_j(I(j) - T(j))^{-1}Q_j$$

and

$$D = I - Q_N + \sum_{j=1}^N (Q_j - Q_{j-1})(I(j) - T(j))^{-1}(Q_j - Q_{j-1}).$$

Regarding now $Q_j - Q_{j-1}$ as a $j \times j$ matrix with (j, j)th element $= 1$ and all other elements $= 0$, we have

$$D = I - Q_N + \sum_{j=1}^N (d_{j-1}/d_j)(Q_j - Q_{j-1}).$$

Therefore

$$\Delta = D^{-1} = \sum_{j=1}^N (d_j/d_{j-1})(Q_j - Q_{j-1}) + I - Q_N$$

and

$$\Delta^{\frac{1}{2}} = \sum_{j=1}^{N} (d_j/d_{j-1})^{\frac{1}{2}}(Q_j - Q_{j-1}) + I - Q_N .$$

That $d_j > 0, j = 1, 2, \cdots, N$, follows from the positive definiteness of $S = I - T$.

In the matrix form $\Delta^{\frac{1}{2}}$ can be represented as

$$\Delta^{\frac{1}{2}} = \begin{bmatrix} \Delta_{11} & 0 \\ 0 & I \end{bmatrix}$$

where Δ_{11} is the diagonal matrix with diagonal elements $(d_j/d_{j-1})^{\frac{1}{2}}$.

Thus we have

$$\begin{aligned} &\Delta^{\frac{1}{2}}(I + W_+)\{(f_j(t)); \chi(t, u)F(t, u)\} \\ &\quad = \{(d_j/d_{j-1})^{\frac{1}{2}}[f_j(t) + \textstyle\sum_{k=j+1}^{N} w_{jk} f_k(t) + \int_0^t W_{12}(j, v)F(t, v)\, dv]; \\ &\qquad \chi(t, u)F(t, u) + \textstyle\int_0^1 W_{22}(u, v)\chi(t, v)F(t, v)\, dv\} \end{aligned}$$

and hence, from (6.2),

$$\begin{aligned} Y(t) &= \tilde{\Delta}^{\frac{1}{2}}(I + \tilde{W}_+)X(t) \\ &= \textstyle\sum_{j=1}^{N} c_j(t)\eta_j + \int_0^t F(t, u)\, dB(u) \\ &\quad + \textstyle\int_0^t \{\int_0^v W_{22}(u, v)\, dB(u)\}F(t, v)\, dv , \end{aligned} \tag{6.6}$$

where

$$c_j(t) = (d_j/d_{j-1})^{\frac{1}{2}}\{f_j(t) + \textstyle\sum_{k=j+1}^{N} w_{jk} f_k(t) + \int_0^t W_{12}(j, v)F(t, v)\, dv\} ,$$

$$W_{12}(j, v) \in L^2[0, 1], j = 1, 2, \cdots, N ,$$

and $W_{22}(u, v)$ is a Volterra kernel in $L^2([0, 1] \times [0, 1])$.

REMARK. (6.6) may be written in the form

$$\begin{aligned} Y(t) &= \textstyle\sum_{j=1}^{N} b_j(t)X^{(j-1)}(0) + \int_0^t F(t, u)\, dB(u) \\ &\quad + \textstyle\int_0^t \{\int_0^v W_{22}(u, v)\, dB(u)\}F(t, v)\, dv , \end{aligned} \tag{6.7}$$

where $X^{(j)}(0)$ are the derivatives of $X(t)$ at $t = 0$. (6.7) may be obtained directly using a different maximal chain containing $\{P(t), 0 < t \leqq 1\}$.

7. Radon-Nikodym derivatives. Hitsuda [3] showed also that if $\{X(t), 0 \leqq t \leqq 1, Q\}$ is the standard Wiener process and if a Gaussian measure P is equivalent to Q, then the Radon–Nikodym (RN) derivative dP/dQ is given by

$$dP/dQ = \exp\{\textstyle\int_0^1 (\int_0^v k(u, v)\, dX(u))\, dX(v) - \frac{1}{2} \int_0^1 (\int_0^v k(u, v)\, dX(u))^2\, dv\} \tag{7.1}$$

where $k(u, v)$ is a Volterra kernel in $L^2([0, 1] \times [0, 1])$. In [4] Kailath derived the above form of RN derivative from Shepp's result [8] using a certain identity for Carleman–Fredholm determinants. In this section we shall obtain a similar form of RN derivatives for the case considered in Section 5.

First we note that, for any equivalent Gaussian measures P and Q corresponding to general Gaussian processes $\{X(t), 0 \leqq t \leqq 1, P\}$ and $\{X(t), 0 \leqq t \leqq 1, Q\}$ with $E_P X(t) = E_Q X(t) = 0$ and covariance functions Γ_P, Γ_Q, the RN derivative

dP/dQ is given by

$$dP/dQ = \lim_n \prod_{j=1}^n (1-\lambda_j)^{-\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\sum_{j=1}^n \frac{\lambda_j}{1-\lambda_j} X_j^2\right\}, \tag{7.2}$$

where $\{\lambda_j\}$ are the eigenvalues of the operator $T = I - S$ on the RKHS $H(\Gamma_Q)$ defined in Theorem 4.2 and $X_j = \psi^{-1}\varphi_j$, φ_j denoting the eigenfunction corresponding to λ_j (cf. e.g. [7]).

Suppose now that $\{X(t), 0 \leqq t \leqq 1, Q\}$ is a Wiener process (see Section 5). Then, because of the isometric isomorphism between $H(\Gamma_Q)$ and $L^2([0,1], \mu)$, $\{\lambda_j\}$ and $\{\varphi_j\}$ can be taken to be the eigenvalues and eigenfunctions of the operator $T \in \mathscr{S}_2$ on $L^2([0,1],\mu)$, and

$$X_j = \varphi_j(0)X(0) + \int_0^1 \varphi_j(u)\, dB(u) .$$

Let U be the Fredholm resolvent operator of T at 1, i.e.,

$$(I+U)(I-T) = I$$

and let $U(u,v)$ be the corresponding kernel. (We shall denote by $A(u,v)$ the $L^2([0,1]\times[0,1], \mu\times\mu)$ kernel corresponding to an operator $A \in \mathscr{S}_2$ on $L^2([0,1], \mu)$.) Then the kernel $U(u,v)$ has the expansion

$$U(u,v) = \sum_{j=1}^\infty \frac{\lambda_j}{1-\lambda_j}\varphi_j(u)\varphi_j(v) .$$

We define the double integral $\int_0^1\int_0^1 H(u,v)\,dX(u)\,dX(v)$ for any symmetric kernel $H(u,v) \in L^2([0,1]\times[0,1], \mu\times\mu)$ by

$$\begin{aligned}\int_0^1\int_0^1 & H(u,v)\,dX(u)\,dX(v) \\ &= H(0,0)[X^2(0)-\sigma^2] + X(0)\int_0^1 H(u,0)\,dB(u) \\ &\quad + X(0)\int_0^1 H(0,v)\,dB(v) + \int_0^1\int_0^1 H(u,v)\,dB(u)\,dB(v) ,\end{aligned}$$

where the last term is the usual Itô's double Wiener integral. This double integral is quite similar to the usual one and we have

$$\lim_n \sum_{j=1}^n \frac{\lambda_j}{1-\lambda_j}(X_j^2 - 1) = \int_0^1\int_0^1 U(u,v)\,dX(u)\,dX(v) .$$

Let $\delta_A(\lambda)$ denote the Carleman-Fredholm determinant of an operator $A \in \mathscr{S}_2$, i.e.,

$$\delta_A(\lambda) = \prod_{j=1}^\infty [(1-\lambda\lambda_j(A))\exp(\lambda\lambda_j(A))] ,$$

where $\lambda_j(A)$ are the eigenvalues of A. If A is Volterra, then, by definition, $\delta_A(\lambda) = 1$. Since

$$\begin{aligned}\lim_n & \prod_{j=1}^n (1-\lambda_j)^{-\frac{1}{2}}\exp\left\{-\tfrac{1}{2}\sum_{j=1}^n \frac{\lambda_j}{1-\lambda_j}\right\} \\ &= \lim_n \prod_{j=1}^n \left[\left(1+\frac{\lambda_j}{1-\lambda_j}\right)\exp\left\{-\frac{\lambda_j}{1-\lambda_j}\right\}\right]^{\frac{1}{2}} \\ &= [\delta_U(-1)]^{\frac{1}{2}} ,\end{aligned}$$

(7.2) can be written in the form

(7.3) $$dP/dQ = [\delta_U(-1)]^{\frac{1}{2}} \exp\{-\tfrac{1}{2} \textstyle\int_0^1 \int_0^1 U(u, v)\, dX(u)\, dX(v)\}\,.$$

We have, by Theorem 2.2,

$$I + U = S^{-1} = (I + X_+)D(I + X_-)\,.$$

Define the operator V by

$$I + V = (I + X_+)(I + X_-)\,.$$

Then the operator

$$U - V = (I + X_+)(D - I)(I + X_-)$$

is 0 or one-dimensional according as $X(0) = 0$ or $X(0) \neq 0$, because

$$D - I = P(0)[(I - P(0)TP(0))^{-1} - I]P(0)$$

and

$$U - V = P(0)(I + X_+)(D - I)(I + X_-)P(0)\,.$$

We apply the following relation to U and V (see [1] page 172): if $A, B \in \mathscr{S}_2$ and $A - B$ is nuclear, then, for λ such that $(I - \lambda A)^{-1}$ exists,

$$\delta_A(\lambda) = \delta_B(\lambda)[D_{B/A}(\lambda)]^{-1} \exp\{\lambda \operatorname{tr}(A - B)\}\,,$$

where $D_{B/A}(\lambda)$ is the perturbation determinant of A by $B - A$, i.e.,

$$D_{B/A}(\lambda) = \textstyle\prod_{j=1}^{\infty} (1 - \nu_j)$$

and $\{\nu_j\}$ denote the eigenvalues of the operator $\lambda(B - A)(I - \lambda A)^{-1}$. Then we have

$$\delta_U(-1) = \delta_V(-1)[D_{V/U}(-1)]^{-1} \exp\{-\operatorname{tr}(U - V)\}\,.$$

Since the eigenvalue of the one-dimensional operator $U - V$ (for the case $X(0) \neq 0$) is $\sigma^2 T(0, 0)/(1 - \sigma^2 T(0, 0))$, which is easily found by applying Lemma 5.1 to X_+ and $X_- = X_+{}^*$,

$$\begin{aligned} \operatorname{tr}(U - V) &= \textstyle\int_0^1 (U - V)(u, u)\mu(du) \\ &= (U - V)(0, 0) \cdot \sigma^2 \\ &= \frac{\sigma^2 T(0, 0)}{1 - \sigma^2 T(0, 0)}\,, \end{aligned}$$

and the eigenvalue of the one-dimensional operator $(U - V)S$ is $\sigma^2 T(0, 0)$, and hence

$$D_{V/U}(-1) = 1 - \sigma^2 T(0, 0)\,.$$

Furthermore, from the identity (see [1] page 169): for $A, B \in \mathscr{S}_2$ and $I - C = (I - A)(I - B)$,

$$\delta_C(1) \exp\{\operatorname{tr}(AB)\} = \delta_A(1)\delta_B(1)\,,$$

it follows that

$$\delta_V(-1) = \exp\{-\mathrm{tr}\,(X_+X_-)\}\,.$$

Thus we have

$$(7.4)\qquad [\delta_U(-1)]^{\frac{1}{2}} = (1 - \sigma^2 T(0,0))^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\,\frac{\sigma^2 T(0,0)}{1-\sigma^2 T(0,0)}\right\} \times \exp\{-\tfrac{1}{2}\,\mathrm{tr}\,(X_+X_-)\}\,.$$

Now consider

$$\begin{aligned}\int_0^1\int_0^1 U(u,v)\,dX(u)\,dX(v) = {} & \int_0^1\int_0^1 (U-V)(u,v)\,dX(u)\,dX(v)\\ & + \int_0^1\int_0^1 (X_+ + X_-)(u,v)\,dX(u)\,dX(v)\\ & + \int_0^1\int_0^1 (X_+X_-)(u,v)\,dX(u)\,dX(v)\,.\end{aligned}$$

Since $(U-V)(u,v) = 0$ a.e. $\mu\times\mu$ for $(u,v)\neq(0,0)$, the first term is

$$(U-V)(0,0)[X^2(0)-\sigma^2] = \frac{T(0,0)}{1-\sigma^2T(0,0)}\,X^2(0) - \frac{\sigma^2T(0,0)}{1-\sigma^2T(0,0)}$$

The second term can be written as the iterated integral

$$2\int_0^1\left\{\int_0^v X_+(u,v)\,dX(u)\right\}dX(v)$$

using properties of stochastic integrals and noting that $X_+(0,0) = X_-(0,0) = 0$ if $\sigma^2 > 0$. The third term is equal to

$$\int_0^1\left\{\int_0^v X_+(u,v)\,dX(u)\right\}^2\mu(dv) - \mathrm{tr}\,(X_+X_-)\,,$$

which follows easily from the definition of the double integral. Therefore we have

$$\begin{aligned}&\exp\left\{-\tfrac{1}{2}\int_0^1\int_0^1 U(u,v)\,dX(u)\,dX(v)\right\}\\ (7.5)\qquad &= \exp\left\{-\frac{1}{2}\,\frac{T(0,0)}{1-\sigma^2T(0,0)}\,X^2(0)\right\}\exp\left\{\frac{1}{2}\,\frac{\sigma^2T(0,0)}{1-\sigma^2T(0,0)}\right\}\\ &\quad\times\exp\left[-\int_0^1\left\{\int_0^v X_+(u,v)\,dX(u)\right\}dX(v)\right.\\ &\quad\left.-\tfrac{1}{2}\int_0^1\left\{\int_0^v X_+(u,v)\,dX(u)\right\}^2\mu(dv)\right]\exp\{\tfrac{1}{2}\,\mathrm{tr}\,(X_+X_-)\}\,.\end{aligned}$$

Substituting (7.4) and (7.5) in (7.3), we obtain the following form for the RN derivative.

$$\begin{aligned}dP/dQ = {} & (1-\sigma^2T(0,0))^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\,\frac{T(0,0)}{1-\sigma^2T(0,0)}\,X^2(0)\right\}\\ & \times\exp\left[-\int_0^1\left\{\int_0^v X_+(u,v)\,dX(u)\right\}dX(v)\right.\\ & \left.-\tfrac{1}{2}\int_0^1\left\{\int_0^v X_+(u,v)\,dX(u)\right\}^2\mu(dv)\right].\end{aligned}$$

In [8] Shepp has derived the RN derivative for the more general case when $X(t)$ is a so-called "free" Wiener process. Our method is different and, we feel, can be used to deduce his result in its full generality.

Note added. After this paper had been submitted the authors learned from Professor T. Kailath that he and Dr. D. Duttweiler have also obtained Theorem 4.1 of this paper. They also derive a general likelihood ratio formula and consider several applications from the engineering standpoint.

REFERENCES

[1] GOHBERG, I. C. and KREIN, M. G. (1969). *Introduction to Linear Nonself-Adjoint Operators* (Eng. trans.). American Mathematical Society, Providence.

[2] GOHBERG, I. C. and KREIN, M. G. (1970). *Theory and Applications of Volterra Operators in Hilbert Space* (Eng. trans.). American Mathematical Society, Providence.

[3] HITSUDA, M. (1968). Representation of Gaussian processes equivalent to Wiener process. *Osaka J. Math.* **5** 299–312.

[4] KAILATH, T. (1970). Likelihood ratios of Gaussian processes. *IEEE Trans. Information Theory* **IT-16** 276–288.

[5] KALLIANPUR, G. and OODAIRA, H. (1963). The equivalence and singularity of Gaussian measures. *Proceedings of Symposium on Time Series Analysis* Wiley, New York, 279–291.

[6] RIESZ, F. and SZ.-NAGY, B. (1955). *Functional Analysis*. Ungar, New York.

[7] ROZANOV, YU. A. (1971). Infinite-dimensional Gaussian distributions (Eng. trans.). American Mathematical Society, Providence.

[8] SHEPP, L. A. (1966). Radon-Nikodym derivatives of Gaussian measures. *Ann. Math. Statist.* **37** 321–354.

[9] KAILATH, T. and DUTTWEILER, D. (1972). An RKHS approach to detection and estimation problems, Part III: Generalized innovations representations and a likelihood-ratio formula. To appear.