A CHARACTERIZATION OF THE KERNELS ASSOCIATED WITH
THE MULTIPLE INTEGRAL REPRESENTATION OF SOME
FUNCTIONAL OF THE WIENER PROCESS

by

Eugene Wong and Moshe Zakai

Memorandum No. UCB/ERL M82/2

5 January 1982

ELECTRONICS RESEARCH LABORATORY
College of Engineering
University of California, Berkeley
94720
A Characterization of the Kernels Associated with the Multiple Integral Representation of Some Functionals of the Wiener Process

Eugene Wong and Moshe Zakai

Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory
University of California, Berkeley, California 94720

Abstract

In this paper we present a characterization of those Wiener functionals that are the likelihood ratio for a "signal plus independent noise" model. The characterization is expressed in terms of the representation of such functionals in a series of multiple Wiener integrals.

Keywords

Wiener functionals, multiple Wiener integrals, likelihood ratio, additive noise model, Radon-Nikodym derivative
Introduction

Let \( \{y_s, 0 \leq s \leq T\} \) be a random process with measurable sample functions satisfying \( |y_s| \leq K \) a.s. Let \( \{W_s, 0 \leq s \leq T\} \) be a Wiener process which is independent of the \( \{y_s, 0 \leq s \leq T\} \) process. Let

\[
A_t = \exp\left\{ \int_0^t y_s \, dW_s - \frac{1}{2} \int_0^t y_s^2 \, ds \right\}
\]

Then \( A_t \) is the unique solution to the integral equation

\[
A_t = 1 + \int_0^t A_s \, y_s \, dW_s
\]

and admits the series representation

\[
A_t = \lim_{N \to \infty} \sum_{n=0}^{N} u_n(t)
\]

where \( u_0(t) = 1 \) and

\[
u_n(t) = \int_0^t u_{n-1}(s) \, y_s \, dW_s
\]

Let \( W_t^\sigma \) denote the \( \sigma \)-field generated by \( \{W_\theta, 0 \leq \theta < t\} \). Since convergence on quadratic mean commutes with conditional expectation, we have (as was observed in [1] and [3])

\[
E(A_t | F_s^w) = \sum_{n=0}^{\infty} E(u_n(t) | F_s^w)
\]

(1)

\[
= \sum_{n=0}^{\infty} \int_0^t (\int_0^{t_n} (\cdots \int_0^{t_2} v_n(t_1, \cdots, t_n) \, dW_{t_1} \cdots) \, dW_{t_{n-1}}) \, dW_{t_n}
\]

where

(2) \( v_n(t_1, \cdots, t_n) = E(y(t_1) \, y(t_2) \cdots \, y(t_n)) \).
and the integrals are iterated Ito integrals. The functional 
\[ g(W) = E(A_T | F_T^W) \]
is a nonnegative functional of the Brownian motion, its 
Wiener-Ito representation is given by (1) with the kernels \( v_n(\cdot) \) 
satisfying (2); that is, the \( n \)-th order kernel \( v_n(t_1, \ldots, t_n) \) is the 
\( n \)-th order moment of a process which is independent of \( W \). In this note 
we consider the converse problem: Let \( g(W) \) be a square integrable 
functional of the Wiener process \( W \) with the Wiener-Ito representation

\[
(3) \quad g(W) = C + \sum_n \int_0^T \left( \int_0^{t_n} \cdots \int_0^{t_2} h_n(t_1, \ldots, t_n) \, dW_{t_1} \cdots dW_{t_n} \right) \, dW_{t_n} 
\]

where the integrals are iterated Ito integrals. This representation will
be abbreviated by

\[
(3)' \quad g(W) = C + \sum_n h_n = W^n 
\]

with \( h_n = W^n \) denoting the \( n \)-th order iterated stochastic integral. The 
problem that we consider is the following: given a square integrable 
functional of the Brownian motion with representation (3) or (3)', what 
conditions would ensure the existence of a process \( \{y_t, 0 < t < T\} \), 
independent of \( W \), such that

\[
(4) \quad h_n(t_1, \ldots, t_n) = E(y(t_1) y(t_2) \cdots y(t_n)) 
\]

for all \( n \) and all \( 0 \leq t_1 < t_2 \ldots < t_n \leq T \)? Another way to state the 
problem is the following: the functional (1) is the likelihood ratio of 
a "signal plus independent noise" with respect to the "noise only" 
hypothesis and the problem is to characterize the nonnegative functionals 
of the Brownian motion that represent the likelihood ratio of a "signal 
plus independent noise" with respect to the "noise only" hypothesis.
Let \( g(\lambda, W) \) denote

\[
g(\lambda, W) = C + \sum_{n} \lambda^{n} h_{n} \ominus W^{n}
\]

It will be shown that \( g(W) \) has the "signal plus independent noise" representation, i.e., \( h_{n} \) satisfy (4) if and only if \( g(\lambda, W) \) as defined by (5) is a nonnegative random variable for every nonnegative \( \lambda \).

Notation: For the representation of \( g(W) \), \( h_{n}(t_{1}, \ldots, t_{n}) \) has to be defined for ordered n-tuples only (cf. (3)), namely for \( t_{1}, \ldots, t_{n} \) satisfying \( 0 \leq t_{1} < t_{2} < \cdots < t_{n} \leq T \). Define \( h_{n}(t_{1}, \ldots, t_{n}) \) for unordered n-tuples of distinct times by

\[
h_{n}(t_{1}, \ldots, t_{n}) = h_{n}(t_{i_{1}}, \ldots, t_{i_{n}})
\]

where \( (t_{i_{1}}, \ldots, t_{i_{n}}) \) is the rearrangement of \( (t_{1}, \ldots, t_{n}) \) which yields an increasing sequence. We will not distinguish between two kernels \( h_{n}(t_{1}, \ldots, t_{n}) \) and \( h_{n}(t_{1}, \ldots, t_{n}) \) which are equal almost everywhere (Lebesgue) on \([0, T]^{n}\). Note that (6) leaves \( h_{n}(t_{1}, \ldots, t_{n}) \) undefined on a Lebesgue set of measure zero in \([0, T]^{n}\).

Let \( \Pi(n, m) \) denote a multinomial of degree \( m \) in \( n \) variables \( x_{1}, \ldots, x_{n} \)

\[
\Pi(n, m) = \sum_{|p| \leq m} C_{P_{1}'P_{2}'\ldots'P_{n}'} x_{1}^{p_{1}} x_{2}^{p_{2}} \ldots x_{n}^{p_{n}}
\]

where \( p_{i}, i = 1, \ldots, n \) are integers and \( |p| = \Sigma_{i} p_{i} \). A multinomial \( \Pi(n, m) \) is said to be nonnegative if it is nonnegative for all values of its arguments \( x_{1}, \ldots, x_{n} \). Let \( W \) be a Wiener process and \( t^{n} \) an n-tuple of real numbers satisfying \( 0 \leq t_{i} \leq T \), \( i = 1, \ldots, n \). \( \Pi(n, m)(W, t^{n}) \)
will denote the multinomial \( \Pi^{(n,m)} \) evaluated at \( x_i = W(t_i), i = 1, \ldots, n: \)

\[
\Pi^{(n,m)}(W, t^n) = \sum_{|p| \leq m} c_{p_1, p_2, \ldots, p_n} x_1^{p_1} x_2^{p_2} \cdots x_n^{p_n}
\]

where \( p_i, i = 1, \ldots, n \) are integers and \( |p| = \sum p_i \). A multinomial \( \Pi^{(n,m)} \) is said to be nonnegative if it is nonnegative for all values of its arguments \( x_1, \ldots, x_n \). Let \( W \) be a Wiener process and \( t^n \) an \( n \)-tuple of real numbers satisfying \( 0 \leq t_i \leq T, i = 1, \ldots, n \). \( \Pi^{(n,m)}(W, t^n) \) will denote the multinomial \( \Pi^{(n,m)} \) evaluated at \( x_i = W(t_i), i = 1, \ldots, n: \)

\[
\Pi^{(n,m)}(W, t^n) = \sum_{|p| \leq m} c_{p_1, \ldots, p_n(W(t_1))^p_1 \ldots (W(t_n))^p_n}
\]

\( H_n(t^n) \) is defined as

\[ H_n(t^n) = H_n(t_1, \ldots, t_n) = \int_0^{t_1} \cdots \int_0^{t_n} h_n(\theta_1, \ldots, \theta_n) d\theta_1 \cdots d\theta_n \]

Finally, \( \overline{\Pi}^{(n,m)}(H, t^n) \) is defined as

\[
\overline{\Pi}^{(n,m)}(H, t^n) = \sum_{|p| \leq m} c_{p_1, \ldots, p_n} H_{|p|}(t_1, t_1, \ldots, t_1, t_2, t_2, \ldots, t_n, \ldots, t_n)
\]

**Theorem.** Let \( g(W) = C + \sum h_n = \overline{W}^n \) be the Wiener Ito representation of a square integrable functional of the Wiener process over \([0,T]\). Assume that for all \( n \) and all \( n \)-tuples \( t^n, |h_n(t^n)| \leq K \). Then the following are equivalent.

(a) \( g(\lambda, W) \) is a nonnegative square integrable random variable for every positive real \( \lambda \).
(b) There exists a sequence of positive real numbers \( \lambda_r \) such that \( \lambda_r \to \infty \) as \( r \to \infty \) and \( g(W) \) and \( g(\lambda_r, W) \) are nonnegative random variables.

(c) For every nonnegative multinomial \( \Pi(n, m) \) it holds that \( \Pi(n, m)(H, t^n) \) is nonnegative.

(d) There exists a random process \( \{y_s, 0 \leq s \leq T\} \) such that \( |y_s| \leq K \) and

\[
E(y(t_1)y(t_2) \cdots y(t_n)) = h_n(t_1, \ldots, t_n)
\]

for almost all (Lebesgue) points \( (t_1, \ldots, t_n) \) in \([0, T]^n\) (the probability space on which the \( y \) process is defined is unrelated to the probability space on which the Wiener process \( W \) is defined).

(e) Let \( P^X_W \) denote the probability measure on the space of continuous functions induced by \( W \). Let \( \{z_t, 0 \leq t \leq T\} \) be a random process with measurable sample paths on \([0, T]\). Let \( P^{X+W}_Z \) denote the probability measure induced by \( X = \int_0^T z_s \text{d}s + W_t \) on the space of continuous functions. There exists a process \( \{z_s, 0 \leq s \leq T\} \) such that \( |z_s| \leq K \) a.s. and \( \{z_s, 0 \leq s \leq T\} \) is independent of the Wiener process \( W \) such that \( P^{X+W}_Z \) is equivalent to \( P^X_W \) and the Radon-Nikodym derivative of \( P^{X+W}_Z \) with respect to \( P^X_W \) satisfies

\[
\frac{dP^{X+W}_Z}{dP^X_W}(W) = \frac{1}{C} g(W). 
\]

**Proof.** Obviously (a) implies (b). We turn now to the proof that (b) implies (c). Recall first that for a deterministic square integrable function \( \phi(\cdot) \) we have (cf. equation (3.4) of [2]):

\[
\int_0^T \phi(\theta)d\theta \int_0^T \cdots \int_0^T h_n(\theta_1, \ldots, \theta_n)d\theta_1 \cdots d\theta_n
\]

\[
= \int_0^T \cdots \int_0^T \phi(\theta_1) h_n(\theta_2, \ldots, \theta_{n+1})d\theta_1 \cdots d\theta_{n+1} + \sum_{q=0}^{n-1} \psi_q \equiv \psi^q
\]

(8)
the exact form of \( \psi \) will not interest us. It follows by repeated applications of this result that

\[
W(t_1) W(t_2) \ldots W(t_n) = \int_0^T \cdots \int_0^T x_{t_1}(\theta_1) x_{t_2}(\theta_2) \cdots x_{t_n}(\theta_n) dW_{\theta_1} \cdots dW_{\theta_n}
\]

\[
+ \sum_{q=0}^{n-2} \psi_q = W^q
\]

for some deterministic \( \psi_q \), \( 0 \leq q \leq n-2 \) where \( x_t(\theta) \) denotes the characteristic function (\( x_t(\theta) = 1 \) for \( \theta \leq t \) and zero otherwise). Note that equation (8) and the last equation the integrals over [0,T]^n are multiple Wiener-Ito integrals. Rewriting \( h_n = W^n \) as a multiple Wiener-Ito integral we have

\[
h_n = W^n = \frac{1}{n!} \int_0^T \cdots \int_0^T h_n(\theta_1, \ldots, \theta_2) dW_{\theta_1} \cdots dW_{\theta_n}
\]

with \( h_n \) extended to [0,T]^n by (6). Consequently, by the orthogonality properties of the Wiener-Ito integrals

(9) \( E\{W(t_1)W(t_2)\ldots W(t_n) \ (h_n = W^n)\} \)

\[
= \int_0^T \cdots \int_0^T x_{t_1}(\theta_1) \cdots x_{t_n}(\theta_n) h_n(\theta_1, \ldots, \theta_n) d\theta_1 \cdots d\theta_n
\]

\[
= H_n(t_1, \ldots, t_n)
\]

Now, let \( \Pi^{(n,m)} \) be a nonnegative multinomial, that is,

\[
\sum_{|p| \leq m} c_{p_1, \ldots, p_n} x_1^{p_1} \cdots x_n^{p_n} \geq 0
\]

for all values of \( x_1, \ldots, x_n \), then, replacing \( x_i \) by \( x_i/\lambda \) it follows that
for all $\lambda > 0$. Therefore, for all values of $\lambda$ for which (b) is satisfied we have

$$E(g(z, W) = \sum_{|\rho| \leq m} \lambda^{m-|\rho|} c_{\rho_1} \cdots c_{\rho_n} \prod_{i=1}^{n} p_i(t_i)^{\rho_i} \prod_{i=1}^{n} (w(t_i))^{\rho_i} \geq 0.$$ 

Denoting by $t = (t_1, t_2, \ldots, t_n)$ the $|\rho|$-tuple

$$t_1, \ldots, t_n = (t_1, t_1, \ldots, t_2, \ldots, t_n, \ldots t_n),$$

we can rewrite the last equation as

$$E\left(1 + \sum_{q} \lambda^{q} h_q = w^q\right) \cdot \left(\sum_{|\rho| \leq m} \lambda^{m-|\rho|} c_{\rho_1} \cdots c_{\rho_n} \int_0^T \cdots \int_0^T \prod_{i=1}^{n} p_i(t_i)^{\rho_i} \prod_{i=1}^{n} (w(t_i))^{\rho_i} \right) + \sum_{i \leq m} \lambda^{m-i} \psi_i = w^{i-2}\right) \geq 0.$$

The result will be a polynomial of order $m$ in $\lambda$. Note that terms of the form

$$E\left((\lambda^{q} h_q = w^q)(\lambda^{m-i} \psi_i = w^{i-2})\right)$$

will contribute to the coefficient of $\lambda^{m-2}$ but not to the coefficient of $\lambda^m$. The coefficient of $(\lambda)^m$ will, therefore, be

$$\sum_{|\rho| \leq m} \lambda^{-|\rho|} c_{\rho_1} \cdots c_{\rho_n} H_{|\rho|}(t_1, \ldots, t_n).$$
Since this is the coefficient of the highest term of a nonnegative polynomial, it must be nonnegative and this proves (c). The proof that (c) implies (d) is based on an infinite dimensional extension of the fundamental result on the existence of a solution to the moment problem (theorem 1.1 of [4]). L. A. Shepp, in an unpublished memorandum, extended theorem 1.1 of [4] and derived conditions for the existence of a probability measure on function space with given moments. His arguments will be repeated here. Let \( X = \{X(t)\} \) denote the space of real valued functions on \([0,T]\) satisfying \( X(t) = \int_0^t x_s ds \) where \( x_s \) is measurable on \([0,T]\) and ess-sup \( |x_t| \leq K \) (the ess. is with respect to the Lebesgue measure). Let \( \Gamma = \{\gamma\} \) be the collection of bounded continuous functions on \( X \) with the norm

\[
|\gamma| = \sup_{X \in X} |\gamma(X)|
\]

For a multinomial \( \Pi^{(m,n)} \) and \( n \)-tuple \( t^n \) set

\[
\gamma_n(X) = \sum_{|p| \leq m} c_{p_1, \ldots, p_n}(X(t_1))^{p_1} \cdots (X(t_n))^{p_n}
\]

Let \( \Gamma_n \) denote the collection of functions on \( X \) which are of the form \( \gamma_n \), note that this is a linear collection of bounded and continuous functions on \( X \). To each \( \gamma_n \) in \( \Gamma_n \) associate the functional

\[
F(\gamma_n) = \Pi^{(m,n)}(H, t^n)
\]

This functional is linear and continuous hence bounded. Therefore, by the Hahn-Banach theorem, there exists a bounded linear extension of \( F(\cdot) \) to all functions \( \gamma \) in \( \Gamma \) and this extension is nonnegative since \( \Pi^{(m,n)}(H, t^n) \)
was assumed to be nonnegative. By the Riesz representation theorem there exists a nonnegative measure \( \mu \) on \( X \) such that

\[
F(\gamma) = \int_X \gamma(x) \, d\mu(x) .
\]

Since \( \mu(\cdot) \) is nonnegative with \( \mu(X) = 1 \), \( \mu \) is a probability measure on the space of functions which are differentiable with a derivative essentially bounded by \( K \). The measure \( \mu(\cdot) \) therefore defines a process \( \{ Y(t), 0 \leq t \leq T \} \) such that

\[
E(Y(t_1) \cdots Y(t_n)) = H_n(t_1, \cdots, t_n)
\]

and this measure induces a measure on the space of bounded measurable functions \( \{ y_s, 0 \leq s \leq T \}, |y_s| \leq K \) and

\[
E(y(t_1) \cdots y(t_n)) = h_n(t_1, \cdots, t_n)
\]

which completes the proof of (d). The proof that (d) implies (e) is given in the introduction and (a) follows by replacing the "signal"

\[
\int_0^t y_s \, ds + W_t \quad \text{by} \quad \lambda \int_0^t \lambda_s \, ds + W_t .
\]

Remarks (a) The questions arises whether every "reasonable" nonnegative functional of the Brownian motion satisfies condition (a) or (b) of the theorem. The answer is negative as the following two simple examples show. The first example, due to L. A. Shepp is as follows: Let

\[
g(W) = \frac{1}{T} W^2(T) = 1 + \frac{2}{T} \int_0^T \int_0^\theta dW_\theta dW_	heta
\]

-10-
In this case $h_n = 0$ for $n > 2$ and therefore part (d) of the theorem cannot be true for this $g(W)$. Note that $g(W) = \frac{1}{T} W^2(T)$ is a continuous nonnegative functional on the space of continuous functions and $\frac{1}{T}(\lambda W(T))^2$ is also nonnegative, however, in this case

$$g(\lambda \cdot W) \neq g(\lambda, W)$$

(the representation (3) is not a continuous functional on the space of continuous functions). The second example is due to B. Hajek: Condition (e) of the theorem implies that

$$E_1(W(T))^2 = E_0(\int_0^T y_s ds + W_T)^2$$

(10) $$= E(\int_0^T y_s ds)^2 + E_0 W^2(T) \geq E_0 W^2(T)$$

where $E_0$ denotes expectation with respect to the $P_W$ measure and $E_1$ denotes expectation with respect to $P_{Y+W}$. On the other hand, if $P_1$ is the measure induced by $dX_t = -\alpha X_t dt + dW_t$, $\alpha > 0$, $X_0 = 0$; then, by Ito's formula

$$E(X(T))^2 = -\alpha 2E[\int_0^T X_s^2 ds + T]$$

hence $E(X(T))^2 < E W^2(T)$ which contradicts (10). Therefore $dP_1/dP_W$ does not satisfy condition (e) of the theorem.

(b) The extension of the results of this note to the case of multiparameter Wiener processes $W(t_1, \ldots, t_n)$ is straightforward and therefore omitted (cf. [1] and [2]).
Acknowledgement: The authors would like to thank L. A. Shepp for fruitful conversations and in particular for explaining to them his extension to function space of the conditions for the existence of a solution to the moment problem and to B. Hajek for interesting discussions. Research sponsored by the Office of Naval Research Contract N00014-80-C-0507.
References


