

Multiple Stochastic Integrals: Projection and Iteration*

Bruce Hajek¹ and Eugene Wong²

¹ Department of Electrical Engineering and Coordinated Science Laboratory,
University of Illinois at Urbana, Illinois 61801, USA

² Department of Electrical Engineering and Computer Sciences and the Electronics
Research Laboratory, University of California, Berkeley, California 94720, USA

Summary. Multiple stochastic integrals are defined relative to a class of sets. The classic cases of multiple Wiener integral and Ito integral (as well as its generalization by Wong-Zakai-Yor) are recovered by specializing the class of sets appropriately. Any square-integrable functional of the Wiener process has a canonical representation in terms of the integrals.

Formulas are given for projecting a stochastic integral onto the space of Wiener functionals and for representing multiple stochastic integrals as iterated integrals. Applications to a change in probability measure arising in a signal detection problem are given.

1. Introduction

Let \mathcal{A} denote the collection of all Borel sets in \mathbb{R}^n with finite Lebesgue measure (denoted by μ). Define a *Wiener process* $\{W(A), A \in \mathcal{A}\}$ as a family of Gaussian random variables with zero mean and

$$EW(A)W(B) = \mu(A \cap B). \quad (1.1)$$

As a set-parameter process, $W(A)$ is additive, i.e.,

$$W(A+B) = W(A) + W(B), \quad \text{a.s.} \quad (1.2)$$

where $A+B$ denotes the union of disjoint sets, and intuitively, we can view $W(A)$ as the integral over A of a Gaussian white noise.

The connection with white noise renders the Wiener process important in applications as well as theory. Consider for example, the following signal detection problem.

A process ξ_t is observed on $t \in T$ where T is a fixed rectangle in \mathbb{R}^n , and we

* Research sponsored by the U.S. Army Research Office under Contracts DAAG29-79-G-0186 and DAAG29-82-K-0091

have to decide between the possibilities: (a) ξ_t contains a random signal Z_t , plus an additive Gaussian white noise and (b) ξ_t contains only noise.

Formulated so as to avoid the pathologies of "white noise", the problem can be stated as follows: Let $\{W(A), A \in \mathcal{R}(T)\}$ be a set-parameter process, with parameter space $\mathcal{R}(T) = \{\text{Borel subsets of } T\}$, and defined on a fixed measurable space (Ω, \mathcal{F}) . Let \mathcal{P}' and \mathcal{P} be two probability measures such that (a) under \mathcal{P}' $W(A) - \int_A Z_t dt$ is a Wiener process independent of $\{Z_t, t \in T\}$, (b) under \mathcal{P} $W(A)$ is a Wiener process.

Now, let \mathcal{F}_W denote the σ -algebra generated by the process W , and let \mathcal{P}'_W and \mathcal{P}_W denote the respective probability measures restricted to \mathcal{F}_W . If $\int_T Z_t^2 dt < \infty$; a.s., then $\mathcal{P}'_W \ll \mathcal{P}_W$ and the detection problem in most cases reduces to one of computing the likelihood ratio

$$\Lambda = \frac{d\mathcal{P}'_W}{d\mathcal{P}_W} \quad (1.3)$$

in terms of the observed process W .

With respect to the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ $\{W(A), A \in \mathcal{R}(T)\}$ is a Wiener process. Hence, Λ is a positive integrable functional of a Wiener process. Computing Λ in terms of W is a problem that can be embedded in a more general one of finding representations of a Wiener functional, which in turn can be embedded (and illuminated in the process) in a still more general problem of representing martingales generated by a Wiener process.

For a random variable Y that is a square-integrable functional of a Wiener process $\{W(A), A \in \mathcal{R}(T)\}$, several representations already exist. The first is the Hermite-Wiener series of Cameron and Martin [1]. The second is in terms of the multiple Wiener integrals as defined by Ito [6]. The third is in terms of the Ito integral [5], and its generalization as defined by Wong and Zakai [9] and Yor [11]. In the last representation the concept of martingales plays a crucial role.

For processes with a multidimensional parameter, it is both more natural and more general to define *martingales* for processes parameterized by sets rather than by points in \mathbb{R}^n . Let $\mathcal{C} \subset \mathcal{R}(T)$ be a collection of closed sets, let $\{\mathcal{F}(A), A \in \mathcal{C}\}$ be a family of σ -algebras such that $A \supset B \Rightarrow \mathcal{F}(A) \supset \mathcal{F}(B)$, and let $\{M(A), A \in \mathcal{C}\}$ be a set-parameter process. We say that $\{M(A), \mathcal{F}(A), A \in \mathcal{C}\}$ is a *martingale* if

$$E(M(A) | \mathcal{F}(B)) = M(B) \quad \text{a.s.}$$

whenever $A \supset B$. Let $\{W(A), A \in \mathcal{R}(T)\}$ be a Wiener process and denote

$$\mathcal{F}_W(A) = \sigma(\{W(B), B \subset A \text{ and } B \in \mathcal{R}(T)\}) \quad (1.4)$$

One of the main objects of this paper is to show that under very general conditions on \mathcal{C} , there is a canonical representation of all square-integrable martingales with respect to $\{\mathcal{F}_W(A), A \in \mathcal{C}\}$, and hence representation for square integrable Wiener functionals. For $\mathcal{C} = \{\text{all closed sets}\}$ the representation reduces to that of multiple Wiener integrals. For $\mathcal{C} = \{\text{all closed rectangles in}$

\mathbb{R}_+^n with the origin as one corner} the representations of Ito, Wong-Zakai, and Yor are recovered. These two are in a sense limiting cases, and between them lies a vast spectrum of choices for \mathcal{C} , giving rise to an equally large array of representations for \mathcal{C} -martingales and Wiener functionals.

The key to these representations is to define multiple stochastic integrals of the form

$$\int_{T^m} \phi(t_1, t_2, \dots, t_m) W(dt_1) \dots W(dt_m)$$

where ϕ is (in general) a random integrand \mathcal{C} -adapted in a suitable sense to be defined later. The integrand in such a stochastic integral is then identified as a certain density of conditional moments.

Next, formulae are found for transformation of multiple stochastic integrals under two operations. The first is a projection formula for the projection of a multiple stochastic integral onto $L^2(\Omega, \mathcal{F}_W, \mathcal{P})$ (equivalently, this is a formula for the conditional expectation of a multiple stochastic integral given \mathcal{F}_W). The second is an iterated integral formula for expressing multiple stochastic integrals defined relative to \mathcal{C} in terms of stochastic integrals defined relative to another class of sets \mathcal{C}' .

Finally, the transformation formulae are applied to the signal detection problem noted above. The projection formula is relevant since the likelihood ratio is the projection of the Radon-Nikodym derivative $d\mathcal{P}/d\mathcal{P}$, and the iterated integral formula is relevant as a first step towards a stochastic calculus in a general framework.

Portions of this paper appear in [4]. This work is an outgrowth of ideas first introduced in the dissertation [3]. The present paper is self-contained except for the omission of two proofs for which the reader is referred to [4].

2. Multiple Stochastic Integrals

Let \mathcal{C} be a collection of closed subsets of a fixed rectangle T in \mathbb{R}^n . Given sets $A_1, A_2, \dots, A_m \in \mathcal{R}(T)$, we shall define their *support* relative to \mathcal{C} to be the following subset of \mathcal{C} :

$$S_{A_1, A_2, \dots, A_m} = \bigcap \{B: B \in \mathcal{C} \text{ and } B \cap A_i \neq \emptyset \text{ for } 1 \leq i \leq m\} \quad (2.1)$$

with the convention that if no such sets B exist then the support is taken to be all of T . Also, the support of the empty collection of sets (i.e. $m=0$) is simply the intersection of all sets in \mathcal{C} and is denoted by S . (Note that $S=S_T$). It will be assumed that the support of any collection of sets $A_1, \dots, A_m \in \mathcal{R}(T)$ is contained in \mathcal{C} . This assumption can be met by enlarging a given collection of sets \mathcal{C} .

If t_1, t_2, \dots, t_m are points in T , their support will be written as S_{t_1, t_2, \dots, t_m} . We say t_1, t_2, \dots, t_m are \mathcal{C} -independent if no point is contained in the support of the remaining ones.

For $\mathcal{C} = \{\text{all closed sets in } T\}$, S_{t_1, t_2, \dots, t_m} is just $\{t_1, \dots, t_m\}$ so that \mathcal{C} -independent means distinct. For $\mathcal{C} = \{\text{all convex sets in } T\}$, the support of m points

is their convex hull and the points are \mathcal{C} -independent if and only if they are extreme points of their convex hull. When $T \subset \mathbb{R}_+^n$ and $\mathcal{C} = \{R_t: t \in T\}$ where R_t denotes the closed rectangle bounded by the origin and t , then S_{t_1, t_2, \dots, t_m} is the smallest set in \mathcal{C} which contains t_1, t_2, \dots, t_m .

Another example is when $T \subset \mathbb{R}_+^n$ and \mathcal{C} is generated by $\{Q_i: i \in T\}$ where $Q_i = \{s \in T: s_i \leq t_i \text{ for some } i\}$. Then for $t_1, t_2, \dots, t_m \in T$, $S_{t_1, t_2, \dots, t_m} = \bigcup_i R_{t_i}$. Moreover

$$\mathcal{C} = \left\{ \bigcup_{i=1}^m R_{t_i}: m < +\infty \text{ and } t_1, t_2, \dots, t_m \in T \right\}$$

For this example, m points are unordered if and only if they are pairwise unordered.

Let \hat{T}^m denote the subset of \mathcal{C} -independent points in T^m . For a given collection \mathcal{C} , \hat{T}^m may be vacuous for sufficiently large m . For example, if $\mathcal{C} = \{R_t\}$ is the collection of rectangles bounded by the origin and $t \in T \subset \mathbb{R}_+^n$, then \hat{T}^m is empty for $m > n$. That is, no more than n points can be \mathcal{C} -independent. In the extreme case $\mathcal{C} = \{T\}$, \hat{T}^m is empty for all $m \geq 1$.

For a subset A of T define $B(\varepsilon, A)$ to be the set of points in T of Euclidean distance at most ε from A . For $\varepsilon > 0$ define the ε -support relative to \mathcal{C} of $A_1, A_2, \dots, A_m \in \mathcal{R}^n(T)$ by

$$S_{A_1, A_2, \dots, A_m}^\varepsilon = S_{B(\varepsilon, A_1) B(\varepsilon, A_2) \dots B(\varepsilon, A_m)},$$

and let $S_{A_1, A_2, \dots, A_m}^{(-)}$ denote the union over all $\varepsilon > 0$ of the ε -support of A_1, A_2, \dots, A_m . Note that the ε -support of A_1, A_2, \dots, A_m increases to $S_{A_1, A_2, \dots, A_m}^{(-)}$ as ε decreases to zero and $S_{A_1, A_2, \dots, A_m}^{(-)}$ is contained in the support of A_1, A_2, \dots, A_m .

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a fixed probability space, let $\{\mathcal{F}(A): A \in \mathcal{R}(T)\}$ be a family of sub- σ -algebras of \mathcal{F} which is increasing in the sense that $A \subset B$ implies that $\mathcal{F}(A) \subset \mathcal{F}(B)$, and let $\{W(A): A \in \mathcal{R}(T)\}$ be a Wiener process such that $\mathcal{F}_W(A) \subset \mathcal{F}(A)$ and $\mathcal{F}_W(A^c)$ is independent of $\mathcal{F}(A)$ for all A in $\mathcal{R}(T)$, where $\mathcal{F}_W(A)$ is defined by Eq. (1.4). These conditions are true, for example, if $\mathcal{F}(A) = \mathcal{F}_W(A)$ for all A .

We shall assume the following conditions on \mathcal{C} and $\{\mathcal{F}(A): A \in \mathcal{R}(T)\}$:

(c₁) For every collection of rectangles A_1, A_2, \dots, A_m such that $\prod_{i=1}^m A_i \subset \hat{T}^m$,

$$\mu(A_i \cap S_{A_1, A_2, \dots, A_m}) = 0;$$

(c₂) For each $m \geq 1$, the mapping

$$t = (t_1, t_2, \dots, t_m) \mapsto S_t$$

is a continuous map from T^m to the collection of compact sets under the Hausdorff metric:

$$\rho(A, B) = (\max_{x \in A} \min_{y \in B} |x - y| + \max_{x \in B} \min_{y \in A} |x - y|); \quad (2.2)$$

(c₃) For every collection of rectangles A_1, A_2, \dots, A_m in T ,

$$\bigvee_{\varepsilon > 0} \mathcal{F}(S_{A_1, A_2, \dots, A_m}^\varepsilon) = \mathcal{F}(S_{A_1, A_2, \dots, A_m}).$$

Since $\mathcal{F}_W(A) \subset \mathcal{F}(A)$ for all A in $\mathcal{R}(T)$, condition (c₃) implies the following condition:

(c'₃) For every collection of rectangles A_1, A_2, \dots, A_m in T ,

$$\mu(S_{A_1 A_2 \dots A_m} - S_{A_1 A_2 \dots A_m}^{(-)}) = 0, \quad (2.3)$$

and if $\mathcal{F}_W(A) = \mathcal{F}(A)$ for all A then conditions (c₃) and (c'₃) are equivalent. Condition (c₃), as well as condition (c₂), is a continuity condition. Note that since the sets in \mathcal{C} are closed, condition (c₂) insures that \hat{T}^m is an open subset of T^m in the relative topology on T^m .

For a \mathcal{C} satisfying conditions c₁-c₃, we shall define multiple stochastic integrals of order m

$$\phi \circ W^m = \int_{\hat{T}^m} \phi_t W(dt_1) \dots W(dt_m)$$

for integrands $\phi(\omega, t)$, $(\omega, t) \in \Omega \times \hat{T}^m$, satisfying

(h₁) ϕ is $\mathcal{F} \times \mu^m$ -measurable.

(h₂) For each $t \in \hat{T}^m$ ϕ_t is $\mathcal{F}(S_t)$ -measurable.

(h₃) $\int_{\hat{T}^m} E \phi_t^2 dt < \infty$.

The space of functions satisfying h₁-h₃ will be denoted by $L_a^2(\Omega \times \hat{T}^m)$. For ϕ and θ in $L_a^2(\Omega \times \hat{T}^m)$ define

$$\langle \phi, \theta \rangle = E \int_{\hat{T}^m} \phi_t \theta_t dt$$

and let $\tilde{\phi}$ denote the symmetrization of ϕ , i.e., $\tilde{\phi}_t = \frac{1}{m!} \sum_{\pi} \phi_{\pi(t)}$, $\pi(t)$ = permutation. Call ϕ atomic if $\phi(\omega, t) = \alpha(\omega) I_A(t)$ where I_A is the indicator function of a product of rectangles $A = \prod_{i=1}^m A_i$ such that $A \subset \hat{T}^m$. A proof of the following theorem is given in [4]. The proof parallels Ito's original construction [5] of the stochastic integral.

Theorem 2.1. *There is a unique linear map denoted by $\phi \circ W^m$ of $\phi \in L_a^2(\Omega \times \hat{T}^m)$ into the space of square-integrable random variables such that*

(a) For an atomic function $\phi = \alpha I_A$

$$\phi \circ W^m = \alpha \prod_i W(A_i).$$

(b) Symmetry:

$$\phi \circ W^m = \tilde{\phi} \circ W^m.$$

(c) Isometry:

$$E(\phi \circ W^m)(\theta \circ W^m) = \langle \tilde{\phi}, \tilde{\theta} \rangle \delta_{m,p}.$$

Remark. Observe that the isometry property of the multiple stochastic integral implies uniqueness up to equivalence of the integrand. That is, if $\phi \circ W^m = \theta \circ W^m$ then

$$\|\tilde{\phi} - \tilde{\theta}\|^2 = \int_{\hat{T}^m} E(\tilde{\phi}_t - \tilde{\theta}_t)^2 dt = 0.$$

Let $\{(\phi \circ W^m)_B, B \in \mathcal{C}\}$ be the set-parameterized process defined by

$$(\phi \circ W^m)_B = \phi I_{B^m} \circ W^m.$$

We shall call $(\phi \circ W^m)_B$ the *indefinite integral* of $\phi \circ W^m$.

Proposition 2.2. *The process $\{(\phi \circ W^m)_B, \mathcal{F}(B): B \in \mathcal{C}\}$ is a martingale.*

Proof. It is enough to establish the proposition when ϕ is atomic. Let $\phi = \alpha I_{A_1 \times A_2 \times \dots \times A_m}$. Then for $B \in \mathcal{C}$,

$$\begin{aligned} E(\phi \circ W^m | \mathcal{F}(B)) &= E\left(\alpha \prod_{i=1}^m W(A_i) | \mathcal{F}(B)\right) \\ &= E\left(\alpha E\left[\prod_{i=1}^m W(A_i) | \mathcal{F}(B \cup S_{A_1, A_2, \dots, A_m})\right] | \mathcal{F}(B)\right) \\ &= E\left(\alpha \prod_{i=1}^m W(A_i \cap B) | \mathcal{F}(B)\right) \\ &= E(\alpha | \mathcal{F}(B)) \prod_{i=1}^m W(A_i \cap B). \end{aligned}$$

Now since $B \in \mathcal{C}$, if $A_i \cap B \neq \emptyset$ for $1 \leq i \leq m$ then $B \supset S_{A_1, A_2, \dots, A_m}$ and in that case $E(\alpha | \mathcal{F}(B)) = \alpha$ a.s. On the other hand, if $A_i \cap B = \emptyset$ for some i , then $\prod_{i=1}^m W(A_i \cap B) = 0$. Hence in either case

$$E(\alpha | \mathcal{F}(B)) \prod_{i=1}^m W(A_i \cap B) = \alpha \prod_{i=1}^m W(A_i \cap B) = (\psi \circ W^m)_B.$$

Thus, $\{(\phi \circ W^m)_B, \mathcal{F}(B): B \in \mathcal{C}\}$ is indeed a martingale. \square

3. Integrands as Moment Densities and a Projection Formula

The isometry property of the multiple stochastic integrals can be given the following interpretation. Suppose that for each $m \geq 1$ and $t \in T^m$ that $\{\phi_{m,k}(t): k \geq 1\}$ is a complete orthogonal basis for the space of square integrable, $\mathcal{F}(S_t)$ -measurable random variables, and suppose that $\phi_{m,k}(t)$ is a symmetric function of t . Then, formally, the isometry property of multiple stochastic integrals means that the set of "incremental" random variables

$$\{\phi_{m,k}(t) W(dt_1) W(dt_2) \dots W(dt_m): m \geq 0, k \geq 1, t \in \hat{T}^m\} \quad (3.1)$$

is an orthogonal collection of random variables which are also orthogonal to the $\mathcal{F}(S)$ -measurable random variables. (Of course, the increments dt_i in (3.1) are "outward" from S_t .) This fact is reflected in the next proposition which states that the symmetrized integrands are uniquely determined as moment densities. The completeness property proven in Sect. 5 formally means that the collection of variables in (3.1) together with the $\mathcal{F}(S)$ -measurable variables are complete in $L^2(\Omega, \mathcal{F}_W(T), \mathcal{P})$ if $\mathcal{F}(A) = \mathcal{F}_W(A)$ for all A in $\mathcal{A}(T)$.

Proposition 3.1. Let $\gamma \in L^2_a(\Omega \times \hat{T}^m)$. Then for $t \in \hat{T}^k$,

$$E[W(dt_1)W(dt_2)\dots W(dt_k)\gamma \circ W^m | \mathcal{F}(S_t)]/dt_1 dt_2 \dots dt_k = m! \hat{\gamma}(t) \delta_{mk} \quad (3.2)$$

in the sense that the linear functional

$$f \rightsquigarrow E \int_{\hat{T}^k} f(t) W(dt_1) W(dt_2) \dots W(dt_k) \gamma \circ W^m$$

defines a symmetric finite signed measure on the σ -algebra of subsets of $\Omega \times \hat{T}^k$ generated by \mathcal{C} -adapted atomic functions, the measure is absolutely continuous with respect to $\mathcal{P} \times \mu^k$ measure, and the Radon-Nikodym derivative is $m! \gamma \delta_{mk}$.

Proof. In view of the definition of Radon-Nikodym derivatives, Proposition 3.1 is simply a restatement of the isometry property of the multiple stochastic integrals. \square

In the following proposition, $L^2_a(\Omega \times \hat{T}^m, \mathcal{F}_W(\cdot))$ is defined in the same way as $L^2_a(\Omega \times \hat{T}^m)$ except with the σ -algebras $\mathcal{F}(A)$ replaced by $\mathcal{F}_W(A)$ for all $A \in \mathcal{A}(T)$.

Proposition 3.2. (Projection formula.) For each $\gamma \in L^2_a(\Omega \times \hat{T}^m)$ there is a $\hat{\gamma} \in L^2_a(\Omega \times \hat{T}^m, \mathcal{F}_W(\cdot))$ such that

$$\hat{\gamma}(t) = E[\hat{\gamma}(t) | \mathcal{F}_W(S_t)] \quad \text{for a.e. } t \in \hat{T}^m \quad (3.3)$$

and for such $\hat{\gamma}$ and all $A \in \mathcal{C}$,

$$E[\gamma \circ W^m | \mathcal{F}_W(A)] = (\hat{\gamma}_m \circ W^m)(A) \quad (3.4)$$

Proof. By the completeness of multiple stochastic integrals in $L^2(\Omega, \mathcal{F}_W(T), \mathcal{P})$ (see Proposition 5.1 below) and the fact that $E[\gamma \circ W^m | \mathcal{F}_W(S)] = 0$, there exists a collection $\{\phi_k: k \geq 1\}$ with $\phi_k \in L^2_a(\Omega \times \hat{T}^m, \mathcal{F}_W(\cdot))$ such that

$$E[\gamma \circ W^m | \mathcal{F}_W(T)] = \sum_{k=1}^{\infty} \phi_k \circ W^k.$$

Now by Proposition 3.1 with \mathcal{F} replaced by \mathcal{F}_W ,

$$E[W(dt_1)W(dt_2)\dots W(dt_k)E[\gamma \circ W^m | \mathcal{F}_W(T)] | \mathcal{F}_W(S_t)]/dt_1 dt_2 \dots dt_k = k! \tilde{\phi}_k(t)$$

so that

$$E[W(dt_1)W(dt_2)\dots W(dt_k)\gamma \circ W^m | \mathcal{F}_W(S_t)]/dt_1 dt_2 \dots dt_k = k! \tilde{\phi}_k(t) \quad (3.5)$$

on \hat{T}^k . Comparison of Eqs. (3.2) and (3.5) reveals that

$$\tilde{\phi}_k(t) = E[\hat{\gamma}(t) \delta_{mk} | \mathcal{F}_W(S_t)] \quad \text{a.e. } t \in \hat{T}^k.$$

Thus $\tilde{\phi}_k(t) = 0$ for a.e. $t \in \hat{T}^k$ unless $k = m$. So if $\hat{\gamma}$ is defined by $\hat{\gamma} = \phi_m$ then $\hat{\gamma}$ satisfies Eqs. (3.3) and (3.4) is true for $A = T$. Since each side of Eq. (3.4) is a martingale relative to $\{\mathcal{F}_W(A): A \in \mathcal{C}\}$, (3.4) is true for all $A \in \mathcal{C}$. Finally, since $\hat{\gamma}$ is uniquely determined on \hat{T}^m up to a set of $\mathcal{P} \times \mu^m$ measure zero by Eq. (3.3), any $\hat{\gamma} \in L^2_a(\Omega \times \hat{T}^m, \mathcal{F}_W(\cdot))$ satisfying (3.3) also satisfies Eq. (3.4). \square

4. Nested Classes of Sets \mathcal{C} and the Iterated Integration Formula

Let \mathcal{C} and \mathcal{C}' with $\mathcal{C} \supset \mathcal{C}'$ be two classes of sets which each satisfy conditions c_1 - c_3 for a Wiener process $\{W(A): A \in \mathcal{R}(T)\}$ and a collection of σ -algebras $\{\mathcal{F}(A): A \in \mathcal{R}(T)\}$ as in Sect. 2. A dot above (or above and to the right) denotes definition relative to \mathcal{C}' so that, for example, \dot{S}_t denotes the \mathcal{C}' -support of t and \dot{T}^m denotes the collection of \mathcal{C}' -independent points in T^m .

An important example which is exploited in the next section is when \mathcal{C}' is any class satisfying conditions c_1 - c_3 with $\mathcal{F}(A) = \mathcal{F}_W(A)$ for all $A \in \mathcal{R}(T)$, and \mathcal{C} is the collection of all closed sets. Another natural way in which nested collections of sets \mathcal{C} arise is given by the following propositions.

Proposition 4.1. *Let \mathcal{C} satisfy conditions c_1 - c_3 . Suppose that $t = (t_1, t_2, \dots, t_k) \in T^k$ is fixed and define a subcollection \mathcal{C}_t of \mathcal{C} by*

$$\mathcal{C}_t = \{C \in \mathcal{C} : \{t_1, t_2, \dots, t_k\} \subset C\}.$$

Then \mathcal{C}_t also satisfies conditions c_1 - c_3 .

Proof. See Appendix A.

Theorem 4.2 (Iterated integration formula). *Suppose that \mathcal{C} and \mathcal{C}' each satisfy conditions c_1 - c_3 and that $\mathcal{C}' \subset \mathcal{C}$. Then for $\theta \in L^2_\alpha(\Omega \times \dot{T}^m)$ the class- \mathcal{C} stochastic integral $\theta \circ W^m$ can be represented as a sum of class- \mathcal{C}' integrals:*

$$\theta \circ W^m = E[\theta \circ W^m | \mathcal{F}(\dot{S})] + \sum_{k=1}^m \binom{m}{k} \phi_k \circ W^k \quad (4.1)$$

where the integrands $\phi_k \in L^2_\alpha(\Omega \times \dot{T}^k)$ satisfy

$$\phi_k(t) = (\tilde{\theta}(t \times \cdot) I_{\dot{S}_t^{m-k}}) \circ W^{m-k} \quad \text{for a.e. } t \in \dot{T}^k. \quad (4.2)$$

For each fixed t the integral on the right side of Eq. (4.2) is defined relative to the collection of sets \mathcal{C}_t .

Proof. Let $\Pi\theta$ denote transformation of θ by a permutation of its arguments. Suppose for some permutation Π that

$$\Pi\theta = \alpha I_{A_1 \times \dots \times A_m}$$

where $A_i \in \mathcal{R}(T)$, $A_1 \times \dots \times A_m \subset \dot{T}^m$, α is a bounded $\mathcal{F}(S_{A_1, \dots, A_m})$ measurable random variable, $A_1 \times \dots \times A_k \subset \dot{T}^k$ and $A_{k+1}, A_{k+2}, \dots, A_m \subset S_{A_1, A_2, \dots, A_k}$. Then, symmetry implies that

$$\begin{aligned} \theta \circ W^m &= \Pi\theta \circ W^m = \left[\alpha \prod_{i=k+1}^m W(A_i) \right] \prod_{i=1}^k W(A_i) \\ &= h_k \circ W^k \end{aligned}$$

where for $t \in \dot{T}^k$

$$\begin{aligned} h_k(t) &= I_{A_1 \times \dots \times A_k}(t) [\alpha I_{A_{k+1} \times \dots \times A_m} \circ W^{m-k}] \\ &= [\Pi\theta(t \times \cdot) I_{\dot{S}_t^{m-k}}] \circ W^{m-k}. \end{aligned}$$

The isometry property of multiple stochastic integrals relative to \mathcal{C} implies that both k and the two sets $\{A_1, A_2, \dots, A_k\}$ and $\{A_{k+1}, A_{k+2}, \dots, A_m\}$ are unique. The integer k is unique because otherwise we would have

$$E(\theta \circ W^m)^2 = E(h_k \circ W^k)(h_{k'} \circ W^{k'}) = 0.$$

The collection $\{A_1, A_2, \dots, A_k\}$ is unique because otherwise we would have

$$\theta \circ W^m = h_k \circ W^k = g_k \circ W^k$$

and $\tilde{h}_k \tilde{g}_k \equiv 0$. It follows that

$$\sum_{\text{all } \Pi} [(\Pi \theta)(t \times \cdot) I_{S_T^{m-k}}(\cdot) \circ W^{m-k}] \circ W^k = k!(m-k)! \theta \circ W^m \\ = m! \phi_k \circ W^k$$

where ϕ_k is given by Eq. (4.2). Hence

$$\theta \circ W^m = \binom{m}{k} \phi_k \circ W^k$$

which is just Eq. (4.2) for the given θ . In Appendix B it is proved that linear combinations of such θ 's are dense in $L_a^2(\Omega \times \hat{T}^m)$. The proof of the theorem is then completed by an application of the isometric property of the stochastic integrals. \square

5. Completeness of Multiple Stochastic Integrals and an Exponential Formula

The iterated integration formula is applied in this section when one of the classes of sets \mathcal{C} consists of all closed subsets of T and $\mathcal{F}(A) = \mathcal{F}_W(A)$ for all $A \in \mathcal{R}(T)$. The associated integrals are then multiple Wiener integrals. Let \hat{T}^m denote the set of m -tuples of distinct points in T and for θ in $L^2(\hat{T}^m)$ let $\theta \square W^m$ denote a multiple Wiener integral of order m .

Proposition 5.1 (Completeness of multiple stochastic integrals). *Let \mathcal{C} be a collection of sets such that \mathcal{C} and $\{\mathcal{F}_W(A)\}$ satisfy conditions c_1 - c_3 . Then every square-integrable $\mathcal{F}_W(T)$ -measurable random variable Z has a representation of the form*

$$Z = E[Z | \mathcal{F}(S)] + \sum_{m=1}^{\infty} Z_m \circ W^m \quad (5.1)$$

where $Z_m \circ W^m$ are stochastic integrals defined relative to \mathcal{C} and $S = \bigcap \{C : C \in \mathcal{C}\}$.

Proof. The proposition is well known [6] in case \mathcal{C} consists of all closed subsets of T , for then the integrals are multiple Wiener integrals. Since by the iterated integration formula any multiple Wiener integral can be represented as a sum of multiple stochastic integrals relative to the smaller class of sets \mathcal{C} , Proposition 5.1 is true in general. \square

Proposition 5.2. For f in $L^2(T)$, define

$$\dot{f}^m(t_1, t_2, \dots, t_m) = \prod_{i=1}^m f(t_i) \quad (5.2)$$

and set

$$W_m(f, A) = (\dot{f}^m \square W^m)_A. \quad (5.3)$$

If \mathcal{C} and $\{\mathcal{F}_W(A)\}$ satisfy conditions c_1 - c_3 then for $A \in \mathcal{C}$,

$$W_m(f, A) = W_m(f, S \cap A) + \sum_{k=1}^m \binom{m}{k} [\dot{f}^k(\cdot) W_{m-k}(f, S.) \circ W^k]_A \quad (5.4)$$

Proof. Observe that \dot{f}^k is symmetric and

$$\dot{f}^m(t_1, t_2, \dots, t_m) = \dot{f}^k(t_1, t_2, \dots, t_k) \dot{f}^{m-k}(t_{k+1}, \dots, t_m)$$

Hence, Eq. (5.4) for $A=T$ is obtained by applying the iterated integration formula to express the multiple Wiener integral $W_m(f, T)$ in terms of stochastic integrals relative to \mathcal{C} . Then Eq. (5.4) is true in general since each side is a martingale relative to $\{\mathcal{F}_W(A): A \in \mathcal{C}\}$.

Proposition 5.3. Let \mathcal{C} and $\{\mathcal{F}(A): A \in \mathcal{A}(T)\}$ satisfy the conditions of Sect. 2. Then if either $f \in L^2(T)$ or if f is a bounded function in $L^2_a(\Omega \times T)$ define

$$L(f, A) = \exp((f \circ W)_A - \frac{1}{2}(f^2 \circ \mu)_A) \quad (5.5)$$

where $(f^2 \circ \mu)_A$ denotes the Lebesgue integral of f^2 over A . Then for $A \in \mathcal{C}$,

$$L(f, A) = L(f, S \cap A) + \sum_{m=1}^{\infty} \frac{1}{m!} [\dot{f}^m(\cdot) L(f, S.) \circ W^m]_A. \quad (5.6)$$

Proof. Suppose first that $f \in L^2(T)$. For multiple Wiener integrals ($\mathcal{C} = \{\text{all closed sets}\}$) Eq. (5.6) reduces to

$$L(f, A) = 1 + \sum_{m=1}^{\infty} \frac{1}{m!} W_m(f, A) \quad (5.7)$$

which is well known [6]. For the case of general \mathcal{C} , we use (5.4) in (5.7) and write

$$\begin{aligned} L(f, A) &= 1 + \sum_{m=1}^{\infty} \frac{1}{m!} \left(W_m(f, S \cap A) + \sum_{k=1}^m \binom{m}{k} [\dot{f}^k W_{m-k}(f, S.) \circ W^k]_A \right) \\ &= L(f, S \cap A) + \sum_{k=1}^{\infty} \frac{1}{k!} \left[\dot{f}^k \sum_{j=0}^{\infty} \frac{1}{j!} W_j(f, S.) \circ W^k \right]_A \\ &= L(f, S \cap A) + \sum_{k=1}^{\infty} \frac{1}{k!} [\dot{f}^k L(f, S.) \circ W^k]_A \end{aligned}$$

which establishes (5.6) for f in $L^2(T)$. The Eq. (5.6) can then be extended to bounded f in $L^2_a(\Omega \times \hat{T})$ by an approximation argument (see [4], Proposition 3.2). \square

6. Change of Measure and Likelihood Ratio Formulas

Let $\{W(A), \mathcal{F}(A): A \in \mathcal{D}(T)\}$ on $(\Omega, \mathcal{F}, \mathcal{P})$ and a collection of sets \mathcal{C} satisfy the assumptions of Sect. 2. Suppose that \mathcal{P}' is another probability measure on (Ω, \mathcal{F}) which is mutually absolutely continuous relative to \mathcal{P} and is such that the Radon-Nikodym derivative $d\mathcal{P}'/d\mathcal{P}$ is \mathcal{P} -square-integrable and has the representation

$$\frac{d\mathcal{P}'}{d\mathcal{P}} = E \left[\frac{d\mathcal{P}'}{d\mathcal{P}} \middle| \mathcal{F}(S) \right] + \sum_{m=1}^{\infty} \gamma_m \circ W^m \quad (6.1)$$

in terms of \mathcal{C} -stochastic integrals. If $L(A)$ denotes the Radon-Nikodym derivative of $(\mathcal{P}'$ restricted to the σ -algebra $\mathcal{F}(A)$) relative to $(\mathcal{P}$ restricted to $\mathcal{F}(A))$ then $\{L(A), \mathcal{F}_W(A): A \in \mathcal{C}\}$ is a martingale with the representation

$$L(A) = E \left[\frac{d\mathcal{P}'}{d\mathcal{P}} \middle| \mathcal{F}(S) \right] + \sum_{m=1}^{\infty} (\gamma_m \circ W^m)(A).$$

Now replacing γ by γ_m in each side of Eq. (3.2) and summing over m yields that for each $m \geq 1$ and for $t \in \hat{T}^m$,

$$E[W(dt_1)W(dt_2) \dots W(dt_m)L(T)|\mathcal{F}(S_t)]/dt_1 dt_2 \dots dt_m = m! \tilde{\gamma}_m(t).$$

Dividing each side of this equation by $L(S_t)$ and defining $r_m(t) = m! \tilde{\gamma}_m(t)/L(S_t)$ yields that

$$E'[W(dt_1)W(dt_2) \dots W(dt_m)|\mathcal{F}(S_t)]/dt_1 \dots dt_m = r_m(t) \quad (6.2)$$

where E' denotes (conditional) expectation relative to measure \mathcal{P}' . Thus, the Radon-Nikodym derivative $L(A)$ for $A \in \mathcal{C}$ has the representation

$$L(A) = E \left[\frac{d\mathcal{P}'}{d\mathcal{P}} \middle| \mathcal{F}(S) \right] + \sum_{m=1}^{\infty} \frac{1}{m!} [(r_m(\cdot)L(S_\cdot)) \circ W^m]_A \quad (6.3)$$

where the functions r_m have been identified in Eq. (6.2) as the density of conditional m -th moments of W under measure \mathcal{P}' .

Next, define $\Lambda(A) = E[L(A)|\mathcal{F}_W(A)]$. $\Lambda(A)$ is called a likelihood ratio. By an application of the projection formula to each term on the right side of Eq. (6.1),

$$\Lambda(A) = E \left[\frac{d\mathcal{P}'}{d\mathcal{P}} \middle| \mathcal{F}_W(S) \right] + \sum_{m=1}^{\infty} (\hat{\gamma}_m \circ W^m)(A)$$

where the integrands $\hat{\gamma}_m \in L^2_{\mathcal{P}'}(\Omega \times \hat{T}^m, \mathcal{F}_W(\cdot))$ satisfy

$$\hat{\gamma}_m(t) = E[\tilde{\gamma}_m(t)|\mathcal{F}_W(t)] \quad \text{a.e. } t \in \hat{T}^m. \quad (6.4)$$

Now $\Lambda(A)$ is the Radon-Nikodym derivative of $(\mathcal{P}'$ restricted to $\mathcal{F}_W(A))$ relative to $(\mathcal{P}$ restricted to $\mathcal{F}_W(A))$ and thus $(\Lambda(A), \mathcal{F}_W(A))$ has the same structure assumed of $(L(A), \mathcal{F}(A))$. Thus, if \hat{r}_m is defined by $\hat{r}_m(t) = m! \hat{\gamma}_m(t)/\Lambda(S_t)$ for $t = (t_1, t_2, \dots, t_m) \in \hat{T}^m$ then

$$\hat{r}_m(t) = E'[W(dt_1)W(dt_2) \dots W(dt_m)|\mathcal{F}_W(S_t)]/dt_1 dt_2 \dots dt_m \quad (6.5)$$

and the likelihood ratio $\Lambda(A)$ for $A \in \mathcal{C}$ has the representation

$$\Lambda(A) = E \left[\frac{d\mathcal{P}'}{d\mathcal{P}} \middle| \mathcal{F}_W(S) \right] + \sum_{m=1}^{\infty} \frac{1}{m!} [(\hat{r}(\cdot) \Lambda(S, \cdot)) \circ W^m]_A. \quad (6.6)$$

Also, comparing (6.2) and (6.5) (or using Eq. (6.4)) yields that

$$\hat{r}_m(t) = E[r_m(t) | \mathcal{F}_W(S_t)] \quad \text{a.e. } t \in \hat{T}^m. \quad (6.7)$$

Remark. Equation (6.3) (resp. Eq. (6.6)) can be viewed as an integral equation for $L(\cdot)$ (resp. $\Lambda(\cdot)$) in terms of the moment densities r_m (resp. \hat{r}_m) and the Wiener process. As shown in examples below, it is sometimes possible to explicitly solve these equations (also see [2], [7] and [8]).

Since the measure \mathcal{P}' is thus at least formally determined by the functions $\{r_m\}$, it should be possible to express other moments under measure \mathcal{P}' in terms of $\{r_m\}$. In this direction, we consider next moments as in Eq. (6.2) but with S_t replaced by a larger set.

Let \mathcal{C}' be another class of sets with $\mathcal{C}' \subset \mathcal{C}$ so that the assumptions of Sect. 2 are also satisfied by $\{W(A), \mathcal{F}(A): A \in \mathcal{R}(T)\}$ and \mathcal{C}' . The notation introduced in Sect. 4 will be used in what follows.

By the iterated integration formula for \mathcal{C} -stochastic integrals in terms of \mathcal{C}' -stochastic integrals, Eq. (6.1) yields that

$$\frac{d\mathcal{P}'}{d\mathcal{P}} = E \left[\frac{d\mathcal{P}'}{d\mathcal{P}} \middle| \mathcal{F}(S) \right] + \sum_{k=1}^{\infty} \dot{\gamma}_k \circ W^k \quad (6.8)$$

where the \mathcal{C}' -adapted integrands $\dot{\gamma}_k$ satisfy

$$\dot{\gamma}_k(t) = \sum_{m=k}^{\infty} \binom{m}{k} (\dot{\gamma}_m(t \times \cdot) I_{S_t^{m-k}})^{\dagger} W^{m-k} \quad \text{for a.e. } t \in \hat{T}^k. \quad (6.9)$$

Now Eq. (6.8) is the same as Eq. (6.1) with \mathcal{C} and $\{\gamma_m: m \geq 1\}$ replaced by \mathcal{C}' and $\{\dot{\gamma}_m: m \geq 1\}$. Thus, $\{\dot{L}(A): A \in \mathcal{C}'\}$ and \dot{r}_m can be defined relative to \mathcal{C}' in the same way as the corresponding quantities were defined relative to \mathcal{C} . In particular, Eq. (6.2) yields that for $t \in \hat{T}^m$,

$$E'[W(dt_1)W(dt_2)\dots W(dt_m) | \mathcal{F}(S_t)] / dt_1 dt_2 \dots dt_m = \dot{r}_m(t).$$

That is, \dot{r}_m is a conditional moment density for W under \mathcal{P}' just as r_m is, except that S_t is replaced by the larger set \dot{S}_t .

Multiplying each side of Eq. (6.9) by $k!/L(\dot{S}_t)$ yields that for a.e. $t \in \hat{T}^k$,

$$\dot{r}_k(t) = \frac{1}{L(\dot{S}_t)} \sum_{m=k}^{\infty} \frac{1}{(m-k)!} (L(S_{t \times \cdot}) r_m(t \times \cdot) I_{S_t^{m-k}})^{\dagger} W^{m-k} \quad (6.10)$$

This equation represents the moment density \dot{r}_k in terms of the moment densities $\{r_m: m \geq k\}$, the Wiener process, and L . Of course a similar representation holds for \dot{r}_k in terms of $\{\hat{r}_k: k \geq m\}$, the Wiener process, and Λ .

At this point more structure will be assumed on the Radon-Nikodym derivative $d\mathcal{P}'/d\mathcal{P}$. Suppose that $\{Z(t): t \in T\}$ is a bounded, measurable process

such that $Z(t)$ is $\mathcal{F}(S_t)$ -measurable for each $t \in T$. Then it will be assumed that, in the notation of Propositions 5.3,

$$\frac{d\mathcal{P}'}{d\mathcal{P}} = L(Z, T) \quad (6.11)$$

Thus $L(A) = L(Z, A)$ and $r_m = \dot{Z}^m$. By Eq. (6.7), the moment density \hat{r}_m in the likelihood ratio representation (6.6) satisfies

$$\hat{r}_m(t) = E'[\dot{Z}^m(t) | \mathcal{F}(S_t)] \quad \text{a.e. } t \in T^m. \quad (6.12)$$

Therefore \hat{r}_m is now actually a conditional m -th moment rather than just a moment density as in (6.5).

The assumption (6.11) arises in a detection problem for which a signal is observed in white Gaussian noise. Indeed, define a process $\{X(A): A \in \mathcal{R}(T)\}$ by

$$X(A) = W(A) - (Z \circ \mu)(A)$$

Then trivially $W(A) = X(A) + (Z \circ \mu)(A)$, and the following proposition is true:

Proposition 6.1. $\{X(A): A \in \mathcal{R}(T)\}$ is a Wiener process under \mathcal{P}' and for each $t \in \hat{T}^m$, the collection of random variables $\{X(A): A \cap S_t = \emptyset\}$ is \mathcal{P}' -independent of $\mathcal{F}(S_t)$.

Proof. It suffices to prove that for $t \in T^m$, if A_1, A_2, \dots, A_k are disjoint rectangles contained in $T - S_t$ and if $\alpha_1, \alpha_2, \dots, \alpha_k$ are bounded, $\mathcal{F}(S_t)$ -measurable random variables, then $E' \Phi = 1$ where

$$\Phi = \exp\left(\sum_i \alpha_i X(A_i) - \frac{1}{2} \sum_i \alpha_i^2 \mu(A_i)\right).$$

Define a function h on $\Omega \times T$ by $h = \sum \alpha_i I_{A_i}$. Then

$$\Phi = \exp(h \circ W + (hZ - \frac{1}{2} h^2) \circ \mu) \quad (6.13)$$

where $h \circ W$ is a stochastic integral defined relative to the class of sets \mathcal{C}_t . By the fact that $L(Z, S_t) = E\left[\frac{d\mathcal{P}'}{d\mathcal{P}} \middle| \mathcal{F}(S_t)\right]$ we have

$$E'[\Phi | \mathcal{F}(S_t)] = E[\Phi L(Z, T) | \mathcal{F}(S_t)] / L(Z, S_t) = E[\Phi L(Z, T - S_t) | \mathcal{F}(S_t)]. \quad (6.14)$$

Now the integral $(ZI_{S_t}) \circ W$ in the definition of $L(Z, T - S_t)$ can be defined relative to \mathcal{C}_t with the same result as its definition relative to \mathcal{C} . Thus, using Eq. (6.13), Eq. (6.14) becomes

$$E'[\Phi | \mathcal{F}(S_t)] = E[L'(h + ZI_{S_t}, T) | \mathcal{F}(S_t)] \quad (6.15)$$

where L' is defined in the same way as L except relative to the class of sets \mathcal{C}_t instead of \mathcal{C} . Finally, by the martingale property of L' relative to the class of sets \mathcal{C}_t (which contains S_t), the right side of Eq. (6.15) is equal to $L'(h + ZI_{S_t}, S_t) = 1$. Thus $E' \Phi = 1$. \square

Four examples are considered in the remainder of this section. First, let $a \in \mathbb{R}^n$ be a fixed unit vector (i.e., $\|a\| = 1$) and let H_a denote the half space $\{t \in \mathbb{R}^n: (t, a) \geq \alpha\}$. Then the collection $\mathcal{C} = \{H_a \cap T\}$ is a one-parameter family of sets such that \hat{F}^m is vacuous for $m > 1$. That is, two or more points are always \mathcal{C} -dependent. For this choice of \mathcal{C} if $d\mathcal{P}/d\mathcal{P}$ has the form (6.1) then $\gamma_m = 0$ for $m \geq 2$ so that the structure assumption (6.11) is then also satisfied for $Z = \gamma_1$. In this case the likelihood ratio formula given by Eqs. (6.6) and (6.12) reduces to

$$\Lambda(A) = 1 + [\hat{Z}(\cdot) \Lambda(S_\cdot) \circ W]_A, \quad A \in \mathcal{C}$$

and an application of (5.6) yields

$$\Lambda(A) = L(\hat{Z}, A) = \exp\{(\hat{Z} \circ W - \frac{1}{2} \hat{Z}^2 \circ \mu)_A\}, \quad A \in \mathcal{C} \quad (6.16)$$

where

$$\begin{aligned} \hat{Z}(t) &= E'(Z(t) | \mathcal{F}_W(S_t)) \\ &= E'(Z(t) | \mathcal{F}_W(H_{(t,a)} \cap T)) \end{aligned}$$

In this case we see that the likelihood ratio is expressible as an exponential of the conditional mean.

For the second example take $\mathcal{C} = \{\text{all closed sets}\}$ and assume that Eq. (6.11) holds for some bounded \mathcal{C} -adapted process Z . Then for $t = (t_1, t_2, \dots, t_m) \in T^m$,

$$(Z_{t_1}, Z_{t_2}, \dots, Z_{t_m}) \text{ is } \mathcal{F}(S_t)\text{-measurable.} \quad (6.17)$$

By our standing assumptions on the σ -algebras, $\mathcal{F}(S_t)$ is independent of $\mathcal{F}_W(S_t^c)$. But for this example

$$S_t = \{t_1, t_2, \dots, t_m\} \quad (6.18)$$

so that, up to events of \mathcal{P} -measure zero, $\mathcal{F}_W(S_t^c) = \mathcal{F}_W(T)$. Thus the processes Z and W are \mathcal{P} -independent. This implies that Z is identically distributed under \mathcal{P} and \mathcal{P}' . Then, again using (6.18), $\Lambda(S_t) = 1$ \mathcal{P} -almost surely and $\hat{\rho}_m = \rho_m$ where ρ_m is the m -th moment

$$\rho_m(t_1, t_2, \dots, t_m) = E[Z(t_1)Z(t_2)\dots Z(t_m)]. \quad (6.19)$$

Thus, Eq. (6.6) becomes

$$\Lambda(A) = 1 + \sum_{m=1}^{\infty} \frac{1}{m!} (\rho_m \square W^m)(A). \quad (6.20)$$

Equation (6.20) provides a martingale representation of the likelihood ratio for the "additive white Gaussian noise" model under very general conditions. In the one-dimensional case, it was recently obtained in [8].

For the next example let T be the unit interval $[0, 1]$ and, for $k \geq 1$, consider the class of sets

$$\mathcal{C}^k = \{[0, t_1] \cup \{t_2, t_3, \dots, t_k\}: 0 \leq t_1 \leq \dots \leq t_k \leq 1\}.$$

Relative to \mathcal{C}^k , if $0 \leq t_1 \leq t_2 \leq \dots \leq t_m \leq 1$ then

$$S_{t_1, t_2, \dots, t_m} = \begin{cases} \{t_1, t_2, \dots, t_m\} & \text{if } m < k \\ \{t_{m-k+1}, \dots, t_m\} \cup [0, t_{m-k+1}] & \text{if } m \geq k \end{cases} \quad (6.21)$$

and \hat{T}^m is vacuous for $m > k$. Now suppose that for each $k \geq 1$, $d\mathcal{P}/d\mathcal{P}$ has a representation of the form (6.1) relative to $\mathcal{C} = \mathcal{C}^k$. Let r_m^k denote the m^{th} order moment densities when \mathcal{C} is equal to \mathcal{C}^k . Then by (6.21), it is clear that $r_m^k = r_m^j \mathcal{P} \times \mu^m$ a.e. on $[0, 1]^m$ if k and j are each larger than m . Thus, we can define functions r_m by

$$r_m(t_1, t_2, \dots, t_m) = r_m^k(t_1, t_2, \dots, t_m) \text{ for some } k > m$$

and

$$r_m(t_1, t_2, \dots, t_m | [0, t_1]) = r_m^m(t_1, \dots, t_m).$$

Moment densities \hat{r}_m are defined analogously. Thus

$$\hat{r}_m(t_1, t_2, \dots, t_m) = E' [W(dt_1)W(dt_2) \dots W(dt_m)] / dt_1 dt_2 \dots dt_m$$

and

$$\begin{aligned} \hat{r}_m(t_1, t_2, \dots, t_m | [0, t_1]) \\ = E' [W(dt_1)W(dt_2) \dots W(dt_m) | \mathcal{F}_W([0, t_1])] / dt_1 dt_2 \dots dt_m. \end{aligned}$$

Let $A_{t_1} = E \left[\frac{d\mathcal{P}}{d\mathcal{P}} \middle| \mathcal{F}_W([0, t_1]) \right]$ for $t_1 \in T$. Then by (6.21),

$$A(S_{t_1, t_2, \dots, t_m}^k) = \begin{cases} 1 & \text{if } k < m \\ A_{t_{m-k+1}} & \text{if } k \geq m \end{cases} \mathcal{P} \text{ a.s.}$$

if $t_1 < t_2 < \dots < t_m$.

Thus, Eq. (6.6) for the likelihood ratio when $\mathcal{C} = \mathcal{C}^m$ becomes

$$\begin{aligned} A_t = 1 + \int_0^t \hat{r}_1(t_1) W(dt_1) + \int_0^t \int_0^{t_2} \hat{r}_2(t_1, t_2) W(dt_1) W(dt_2) + \dots \\ + \int_0^t \int_0^{t_{m-1}} \dots \int_0^{t_2} \hat{r}_{m-1}(t_1, \dots, t_{m-1}) W(dt_1) \dots W(dt_{m-1}) \\ + \int_0^t \int_0^{t_m} \dots \int_0^{t_2} A_{t_1} \hat{r}_m(t_1, \dots, t_m | [0, t_1]) W(dt_1) W(dt_2) \dots W(dt_m) \end{aligned} \quad (6.22)$$

and the moment Eq. (6.10) when $\mathcal{C} = \mathcal{C}_m$ and $\mathcal{C} = \mathcal{C}_{m+1}$ becomes

$$A_{t_1} \hat{r}_m(t_1, \dots, t_m | [0, t_1]) = \hat{r}_m(t_1, \dots, t_m) + \int_0^t A_\tau \hat{r}_{m+1}(\tau, t_1, \dots, t_m | [0, \tau]) W(d\tau). \quad (6.23)$$

Now suppose that $d\mathcal{P}/d\mathcal{P}$ satisfies (6.11) where $\{Z_t; t \in T\}$ is bounded and \mathcal{C}^k -adapted for all $k \geq 1$. By reasoning similar to that in the second example, the processes Z and W must be independent under \mathcal{P} . Thus

$$\hat{r}_m(t_1, \dots, t_m | [0, t_1]) = E' [Z_{t_1} Z_{t_2} \dots Z_{t_m} | \mathcal{F}_W([0, t_1])]$$

and

$$r_m(t_1, \dots, t_m) = \rho_m(t_1, \dots, t_m)$$

where ρ_m is the m -th moment defined in Eq. (6.19). Defining $\hat{Z}_t = E[Z_t | \mathcal{F}_W([0, t])]$, Eqs. (6.22) and (6.23) for $m=1$ become

$$A_t = 1 + \int_0^t A_\tau \hat{Z}_\tau W(d\tau) \quad (6.24)$$

and

$$A_t \hat{Z}_t = E[Z_t] + \int_0^t A_\tau E[Z_t Z_\tau | \mathcal{F}_W([0, \tau])] W(d\tau) \quad (6.25)$$

respectively.

Note that (6.25) is a well-known representation for the first "unnormalized" conditional moment $\hat{Z}_t A_t$. Using Ito's formula for one-parameter process, Eq. (6.24) and (6.25) together yield a representation for the conditional moment $\hat{Z}_t = (A_t \hat{Z}_t) / A_t$ itself. As an intermediate step, an integral representation for A_t^{-1} would also be derived via Ito's formula. In contrast, in the setting of a general class of sets \mathcal{C} , a suitable analogue of Ito's formula is not yet available. Such a formula might provide, for example, an analogue of the moment Eq. (6.10) for \hat{r}_k which does not contain the factor $1/L(\hat{S}_t)$. (An instance of such equations is provided in [10].)

For the final example suppose that T is a rectangle in \mathbb{R}^n with $\mu(T)=1$, suppose that $\{W(A), \mathcal{F}(A): A \in \mathcal{R}(T)\}$ is as in Sect. 2, and consider \mathcal{P} defined by

$$\frac{d\mathcal{P}}{d\mathcal{P}} = W(T)^2.$$

In terms of multiple Wiener integrals,

$$\frac{d\mathcal{P}}{d\mathcal{P}} = 1 + 1 \square W^2$$

so that the moment densities r_m corresponding to $\mathcal{C} = \{\text{all closed sets}\}$ are given by $r_m \equiv \delta_{m2}$. Also,

$$L(A) = 1 + [W(A)^2 - \mu(A)].$$

Now let \mathcal{C} be a class of sets satisfying the conditions c_1 - c_3 . Then by the moment Eq. (6.10), the moment densities \hat{r}_m satisfy

$$\begin{aligned} \hat{r}_1(t) &= W(\hat{S}_t) / L(\hat{S}_t) \\ \hat{r}_2(t_1, t_2) &= 1 / L(\hat{S}_t) \end{aligned}$$

and $\hat{r}_k \equiv 0$ for $k \geq 2$. Thus, in terms of \mathcal{C} -stochastic integrals,

$$\frac{d\mathcal{P}}{d\mathcal{P}} = L(\hat{S}) + (L(\hat{S} \cdot) \hat{r}_1(\cdot)) \diamond W + (L(\hat{S} \cdot) \hat{r}_2(\cdot)) \diamond W^2.$$

Appendix A

Proposition 4.1 is proved in this appendix. For the proof fix $t \in T^k$ and let $\mathcal{C} = \mathcal{C}_t$. Clearly for subsets A_1, A_2, \dots, A_m of T ,

$$\hat{S}_{A_1 A_2 \dots A_m} = S_{A_1 A_2 \dots A_m t_1 t_2 \dots t_k},$$

so that conditions c_2 and c_3 for \mathcal{C} are immediate consequences of these conditions for \mathcal{C} . It remains to establish condition c_1 for \mathcal{C} .

To begin we will establish that

$$\hat{T}^m \subset \bigcup_{\tau} \{s \in T^m: s \times \tau \in \hat{T}^{m+j} \text{ and } t \in (S_{s \times \tau})^k\} \quad (\text{A.1})$$

where the union is over all $\tau = (\tau_1, \tau_2, \dots, \tau_j)$ such that $\{\tau_1, \tau_2, \dots, \tau_j\} \subset \{t_1, t_2, \dots, t_k\}$ and $s \times \tau$ denotes the point $(s_1, s_2, \dots, s_m, \tau_1, \tau_2, \dots, \tau_k)$ in T^{m+k} . To see (A.1) suppose that $s \in \hat{T}^m$. Then s is contained in any of the sets on the right side of relation (A.1) which correspond to a minimal subset $\{\tau_1, \tau_2, \dots, \tau_j\}$ with satisfies $S_{s \times \tau} = S_{s \times t}$. This establishes relation (A.1).

Now by relation (A.1) if $\varepsilon > 0$ then

$$\hat{T}^m \subset \bigcup_{\tau} \{s \in T^m: s \times \tau \in \hat{T}^{m+j} \text{ and } d(t_i, S_{s \times \tau}) < \varepsilon \text{ for } 1 \leq i \leq k\} \quad (\text{A.2})$$

where d denotes the usual Euclidean distance between subsets of T , and the union is over τ as in (A.1).

Now choose any product of rectangles $A = A_1 \times \dots \times A_m$ so that $A \in \hat{T}^m$. Since each of the sets in the union on the right side of relation (A.2) is open (by condition c_1), the set A can be expressed as a countable union of sets of the form $B = B_1 \times \dots \times B_m$ where B_1, B_2, \dots, B_m are rectangles such that

$$B \subset \{s \in T^m: s \times \tau \in \hat{T}^{m+j} \text{ and } d(t_i, S_{s \times \tau}) < \varepsilon \text{ for } 1 \leq i \leq k\} \quad (\text{A.3})$$

for some τ as in relation (A.2), and where τ depends on B . Condition c_1 for \mathcal{C} applied to such B implies that

$$\mu(B_i \cap S_{B \times \tau}) = 0 \quad \text{for } 1 \leq i \leq m. \quad (\text{A.4})$$

By relation (A.3) it is clear that $S_{B \times \tau} \supset S_{B \times t}^{\varepsilon}$, and since $B \subset A$ we also have that $S_{B \times t}^{\varepsilon} \supset S_{A \times t}^{\varepsilon}$. Thus (A.4) implies that $\mu(B_i \cap S_{A \times t}^{\varepsilon}) = 0$ for $1 \leq i \leq m$, and since A_i is a countable union of such sets B_i ,

$$\mu(A_i \cap S_{A \times t}^{\varepsilon}) = 0 \quad \text{for } 1 \leq i \leq m.$$

Now sending ε to zero and applying condition c_3 for \mathcal{C} shows that $\mu(A_i \cap S_{A \times t}) = 0$. Finally, since $S_{A \times t} = \hat{S}_A$, this establishes condition c_1 for $\mathcal{C} = \mathcal{C}_t$. \square

Appendix B

Let \mathcal{C} and \mathcal{C}' with $\mathcal{C} \supset \mathcal{C}'$ each satisfy conditions c_1 - c_3 . Let \mathcal{J}_m denote the collection of subsets of T^m of the form $A_1 \times \dots \times A_m$ such that each $A_i \in \mathcal{R}^m(T)$ and for some permutation Π ,

- 1) $A_{\Pi(1)}, \dots, A_{\Pi(k)}$ are \mathcal{C} -independent, and
- 2) $A_{\Pi(k+1)}, \dots, A_{\Pi(m)} \subset S_{A_{\Pi(1)}A_{\Pi(2)}\dots A_{\Pi(k)}}$.

Define \mathcal{J}_m relative to \mathcal{C} similarly. The purpose of this appendix is to prove the following proposition:

Proposition B. *The linear span of $\{\alpha I_B: B \in \hat{I}_m, B \in \hat{T}^m, \alpha \text{ is bounded, } \mathcal{F}(S_{A_1 A_2 \dots A_m}) \text{ meas.}\}$ is dense in $L^2_\alpha(\Omega \times \hat{T}^m)$ for each $m \geq 1$.*

Proof. Consider the following two conditions:

- (b₁) There is a countable subcollection of \mathcal{J}_m which covers T^m a.e.
- (b₂) There is a countable subcollection \mathcal{J}_m^d of disjoint sets in \mathcal{J}_m which covers T^m a.e.

By a sequence of lemmas it is shown below that condition b₁ is satisfied and then that condition b₁ \Rightarrow condition b₂ and finally that condition b₂ (but with \mathcal{J}_m^d replaced by \mathcal{J}_m^d) implies Proposition B. \square

Lemma B.1.

$$\bigcup_{l=1}^m \bigcup_{\Pi \in \mathcal{P}(m)} \Pi \circ \{(x, y): x \in \hat{T}^l, y \in (S_x)^{m-l}\} = T^m \quad (*)$$

Proof. Let $\mathbf{q} = (q_1, \dots, q_m) \in T^m$. Choose a permutation $\mathbf{p} = (p_1, \dots, p_m) = \Pi(q_1, \dots, q_m)$ so that for some l with $1 \leq l \leq m$,

$$S_{\mathbf{q}} = S_{p_1, \dots, p_l} \neq S_{p_1, \dots, \hat{p}_i, \dots, p_l} \quad \text{for } 1 \leq i \leq l$$

where " \hat{p}_i " denotes that p_i is to be omitted. That is, the permutation is chosen so that p_1, \dots, p_l is a minimal set from q_1, \dots, q_m with the same support as q_1, \dots, q_m . Now $p_{l+1}, \dots, p_m \in S_{p_1, \dots, p_l}$ since $q_1, \dots, q_m \in S_{\mathbf{q}} = S_{p_1, \dots, p_l}$.

To show that \mathbf{q} is contained in the left side of (*), it remains to show that p_1, \dots, p_l are \mathcal{C} -independent. Now, if p_1, \dots, p_l were not \mathcal{C} -independent, then $p_i \in S_{p_1, \dots, \hat{p}_i, \dots, p_l}$ for some i . Then

$$\{A \in \mathcal{C}: p_1, \dots, p_l \in A\} = \{A \in \mathcal{C}: p_1, \dots, \hat{p}_i, \dots, p_l \in A\}.$$

Intersecting all the sets contained in this collection of sets yields that

$$S_{p_1, \dots, p_l} = S_{p_1, \dots, \hat{p}_i, \dots, p_l}$$

which contradicts our choice of p_1, \dots, p_l . Thus p_1, \dots, p_l are \mathcal{C} -independent so that \mathbf{p} , and hence \mathbf{q} , is contained in the left side of (*). \square

Lemma B.2. Condition b₁ is satisfied.

Proof. Let \mathcal{J}_m^0 denote the subsets of T^m of the form $A_1 \times \dots \times A_m$ such that, for some $\Pi \in \mathcal{P}(m)$ and some $l > 0$,

a) $A_{\Pi(1)}, \dots, A_{\Pi(l)}$ are \mathcal{C} -independent, closed rectangles whose vertices have rational coordinates in $T \subset \mathbb{R}^n$, and

b) $A_{\Pi(l+1)} = \dots = A_{\Pi(m)} = S_{A_{\Pi(1)}A_{\Pi(2)}\dots A_{\Pi(l)}}$.

Then \mathcal{J}_m^0 is a countable subset of \mathcal{J}_m and

$$\begin{aligned}
\bigcup_{A \in \mathcal{F}_m} A &\supset \bigcup_{l=1}^m \bigcup_{\Pi \in \mathcal{P}(m)} \Pi \circ \{(x, y): x \in \hat{T}^l, y \in (S_x^{(-)})^{m-l}\} \\
&= \bigcup_{l=1}^m \bigcup_{\Pi \in \mathcal{P}(m)} \Pi \circ \{(x, y): x \in \hat{T}^l, y \in (S_x)^{m-l}\} \\
&\quad - \bigcup_{l=1}^m \bigcup_{\Pi \in \mathcal{P}(m)} \Pi \circ S_{m,l}
\end{aligned} \tag{B.1}$$

where

$$S_{m,l} = \{(x, y): x \in \hat{T}^l, y \in (S_x)^{m-l} - (S_x^{(-)})^{m-l}\}.$$

The first term on the right side of (B.1) is equal to T^m by Lemma B.1. Thus, to complete the proof it must be shown that $\mu^m(S_{m,l})=0$ for all $m \geq 1$ and $1 \leq l \leq m$.

By condition c_2 ,

$$F_\varepsilon = \{(x, y): x \in \hat{T}^l, y \in (S_x^\varepsilon)^{m-l}\}$$

is a closed subset of $\hat{T}^l \times T^{m-l}$ which increases as ε decreases to zero. Since $S_{m,l} = F_0 - \bigcup_{\varepsilon > 0} F_\varepsilon$, it follows that $S_{m,l}$ is a Borel subset of T^m . By condition c_3 , the section

$$\{y: (x, y) \in S_{m,l}\} \subset T^{m-l}$$

of $S_{m,l}$ at x has Lebesgue measure zero for a.e. $y \in \hat{T}^m$. Hence, by Fubini's theorem, $\mu^m(S_{m,l})=0$ for $1 \leq l \leq m$. \square

Lemma B.3. Condition b_1 implies condition b_2 .

Proof. Let F_1, F_2, \dots be a countable subcollection of \mathcal{F}_m which covers T^m a.e. Then the disjoint sets $D_i = F_i - \bigcup_{j=1}^{i-1} F_j$, $i \geq 1$ cover T^m a.e. We claim that for each $i \geq 1$ there is a finite collection of disjoint sets D_{i1}, \dots, D_{in} in \mathcal{F}_m such that $D_i = \bigcup_{j=1}^{n_i} D_{ij}$. Condition b_2 is then satisfied with $\mathcal{F}_m^d = \{D_{ij}: i \geq 1, 1 \leq j \leq n_i\}$. It remains to prove the claim.

By induction, it suffices to establish the claim for $i=2$. Now $F_1 = A_1 \times \dots \times A_m$ for some Borel sets $A_1, \dots, A_m \subset T$. Thus, $F_1^c = \bigcup_{j=1}^r K_j$ where K_1, \dots, K_r are disjoint and each K_j is the product of m Borel subsets of T . In fact, F_1^c is the union of all sets of the form $B_1 \times \dots \times B_m$ such that $B_i = A_i$ or $B_i = A_i^c$ for each i and such that $B_i = A_i^c$ for at least one i , and these sets are disjoint. So $D_2 = \bigcup_{j=1}^k K_j \cap F_2$. The sets $K_j \cap F_2$ are disjoint sets in \mathcal{F}_m as required so the claim is established. \square

Lemma B.4. Condition b_2 with \mathcal{F}_m^d replaced by \mathcal{F}_m^d implies that the linear span of $\{\alpha 1_B: B \in \mathcal{F}_m, B \subset \hat{T}^m, \alpha \text{ is bounded, } \mathcal{F}(S_{A_1, \dots, A_m}) \text{ measurable}\}$ is dense in $L_a^2(\Omega \times \hat{T}^m)$.

Proof. Let $F = F_1 \times \dots \times F_m$ where each $F_i \in \mathcal{R}^n(T)$ such that $F \in \hat{T}^m$ and let α be a bounded, $\mathcal{F}(S_{F_1, F_2, \dots, F_m})$ -measurable random variable. If $A = A_1 \times \dots \times A_m \in \mathcal{F}_m^d$ then $B = A \cap F$ satisfies $B \in \mathcal{F}_m$ and $B \subset \hat{T}^m$, and α is $\mathcal{F}(S_{A_1 \cap F_1, \dots, A_m \cap F_m})$ -

measurable since \mathcal{G} -supports are no smaller than \mathcal{C} -supports. By condition b_2 with \mathcal{F}_m^d replaced by \mathcal{F}_m^c ,

$$\alpha 1_F = \sum_{A \in \mathcal{F}_m^d} \alpha 1_{A \cap F} \quad \text{a.e. in } T^m.$$

Since the linear span of functions of the form $\alpha 1_F$ is dense in $L^2_a(\Omega \times \hat{T}^m)$, the lemma is established by considering sets of the form $B = A \cap F$. \square

References

1. Cameron, R.H., Martin, W.T.: The orthogonal development of nonlinear functionals in a series of Fourier-Hermite functions. *Ann. Math.* **48**, 385-392 (1947)
2. Duncan, T.E.: Likelihood functions for stochastic signals in white noise. *Inform. and Contr.* **16**, 303-310 (1970)
3. Hajek, B.: Stochastic Integration, Markov Property and Measure Transformation of Random Fields. Ph.D. dissertation, Berkeley, 1979
4. Hajek, B., Wong, E.: Set parametered martingales and multiple stochastic integration. In *Stochastic Integrals*, D. Williams ed. New York: Springer 1981
5. Ito, K.: Stochastic integrals. *Proc. Imp. Acad. Tokyo* **20**, 519-524 (1944)
6. Ito, K.: Multiple Wiener Integral. *J. Math. Soc. Japan* **3**, 157-169 (1951)
7. Kailath, T.: A general likelihood-ratio formula for random signals in Gaussian noise. *IEEE Trans. Inform. Th.* **15**, 350-361 (1969)
8. Mitter, S.K., Ocone, D.: Multiple integral expansion for nonlinear filtering. *Proc. 18th IEEE Conference on Decision and Control* (1979)
9. Wong, E., Zakai, M.: Martingales and stochastic integrals for processes with a multi-dimensional parameter. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **29**, 109-122 (1974)
10. Wong, E., Zakai, M.: Likelihood ratios and transformation of probability associated with two-parameter Wiener processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **40**, 283-309 (1977)
11. Yor, M.: Representation des martingales de carré integrable relative aux processus de Wiener et de Poisson à n paramètres. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **35**, 121-129 (1976)

Received March 8, 1982; in revised form November 13, 1982