MULTIPLE STOCHASTIC INTEGRALS: PROJECTION AND ITERATION

by

Bruce Hajek and Eugene Wong

Memorandum No. UCB/ERL M81/100

8 October 1981

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

MULTIPLE STOCHASTIC INTEGRALS:

PROJECTION AND ITERATION

Bruce Hajek and Eugene Wong

Coordinated Science Laboratory
University of Illinois at Urbana

and

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

ABSTRACT

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.

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

1. Introduction

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

$$EW(A)W(B) = \mu(A \cap B) \tag{1.1}$$

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

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

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 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 $\mathcal{R}(T)$ = {Borel subsets of T}, and defined on a fixed measurable space $(\Omega, \mathcal{F})$. 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

$$\Lambda = \frac{dP'_W}{dP_W} \tag{1.3}$$

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 it 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 <u>martingales</u> 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 $A \supseteq B \Rightarrow \mathcal{F}(A) \supseteq \mathcal{F}(B)$, 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 <u>martingale</u> if

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

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

$$F_W(A) = \sigma(\{W(B),\ B \subset A \text{ and } B \in R(T)\}) \tag{1.4}$$

One of the main objects of this paper is to show that under very general conditions on C, there is a canonical representation of all square-integrable martingales with respect to $\{F_W(A),\ A \in C\}$, and hence representation for square integrable Wiener functionals. For C = {all closed sets} the representation reduces to that of multiple Wiener integrals. For C = {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 C, giving rise to an equally large array of representations for C-martingales and Wiener functionals.

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

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

where ϕ is (in general) a random integrand 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 $L^2(\Omega,F_W,P)$ (equivalently, this is a formula for the conditional expectation of a multiple stochastic integral given F_W.) The second is an iterated integral formula for expressing multiple stochastic integrals defined relative to C in terms of stochastic integrals defined relative to another class of sets C'.

Finally, the transformation formulae are applied to the signal detection problem noted above. The projection formula is relevant since the liklihood ratio is the projection of the Radon-Nikodym derivative $\frac{dP'}{dP}$, 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 ommission of two technical proofs for which the reader is referred to [4].

2. Multiple Stochastic Integrals

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

$$S_{A_1 A_2 \ldots A_m} = \cap\{B : B \in C \text{ and } B \cap A_i \neq \emptyset \text{ for } 1 \leq i \leq m\} \qquad (2.1)$$

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 C and is denoted by S. (Note that $S=S_T$). It will be assumed that the support of any collection of sets $A_1, \ldots, A_m \in R(T)$ is contained in C. This assumption can be met by enlarging a given collection of sets C.

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

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

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

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

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

For this example, m points are unordered if and only if they are pair-wise unordered.

Let $\hat{T}^m$ denote the subset of C-independent points in T^m. For a given collection C, $\hat{T}^m$ may be vacuous for sufficiently large m. For example, if $C = \{R_t\}$ 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 C-independent. In the extreme case $C = \{T\}$, $\hat{T}^m$ 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 C of $A_1, A_2, \ldots, A_m \in \mathcal{R}^n(T)$ by

$$S^{\varepsilon}_{A_1 A_2 \ldots A_m} = S_{B(\varepsilon, A_1) B(\varepsilon, A_2) \ldots B(\varepsilon, A_m)}.$$

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

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

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

(c_1) For every collection of rectangles $A_1, A_2, .., A_m$ such that $\prod_{i=1}^{m} A_i \subset \hat{T}^m$,

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

(c_2) For each $m \geq 1$, the mapping

$$t = (t_1, t_2, .., t_m) \rightsquigarrow S_t$$

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

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

(c_3) For every collection of rectangles $A_1, A_2, .., A_m$ in T,

$$\bigvee_{\varepsilon > 0} F(S^{\varepsilon}_{A_1 A_2 .. A_m}) = F(S_{A_1 A_2 .. A_m}) \quad .$$

Since $F_W(A) \subset F(A)$ for all A in $R(T)$, condition (c_3) implies the following condition:

(c_3') For every collection of rectangles $A_1, A_2, .., A_m$ in T,

$$\mu(S_{A_1 A_2 .. A_m} - S^{(-)}_{A_1 A_2 .. A_m}) = 0 \quad , \tag{2.3}$$

and if $F_W(A) = F(A)$ for all A then conditions (c_3) and (c_3') are equivalent.

Condition (c_3), as well as condition (c_2), is a continuity condition. Note that since the sets in C are closed, condition (c_2) insures that $\hat{T}^m$ is an open subset of T^m in the relative topology on T^m.

For a C satisfying conditions $c_1 - c_3$, we shall define multiple stochastic integrals of order m

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

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

(h_1) ϕ is $F \times \mu^m$-measurable

(h_2) For each $t \in \hat{T}^m$ ϕ_t is $F(S_t)$-measurable.

(h_3) $\int_{\hat{T}^m} E\phi_t^2 \, dt < \infty$

The space of functions satisfying $h_1 - h_3$ will be denoted by $L_a^2(\Omega x\hat{T}^m)$.

Call ϕ <u>atomic</u> 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 $A \subset \hat{T}^m$. Two atomic functions

$$\begin{aligned} \phi(\omega,t) &= \alpha(\omega) \, I_A(t) \quad , A \subset \hat{T}^m \\ \theta(\omega,t) &= \beta(\omega) \, I_B(t) \quad , B \subset \hat{T}^p \end{aligned} \tag{2.4}$$

are said to be <u>comparable</u> if each pair (A_i,B_j) is either equal or disjoint module sets of zero Lebesgue measure, and <u>similar</u> if $m = p$ and $(B_1,B_2,\ldots,B_m)$ is a permutation of $(A_1,A_2,\ldots,A_m)$. Call ϕ <u>simple</u> if $\phi = \sum_{k=1}^{K} \phi_k$ and each ϕ_k is atomic.

For an atomic function ϕ define

$$\phi \circ W^m = \alpha \prod_{i=1}^{m} W(A_i) \tag{2.5}$$

So defined, $\phi \circ W^m$ has the following property:

Lemma 2.1. Let ϕ and θ be comparable atomic functions in $L_a^2(\Omega x \hat{T}^m)$ and $L_a^2(\Omega x \hat{T}^p)$ of the form (2.4). Then

$$E(\phi \circ W^m)\ (\theta \circ W^p) = 0 \tag{2.6}$$

unless ϕ and θ are similar. In the latter case,

$$E(\phi \circ W^m)\ (\theta \circ W^m) = \int_{\hat{T}^m} E\tilde{\phi}_t \tilde{\theta}_t\ dt \overset{\text{def.}}{=} \langle \tilde{\phi}, \tilde{\theta} \rangle \tag{2.7}$$

where $\tilde{\phi}$ denotes the symmetrization of ϕ, i.e.,

$$\tilde{\phi}_t = \frac{1}{m!} \sum_{\Pi} \phi_{\Pi(t)} \quad , \ \Pi(t) = \text{permutation of } t \tag{2.8}$$

Proof: First, assume ϕ and θ to be similar. Then

$$(\phi \circ W^m)\ (\theta \circ W^m) = \alpha\beta \prod_{i=1}^{m} W^2(A_i)$$

and $\alpha\beta$ is measurable with respect to $F(S_{A_1 A_2 \ldots A_m})$. Therefore, condition c_1 implies that

$$E[(\phi \circ W^m)(\theta \circ W^m) | F(S_{A_1 A_2 \ldots A_m})]$$

$$= \alpha\beta \prod_{i=1}^{m} E\ W^2(A_i)$$

$$= \alpha\beta \prod_{i=1}^{m} \mu(A_i)$$

and (2.7) follows.

Next, suppose that ϕ and θ are comparable but not similar. With no loss of generality assume $m \geq p$. Consider two possibilities:

(a) There exists a B_j (say B_1) such that

$$B_1 \cap [\bigcup_{i=1}^{m} A_i \cup S_{A_1A_2..A_m}] = \emptyset$$

(b) For every $j \leq p$

$$B_j \cap [\bigcup_{i=1}^{m} A_i \cup S_{A_1A_2..A_m}] \neq \emptyset$$

For case (a), let

$$D = \bigcup_{i=1}^{m} A_i \bigcup_{j=2}^{p} B_j \cup S_{A_1A_2..A_m} \cup S_{B_1B_2...B_p}$$

Then, with probability 1

$$E[\phi\circ W^m)(\theta\circ W^p)|F(D)] = \alpha\beta \prod_{i=1}^{m} W(A_i) \prod_{j=2}^{p} W(B_j)[EW(B_1)] = 0$$

and (2.6) is verified.

For case (b) we shall prove that $S_{A_1A_2..A_m} \supset S_{B_1B_2..B_p}$. Since ϕ and θ are comparable but not similar and $m \geq p$, there must exist an A_i(say A_1) such that $\mu(A_1 \cap B_j) = 0$ for every j. Hence, $W(A_1)$ is independent of $\alpha\beta \prod_{i=2}^{m} W(A_i) \prod_{j=1}^{p} W(B_j)$ and (2.6) is again proved.

To prove $S_{A_1A_2..A_m} \supset S_{B_1B_2..B_p}$ for case (b), let $D \in C$ be any set such that

$D \cap A_i \neq \phi$ for every i

then, $D \supset S_{A_1A_2..A_m}$ by definition. The defining condition for case (b) implies that for each j

either $B_j \cap \bigcup_i A_i \neq \phi$

which implies $B_j = A_i$ for some i

which in turn implies $D \cap B_j \neq \phi$

or $B_j \cap S_{A_1A_2..A_m} \neq \phi$

which implies $D \cap B_j \neq \phi$.

Therefore,

$$D \cap A_i \neq \phi \text{ for every } i \Rightarrow D \cap B_j \neq \phi \text{ for every } j$$

and thus indeed $S_{A_1A_2..A_m} \supset S_{B_1B_2..B_p}$. ¤

Lemma 2.2. For atomic functions ϕ and θ that are not necessarily comparable, we can write

$$\phi = \sum_{k=1}^{K} \phi_k \tag{2.9}$$

$$\theta = \sum_{\lambda=1}^{L} \theta_\lambda$$

where ϕ_k, ϕ_λ are atomic and the set $\{\phi_k, \theta_\lambda\}$ is pairwise comparable. For any atomic ϕ and θ in L_a^2 the isometry

$$E(\phi \circ W^m)\,(\theta \circ W^p) = \delta_{mp} \langle \tilde{\phi}, \tilde{\theta} \rangle \tag{2.10}$$

holds.

Proof: ϕ and θ, being atomic, are of the form

$$\phi = \alpha\, I_{A_1 x A_2 x..x A_m}$$

$$\theta = \beta\, I_{B_1 x B_2 x..x B_p} .$$

where $A_1, A_2, .., A_m, B_1, .., B_p$ are rectangles in T. Since a union of rectangles is always a union of disjoint rectangles, there exist disjoint rectangles $D_1, D_2, .., D_q$ such that each A_i or B_j is the union of some of the D_ν's. Hence (2.9) follows, with

$$\phi_k = \alpha\, I_{D_{k1} x D_{k2} x \ldots x D_{km}}$$

$$\theta_\lambda = \beta\, I_{D_{\lambda 1} x D_{\lambda 2} x \ldots x D_{\lambda p}}$$

where $D_{ki} \subset A_i$ and $D_{\lambda j} \subset B_j$ for every i and j. It follows that α is $\mathcal{F}(S_{D_{k1} D_{k2} \ldots D_{km}})$-measurable and β is $\mathcal{F}(S_{D_{\lambda 1} D_{\lambda 2} \ldots D_{\lambda m}})$-measurable for each k and λ. From lemma 2.1 we have

$$E(\phi_k \circ W^m)\ (\phi_\lambda \circ W^p) = \delta_{mp} \langle \tilde{\phi}_k, \tilde{\theta}_\lambda \rangle$$

and (2.10) follows from the bilinearity of $\langle \quad \rangle$. ¤

<u>Lemma 2.3</u>. Under conditions c_2 and c_3 the subset of simple functions is dense in $L_a^2(\Omega x \hat{T}^m)$.

A proof of this result is given in appendix A of [4].

<u>Theorem 2.1</u>. There is a unique linear map denoted by $\phi \circ W^m$ of $\phi \in L_a^2(\Omega x \hat{T}^m)$ 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:

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

Proof: First, any simple function ϕ is by definition of the form $\phi = \sum_{k=1}^{K} \phi_k$, where ϕ_k are atomic. Bilinearity of $\langle \quad \rangle$ then implies the isometry property (2.10) for simple functions ϕ and θ. Let ϕ be any function from $L_a^2(\Omega x\hat{T}^m)$. Lemma 2.3 implies that there exists a sequence $\{\phi^{(n)}\}$ of simple functions such that

$$\phi^{(n)} \xrightarrow[n\to\infty]{L_a^2} \phi$$

Hence, $\{\phi^{(n)}\}$ is Cauchy. The isometry property (2.10) then implies that $\{\phi^{(n)}\circ W^m\}$ is mean-square convergent as a sequence of random variables, and we take the limit to be $\phi \circ W^m$. Verification of the properties follow from the isometric property in a straightforward way. ¤

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

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

$$E(\phi\circ W^m | F(B)) = E(\alpha \prod_{i=1}^{m} W(A_i) | F(B))$$

$$= E(\alpha E[\prod_{i=1}^{m} W(A_i) | F(B \cup S_{A_1\, A_2 .. A_m})] | F(B))$$

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

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

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

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

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

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, $F(S_t)$-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

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

is an orthogonal collection of random variables which are also orthogonal to the $F(S)$-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 Section 5 formally means that the collection of variables in (3.1) together with the $F(S)$-measurable variables are complete in $L^2(\Omega, F_W(T), P)$ if $F(A) = F_W(A)$ for all A in $R(T)$.

<u>Proposition 3.1</u>. Let $\gamma \in L_a^2(\Omega x \hat{T}^m)$. Then for $t \in \hat{T}^k$,

$$E[W(dt_1)W(dt_2)..W(dt_k)\ \gamma \circ W^m | F(S_t)]/dt_1 dt_2..dt_k = m!\tilde{\gamma}(t)\delta_{mk} \tag{3.2}$$

in the sense that the linear functional

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

defines a symmetric finite signed measure on the σ-algebra of subsets of $\Omega \times \hat{T}^k$ generated by C-adapted atomic functions, the measure is absolutely continuous with respect to $P \times \mu^k$ 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. ¤

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

<u>Proposition 3.2</u>. (Projection formula) For each $\gamma \in L_a^2(\Omega x \hat{T}^m)$ there is a $\hat{\gamma} \in L_a^2(\Omega x \hat{T}^m, F_W(\cdot))$ such that

$$\hat{\gamma}(t) = E[\tilde{\gamma}(t) | F_W(S_t)] \text{ for a.e. } t \in \hat{T}^m \tag{3.3}$$

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

$$E[\gamma \circ W^m | F_W(A)] = (\hat{\gamma}_m \circ W^m)(A) \tag{3.4}$$

Proof: By the completeness of multiple stochastic integrals in $L^2(\Omega, F_W(T), P)$ (see Proposition 5.1 below) and the fact that

$E[\gamma \circ W^m | F_W(S)] = 0$, there exists a collection $\{\phi_k : k \geq 1\}$ with $\phi_k \in L_a^2(\Omega x \hat{T}^m, F_W(\cdot))$ such that

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

Now by Proposition 3.1 with F replaced by F_W,

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

so that

$$E[W(dt_1)W(dt_2)..W(dt_k)\gamma \circ W^m | F_W(S_t)]/dt_1 dt_2..dt_k = k!\tilde{\phi}_k(t) \qquad (3.5)$$

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

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

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 equation (3.3) and equation (3.4) is true for $A = T$. Since each side of equation (3.4) is martingale relative to $\{F_W(A) : A \in C\}$, (3.4) is thus true for all $A \in C$. Finally, since $\hat{\gamma}$ is uniquely determined on $\hat{T}^m$ up to a set of $P \times \mu^m$ measure zero by equation (3.3), any $\hat{\gamma} \in L_a^2(\Omega x \hat{T}^m, F_W(\cdot))$ satisfying (3.3) also satisfies equation (3.4). ¤

4. Nested Classes of Sets C and the Iterated Integration Formula

Let C and $\dot{C}$ with $C \supset \dot{C}$ be two classes of sets which each satisfy conditions $c_1 - c_3$ for a Wiener process $\{W(A) : A \in R(T)\}$ and a collection of σ-algebras $\{F(A) : A \in R(T)\}$ as in Section 2. A dot above (or above and to the right) denotes definition relative to $\dot{C}$ so that, for example, $\dot{S}_t$ denotes the $\dot{C}$-support of t and $\hat{T}^{m\cdot}$ denotes the collection of $\dot{C}$-independent points in T^m.

An important example which is exploited in the next section is when $\dot{C}$ is any class satisfying conditions $c_1 - c_3$ with $F(A) = F_W(A)$ for all $A \in R(T)$, and C is the collection of all closed sets. Another natural way in which nested collections of sets C arise is given by the following propositions.

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

$$C_t = \{C \in C_t : \{t_1, t_2, .., t_k\} \subset C\} \quad .$$

Then C_t also satisfies conditions $c_1 - c_3$.

Proof: See Appendix A.

Theorem 4.2. (Iterated integration formula) Suppose that C and $\dot{C}$ each satisfy conditions $c_1 - c_3$ and that $\dot{C} \subset C$. Then for $\theta \in L_a^2(\Omega x \hat{T}^m)$ the class-C stochastic integral $\theta \circ W^m$ can be represented as a sum of class-C integrals:

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

where the integrands $\phi_k \in \dot{L}_a^2(\Omega x \hat{T}^{k\cdot})$ satisfy

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

For each fixed t the integral on the right side of equation (4.2) is defined relative to the collection of sets C_t.

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

$$\Pi\theta = \alpha I_{A_1 x .. x A_m}$$

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

$$\theta \circ W^m = \Pi\theta \circ W^m = [\alpha \prod_{i=k+1}^{m} W(A_i)] \prod_{i=1}^{k} W(A_i)$$

$$= h_k \overset{\cdot}{\circ} W^k$$

where for $t \in \hat{T}^{k\cdot}$

$$h_k(t) = I_{A_1 x..A_k}(t)[\alpha I_{A_{k+1} x..x A_m} \overset{t}{\circ} W^{m-k}]$$

$$= [\Pi\theta(tx\cdot) I_{\dot{S}_t^{m-k}}] \overset{t}{\circ} W^{m-k} .$$

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

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

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

$$\theta \circ W^m = h_k \dot{\circ} W^k = g_k \dot{\circ} W^k$$

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

$$\sum_{\text{all } \Pi} [(\Pi\theta)(tx\cdot) I_{\dot{S}_t^{m-k}}(\cdot) \overset{t}{\circ} W^{m-k}] \dot{\circ} W^k$$

$$= k!(m-k)!\ \theta \circ W^m$$

$$= m!\ \phi_k \dot{\circ} W^k$$

where ϕ_k is given by equation (4.2). Hence

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

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

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 C consists of all closed subsets of T and $F(A) = F_W(A)$ for all $A \in R(T)$. The associated integrals are then multiple Wiener integrals. Let $\tilde{T}^m$ denote the set of m-tuples of distinct points in T and for θ in $L^2(\tilde{T}^m)$ let $\theta ¤ W^m$ denote a multiple Wiener integral of order m.

Proposition 5.1. (Completeness of multiple stochastic integrals)

Let C be a collection of sets such that C and $\{F_W(A)\}$ satisfy conditions $c_1 - c_3$. Then every square-integrable $F_W(T)$-measurable random variable Z has a representation of the form

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

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

Proof: The proposition is well known [6] in case 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 C, Proposition 5.1 is true in general. ¤

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

$$\overset{*}{f}{}^{m}(t_1,t_2,\ldots,t_m) = \prod_{i=1}^{m} f(t_i) \tag{5.2}$$

and set

$$W_m(f,A) = (\overset{*}{f}{}^{m} \square W^m)_A \tag{5.3}$$

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

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

Proof: Observe that $\overset{*}{f}{}^{k}$ is symmetric and

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

Hence, equation (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 C. Then equation (5.4) is true in general since each side is a martingale relative to $\{F_W(A) : A \in C\}$.

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

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

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

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

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

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

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

$$L(f,A) = 1 + \sum_{m=1}^{\infty} \frac{1}{m!} \left(W_m(f,S\cap A) + \sum_{k=1}^{m} \binom{m}{k} [\overset{*k}{f} W_{m-k}(f,S.) \circ W^k]_A\right)$$

$$= L(f,S\cap A) + \sum_{k=1}^{\infty} \frac{1}{k!} [\overset{*k}{f} \sum_{j=0}^{\infty} \frac{1}{j!} W_j(f,S.) \circ W^k]_A$$

$$= L(f,S\cap A) + \sum_{k=1}^{\infty} \frac{1}{k!} [\overset{*k}{f} L(f,S.) \circ W^k]_A$$

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

6. Change of Measure and Liklihood Ratio Formulas

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

$$\frac{dP'}{dP} = E[\frac{dP'}{dP} \mid F(S)] + \sum_{m=1}^{\infty} \gamma_m \circ W^m \tag{6.1}$$

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

$$L(A) = E[\frac{dP'}{dP} \mid F(S)] + \sum_{m=1}^{\infty} (\gamma_m \circ W^m)(A).$$

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

$$E[W(dt_1)W(dt_2)..W(dt_m)L(T)|F(S_t)]/dt_1dt_2..dt_m = m!\tilde{\gamma}_m(t) .$$

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

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

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

$$L(A) = E[\frac{dP'}{dP} | F(S)] + \sum_{m=1}^{\infty} \frac{1}{m!} [(r_m(\cdot)L(S.)) \circ W^m]_A \tag{6.3}$$

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

Next, define $\Lambda(A) = E[L(A)|F_W(A)]$. $\Lambda(A)$ is called a liklihood ratio. By an application of the projection formula to each term on the right side of equation (6.1),

$$\Lambda(A) = E[\frac{dP'}{dP} | F_W(S)] + \sum_{m=1}^{\infty} (\hat{\gamma}_m \circ W^m)(A)$$

where the integrands $\hat{\gamma}_m \in L_a^2(\Omega x \hat{T}^m, F_W(\cdot))$ satisfy

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

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

$$\hat{r}_m(t) = E'[W(dt_1)W(dt_2)..W(dt_m)|F_W(S_t)]/dt_1dt_2..dt_m \tag{6.5}$$

and the liklihood ratio $\Lambda(A)$ for $A \in C$ has the representation

$$\Lambda(A) = E[\frac{dP'}{dP}|F_W(S)] + \sum_{m=1}^{\infty} \frac{1}{m!} [(r(\cdot)\Lambda(S.))\circ W^m)]_A \qquad (6.6)$$

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

$$\hat{r}_m(t) = E'[r_m(t)|F_W(S_t)] \quad \text{a.e. } t \in \hat{T}^m \qquad (6.7)$$

Remark: Equation (6.3) (resp. equation (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 P' is thus at least formally determined by the functions $\{r_m\}$, it should be possible to express other moments under measure P' in terms of $\{r_m\}$. In this direction, we consider next moments as in equation (6.2) but with S_t replaced be a larger set.

Let $\dot{C}$ be another class of sets with $\dot{C} \subset C$ so that the assumptions of Section 2 are also satisfied by $\{W(A), F(A) : A \in R(T)\}$ and $\dot{C}$. The notation introduced in Section 4 will be used in what follows.

By the iterated integration formula for C-stochastic integrals in terms of $\dot{C}$-stochastic integrals, equation (6.1) yields that

$$\frac{dP'}{dP} = E[\frac{dP'}{dP} | F(\dot{S})] + \sum_{k=1}^{\infty} \dot{\gamma}_k \dot{\circ} W^k \qquad (6.8)$$

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

$$\dot{\gamma}_k(t) = \sum_{m=k}^{\infty} \binom{m}{k} (\tilde{\gamma}_m(tx\cdot)I_{\dot{S}_t^{m-k}}) \overset{t}{\circ} W^{m-k} \text{ for a.e. } t \in \hat{T}^{k\cdot} . \qquad (6.9)$$

Now equation (6.8) is the same as equation (6.1) with C and $\{\gamma_m : m \geq 1\}$

replaced by $\dot{C}$ and $\{\dot{\gamma}_m : m \geq 1\}$. Thus, $\{\dot{L}(A) : A \in \dot{C}\}$ and $\dot{r}_m$ can be defined relative to $\dot{C}$ in the same way as the corresponding quantities were defined relative to C. In particular, equation (6.2) yields that for $t \in \hat{T}^{m\cdot}$,

$$E'[W(dt_1)W(dt_2)..W(dt_m)|F(\dot{S}_t)]/dt_1dt_2..dt_m = \dot{r}_m(t).$$

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

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

$$\dot{r}_k(t) = \frac{1}{L(\dot{S}_t)} \sum_{m=k}^{\infty} \frac{1}{(m-k)!} (L(S_{tx\cdot})r_m(tx\cdot)I_{\dot{S}_t^{m-k}}) \overset{t}{\circ} W^{m-k} \tag{6.10}$$

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 $\dot{\hat{r}}_k$ 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 $\frac{dP'}{dP}$. Suppose that $\{Z(t) : t \in T\}$ is a bounded, measurable process such that $Z(t)$ is $F(S_t)$-measurable for each $t \in T$. Then it will be assumed that, in the notation of Proposition 5.3,

$$\frac{dP'}{dP} = L(Z,T) \tag{6.11}$$

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

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

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

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

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

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

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

where $h \overset{t}{\circ} W$ is a stochastic integral defined relative to the class of sets C_t. By the fact that $L(Z,S_t) = E[\frac{dP'}{dP} | F(S_t)]$ we have

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

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

$$E'[\Phi|F(S_t)] = E[L^t(h+ZI_{S_t^c},T)|F(S_t)] \tag{6.15}$$

where L^t is defined in the same way as L except relative to the class of sets C_t instead of C. Finally, by the martingale property of L^t relative to the class of sets C_t (which contains S_t), the right side of equation (6.15) is equal to $L^t(h+ZI_{S_t^c},S_t) = 1$. Thus $E'\Phi = 1$. ¤

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_α denote the half space $\{t \in \mathbb{R}^n : (t,a) \geq \alpha\}$. Then the collection $C = \{H_\alpha \cap T\}$ is a one-parameter family of sets such that $\hat{T}^m$ is vacuous for $m > 1$. That is, two or more points are always C-dependent. For this choice of C if $\frac{dP'}{dP}$ 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 equations (6.6) and (6.12) reduces to

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

and an application of (5.6) yields

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

where

$$\hat{Z}(t) = E'(Z(t)|F_W(S_t))$$

$$= E'(Z(t)|F_W(H_{(t,a)}\cap T))$$

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

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

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

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

$$S_t = \{t_1, t_2, .., t_m\} \tag{6.18}$$

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

$$\rho_m(t_1, t_2, .., t_m) = E[Z(t_1)Z(t_2)..Z(t_m)]. \tag{6.19}$$

Thus, equation (6.6) becomes

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

Equation (6.20) provides a martingale representation of the liklihood 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

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

Relative to C^k, if $0 \leq t_1 \leq t_2 \leq .. \leq t_m \leq 1$ then

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

and $\hat{T}^m$ is vacuous for $m > k$. Now suppose that for each $k \geq 1$, $\frac{dP'}{dP}$ has a representation of the form (6.1) relative to $C = C^k$. Let r_m^k denote the m^{th} order moment densities when C is equal to C^k. Then by (6.21), it is clear that $r_m^k = r_m^j$ $P \times \mu^m$ a.e. on $[0,1]^m$ if k and j are each larger than m. Thus, we can define functions r_m by

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

and

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

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

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

and

$$\hat{r}_m(t_1, t_2, .., t_m | [0, t_1]) = E'[W(dt_1) W(dt_2)..W(dt_m) | F_W([0, t_1])]/dt_1 dt_2 .. dt_m$$

Let $\Lambda_{t_1} = E[\frac{dP'}{dP} | F_W([0, t_1])]$ for $t_1 \in T$. Then by (6.21),

$$\Lambda(S^k_{t_1 t_2 .. t_m}) = \begin{cases} 1 \text{ if } k < m \\ \Lambda_{t_{m-k+1}} \quad \text{if } k \geq m \end{cases} \quad P \quad \text{a.s.}$$

if $t_1 < t_2 < .. < t_m$.

Thus, equation (6.6) for the liklihood ratio when $C = C^m$ becomes

$$\Lambda_t = 1 + \int_0^t \hat{r}_1(t_1) W(dt_1) + \int_0^t \int_0^{t_2} \hat{r}_2(t_1, t_2) W(dt_1) W(dt_2) + \ldots$$

$$+\int_0^t \int_0^{t_{m-1}} \cdots \int_0^{t_2} r_{m-1}(t_1,\ldots,t_{m-1})W(dt_1)\ldots W(dt_{m-1})$$

$$+\int_0^t \int_0^{t_m} \cdots \int_0^{t_2} \hat{r}_m(t_1,\ldots,t_m|[0,t_1])W(dt_1)W(dt_2)\ldots W(dt_m) \quad (6.22)$$

and the moment equation (6.10) when $\dot{C} = C_m$ and $C = C_{m+1}$ becomes

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

(6.23)

Now suppose that $\frac{dP'}{dP}$ satisfies (6.11) where $\{Z_t : t \in T\}$ is bounded and 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 P'. Thus

$$\hat{r}_m(t_1,\ldots,t_m|[0,t_1]) = E'[Z_{t_1}Z_{t_2}\ldots Z_{t_m}| F_W([0,t])]$$

and

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

where ρ_m is the m-th moment defined in equation (6.19). Defining $\hat{Z}_t = E'[Z_t|F_W([0,t])]$, equations (6.22) and (6.23) for m = 1 become

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

and

$$\Lambda_t\hat{Z}_t = E[Z_t] + \int_0^t \Lambda_\tau E'[Z_t Z_\tau|F^W([0,\tau])]W(d\tau) \quad (6.25)$$

respectively.

Note that equation (6.25) is a well-known representation for the first "unnormalized" conditional moment $\hat{Z_t\Lambda_t}$. Using Ito's formula for one-parameter process, equations (6.24) and (6.25) together yield a representation for the conditional moment $\hat{Z}_t = (\Lambda_t\hat{Z}_t)/\Lambda_t$ itself. As an intermediate step, an integral representation for Λ_t^{-1} would also be derived via Ito's formula. In constrast, in the setting of a general class of sets C, a suitable analogue of Ito's formula is not yet available. Such a formula might provide, for example, an analogue of the moment equation (6.10) for $\dot{r}_k$ which does not contain the factor $1/L(\dot{S}_t)$. (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 $\{W(A), F(A) : A \in R(T)\}$ is as in Section 2, and consider P' defined by

$$\frac{dP'}{dP} = W(T)^2$$

In terms of multiple Wiener integrals,

$$\frac{dP'}{dP} = 1 + 1 \circ W^2$$

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

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

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

$$\dot{r}_1(t) = W(\dot{S}_t)/L(\dot{S}_t)$$

$$\dot{r}_2(t_1,t_2) = 1/L(\dot{S}_t)$$

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

$$\frac{dP'}{dP} = L(\dot{S}) + (L(S.)\dot{r}_1(\cdot)) \dot{\circ} W + (L(S.)\dot{r}_2(\cdot)) \dot{\circ} W^2.$$

ACKNOWLEDGEMENT

We would like to thank Dr. Moshe Zakai for many helpful suggestions and comments.

References

1. Cameron, R.H., Martin, W.T.: The orthogonal development of non-linear functionals in a series of Fourier-Hermite functions. Ann. of Math. 48, 385-392 (1947).
2. Duncan, T.E.: Likelihood functions for stochastic signals in white noise. Inform. 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., Springer-Verlag, New York (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. A. Wahrscheinlichkeitstheorie 29, 109-122 (1974).
10. Wong, E., Zakai, M.: Likelihood ratios and transformation of probability associated with two-parameter Wiener processes. Z. Wahrscheinlichkeitstheorie 40, 283-309 (1977).
11. Yor, M.: Representation des martingales de carré integrable relative aux processus de Wiener et de Poisson à n paramétres. Z. Wahrscheinlichkeitstheorie 35, 121-129 (1976).

Appendix A

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

$$\dot{S}_{A_1A_2..A_m} = S_{A_1A_2..A_m t_1 t_2 .. t_k} ,$$

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

To begin we will establish that

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

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

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

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

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 $A = A_1 x..x A_m$ so that $A \in \hat{T}^{m\cdot}$. 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 $B = B_1 x..x B_m$ where $B_1, B_2, .., B_m$ are rectangles such that

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

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

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

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

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

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

Appendix B

Let C and $\dot{C}$ with $C \supset \dot{C}$ each satisfy conditions $c_1 - c_3$. Let I_m denote the collection of subsets of T^m of the form $A_1 x \ldots x A_m$ such that each $A_i \in R^n(T)$ and for some permutation Π,

1) $A_{\Pi(1)}, \ldots, A_{\Pi(k)}$ are C-independent, and

2) $A_{\Pi(k+1)}, \ldots, A_{\Pi(m)} \subset S_{A_{\Pi(1)} A_{\Pi(2)} \cdots A_{\Pi(k)}}$.

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

Proposition B. The linear span of $\{\alpha I_B : B \in \dot{I}_m, B \in \hat{T}^m, \alpha$ is bounded, $F(\dot{S}_{A_1 A_2 \ldots A_m})$ meas.$\}$ is dense in $L^2_a(\Omega x \hat{T}^m)$ for each $m \geq 1$.

Proof: Consider the following two conditions:

(b_1) There is a countable subcollection of I_m which covers T^m a.e.

(b_2) There is a countable subcollection I^d_m of disjoint sets in I_m which covers T^m a.e.

By a sequence of lemmas it is shown below that condition b_1 is satisfied and then that condition $b_1 \Rightarrow$ condition b_2 and finally that condition b_2 (but with I^d_m replaced by $\dot{I}^d_m$) implies Proposition B.

¤

Lemma B.1.

$$\bigcup_{\ell=1}^{m} \bigcup_{\Pi \in P(m)} \Pi \circ \{(\underline{x}, \underline{y}) : \underline{x} \in \hat{T}^{\ell}, \underline{y} \in (S_{\underline{x}})^{m-\ell}\} = T^m \qquad (*)$$

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

$$S_{\underline{q}} = S_{p_1, \ldots, p_\ell} \neq S_{p_1, \ldots, \hat{p}_i, \ldots, p_\ell} \quad \text{for } 1 \leq i \leq \ell$$

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

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

$$\{A \in C : p_1,\ldots,p_\ell \in A\} = \{A \in C : p_1,\ldots,\tilde{p}_i,\ldots,p_\ell \in A\} .$$

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

$$S_{p_1,\ldots,p_\ell} = S_{p_1,\ldots,\tilde{p}_i,\ldots,p_\ell}$$

which contradicts our choice of $p_1,\ldots,p_\ell$. Thus $p_1,\ldots,p_\ell$ are C-independent so that $\underline{p}$, and hence $\underline{q}$, is contained in the left side of (*). ¤

Lemma B.2. Condition b_1 is satisfied.

Proof: Let I_m^0 denote the subsets of T^m of the form $A_1 x \ldots x A_m$ such that, for some $\Pi \in \mathcal{P}(m)$ and some $\ell > 0$,

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

b) $A_{\Pi(\ell+1)} = \ldots = A_{\Pi(m)} = S_{A_{\Pi(1)}A_{\Pi(2)}\ldots A_{\Pi(\ell)}}$

Then I_m^0 is a countable subset of I_m and

$$\bigcup_{A\in\mathcal{I}_m} A \supset \bigcup_{\ell=1}^{m} \bigcup_{\Pi\in\mathcal{P}(m)} \Pi \circ \{(\underline{x},\underline{y}) : \underline{x} \in \hat{T}^{\ell}, \underline{y} \in (S_{\underline{x}}^{(-)})^{m-\ell}\}$$

$$= \bigcup_{\ell=1}^{m} \bigcup_{\Pi\in\mathcal{P}(m)} \Pi \circ \{(\underline{x},\underline{y}) : \underline{x} \in \hat{T}^{\ell}, y \in (S_{\underline{x}})^{m-\ell}\}$$

$$- \bigcup_{\ell=1}^{m} \bigcup_{\Pi\in\mathcal{P}(m)} \Pi \circ S_{m,\ell} \tag{B.1}$$

where

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

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 $\mu^m(S_{m,\ell}) = 0$ for all $m \geq 1$ and $1 \leq \ell \leq m$.

By condition c_2,

$$F_\varepsilon = \{(\underline{x},\underline{y}) : \underline{x} \in \hat{T}^{\ell}, \underline{y} \in (S_{\underline{x}}^{\varepsilon})^{m-\ell}\}$$

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

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

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

<u>Lemma B.3</u> Condition b_1 implies condition b_2.

<u>Proof:</u> Let $F_1, F_2, \ldots$ be a countable subcollection of $\mathcal{I}_m$ 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 $D_{i1},\ldots,D_{in}$ in $\mathcal{I}_m$ such that $D_i = \bigcup_{j=1}^{n_i} D_{ij}$. Condition b_2 is then satisfied with $\mathcal{I}_m^d = \{D_{ij} : i \geq 1, 1 \leq j \leq n_i\}$. It remains to prove the claim.

By induction, it suffices to establish the claim for $i = 2$. Now $F_1 = A_1 x \ldots x A_m$ for some Borel sets $A_1,\ldots,A_m \subset T$. Thus, $F_1^c = \bigcup_{j=1}^{r} K_j$ where $K_1,\ldots,K_r$ are disjoint and each K_j is the product of m Borel subsets of T. In fact, F_1^i is the union of all sets of the form $B_1 x \ldots x B_m$ 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 $\mathcal{I}_m$ as required so the claim is established. ¤

<u>Lemma B.4</u>. Condition b_2 with $\mathcal{I}_m^d$ replaced by $\dot{\mathcal{I}}_m^d$ implies that the linear span of $\{\alpha 1_B : B \in \dot{\mathcal{I}}_m, B \subset \hat{T}^m, \alpha$ is bounded, $\mathcal{F}(\dot{S}_{A_1\ldots A_m})$ measurable$\}$ is dense in $L_a^2(\Omega x \hat{T}^m)$.

Proof: Let $F = F_1 x \ldots x F_m$ where each $F_i \in \mathcal{R}^n(T)$ such that $F \in \hat{T}^m$ and let α be a bounded, $\mathcal{F}(S_{F_1 F_2 \ldots F_m})$-measurable random variable. If $A = A_1 x \ldots x A_m \in \dot{\mathcal{I}}_m^d$ then $B = A \cap F$ satisfies $B \in \dot{\mathcal{I}}_m$ and $B \subset \hat{T}^m$, and α is $\mathcal{F}(\dot{S}_{A_1 \cap F_1 \ldots A_m \cap F_m})$-measurable since $\dot{C}$-supports are no smaller than C-supports. By condition b_2 with $\mathcal{I}_m^d$ replaced by $\dot{\mathcal{I}}_m^d$,

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

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