The Annals of Probability
1989, Vol. 17, No. 1, 194–207

INTEGRATION BY PARTS, HOMOGENEOUS CHAOS EXPANSIONS AND SMOOTH DENSITIES[1]

BY ROBERT J. ELLIOTT AND MICHAEL KOHLMANN

University of Alberta and Universität Konstanz

By iterating a martingale representation result a homogeneous chaos expansion is obtained. Using the martingale representation, the integration-by-parts formula of the Malliavin calculus is derived using properties of stochastic flows. The infinite-dimensional calculus of variations is not required.

1. Introduction. Since Malliavin's outstanding breakthrough [9] there have been other treatments and simplifications of the Malliavin calculus, including those of Bismut [2], Stroock [11], Bichteler and Fonken [1] and Norris [10]. In this paper we apply a very simple representation of the integrand in a stochastic integral, Theorem 3.1, to first derive the homogeneous chaos expansion of a certain random variable. An integration-by-parts formula is obtained and, if the Malliavin matrix M has an inverse which belongs to every $L^p(\Omega)$ (a condition guaranteed by Hörmander's H_1 hypothesis), it is shown the diffusion has a smooth density. The principle simplification in this paper is the observation that by considering an enlarged Markov system only the simple stochastic integral representation of Theorem 3.1 is needed. No infinite-dimensional calculus is required.

2. Flows. In this section we recall some definitions and properties of stochastic flows on d-dimensional Euclidean space. Suppose $w_t = (w_t^1, \ldots, w_t^m)$, $0 \le t$, is an m-dimensional Brownian motion on (Ω, F, P). Write $\{F_t\}$ for the right-continuous complete filtration generated by w. Let $X_0, X_1, \ldots, X_m$ be smooth vector fields on $[0, \infty) \times R^d$ all of whose derivatives are bounded. Then from Bismut [2] or Carverhill and Elworthy [4] we quote the following result.

THEOREM 2.1. *There is a map $\xi\colon \Omega \times [0,\infty) \times [0,\infty) \times R^d \to R^d$ such that*:

(i) *For $0 \le s \le t$ and $x \in R^d$ $\xi_{s,t}(x)$ is the essentially unique solution of the stochastic differential equation*

$$d\xi_{s,t}(x) = X_0\big(t, \xi_{s,t}(x)\big)\,dt + X_i\big(t, \xi_{s,t}(x)\big)\,dw_t^i, \tag{2.1}$$

with $\xi_{s,s}(x) = x$. (Note the Einstein summation convention is used.)

Received May 1987; revised January 1988.

[1]Supported by the Natural Sciences and Engineering Research Council of Canada under Grant A7964.

AMS 1980 *subject classifications*. 60H07, 60H10, 60J60.

Key words and phrases. Martingale representation, stochastic flow, homogeneous chaos, Malliavin calculus, integration by parts, smooth densities.

(ii) *For each* ω, s, t *the map* $\xi_{s,t}(\cdot)$ *is* C^∞ *on* $R^d \to R^d$ *with a first derivative, the Jacobian,* $\partial\xi_{s,t}/\partial x = D_{s,t}$, *which satisfies*

$$dD_{s,t} = \frac{\partial X_0}{\partial \xi}\big(t, \xi_{s,t}(x)\big)D_{s,t}\,dt + \frac{\partial X_i}{\partial \xi}\big(t, \xi_{s,t}(x)\big)D_{s,t}\,dw_t^i, \tag{2.2}$$

with initial condition $D_{s,s} = I$, *the* $d \times d$ *identity matrix.*

(iii) *If* $W_{s,t} = \partial^2\xi_{s,t}/\partial x^2$ *is the second derivative, then*

$$\begin{aligned} dW_{s,t} = {} & \frac{\partial X_0}{\partial \xi}\big(t, \xi_{s,t}(x)\big)W_{s,t}\,dt + \frac{\partial X_i}{\partial \xi}\big(t, \xi_{s,t}(x)\big)W_{s,t}\,dw_t^i \\ & + \frac{\partial^2 X_0}{\partial \xi^2}\big(t, \xi_{s,t}(x)\big)D_{s,t} \otimes D_{s,t}\,dt + \frac{\partial^2 X_i}{\partial \xi^2}\big(t, \xi_{s,t}(x)\big)D_{s,t} \otimes D_{s,t}\,dw_t^i, \end{aligned} \tag{2.3}$$

with $W_{s,s} = 0 \in (R^d \otimes R^d) \otimes R^d$.

REMARKS 2.2. Note that (2.2) and (2.3) are obtained formally by differentiating (2.1). However, if we consider the enlarged stochastic system given by (2.1)–(2.3) for $(\xi_{s,t}, D_{s,t}, W_{s,t})$, the coefficients are not bounded. Nevertheless, Norris [10] has extended the results of Theorem 2.1. to such systems. To state Norris' results we first define a class of "lower triangular" coefficients.

DEFINITION 2.3. For positive integers $\alpha, d, d_1, \ldots, d_k$ write $S_\alpha(d_1, \ldots, d_k)$ for the set of $X \in C^\infty(R^d, R^d)$ of the form

$$X(x) = \begin{pmatrix} X^{(1)}(x^1) \\ X^{(2)}(x^1, x^2) \\ \vdots \\ X^{(k)}(x^1, x^2, \ldots, x^k) \end{pmatrix}, \quad \text{for } x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^k \end{pmatrix}, \tag{2.4}$$

where R^d is identified with $R^{d_1} \times \cdots \times R^{d_k}$, $x^j \in R^{d_j}$ and the X satisfy

$$\|X\|_{S(\alpha, N)} = \sup_{x \in R^d}\left(\sup_{0 \le n \le N} \frac{|D^n X(x)|}{(1 + |x|^\alpha)} \vee \sup_{1 \le j \le k} \left|D_j X^{(j)}(x)\right| \right) < \infty, \tag{2.5}$$

for all positive integers N. Write $S(d_1, \ldots, d_k) = \bigcup_\alpha S_\alpha(d_1, \ldots, d_k)$.

REMARKS 2.4. Note (2.1)–(2.3) can be considered as a single system whose coefficients are not bounded but are in $S(d, d^2, d^3)$. The final supremum on the right of (2.5) implies the first derivatives of $X^{(1)}$ are bounded, as are the first derivatives D_j in the "new" variable x^j of $X^{(j)}(x^1, \ldots, x^j)$. This means $X^{(j)}$ is allowed linear growth in x^j, a situation illustrated in (2.2) and (2.3). We quote from Norris [10] the following result.

THEOREM 2.5. *Let $X_0, X_1, \ldots, X_m \in S_\alpha(d_1, \ldots, d_k)$. Then there is a map $\phi: \Omega \times [0, \infty) \times [0, \infty) \times R^d \to R^d$ such that:*

(i) *For $0 \le s \le t$ and $x \in R^d$, $\phi(\omega, s, t, x)$ is the essentially unique solution of the stochastic differential equation*

$$dx_t = X_0(x_t)\, dt + X_i(x_t)\, dw_t^i, \tag{2.6}$$

with $x_s = x$.

(ii) *For each ω, s, t the map $\phi(\omega, s, t, x)$ is C^∞ in x with derivatives of all orders satisfying stochastic differential equations obtained from (2.6) by formal differentiation.*

(iii)
$$\sup_{|x| \le R} E\left[\sup_{s \le u \le t} |D^n\phi(\omega, s, u, x)|^p\right] \le C\big(p, s, t, R, N, d_1, \ldots, d_k, \alpha, \|X_0\|_{s(\alpha, N)}, \ldots, \|X_m\|_{S(\alpha, N)}\big). \tag{2.7}$$

REMARKS 2.6. Norris proves Theorem 2.5. by induction on j. Write (2.6) as a system of stochastic differential equations for $j = 1, \ldots, k$,

$$\begin{aligned} dx_t^j &= X_0^{(j)}(x_t^1, \ldots, x_t^j)\, dt + X_i^{(j)}(x_t^1, \ldots, x_t^j)\, dw_t^i, \\ x_s^j &= x^j \in R^{d_j}. \end{aligned} \tag{2.8}$$

Suppose the result is true for $1, \ldots, j-1$ and write $\tilde{X}_i^{(j)}(\omega, s, t, x^j) = X_i^{(j)}(x_t^1(\omega), \ldots, x_t^{j-1}(\omega), x^j)$. Then (2.8) can be written in the form

$$dx_t^j = \tilde{X}_0(s, t, x_t^j)\, dt + \tilde{X}_i(s, t, x_t^j)\, dw_t^i$$

and Theorem (2.1) applied. The difficult step is establishing the result for $j = 1$. However, this follows using a stopping argument, a technique employed by Bismut [2, 3]. Using the notation of Theorem 2.1, the following result is well known.

LEMMA 2.7. *For $0 \le s \le t$ write $V_{s,t}$ for the solution of*

$$\begin{aligned} dV_{s,t} = &-V_{s,t}\left(\frac{\partial X_0}{\partial \xi}(t, \xi_{s,t}(x))\right) - \sum_{i=1}^m \left(\frac{\partial X_i}{\partial \xi}(t, \xi_{s,t}(x))^2\right) dt \\ &- V_{s,t}\frac{\partial X_i}{\partial \xi}(t, \xi_{s,t}(x))\, dw_t^i, \end{aligned} \tag{2.9}$$

with $V_{s,s} = I$. Then $D_{s,t}V_{s,t} = I$, the $d \times d$ identity matrix.

PROOF. Applying Itô's rule to $V_{s,t}D_{s,t}$, we see $d(V_{s,t}D_{s,t}) = 0$. However, $V_{s,s}D_{s,s} = I$. □

REMARKS 2.8. An application of Jensen's, Burkholder's and Gronwall's inequalities shows that $\sup_{s \le u \le t}|D_{s,u}|$, $\sup_{s \le u \le t}|W_{s,u}|$ and $\sup_{s \le u \le t}|V_{s,u}|$ are in $L^p(\Omega)$ for all $p < \infty$. Alternatively, this conclusion follows from applying Theorem 2.5 to the system (2.1)–(2.3) and (2.9). For $0 \le s \le t$, by the uniqueness of the solution of (2.1)

$$\begin{aligned} \xi_{0,t}(x_0) &= \xi_{s,t}(\xi_{0,s}(x_0)) \\ &= \xi_{s,t}(x), \quad \text{if } x = \xi_{0,s}(x_0). \end{aligned} \tag{2.10}$$

Differentiating (2.10), using the chain rule,

$$D_{0,t} = D_{s,t}D_{0,s} \tag{2.11}$$

and

$$W_{0,t} = W_{s,t}(D_{0,s} \otimes D_{0,s}) + D_{s,t}W_{0,s}. \tag{2.12}$$

3. Representation and series expansion. Suppose $0 \le t \le T$ and $\xi_{0,t}(x_0)$ is the solution of the stochastic differential equation (2.1). Consider a real-valued twice continuously differentiable function c for which the random variable $c(\xi_{0,T}(x_0))$ and the components of the gradient $c_\xi(\xi_{0,T}(x_0))$ are integrable. Let M_t be the right-continuous version of the martingale

$$E\left[c(\xi_{0,T}(x_0))|F_t\right].$$

We then have the following representation result.

THEOREM 3.1. *For* $0 \le t \le T$, $M_t = E[c(\xi_{0,T}(x_0))] + \int_0^t \gamma_i(s)\,dw_s^i$, *where*

$$\gamma_i(s) = E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}|F_s\right]D_{0,s}^{-1}X_i(s, \xi_{0,s}(x_0)).$$

PROOF. It is well known (see [5], for example) that any F_t-martingale M_t has a representation

$$M_t = M_0 + \int_0^t \gamma_i(s)\,dw_s^i, \tag{3.1}$$

for some predictable integrands γ_i. Because the process $\xi_{0,t}(x_0)$ is Markov

$$\begin{aligned} M_t &= E\left[c(\xi_{0,T}(x_0))|F_t\right] \\ &= E\left[c(\xi_{t,T}(x))|F_t\right] \\ &= E_{t,x}\left[c(\xi_{t,T}(x))\right] \\ &= V(t,x), \quad \text{say, where } x = \xi_{0,t}(x_0). \end{aligned} \tag{3.2}$$

By the chain rule and Theorem 2.1, $c(\xi_{t,T}(x))$ is differentiable, in fact smooth, in x. The differentiability of $E[c(\xi_{t,T}(x))|F_t]$ in t can be established by writing the backward equation for $\xi_{t,T}(x)$ as in Kunita [8]. Consequently, applying the Itô rule to $V(t, x)$, with $x = \xi_{0,t}(x_0)$,

$$V(t, \xi_{0,t}(x_0)) = V(0, x_0) + \int_0^t \left(\frac{\partial V}{\partial s} + LV \right) ds \tag{3.3}$$
$$+ \int_0^t \frac{\partial V}{\partial x}(s, \xi_{0,s}(x_0)) X_i(s, \xi_{0,s}(x_0))\, dw_s^i,$$

where

$$L = \sum_{i=1}^d X_0^i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^d \left(\sum_{k=1}^m X_k^i X_k^j \right) \frac{\partial^2}{\partial x_i\, \partial x_j}.$$

By the uniqueness of the decomposition of special semimartingales, comparing (3.1) and (3.3), we must have (as is well known)

$$\frac{\partial V}{\partial s} + LV = 0,$$

and

$$\gamma_i(s) = \frac{\partial V}{\partial x}(s, \xi_{0,s}(x_0)) X_i(s, \xi_{0,s}(x_0)).$$

From (3.2)

$$\frac{\partial V}{\partial x} = E\left[c_\xi(\xi_{s,T}(x)) D_{s,T} | F_s \right]$$

so by (2.11)

$$\gamma_i(s) = E\left[c_\xi(\xi_{0,T}(x_0)) D_{0,T} | F_s \right] D_{0,s}^{-1} X_i(s, \xi_{0,s}(x_0)). \qquad \square$$

REMARKS 3.2. Note in particular the representation

$$c(\xi_{0,T}(x_0)) = E\left[c(\xi_{0,T}(x_0)) \right] \tag{3.4}$$
$$+ \int_0^T E\left[c_\xi(\xi_{0,T}(x_0)) D_{0,T} | F_s \right] D_{0,s}^{-1} X_i(s, \xi_{0,s}(x_0))\, dw_s^i.$$

Theorem 3.1 can be extended immediately to vector (or matrix) functions c. Finally, it seems the proof of Theorem 3.1 can be extended to the non-Markov case ([6]).

3.3 NOTATION. Write $\xi^{(0)} = \xi$ for the solution flow of Theorem 2.1, and $D^{(0)} = D$ for its Jacobian given by (2.2). Write $\xi^{(1)}$ for the $d + d^2$-dimensional process with components $\xi^{(1)} = (\xi^{(0)}, D^{(0)})$. Write $D^{(1)}$ for the Jacobian of this $d + d^2$-dimensional process. Write $\xi^{(2)}$ for the process $\xi^{(2)} = (\xi^{(1)}, D^{(1)})$ and so on. Then $\xi^{(n+1)} = (\xi^{(n)}, D^{(n)})$. Note $\xi^{(n)}$ is a process for which the stochastic flow results of Theorem 2.5 hold. Write

$$c^{(1)}\big(\xi^{(1)}_{0,T}(x_0)\big) = \frac{\partial c}{\partial \xi}\big(\xi_{0,T}(x_0)\big)D_{0,T},$$

$$c^{(2)}\big(\xi^{(2)}_{0,T}(x^{(2)}_0)\big) = \frac{\partial c^{(1)}}{\partial \xi^{(1)}}\big(\xi^{(1)}_{0,T}(x^{(1)}_0)\big)D^{(1)}_{0,T}$$

and so on, so

$$c^{(n+1)}\big(\xi^{(n+1)}_{0,T}(x^{(n+1)}_0)\big) = \frac{\partial c^{(n)}}{\partial \xi^{(n)}}\big(\xi^{(n)}_{0,T}(x^{(n)}_0)\big)D^{(n)}_{0,T}.$$

Note the initial condition at 0 for the variable $D^{(n)}$ is always the identity matrix of appropriate dimension. Write $X^{(n)}_i$ for the vector field coefficient of w^i in the stochastic differential equation defining $\xi^{(n)}$ and abbreviate

$$X^{(n)}_i\big(s, \xi^{(n)}_{0,s}(x^{(n)}_0)\big) \quad \text{as } X^{(n)}_i(s).$$

Then by iterating Theorem 3.1, we have the following representation of the random variable $c(\xi_{0,T}(x_0))$.

Theorem 3.4. *If c has bounded derivatives of all order, then for any n,*

$$\begin{aligned} c(\xi_{0,T}(x_0)) &= E\left[c(\xi_{0,T}(x_0))\right] + \sum_{k=1}^{n} E\left[c^{(k)}\big(\xi^{(k)}_{0,T}(x^{(k)}_0)\big)\right] \\ &\quad \times \int_0^T\left(\int_0^{s_1}\cdots\left(\int_0^{s_n} D^{(n-1)-1}_{0,s_n}X_k(s_n)\,dw^k_{s_n}\right)\cdots\right)D^{-1}_{0,s}X_i(s)\,dw^i_s \\ &\quad + \int_0^T\left(\int_0^{s_1}\cdots\left(\int_0^{s_n} E\left[c^{(n+1)}|F_{s_{n+1}}\right]D^{(n)-1}_{0,s_{n+1}}X_j(s_{n+1})\,dw^j_{s_{n+1}}\right)\cdots\right) \\ &\quad \times D^{-1}_{0,s}X_i(s)\,dw^i_s. \end{aligned} \tag{3.5}$$

Proof. From (3.4)

$$\begin{aligned} c\big(\xi_{0,T}(x_0)\big) &= E\left[c\big(\xi_{0,T}(x_0)\big)\right] + \int_0^T E\left[c^{(1)}|F_s\right]D^{-1}_{0,s}X_i(s)\,dw^i_s \\ &= E[c] + E\left[c^{(1)}\right]\int_0^T D^{-1}_{0,s}X_i(s)\,dw^i_s \\ &\quad + \int_0^T\left(\int_0^{s_1} E\left[c^{(2)}|F_{s_1}\right]D^{(1)-1}_{0,s_1}X^{(1)}_i(s_1)\,dw^i_{s_1}\right)D^{-1}_{0,s}X_j(s)\,dw^j_s. \end{aligned}$$

The result follows by repeated application of the representation of Theorem 3.1. □

REMARKS 3.5. In principle, it is possible to write the previous expansion in terms of $D_{0,t}$, $W_{0,t}$ and higher derivatives of the diffeomorphism $\xi_{0,t}$, rather than by considering higher and higher dimensional systems. However, this gives rise to very complicated formulae. Consider the case when $d = 1$. Then $\xi^{(1)} = (\xi^{(0)}, D^{(0)})$ is two dimensional and by Theorem 3.1

$$
\begin{aligned}
E\left[c^{(1)}(\xi^{(1)})|F_s\right] &= E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}\right] \\
&\quad + \int_0^s E\left[c_{\xi\xi}(\xi_{u,T}(x))D_{0,T}D_{u,T}\right. \\
&\qquad \left. + c_\xi(\xi_{u,T}(x))W_{u,T}D_{0,u}|F_u\right] X_i(u)\, dw_u^i \\
&\quad + \int_0^s E\left[c_\xi(\xi_{u,T}(x_0))D_{u,T}|F_u\right] \frac{\partial X_i}{\partial \xi}(u)\, dw_u^i.
\end{aligned} \tag{3.6}
$$

Here we are writing $\xi_{0,T}(x_0) = \xi_{u,T}(x)$, where $x = \xi_{0,u}(x_0)$, and $D_{0,T} = D_{u,T}D$, where $D = D_{0,u}$. Note the final integral in (3.6) is a result of differentiating in the D variables. Recalling (2.11) and (2.12), we have

$$
\begin{aligned}
E\left[c^{(1)}(\xi^{(1)})|F_s\right] &= E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}\right] \\
&\quad + \int_0^s E\left[c_{\xi\xi}(\xi_{0,T}(x_0))D_{0,T}^2\right. \\
&\qquad \left. + c_\xi(\xi_{0,T}(x_0))W_{0,T}|F_u\right] D_{0,u}^{-1}X_i(u)\, dw_u^i \\
&\quad - \int_0^s E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}|F_u\right] D_{0,u}^{-2}W_{0,u}X_i(u)\, dw_u^i \\
&\quad + \int_0^s E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}|F_u\right] D_{0,u}^{-1}\frac{\partial X_i}{\partial \xi}(u)\, dw_u^i \\
&= E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}\right] + \int_0^s \gamma_j(u,2)\, dw_u^j,
\end{aligned} \tag{3.7}
$$

where

$$
\begin{aligned}
\gamma_j(u,2) &= \alpha(u,2,1)D_{0,u}^{-1}X_j(u) - \alpha(u,2,2)D_{0,u}^{-2}W_{0,u}X_j(u) \\
&\quad + \alpha(u,2,2)D_{0,u}^{-1}\frac{\partial X_j}{\partial \xi}(u).
\end{aligned} \tag{3.8}
$$

Here

$$
\alpha(u,2,1) = E\left[c_{\xi\xi}(\xi_{0,T}(x_0))D_{0,T}^2 + c_\xi(\xi_{0,t}(x_0))W_{0,T}|F_u\right]
$$

and

$$
\alpha(u,2,2) = E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}|F_u\right].
$$

Substituting (3.7) in (3.4), we have

$$c(\xi_{0,T}(x_0)) = E\left[c(\xi_{0,t}(x_0))\right] + E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}\right]\int_0^T D_{0,s}^{-1}X_i(s)\,dw_s^i \tag{3.9}$$
$$+\int_0^T\left(\int_0^s \gamma_j(u,2)\,dw_u^j\right)D_{0,s}^{-1}X_i(s)\,dw_s^i.$$

In turn, the martingales $\alpha(u,2,1)$ and $\alpha(u,2,2)$ can be expressed as stochastic integrals. Substituting again, we have

$$\begin{aligned}
c(\xi_{0,T}(x_0)) &= E\left[c(\xi_{0,T}(x_0))\right] + E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}\right]\int_0^T D_{0,s}^{-1}X_i(s)\,dw_s^i \\
&\quad + E\left[c_{\xi\xi}(\xi_{0,T}(x_0))D_{0,T}^2 + c_\xi(\xi_{0,t}(x_0))W_{0,T}\right] \\
&\qquad \times\int_0^T\left(\int_0^s D_{0,u}^{-1}X_j(u)\,dw_u^j\right)D_{0,s}^{-1}X_i(s)\,dw_s^i \\
&\quad - E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}\right]\int_0^T\left(\int_0^s D_{0,u}^{-2}W_{0,u}X_j(u)\,dw_u^j\right)D_{0,s}^{-1}X_i(s)\,dw_s^i \\
&\quad + E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}\right]\int_0^T\left(\int_0^s D_{0,u}^{-1}\frac{\partial X_j}{\partial\xi}(u)\,dw_u^j\right)D_{0,s}^{-1}X_i(s)\,dw_s^i \\
&\quad + \int_0^T\left\{\int_0^s\left(\int_0^u \gamma_k(v,3)\,dw_v^k\right)D_{0,s}^{-1}X_j(u)\,dw_s^j\right. \\
&\quad + \int_0^s\left(\int_0^u \gamma_k(v,4)\,dw_v^k\right)D_{0,s}^{-2}W_{0,u}X_j(u)\,dw_s^j \\
&\quad \left.+ \int_0^s\left(\int_0^u \gamma_k(v,5)\,dw_v^k\right)D_{0,s}^{-1}\frac{\partial X_j(u)}{\partial\xi}\,dw_u^j\right\}D_{0,s}^{-1}X_i(s)\,dw_s^i.
\end{aligned}\tag{3.10}$$

REMARKS 3.6. Theorem 3.4 [or (3.10) in the one-dimensional case] indicates how a "Taylor series" expansion for the random variable $c(\xi_{0,T}(x_0))$ can be obtained as the sum of multiple stochastic integrals.

The coefficients of the stochastic integrals are functions of the expected values of $c(\xi_{0,T}(x_0))$ and its derivatives, and the Jacobian $D_{0,T}$ and its derivatives. The integrands in the multiple stochastic integrals do not involve c, but are functions of the Jacobian and its derivatives, and the coefficient functions X_i. By uniqueness the expansion is the same as the homogeneous chaos representation. This expansion can be used to investigate variations about the expected trajectory and large deviation problems ([7]).

COROLLARY 3.7. *Taking* $c(\xi_{0,T}(x_0)) = \xi_{0,T}(x_0) \in R^d$, *so* $c_\xi = I_d$, *the* $d\times d$ *identity matrix, and* $c_{\xi\xi} = 0$, (3.4) *gives*

$$\xi_{0,T}(x_0) = E\left[\xi_{0,T}(x_0)\right] + \int_0^T\left[D_{0,T}|F_s\right]D_{0,s}^{-1}X_i(s)\,dw_s^i,$$

with corresponding higher-order expansions.

LEMMA 3.8. *Write * to denote the transpose. Suppose c and g are real-valued, differentiable functions such that the random variables* $c(\xi_{0,T}(x_0))$, $g(\xi_{0,T}(x_0))$, $c_\xi(\xi_{0,T}(x_0))$, $g_\xi(\xi_{0,T}(x_0))$ *are in* $L^2(\Omega)$. *Then*

$$\begin{aligned} &E\left[c(\xi_{0,T}(x_0))g(\xi_{0,T}(x_0))\right] \\ &\quad = E\left[c(\xi_{0,T}(x_0))\right]E\left[g(\xi_{0,T}(x_0))\right] \\ &\qquad + E\left[\sum_{i=1}^m \int_0^T E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}|F_s\right]\right. \\ &\qquad\qquad \left.\times D_{0,s}^{-1}X_i(s)X_i^*(s)D_{0,s}^{*-1}E\left[g_\xi^*(\xi_{0,T}(x_0))|F_s\right]ds\right]. \end{aligned}$$

PROOF. By Theorem 3.1

$$g(\xi_{0,T}(x_0)) = E\left[g(\xi_{0,T}(x_0))\right] + \int_0^T E\left[g_\xi(\xi_{0,T}(x_0))D_{0,T}|F_s\right]D_{0,s}^{-1}X_i(s)\,dw_s^i.$$

The result follows by taking the expectation of the product with (3.4). (Note $g^* = g$.) □

DEFINITION 3.9. The nonnegative matrix

$$M_{s,t} = \sum_{i=1}^m \int_s^t D_{s,u}^{-1}X_i(u)X_i^*(u)D_{s,u}^{*}\,du$$

will be called the Malliavin matrix for the system (2.1). Note that something similar to $M_{0,T}$ appears in Lemma 3.8. In some references, [11] and [12], the matrix $D_{0,T}M_{0,T}D_{0,T}^*$ is called the Malliavin matrix.

4. Integration by parts.

THEOREM 4.1. *Suppose c is a twice continuously differentiable scalar function such that* $c(\xi_{0,T}(x_0))$ *and* $c_\xi(\xi_{0,T}(x_0))$ *are square integrable. Then for any square-integrable predictable process* $u(s) = (u_1(s), \dots, u_m(s))$,

$$\begin{aligned} &E\left[c(\xi_{0,T}(x_0))\int_0^T u_i(s)\,dw_s^i\right] \\ &\quad = \sum_{i=1}^m E\left[c_\xi(\xi_{0,t}(x_0))D_{0,T}\int_0^T D_{0,s}^{-1}X_i(s)u_i(s)\,ds\right]. \end{aligned}$$

PROOF. Using the representation (3.4),

$$\begin{aligned} &E\left[c(\xi_{0,T}(x_0))\int_0^T u_i(s)\,dw_s^i\right] \\ &\quad = \sum_{i=1}^m E\left[\int_0^T E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}|F_s\right]D_{0,s}^{-1}X_i(s)u_i(s)\,ds\right] \end{aligned}$$

and by Fubini's theorem this is

$$= \sum_{i=1}^{m} E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}\int_0^T D_{0,s}^{-1}X_i(s)u_i(s)\,ds\right]. \qquad \square$$

Corollary 4.2. *The result is still true for vector- (or matrix-) valued functions c.*

Corollary 4.3. *Taking each $u_i(s)$ to be $(D_{0,s}^{-1}X_i(s))^*$, we have*

$$E\left[c(\xi_{0,T}(x_0))\int_0^T (D_{0,s}^{-1}X_i(s))^*\,dw_s^i\right] = E\left[c_\xi(\xi_{0,T}(x_0))D_{0,T}M_{0,T}\right].$$

Corollary 4.4. *Consider a product function*

$$h(\xi_{0,T}(x_0)) = c(\xi_{0,T}(x_0))g(\xi_{0,T}(x_0))$$

satisfying the conditions of the theorem. Then

$$\begin{aligned} &E\left[c(\xi_{0,T}(x_0))g(\xi_{0,t}(x_0))\int_0^T (D_{0,s}^{-1}X_i(s))^*\,dw_s^i\right] \\ (4.1)\qquad &= E\left[\left(c_\xi(\xi_{0,T}(x_0))g(\xi_{0,t}(x_0))\right.\right. \\ &\qquad \left.\left. + c(\xi_{0,T}(x_0))g_\xi(\xi_{0,T}(x_0))\right)D_{0,T}M_{0,T}\right]. \end{aligned}$$

Remarks 4.5. What we would like to do in (4.1) is take

$$g = M_{0,T}^{-1}D_{0,T}^{-1},$$

so that we can obtain a bound for c_ξ. However, $D_{0,T}^{-1}$ and $M_{0,T}^{-1}$ involve the past of the processes $\xi_{0,T}$, $D_{0,T}$, $M_{0,T}$. This difficulty can be circumvented by considering an enlarged system, similar to the technique used in Section 3. However, the sequence of enlarged systems is different to that discussed in Section 3, so different notation will be used. Note that even when the original process ξ is one dimensional the method leads to a discussion of higher-dimensional processes, so not much simplification is obtained by taking $d = 1$.

4.6 **Notation.** Write $\phi^{(0)}(\omega, s, t, x) = \xi_{s,t}(x)$ for the stochastic flow defined by (2.1). Now $D_{s,t}^{(0)}(x) = D_{s,t}(x)$ denotes the Jacobian of the flow $\phi^{(0)}$. From (2.11), if $D = D_{0,s}$ and $x = \xi_{0,s}(x_0)$,

$$D_{0,t}^{(0)}(x_0) = D_{s,t}(x)D,$$

so the system $(\phi^{(0)}, D^{(0)})$ is Markov. Write $R_{s,t}^{(0)}(x) = \int_s^t (D_{s,u}^{-1}X_i(u))^*\,dw_u^i$ and $R = R_{0,s}^{(0)}$. Then $R_{0,t}^{(0)} = R + D^{-1}R_{s,t}^{(0)}(x)$, so the system $(\phi^{(0)}, D^{(0)}, R^{(0)})$ is Markov. Finally, recall the definition (3.9) of $M_{s,t}$ and write $M_{s,t}^{(0)} = M_{s,t}$, $M = M_{0,s}^{(0)}$. Then $M_{0,t}^{(0)} = M + D^{-1}M_{s,t}(x)D^{*-1}$ and the system

$$\phi^{(1)} = (\phi^{(0)}, D^{(0)}, R^{(0)}, M^{(0)})$$

is Markov with coefficients in

$$S(d, d+d^2, 2d+d^2, 2d+d^2).$$

Consequently, Theorem 2.5 applies to this system and its stochastic flow $\phi^{(1)}$. Note that $M_{s,t}$ is the predictable quadratic variation of the tensor product of $R_{s,t}$ with $R^*_{s,t}$. Write $X_i^{(1)}$ for the coefficient vector fields of w^i in $\phi^{(1)}$. Furthermore, write $D_{s,t}^{(1)}$ for the Jacobian of $\phi^{(1)}$, $R_{s,t}^{(1)} = \int_s^t (D_{s,u}^{(1)-1} X_i^{(1)}(u))^* \, dw_u^i$ and $M_{s,t}^{(1)}$ for the predictable quadratic variation of the tensor product of $R_{s,t}^{(1)}$ with $R_{s,t}^{(0)*}$, which we shall denote by

$$M_{s,t}^{(1)} = \langle R_{s,t}^{(1)} \otimes R_{s,t}^{(0)} \rangle.$$

Then define

$$\phi^{(2)} = \left(\phi^{(1)}, D^{(1)}, R^{(1)}, M^{(1)}\right),$$

so $\phi^{(2)}$ is a Markov process for which the results of Theorem 2.5 hold. Proceeding in this way, we inductively define $\phi^{(n+1)} = (\phi^{(n)}, D^{(n)}, R^{(n)}, M^{(n)})$, where $R^{(n)} = \int_s^t (D_{s,u}^{(n)-1} X_i^{(n)}(u)) \, dw_u^i$ and $M^{(n)} = \langle R^{(n)} \otimes R^{(0)} \rangle$. Write ∇_n for the gradient operator in the components of $\phi^{(n)}$.

THEOREM 4.7. *Suppose c is a bounded C^∞ scalar function on R^d with bounded derivatives. Let g be a possibly vector- (or matrix-) valued function on the state space of $\phi^{(n)}$ such that $g(\phi^{(n)}(0, T, x_0))$ and $\nabla_n g(\phi^{(n)}(0, T, x_0))$ are both in some $L^p(\Omega)$. Then*

$$\begin{aligned} E\left[c(\phi^{(0)}(0,T))g(\phi^{(n)}(0,T)) \otimes R_{0,T}^{(0)}\right] &= E\left[(\nabla_0 c)(\phi^{(0)}(0,T))g(\phi^{(n)}(0,T))D_{0,T}M_{0,T}\right] \\ &\quad + E\left[c(\phi^{(0)}(0,T))(\nabla_n g)(\phi^{(n)}(0,T))D_{0,T}^{(n)}M_{0,T}^{(n)}\right]. \end{aligned}$$

PROOF. Applying Theorem 3.1 to cg, we have

$$\begin{aligned} &c(\phi^{(0)}(0,T))g(\phi^{(n)}(0,T)) \\ &\quad = E\left[c(\phi^{(0)}(0,T))g(\phi^{(n)}(0,T))\right] \\ &\qquad + \int_0^T E\left[(\nabla_0 c)(\phi^{(0)}(0,T))g(\phi^{(n)}(0,T))D_{0,T} | F_s\right] D_{0,s}^{-1} X_i(s) \, dw_s^i \\ &\qquad + \int_0^T E\left[c(\phi^{(0)}(0,T))(\nabla_n g)(\phi^{(n)}(0,T))D_{0,T}^{(n)} | F_s\right] D_{0,s}^{(n)-1} X_i^{(n)}(s) \, dw_s^i. \end{aligned}$$

Taking the tensor product with $R_{0,T}^{(0)*}$ and the expected value, the result follows. □

REMARKS 4.8. To write out the preceding result in terms of $D_{0,t}$, $W_{0,t}$ and higher derivatives of the flow involves very involved calculations. Even in dimension 1 it seems better to introduce the sequence $\phi^{(n)}$ of flows. Note Theorem 2.5 implies $\sup_{s \le t} |D_{0,s}^{(n)}|$, $\sup_{s \le t} |M_{0,s}^{(n)}|$ are in every $L^p(\Omega)$. Theorem 4.7 is an integration by parts formula as only one term involves the gradient of derivatives $\nabla_0 c = c_\xi$ of c.

COROLLARY 4.9. *Taking* $g(\phi^{(1)}(0,T)) = M_{0,T}^{-1} D_{0,T}^{-1}$, *if* $M_{0,T}^{-1}$ *is in some* $L^p(\Omega)$,

$$\begin{aligned} E\left[c_\xi\left(\xi_{0,T}(x_0)\right)\right] &= E\left[c\left(\xi_{0,t}(x_0)\right) M_{0,T}^{-1} D_{0,T}^{-1} \otimes R_{0,T}^{(0)}\right] \\ &\quad - E\left[c\left(\xi_{0,T}(x_0)\right)(\nabla_1 g)\left(D_{0,T}, M_{0,t}\right) D_{0,T}^{(1)} M_{0,T}^{(1)}\right]. \end{aligned} \tag{4.2}$$

Because the remaining terms are integrable we have, therefore, proved the following result.

THEOREM 4.10. *Suppose* $\xi_{0,t}(x_0)$ *is the solution of* (2.1) *and c is a bounded smooth function with bounded derivatives. Then if* $M_{0,T}^{-1}$ *is in some* $L^p(\Omega)$,

$$\left|E\left[c_\xi\left(\xi_{0,T}(x_0)\right)\right]\right| \le K \sup_{x \in R^d} |c(x)|. \tag{4.3}$$

REMARKS 4.11. It is well known that inequality (4.3) implies that the random variable $\xi_{0,T}(x_0)$ has a density (see Malliavin [9] or Stroock [11]). The remaining question concerns the existence and integrability properties of $M_{0,T}^{-1}$. These have been carefully studied (see Malliavin [9], Stroock [11] and Norris [10]). In fact, it is known that $M_{0,T}^{-1}$ is in $L^p(\Omega)$ for all $p < \infty$ if the following condition H_1 of Hörmander on the coefficient vector fields $X_0, \dots, X_m$ of (2.1) is satisfied.

CONDITION H_1. $X_1, \dots, X_m, [X_i, X_j]$, for $i, j = 0, \dots, m$, $[X_i,[X_j, X_k]]$ for $i, j, k = 0, \dots, m$, etc. evaluated at $x_0 \in R^d$ span R^d.

Finally, recall that, if u is a nonsingular linear map of R^d to itself, then the map $\phi\colon u \to u^{-1}$ has a derivative $\phi'(u)$, which is a linear map on the space of linear maps of R^d to itself, given by $\phi'(u)h = -u^{-1}hu^{-1}$. Applying this to $g(D_{0,T}, M_{0,T}) = M_{0,T}^{-1} D_{0,T}^{-1}$, we have

$$\begin{aligned} &E\left[c_\xi\left(\xi_{0,T}(x_0)\right)\right] \\ &\quad = E\left[c\left(\xi_{0,T}(x_0)\right) M_{0,T}^{-1} D_{0,T}^{-1} \otimes R_{0,T}^{(0)}\right] \\ &\qquad + E\left[c\left(\xi_{0,T}(x_0)\right) M_{0,T}^{-1}\left(\left(\nabla_1 M_{0,T}\right)\left(D_{0,T}^{(1)} M_{0,T}^{(1)}\right)\right) M_{0,T}^{-1} D_{0,T}^{-1}\right] \\ &\qquad + E\left[c\left(\xi_{0,T}(x_0)\right) M_{0,T}^{-1} D_{0,T}^{-1}\left(\left(\nabla_1 D_{0,T}\right)\left(D_{0,T}^{(1)} M_{0,T}^{(1)}\right)\right) D_{0,T}^{-1}\right]. \end{aligned} \tag{4.4}$$

5. Bounds for higher derivatives. To show the density of $\xi_{0,T}(x_0)$ is differentiable, we must obtain bounds for higher derivatives of the form

$$\left|E\left[\frac{\partial^\alpha c}{\partial \xi^\alpha}\left(\xi_{0,T}(x_0)\right)\right]\right| \le K \sup_{x \in R^d} |c(x)|. \tag{5.1}$$

Here $\alpha = (\alpha_1, \dots, \alpha_d)$ is a multiindex of nonnegative integers and

$$\frac{\partial^{\alpha}}{\partial \xi^{\alpha}} = \frac{\partial^{\alpha_1}}{\partial \xi_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial \xi_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_d}}{\partial \xi_d^{\alpha_d}}.$$

In fact, a well-known argument from harmonic analysis (see [10]) implies that if (5.1) is true for all α with $|\alpha| = \alpha_1 + \cdots + \alpha_d \leq n$, where $n \geq d + 1$, then the random variable $\xi_{0,T}(x_0)$ has a density $d(x) = d(x_1, \dots, x_d)$ which is in $C^{n-d-1}(R^d)$.

To see how to proceed, apply (4.2) to c_ξ in place of c. [If preferred, (4.2) could be applied to just one partial derivative $\partial c / \partial \xi_k$ in place of c; however, the result of Corollary 4.9 is true for vector functions c.] This gives

$$\begin{aligned} E\left[c_{\xi\xi}\left(\xi_{0,T}(x_0)\right)\right] &= E\left[c_\xi\left(\xi_{0,T}(x_0)\right) M_{0,T}^{-1} D_{0,T}^{-1} \otimes R_{0,T}^{(0)}\right] \\ &\quad - E\left[c_\xi\left(\xi_{0,T}(x_0)\right)(\nabla_1 g)\left(D_{0,T}, M_{0,T}\right) D_{0,T}^{(1)} M_{0,T}^{(1)}\right]. \end{aligned} \tag{5.2}$$

Consider the two terms on the right,

$$E\left[c_\xi\left(\xi_{0,T}(x_0)\right) M_{0,T}^{-1} D_{0,T}^{-1} \otimes R_{0,T}\right] \tag{5.3}$$

and

$$E\left[c_\xi\left(\xi_{0,T}(x_0)\right)(\nabla_1 g)\left(D_{0,T}, M_{0,T}\right) D_{0,T}^{(1)} M_{0,T}^{(1)}\right]. \tag{5.4}$$

5.1 NOTATION. Write $M = M_{0,T}$, $D = D_{0,T}$, $D^{(1)} = D_{0,T}^{(1)}$, etc. Let $g_1(\phi^{(1)})$ be the function $M^{-1}D^{-1} \otimes RM^{-1}D^{-1}$ and $g_2(\phi^{(2)})$ be the function $(\nabla_1 g)(D, M) D^{(1)} M^{(1)} M^{-1} D^{-1}$.

Applying Theorem 4.7 to c and g_1, we have

$$\begin{aligned} &E\left[c\left(\xi_{0,T}(x_0)\right) g_1(\phi^{(1)}) \otimes R_{0,T}^{(0)}\right] \\ &\quad = E\left[c_\xi\left(\xi_{0,T}(x_0)\right) M_{0,T}^{-1} D_{0,T}^{-1} \otimes R_{0,T}^{(0)}\right] \\ &\quad\quad + E\left[c\left(\xi_{0,T}(x_0)\right)(\nabla_2 g_1)\left(\phi^{(2)}(0,T)\right) D_{0,T}^{(2)} M_{0,T}^{(2)}\right]. \end{aligned} \tag{5.5}$$

Applying Theorem 4.7 to c and g_2 we have

$$\begin{aligned} &E\left[c\left(\xi_{0,T}(x_0)\right) g_2(\phi^{(2)}) \otimes R_{0,T}\right] \\ &\quad = E\left[c_\xi\left(\xi_{0,T}(x_0)\right)(\nabla_1 g)\left(D_{0,T} M_{0,T}\right) D_{0,t}^{(1)} M_{0,T}^{(1)}\right] \\ &\quad\quad + E\left[c\left(\xi_{0,T}(x_0)\right)(\nabla_3 g_2)\left(\phi^{(3)}(0,T)\right) D_{0,T}^{(3)} M_{0,T}^{(3)}\right]. \end{aligned} \tag{5.6}$$

Substituting in (5.2), we obtain an expression on the right which involves only c and not its derivatives. This procedure can be iterated, using Theorem 4.7. At any stage, to replace a term of the form $E[c_\xi(\xi_{0,T}(x_0)) h(\phi^{(n)}(0,T))]$ by one involving only c define $\tilde{h}(\phi^{(n)}(0,T)) = h(\phi^{(n)}(0,T)) M_{0,T}^{-1} D_{0,T}^{-1}$ and apply Theorem 4.7. Clearly, higher powers of $M_{0,T}^{-1}$ are introduced at each iteration. [From

Theorem 2.5 $D_{0,T}^{-1}$ is in every $L^p(\Omega)$.] Hörmander's condition H_1 is sufficient to ensure that $M_{0,T}^{-1}$ is in every $L^p(\Omega)$, $1 \le p < \infty$. We have, therefore, proved the following result.

THEOREM 5.2. *Suppose Hörmander's condition* H_1 *is satisfied. Then the random variable* $\xi_{0,T}(x_0)$ *has a density* $d(x)$ *which is in* $C^\infty(R^d)$.

Acknowledgments. This paper was completed during a visit by Professor Kohlmann to the University of Alberta in the spring of 1987. Professor Kohlmann would like to express his thanks to the Department of Statistics and Applied Probability for its hospitality and support.

REFERENCES

[1] BICHTELER, K. and FONKEN, D. (1982). A simple version of the Malliavin calculus in dimension one. *Martingale Theory in Harmonic Analysis and Banach Spaces. Lecture Notes in Math.* **939** 6–12. Springer, New York.

[2] BISMUT, J. M. (1981). Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions. *Z. Wahrsch. verw. Gebiete* **56** 469–505.

[3] BISMUT, J.-M. (1981). *Mécanique aléatoire. Lecture Notes in Math.* **866**. Springer, New York.

[4] CARVERHILL, A. P. and ELWORTHY, K. D. (1983). Flows of stochastic dynamical systems: The functional analytic approach. *Z. Wahrsch. verw. Gebiete* **65** 245–267.

[5] ELLIOTT, R. J. (1982). *Stochastic Calculus and Applications*. Springer, New York.

[6] ELLIOTT, R. J. and KOHLMANN, M. (1987). Stochastic integral representation for non-Markov processes. Unpublished.

[7] ELLIOTT, R. J. and KOHLMANN, M. (1988). Martingale representation and large deviations. Unpublished.

[8] KUNITA, H. (1982). Stochastic partial differential equations connected with nonlinear filtering. *Nonlinear Filtering and Stochastic Control. Lecture Notes in Math.* **972** 100–169. Springer, New York.

[9] MALLIAVIN, P. (1978). Stochastic calculus of variations and hypoelliptic operators. *Proc. of the International Symposium on Stochastic Differential Equations, Kyoto, 1976* (K. Itô, ed.) 195–263. Wiley, New York.

[10] NORRIS, J. (1986). Simplified Malliavin calculus. *Séminaire de Probabilités XX, 1984/85. Lecture Notes in Math.* **1204**. Springer, New York.

[11] STROOCK, D. (1983). Some applications of stochastic calculus to partial differential equations. *Ecole d'Eté de Probabilités de Saint-Flour XI—1981. Lecture Notes in Math.* **976** 268–382. Springer, New York.

[12] ZAKAI, M. (1985). The Malliavin calculus. *Acta Appl. Math.* **3** 175–207.

DEPARTMENT OF STATISTICS AND
APPLIED PROBABILITY
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA
CANADA T6G 2G1

FAKULTÄT FÜR WIRTSCHAFTS-WISSENSCHAFTEN
UND STATISTIK
UNIVERSITÄT KONSTANZ
D7750 KONSTANZ
FEDERAL REPUBLIC OF GERMANY