Functionals of diffusion processes as stochastic integrals

By M. H. A. DAVIS

Department of Computing and Control,
Imperial College, 180 Queen's Gate, London SW7 2BZ

(Received 19 December 1978)

1. Introduction. Let $\xi = \{x_t, 0 \leq t \leq 1\}$ be a standard $d$-dimensional Brownian motion and $L$ a smooth functional on the space $C = C_0^d[0, 1]$ of continuous functions from $[0, 1]$ to $\mathbb{R}^d$, such that $E(L(\xi))^2 < \infty$. Denote $F_t = \sigma(x_s, 0 \leq s \leq t)$. Then

$$\eta_t = E[L(\xi)|F_t]$$

is a square integrable martingale which can be represented as a stochastic integral; consequently the random variable $L(\xi)$ has the representation

$$L(\xi) = E[L(\xi)] + \int_0^1 \xi_t \, d\xi_t,$$  \hspace{1cm} (1.1)

For Fréchet differentiable functions $L$, Clark (1) gave the following formula for the integrand $\xi$:

$$\xi_t = E[\mu_x(1) - \mu_x(0)|F_t].$$

Here $\mu_x(.)$ is the function of bounded variation corresponding to the Fréchet derivative of $L$ at $x \in C$.

In (3) (4), Haussmann extended these results to cover the case where $\xi$ is an Itô process, i.e. the solution of a stochastic differential equation

$$d\xi_t = f(t, \xi_t) \, dt + \sigma(t, \xi_t) \, dw_t, \quad \xi_0 = 0,$$  \hspace{1cm} (1.2)

where $w$ is a vector Brownian motion and $f(t, x), \sigma(t, x)$ are smooth non-anticipative functions of $x \in C$. Then $L(\xi)$ has the representation

$$L(\xi) = E[L(\xi)] + \int_0^1 \xi_t \, dw_t,$$  \hspace{1cm} (1.3)

and the formula for $\xi$ is

$$\xi_t = E \left[ \int_{\mu_{t, 1}} \mu_t(\sigma[s, t] | F_s) \sigma(s, \xi) \right],$$  \hspace{1cm} (1.4)

where $\Phi$ is the fundamental matrix solution of equation (1.2) linearized about the path $\xi$ (see (2.7)). The proof in (3) involved applying Clark's original result to a discretized version of (1.2) and then taking limits, whereas in (4), a very neat perturbation argument involving the Girsanov transformation was used. In this paper we give, for the case where $\xi$ is a diffusion process, a completely different proof based on potential-theoretic ideas. While this is certainly no neater than the argument in (4), it is rather transparent in that one sees exactly why the integrand has to be what it is.

\[\textit{+ Work supported by the U.S. Air Force Office of Sponsored Research under Grant AFOSR 77-3281 and by the National Science Foundation, Grant ENG 76-10440.}\]
The method is based on the following very simple idea. Consider the scalar case where $\gamma$ is the solution of
\begin{equation}
\frac{d\gamma}{dt} = f(\gamma) dt + \sigma(\gamma) dw(t),
\end{equation}
$\gamma_0 = 0$.
Here $f, \sigma$ are $C^2$ functions with bounded derivatives and $\sigma^2(\xi) \geq \delta > 0$ for all $\xi \in \mathbb{R}$.
Suppose $L(x) = l(x)$ for some $C^1$ function $l$. Let $V(t, \xi)$ be the solution of the parabolic equation
\begin{equation}
V_t + \frac{1}{2} \sigma^2(\xi) V_{\xi\xi} + f(\xi) V_{\xi} = 0, \quad (t, \xi) \in [0, 1] \times \mathbb{R},
\end{equation}
$V(1, \xi) = l(\xi)$.
(Here $V_t = \partial V/\partial t$, etc.) Expanding $V(t, \xi)$ by the Ito rule shows that
\begin{equation}
V(t, \xi) = E_t l(\xi_0) + \int_0^t V_t(s, \xi_s) \sigma(\xi_s) dw(s),
\end{equation}
and gives a stochastic integral representation for $L(x)$ in the form
\begin{equation}
L(x) = l(x) + \int_0^1 V(s, \xi_s) \sigma(\xi_s) dw(s).
\end{equation}
We can show that this is equivalent to (1.3), (1.4) as follows. Differentiating (1.6) with respect to $t$ shows that
\begin{equation}
V_t + \frac{1}{2} \sigma^2(\xi) V_{\xi\xi} + f(\xi) V_{\xi} = 0,
\end{equation}
$W(1, \xi) = l(\xi)$.
Now let $\eta_t$ be the solution of equation (1.5) linearized about $\gamma$, i.e.
\begin{equation}
d\eta_t = f(\eta_t) dt + \sigma(\eta_t) dw_t,
\end{equation}
Apply the Ito rule to compute the product $\eta_t W(t, \xi_t)$; this gives
\begin{equation}
d(\eta_t) W(t, \xi_t) = \sigma(\eta_t) W_{\xi}(t, \xi_t) dw_t.
\end{equation}
Thus, taking the initial condition $\eta_0 = 1$, we get
\begin{equation}
\Phi(t, \xi) = W(t, \xi) = E_t l(\xi_0) + \int_0^t \sigma(\eta_s) W_{\xi}(s, \xi_s) dw_s.
\end{equation}
Now $\Phi(x, t) = \eta_t$ is precisely the fundamental solution of the linearized equation, and the Fréchet derivative of $L(x) = l(x_1)$ is
\begin{equation}
\mu_t = \left[ l(x_1) \delta_t \right](dx_1),
\end{equation}
where $\delta_t$ is the Dirac measure at $t$. Thus the integrand $V_t \sigma$ coincides with $e$ of (1.4).
The argument is even more direct if $\sigma$ is constant, for it is a standard result (2, Theorem 1.2.3) that the solution of (1.7) with $\sigma_2 = 0$ is
\begin{equation}
W(t, \xi) = E_t \left[ e^{\int_0^t f(\xi_s) ds} \right]
\end{equation}
and the exponential term is, of course, the fundamental solution of (1.8) in this case.
and \( ||\chi|| \leq TV(\mu_2) \) (the total variation of \( \mu_2 \)). We shall denote by \( \mu_x(t) \) the right-
continuous bounded-variation function \( \mu_x((0, t]) \). We make the following assumptions on the functional \( L \).

(i) \( L \) is a Fréchet differentiable functional, with Fréchet derivative \( \mu_x \) at \( x \in C \).

(ii) There exist positive integers \( K, \gamma \) such that

\[
|\mu_x(x)| + TV(\mu_x) \leq K(1 + ||x||)^{\gamma},
\]

(2.4)

(iii) \( \mu_x \) is continuous in \( x \) in the weak topology. Note that under conditions (2.4)

\[
E(L(\chi))^2 < \infty.
\]

so that \( L(\chi) \) certainly has a representation of the form (1.3).

**Theorem 2.5.** Suppose conditions (2.2), (2.4) are satisfied. Then

\[
L(\chi) = E[L(\chi)] + \int_0^1 c_x \, dw_x \quad {\text{a.s.}}
\]

where the integrand \( c_x \) satisfies \( E \int_0^1 |c_x|^2 \, dt < \infty \) and is given by

\[
c_x = E \left[ \int_{[u, 1]} \mu_x(ds) \Phi(s, x) \sigma(t, x) \right].
\]

(2.7)

Here \( \Phi(t, s) \) is the matrix-valued process defined for \( 0 \leq s < t \leq 1 \) by

\[
d_t \Phi(t, s) = A(t, x) \Phi(t, s) \, dt + B(t, x) \Phi(t, s) \, dw_t
\]

(2.8)

\[
\Phi(s, s) = I_d \quad (\text{identity matrix})
\]

(2.9)

Note that (2.8) is equation (2.1) linearized about the random trajectory \( \xi \).

The proof proceeds in the stages outlined in the introduction. In the next section,
the result is established for the 'finite-dimensional' case. In Section 4, we consider
approximation of Fréchet-differentiable functions and use these approximations in
Section 5 to complete the proof.

**3. The finite dimensional case.**

Suppose \( L \) is of the form

\[
L(\chi) = L(x_{t_0}, x_{t_1}, \ldots, x_{t_n}),
\]

where \( 0 = t_0 < t_1 < \ldots < t_n = 1 \). We assume that \( L: R^{n+1d} \to R \) is continuously differentiable and that

\[
|L(\theta)| + \sum_{i=1}^d |\theta_i| \leq K(1 + |\theta|)^{\gamma}, \quad \theta \in R^{n+1d}
\]

(3.1)

Then (2.3) is satisfied. We can obtain a stochastic integral representation for \( L(\chi) \) in this case using the method outlined in Section 1.
This gives a stochastic integral representation for \( L(\xi) \); indeed, as in (3.6) we have, for each \( k \),
\[
V^k-1(\xi_{t_0}, \ldots, \xi_{t_n}, t_{k-1}) - V^{k-1}(\xi_{t_0}, \ldots, \xi_{t_n}, t_{k-1}) = \int_{t_{k-1}}^{t_k} (V^k - 1)'(\xi, \sigma) \, dw,
\]
and hence, in view of the boundary condition (3.7), summing over \( k \) gives
\[
L(\xi) = EL(\xi) + \int_0^t \dot{\xi}_t \, dt,
\]
where
\[
\dot{\xi}_t = (V^k - 1)'(\xi_{t_0}, \ldots, \xi_{t_n}, t_k) \sigma(t_k, \xi_k), \quad t \in [t_{k-1}, t_k].
\]

We now show that this integrand coincides with the one given by (2.8). Now, the Frechet derivative of \( L \) at \( \xi \in C \) is
\[
\mu_\xi = \sum \mu_t(\xi_0, \ldots, \xi_t) \, dt,
\]
and hence the integrand given by (2.7) is
\[
\dot{\xi}_t = \sum \mu_t(\xi_0, \ldots, \xi_t) \dot{\xi}_t.
\]

**Proposition 3-8.** \( \dot{\xi}_t = \dot{\xi}_t \) a.s. for each \( t \in [0, 1] \).

Take first \( t \in [t_{n-1}, t_n] \), regard \( \xi_0 \ldots, \xi_{n-1} \) temporarily as fixed constants and denote
\[
V(\gamma, t) = V^{n-1}(\xi_0, \ldots, \xi_{n-1}, \gamma, t),
\]
so that
\[
V(\gamma, t) = E_k[I(\xi_0 \ldots, \xi_{n-1}, \gamma)]
\]
and hence the integrand given by (2.7) is
\[
\dot{\xi}_t = \sum E_k[I(\xi_0 \ldots, \xi_t) \Phi(t_k, \xi_k) | F_t] \sigma(t_k, \xi_k).
\]

**Proposition 3-8.** \( \dot{\xi}_t = \dot{\xi}_t \) a.s. for each \( t \in [0, 1] \).

Take first \( t \in [t_{n-1}, t_n] \), regard \( \xi_0 \ldots, \xi_{n-1} \) temporarily as fixed constants and denote
\[
V(\gamma, t) = V^{n-1}(\xi_0, \ldots, \xi_{n-1}, \gamma, t),
\]
so that
\[
V(\gamma, t) = E_k[I(\xi_0 \ldots, \xi_{n-1}, \gamma)]
\]
and hence the integrand given by (2.7) is
\[
\dot{\xi}_t = \sum E_k[I(\xi_0 \ldots, \xi_t) \Phi(t_k, \xi_k) | F_t] \sigma(t_k, \xi_k).
\]

**Theorem 4-2.**

(i) \( L \) is continuously differentiable for each \( n \).
Using (4.3) shows that the other two cases.

(i.e. \( L_n \) satisfies (2.3)).

(iii) \( v^2 \rightarrow \mu_x \) weakly as \( n \rightarrow \infty \), for each \( x \in C \).

Proof. We give the proof for the case \( d = 1 \). For \( d > 1 \) the argument is the same but notionally rather cumbersome.

We first show that the derivatives of \( L_n \) are given by the following formula, from which statement (i) in the Theorem is immediate.

\[
\frac{\partial}{\partial x_k} L_n(x_0, x_1, \ldots, x_n) = \begin{cases} 
\frac{1}{h} \int_0^h v(t) dt, & k = 0, \\
\frac{1}{h} \left( \int_{(k-1)h}^{kh} v(t) dt - \int_{kh}^{(k+1)h} v(t) dt \right), & k = 1, 2, \ldots, 2^n - 1.
\end{cases}
\]  

(4.3)

Here \( v(t) = \mu_{\beta_n}(0, t) \) and \( \xi_k = x_k, k = 0, 1, \ldots, 2^n. \) Indeed, for \( x, y \in C \) we have

\[
L_n(x + y) = L(x + y).
\]  

(4.4)

Fix \( \alpha \in R \) and a positive integer \( k \in [1, 2^n - 1] \) and choose \( y \) to be any continuous function such that \( y(t) = 0 \) for \( t \in [(k-1)h, kh] \) and \( y(t) = \alpha \) for \( t \in [kh, (k+1)h] \). Then

\[
L_n(x + y) = L_n(x) + \int_{(0, 1)} \beta_n(y) \mu_{\beta_n}(ds) + o(||y||). \]

(4.5)

On the other hand

\[
\beta_n(y) = \begin{cases} 
\frac{\alpha}{h} (t - (k-1)h), & t \in [(k-1)h, kh], \\
\frac{-\alpha}{h} (t - (k+1)h), & t \in [kh, (k+1)h], \\
0, & \text{elsewhere}
\end{cases}
\]

and an integration by parts shows that

\[
\int_{(0, 1)} \beta_n(y) v(dt) = -\frac{\alpha}{h} \left( \int_{(k-1)h}^{kh} v(t) dt - \int_{kh}^{(k+1)h} v(t) dt \right).
\]

The central formula in (4.3) follows from this and (4.4) and a similar argument gives the other two cases.

In view of the representation (4.1) the Fréchet derivative of \( L_n \) at \( x \in C \) is

\[
v_n^2(ds) = \sum_{k=0}^N \frac{\partial}{\partial x_k} L_n(x_0, x_1, \ldots, x_n) \delta_{x_k}(ds).
\]

Using (4.3) shows that

\[
v_n^2(t) = v_n^2(0, t) = \frac{1}{h} \int_{kh}^{(k+1)h} \mu_{\beta_n}(s) ds, t \in [kh, (k+1)h].
\]

(4.6)
Now in view of conditions (2.2) and part (ii) of Theorem 4.2
\[
\left| \int_t^1 v_n^2(ds) \Phi(s, t) \sigma(s, x_s) \right| \leq \sup_{\ell < n} \left| \Phi(s, t) \right| \left| \sigma(s, x_s) \right| TV(v_n^2) \\
\leq \sup_{\ell < n} \left| \Phi(s, t) \right| K \|1 + \|_\ell^2\|^2.
\]

Hence from (5-2) and the Schwarz inequality
\[
E \left[ \left| \int_t^1 v_n^2(ds) \Phi(s, t) \sigma(t, x_t) \right|^p \right] < K_1
\]
for some constant $K_1$. This shows that the set \( \left| \int_t^1 v_n^2(ds) \Phi(s, t) \sigma(t, x_t) \right|^2, n = 1, 2 \ldots \) is a uniformly integrable subset of $L_2(\Omega \times [0,1], d\mathcal{P}^\varepsilon dt)$, and therefore in view of (5-3) and Jensen's inequality we have\(
E \left[ \left| \int_t^1 v_n^2(ds) \Phi(s, t) \sigma(t, x_t) \right|^p \right] dt \\
\leq E \left[ \left| \int_0^1 \left( \int_t^1 v_n^2(ds) \Phi(s, t) - \int_t^1 \mu_n(ds) \Phi(s, t) \right) \sigma(t, x_t) \right|^p \right] dt \\
\to 0 \quad \text{as} \quad n \to \infty.
\)

This shows that the stochastic integral in (5-1) converges in $L_2$ to the corresponding integral in (2-6), (2-7). Therefore a subsequence converges almost surely and the representation (2-6), (2-7) is established.

6. Concluding remarks

In (3), (4), Haussmann proves Theorem (2-5) for a process $\xi$ satisfying\[
d_{\xi(t)} = f(t, \xi(t)) dt + \sigma(t, \xi(t)) d\mathcal{W}(t),\]
where $f, \sigma : [0,1) \times \mathbb{R} \to \mathbb{R}$ are non-anticipative Fréchet-differentiable functionals. It seems likely that our approach could be extended to cover this case by introducing some form of finite-dimensional approximation for $f, \sigma$ (though this would have to be done with some care – for example $f_n(t, x) = f(t, \beta_n(x))$ is no longer non-anticipative). However, it also seems likely that the details would be very complex.

REFERENCES


