Math. Proc. Camb. Phil. Soc. (1980), **87**, 157
Printed in Great Britain

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 $\mathfrak{x} = \{\mathfrak{x}_t, 0 \leqslant t \leqslant 1\}$ be a standard d-dimensional Brownian motion and L a smooth functional on the space $C = C^d[0, 1]$ of continuous functions from $[0, 1]$ to R^d, such that $E(L(\mathfrak{x}))^2 < \infty$. Denote $F_t = \sigma\{\mathfrak{x}_s, 0 \leqslant s \leqslant t\}$. Then

$$\mathfrak{y}_t = E[L(\mathfrak{x})|F_t]$$

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

$$L(\mathfrak{x}) = EL(\mathfrak{x}) + \int_0^1 \mathfrak{e}_s \, d\mathfrak{x}_s. \tag{1·1}$$

For Fréchet differentiable functions L, Clark (1) gave the following formula for the integrand $\mathfrak{e}$:

$$\mathfrak{e}_t = E[\mu_{\mathfrak{x}}(1) - \mu_{\mathfrak{x}}(t)|F_t].$$

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 $\mathfrak{x}$ is an Ito process, i.e. the solution of a stochastic differential equation

$$d\mathfrak{x}_t = f(t, \mathfrak{x})\, dt + \sigma(t, \mathfrak{x})\, d\mathfrak{w}, \quad \mathfrak{x}_0 = 0, \tag{1·2}$$

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

$$L(\mathfrak{x}) = EL(\mathfrak{x}) + \int_0^1 \mathfrak{e}_t \, d\mathfrak{w}_t \tag{1·3}$$

and the formula for $\mathfrak{e}$ is

$$\mathfrak{e}_t = E\left[\int_{]t,1]} \mu_{\mathfrak{x}}(ds)\, \Phi(s, t)\Big|F_t\right] \sigma(t, \mathfrak{x}), \tag{1·4}$$

where Φ is the fundamental matrix solution of equation (1·2) linearizedmpa about les path $\mathfrak{x}$ (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 $\mathfrak{x}$ 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.

† 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.

0305-0041/80/0000-6650 $03.50 © 1980 Cambridge Philosophical Society

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press

The method is based on the following very simple idea. Consider the scalar case where $\mathfrak{x}$ is the solution of

$$d\mathfrak{x}_t = f(\mathfrak{x}_t)\,dt + \sigma(\mathfrak{x}_t)\,d\mathfrak{w}_t, \qquad (1{\cdot}5)$$
$$\mathfrak{x}_0 = 0.$$

Here f, σ are C^1 functions with bounded derivatives and $\sigma^2(\xi) \geqslant \delta > 0$ for all $\xi \in R$. Suppose $L(x) = l(x_1)$ for some C^1 function l. Let $V(t, \xi)$ be the solution of the parabolic equation

$$V_t + \tfrac{1}{2}\sigma^2(\xi)\,V_{\xi\xi} + f(\xi)\,V_\xi = 0, \ (t, \xi) \in [0, 1] \times R, \qquad (1{\cdot}6)$$
$$V(1, \xi) = l(\xi).$$

(Here $V_\xi = \partial V/\partial \xi$, etc.) Expanding $V(t, \mathfrak{x}_t)$ by the Ito rule shows that

$$V(t, \xi) = E_{t,\xi}[l(\mathfrak{x}_1)]$$

and gives a stochastic integral representation for $L(\mathfrak{x})$ in the form

$$L(\mathfrak{x}) = l(\mathfrak{x}_1) = EL(\mathfrak{x}) + \int_0^1 V_\xi(t, \mathfrak{x}_t)\,\sigma(\mathfrak{x}_t)\,d\mathfrak{w}_t.$$

We can show that this is equivalent to (1·3), (1·4) as follows. Differentiating (1·6) with respect to ξ shows that $W = V_\xi$ satisfies

$$W_t + \tfrac{1}{2}\sigma^2 W_{\xi\xi} + W_\xi(\sigma\sigma_\xi + f) + Wf_\xi = 0, \qquad (1{\cdot}7)$$
$$W(1, \xi) = l_\xi(\xi).$$

Now let $\mathfrak{y}_t$ be the solution of equation (1·5) linearized about $\mathfrak{x}$, i.e.

$$d\mathfrak{y}_t = f_\xi(\mathfrak{x}_t)\,\mathfrak{y}_t\,dt + \sigma_\xi(\mathfrak{x}_t)\,\mathfrak{y}_t\,d\mathfrak{w}_t. \qquad (1{\cdot}8)$$

Apply the Ito rule to compute the product $\mathfrak{y}_t\,W(t, \mathfrak{x}_t)$; this gives

$$d(\mathfrak{y}_t\,W(t,\ \mathfrak{x}_t)) = \mathfrak{y}(\sigma_\xi\,W + \sigma W_\xi)\,d\mathfrak{w}_t.$$

Thus, taking the initial condition $\mathfrak{y}_t = 1$, we get

$$V_\xi(t, \xi) = W(t, \xi) = E_{t,\xi}[\mathfrak{y}_1\,l_\xi(\mathfrak{x}_1)].$$

Now $\Phi(s, t) = \mathfrak{y}_s$ is precisely the fundamental solution of the linearized equation, and the Fréchet derivative of $L(x) = l(x_1)$ is

$$\mu_x(ds) = l_\xi(x_1)\,\delta_1(ds),$$

where δ_t is the Dirac measure at t. Thus the integrand $V_\xi\,\sigma$ coincides with $\mathfrak{e}$ of (1·4).

The argument is even more direct if σ is constant, for it is a standard result (2, Theorem 1·2·3) that the solution of (1·7) with $\sigma_\xi = 0$ is

$$W(t, \xi) = E_{t,\xi}\left[l_\xi(\mathfrak{x}_1)\exp\left(\int_t^1 f_\xi(\mathfrak{x}_s)\,ds\right)\right]$$

and the exponential term is, of course, the fundamental solution of (1·8) in this case.

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press

The purpose of this paper is to show that this method can be used to obtain the formula (1·4) for general functionals L (precise conditions are given below in Section 2). In Section 3, the argument is extended to functionals of the form

$$L(x) = l(x_{t_0}, x_{t_1} \dots x_{t_n}).$$

We then consider, in Section 4, the approximation of a general functional $L(x)$ by functionals $L_n(x) = L[\beta_n(x)]$, where $\beta_n(x)$ is the linear interpolation of x with grid $1/2^n$. These results are then combined with those of Section 3 to complete the proof.

2. *Formulation and assumptions.* [*Notation*: Throughout the paper x, y denote generic elements of $C = C^d[0, 1]$ with coordinates x_t, y_t, $t \in [0, 1]$. The norm of $x \in C$ is $\|x\| = \sup_t |x_t|$. Generic points of R^d are ξ, η with components ξ^i, η^i, $i = 1, 2, \dots d$. ξ' denotes the transpose of a vector ξ. Gradients of functions are denoted by subscripts: for example, $[l_{\xi_k}(\xi_0, \xi_1 \dots, \xi_n)]^i = (\partial l/\partial \xi_k^i)\ (\xi_0, \dots, \xi_n)$. The convention of implied summation from 1 to d over repeated indices is used. Finally, *processes* are denoted by gothic script: $\mathfrak{x}$, $\mathfrak{y}$, etc.]

A d-dimensional process $\mathfrak{x} = \{\mathfrak{x}_t, 0 \leqslant t \leqslant 1\}$ is fixed throughout the paper and is the unique solution of the Ito equation

$$d\mathfrak{x}_t = f(t, \mathfrak{x}_t)\, dt + \sigma(t, \mathfrak{x}_t)\, d\mathfrak{w}_t, \qquad (2\cdot1)$$
$$\mathfrak{x}_0 = 0.$$

Here $\mathfrak{w}$ is a standard d-dimensional Brownian motion and the functions f, σ have the following properties (F denotes any of the components f^i, σ^{ij})

(i) $F : [0, 1] \times R^d \to R$ is bounded and measurable.

(ii) $F(t, \xi)$ is C^1 in ξ with bounded derivative (K will denote a common bound for (i) and (ii)).

(iii) $$a(t, \xi) \geqslant c_1 I \quad \text{for all} \quad (t, \xi), \quad \text{for some} \quad c_1 > 0 \qquad (2\cdot2)$$

(here $a = \sigma\sigma'$ and I is the $d \times d$ identity matrix).

(iv) The functions a^{ij} are Hölder continuous in ξ, uniformly in $(t, \xi) \in [0, 1] \times R^d$.

These conditions are rather restrictive. They are imposed to enable us to use results from Friedman (2, § 6·5) on solutions of partial differential equations and their probabilistic interpretations. Haussmann (3), (4) imposes somewhat weaker conditions.

Conditions (2·2) imply the standard Ito conditions (see Theorem 5·1·1 of (2)) so there is a unique process $\mathfrak{x}$ satisfying (2·1) and according to (2, Theorem 5·2·3),

$$\sup_{t \in [0, 1]} E|\mathfrak{x}_t|^m < \infty \qquad (2\cdot3)$$

for any positive integer m.

Let $L : C \to R$ be a fixed Fréchet differentiable functional. Then there exists a map $\psi : C \to C^*$ such that, for $x, y \in C$,

$$L(x+y) = L(x) + \psi(x)\,(y) + o(\|y\|).$$

According to the Riesz representation theorem there is, for each $x \in C$, an R^d-valued Radon measure μ_x such that

$$\psi(x)\,(y) = \int_{[0, 1]} y(s)\, \mu_x(ds)$$

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press

and $\|\psi(x)\|_{C^*} = TV(\mu_x)$ (the total variation of μ_x). 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 μ_x at $x \in C$.

(ii) There exist positive integers K, q such that

$$|L(x)| + TV(\mu_x) \leqslant K(1 + \|x\|)^q. \tag{2·4}$$

(iii) μ_x is continuous in x in the weak topology. Note that under conditions (2·4)

$$E(L(\mathfrak{x}))^2 < \infty.$$

so that $L(\mathfrak{x})$ certainly has a representation of the form (1·3).

THEOREM 2·5. *Suppose conditions* (2·2), (2·4) *are satisfied. Then*

$$L(\mathfrak{x}) = EL(\mathfrak{x}) + \int_0^1 \mathfrak{e}_s \, d\mathfrak{w}_s \quad \text{a.s.} \tag{2·6}$$

where the integrand $\mathfrak{e}_t$ *satisfies* $E\int_0^1 |\mathfrak{e}_t|^2 dt < \infty$ *and is given by*

$$\mathfrak{e}_t = E\left[\int_{]t,1]} \mu_{\mathfrak{x}}(ds)\, \Phi(s,t) \,\middle|\, F_t\right] \sigma(t, \mathfrak{x}_t). \tag{2·7}$$

Here $\Phi(t,s)$ *is the matrix-valued process defined for* $0 \leqslant s \leqslant t \leqslant 1$ *by*

$$d_t \Phi(t,s) = A(t, \mathfrak{x}_t)\, \Phi(t,s)\, dt + B^l(t, \mathfrak{x}_t)\, \Phi(t,s)\, d\mathfrak{w}_t^l \tag{2·8}$$

$$\Phi(s,s) = I_d \,(= \textit{identity matrix}) \tag{2·9}$$

$$[A(t,\xi)]^{ij} = f^i_{\xi^j}(t,\xi)$$

$$[B^l(t,\xi)]^{ij} = \sigma^{il}_{\xi^j}(t,\xi).$$

Note that (2·8) is equation (2·1) linearized about the random trajectory $\mathfrak{x}$.

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

$$L(x) = l(x_{t_0}, x_{t_1}, \ldots, x_{t_n}),$$

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

$$|l(\theta)| + \sum_{i=1}^{nd} |l_{\theta^i}(\theta)| \leqslant K(1 + |\theta|)^q, \quad \theta \in R^{(n+1)d} \tag{3·1}$$

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.

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press

Let $\xi_0, \xi_1 \dots \xi_{n-1}, \eta \in R^d$ and $t \in [t_{n-1}, t_n]$ and let $V^{n-1}(\xi_0 \dots \xi_{n-1}, \eta, t)$ be the solution of the parabolic equation

$$\frac{\partial}{\partial t} V^{n-1} + \tfrac{1}{2} a^{ij}(t,\eta) \frac{\partial^2 V^{n-1}}{\partial \eta^i \partial \eta^j} + f^i(t,\eta) \frac{\partial V^{n-1}}{\partial \eta^i} = 0, \quad (t,\eta) \in [t_{n-1}, t_n] \times R^d \tag{3·2}$$

$$V^{n-1}(\xi_0, \dots \xi_{n-1}, \eta, t_n) = l(\xi_0, \dots, \xi_{n-1}, \eta) \tag{3·3}$$

(V^{n-1} and its derivatives are evaluated at $(\xi_0, \dots \xi_{n-1}, \eta, t)$ in (3·2), and

$$a(t,\eta) = \sigma(t,\eta)\, \sigma'(t,\eta)).$$

Under conditions (2·2), the equation (3·2), (3·3) has a unique solution V^{n-1} satisfying

$$|V^{n-1}(\xi_0, \dots, \xi_{n-1}, \eta, t)| \leqslant K(\xi_0, \dots, \xi_{n-1})\,(1 + |\eta|)^q \tag{3·4}$$

(see (**2**, pp. 147–148)).

LEMMA 3·5

(i) *For each* $(\xi_0, \dots, \xi_{n-1}, \eta) \in R^{(n+1)d}$ *and* $t \in (t_{n-1}, t_n)$

$$V^{n-1}(\xi_0, \dots \xi_{n-1}, \eta, t) = E_{t,\eta}[l(\xi_0, \dots \xi_{n-1}, \mathfrak{x}_{t_n})],$$

(ii) V^{n-1} *is continuously differentiable in* ξ_j *for* $j = 0, 1 \dots n-1$ *and*

$$V^{n-1}_{\xi_j}(\xi_0 \dots \xi_{n-1}, \eta, t) = E_{t,\eta}[l_{\xi_j}(\xi_0 \dots \xi_n, \mathfrak{x}_{t_n})],$$

(iii) $V^{n-1}(\mathfrak{x}_{t_0} \dots, \mathfrak{x}_{t_{n-1}}, \mathfrak{x}_t, t) = E[l(\mathfrak{x}_{t_0} \dots, \mathfrak{x}_{t_n}) | F_t]$ a.s.

Proof. (i) Applying the Ito rule to calculate $V^{n-1}(\xi_0 \dots \xi_{n-1}, \mathfrak{x}_t, t)$ for $t \in [t_{n-1}, t_n]$, we find that

$$d_t\, V^{n-1}(\xi_0 \dots \xi_{n-1}, \mathfrak{x}_t, t) = (V^{n-1}_\eta)'\, \sigma dw. \tag{3·6}$$

It follows as in (**2**, Theorem 6·5·3) that the stochastic integral is a martingale, and this gives (i) in view of the terminal condition (3·3). Part (ii) follows from the dominated convergence theorem, using (2·3) and (3·1). Part (iii) is a consequence of the Markov property of $\mathfrak{x}$, given that $V^{n-1}(\mathfrak{x}_{t_0}, \dots \mathfrak{x}_t, t)$ is, from part (ii), a measurable function of $(\mathfrak{x}_{t_0}, \dots, \mathfrak{x}_{t_n}, t)$.

From (3·5) (iii) we have, in particular, that

$$V^{n-1}(\mathfrak{x}_{t_0}, \dots, \mathfrak{x}_{t_{n-1}}, \mathfrak{x}_{t_{n-1}}, t_{n-1}) = E[l(\mathfrak{x}_{t_1} \dots \mathfrak{x}_{t_n}) | F_{t_{n-1}}].$$

Now proceed by induction. Supposing that $V^k : R^{(k+1)d} \times [t_k, t_{k+1}] \to R$ has been defined, define $V^{k-1}(\xi_0 \dots \xi_{k-1}, \eta, t)$ for $t \in [t_{k-1}, t_k]$ as the solution of (3·2) (with V^{k-1} and derivatives evaluated at $(\xi_0 \dots \xi_{k-1}, \eta, t)$ replacing V^{n-1}) with terminal condition

$$V^{k-1}(\xi_0 \dots \xi_{k-1}, \eta, t_k) = V^k(\xi_0 \dots \xi_{k-1}, \eta, \eta, t_k).$$

This defines V^k for $k = n-1, n-2 \dots, 0$, and an argument similar to the above using the iteration property of conditional expectations shows that for $t \in [t_{k-1}, t_k]$

$$V(\mathfrak{x}_{t_0}, \dots, \mathfrak{x}_{t_{k-1}}, \mathfrak{x}_t, t) = E[l(\mathfrak{x}_{t_0}, \dots, \mathfrak{x}_{t_n}) | F_t].$$

In particular

$$V(0,0) = E[l(\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_n})].$$

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press

This gives a stochastic integral representation for $L(\mathfrak{x})$; indeed, as in (3·6) we have, for each k,

$$V^{k-1}(\mathfrak{x}_{t_0}, \mathfrak{x}_{t_{k-1}}, \mathfrak{x}_{t_k}, t_k) - V^{k-1}(\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_{k-1}}, \mathfrak{x}_{t_{k-1}}, t_{k-1}) = \int_{t_{k-1}}^{t_k} (V_\eta^{k-1})' \sigma \, d\mathfrak{w}$$

and hence, in view of the boundary condition (3·7), summing over k gives

$$L(\mathfrak{x}) = EL(\mathfrak{x}) + \int_0^1 \mathfrak{s}_t \, d\mathfrak{w}_t,$$

where $$\mathfrak{s}_t = (V_\eta^{k-1}(\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_{k-1}}, \mathfrak{x}_t, t))' \sigma(t, \mathfrak{x}_t), \quad t \in [t_{k-1}, t_k[.$$

We now show that this integrand coincides with the one given by (2·8). Now, the Fréchet derivative of L at $x \in C$ is

$$\mu_x(ds) = \sum_i l_{\xi_i}(x_{t_0} \dots x_{t_n}) \, \delta_{t_i}(ds)$$

and hence the integrand given by (2·7) is

$$\mathfrak{e}_t = \sum_{t_i \geqslant t} E[l_{\xi_i}(\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_n}) \, \Phi(t_i, t) | F_t] \, \sigma(t, \mathfrak{x}_t).$$

PROPOSITION 3·8. $\mathfrak{s}_t = \mathfrak{e}_t$ a.s. *for each* $t \in [0, 1]$.

Take first $t \in [t_{n-1}, t_n]$, regard $\xi_0 \dots \xi_{n-1}$ temporarily as fixed constants and denote $V(\eta, t) = V^{n-1}(\xi_0, \dots, \xi_{n-1}, \eta, t)$, so that

$$V(\eta, t) = E_{\eta,t}[l(\xi_0 \dots \xi_{n-1}, \mathfrak{x}_{t_n})].$$

Define $W'(\eta, t) = (W^1(\eta, t), \dots, W^d(\eta, t))$, where

$$W^k(\eta, t) = \frac{\partial}{\partial \eta_k} V(\eta, t).$$

Differentiating (3·2) by η^k, we see that W^k satisfies the equation

$$W_t^k = \tfrac{1}{2} a_{\eta^k}^{ij} W_{\eta^i}^j + \tfrac{1}{2} a^{ij} W_{\eta^i \eta^j}^k + W_{\eta^i}^k f^i + W^i f_{\eta^i}^k = 0, \quad (t, \eta) \in [t_{n-1}, t_n] \times R^d \tag{3·9}$$

$$W^k(\eta, t_n) = l_{\xi_n^k}(\xi_0 \dots \xi_{n-1}, \eta). \tag{3·10}$$

According to the remarks on pp. 147–148 of (2) this equation has a unique solution satisfying the same growth condition as V^k. An extremely laborious application of the Ito rule, using (2·1), (2·8) and (3·9) shows that

$$d_t[\Phi'(t, s) \, W(\chi_t, t)] = \Phi^{ki}(W_{\eta_j}^k \sigma^{jl} + W^j \sigma_\eta^{jl}) \, d\mathfrak{w}^l,$$

i.e. the process $\Phi'(t, s) \, W(\chi_t, t)$ is a martingale. In view of the boundary conditions (2·9) (3·10) this shows that

$$W(\eta, t) = E_{\eta,t}[\Phi'(1, t) \, l_{\xi_n}(\xi_0 \dots \xi_{n-1}, \mathfrak{x}_1)]$$

and hence that

$$(V_\eta^{n-1})' \, (\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_{n-1}}, \mathfrak{x}_t, t) = E[(l'_{\xi_n}(\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_n}) \, \Phi(1, t) | F_t].$$

Thus $\mathfrak{e}_t$ and $\mathfrak{s}_t$ agree for $t \in [t_{n-1}, t_n]$.

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press

Now suppose that $\mathfrak{e}_t = \hat{\mathfrak{s}}_t$ a.s. for each $t \in [t_k, t_n]$, so that in particular, denoting

$$(V_\eta^k)^i(\xi_0 \dots \xi_k, \eta, t) = \frac{\partial}{\partial \eta^i} V^k(\xi_0 \dots \xi_k, \eta, t),$$

we have

$$(V_\eta^k)'(\mathfrak{x}_{t_0}, \dots, \mathfrak{x}_{t_k}, \mathfrak{x}_{t_k}, t_k) = E[\sum_{i \geqslant k+1} (l'_{\xi_i}(\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_n})\, \Phi(t_i, t_k) | F_{t_k}]. \tag{3·11}$$

Now consider the function V^{k-1}. An exactly similar argument to the above shows that for $t \in [t_{k-1}, t_k]$

$$V_\eta^{k-1}(\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_{k-1}}, \mathfrak{x}_t, t) = E[\Phi'(t_k, t)\, V_\eta^{k-1}(\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_k}, t_k) | F_t]. \tag{3·12}$$

In view of the terminal condition (3·7) we have

$$V_\eta^{k-1}(\xi_0 \dots \xi_{k-1}, \eta, t_k) = V_{\eta_1}^k(\xi_0 \dots \xi_{k-1}, \eta_1, \eta_2, t_k)|_{\eta_1=\eta_2=\eta} + V_{\eta_2}^k(\xi_0 \dots \xi_{k-1}, \eta_1, \eta_2, t_k)|_{\eta_1=\eta_2=\eta}.$$

The second term on the right is given by (3·9) from the induction hypothesis. For the first, we apply part (ii) of Lemma (3·5), which gives:

$$V_{\eta_1}^k(\xi_0 \dots \xi_{k-1}, \eta_1, \eta_2, t_k) = E_{\eta_2, t_k}[l_{\xi_k}(\xi_0 \dots \xi_{k-1}, \eta_1, \mathfrak{x}_{t_{k+1}} \dots \mathfrak{x}_{t_n})]$$

and hence

$$V_{\eta_1}^k(\mathfrak{x}_{t_0} \dots \mathfrak{x}_{t_k}, \mathfrak{x}_{t_k}, t_k) = E[l_{\xi_i}(\mathfrak{x}_{t_0}, \dots \mathfrak{x}_{t_n}) | F_{t_k}]. \tag{3.14}$$

Using (3·11), (3·13), (3·14) in (3·12) gives

$$\begin{aligned} V_\eta^{k-1}(\mathfrak{x}_{t_0}, \dots, \mathfrak{x}_{t_{k-1}}, \mathfrak{x}_t, t) &= E[\Phi'(t_k, t)\{l_{\xi_k} + \sum_{i \geqslant k+1} \Phi'(t_i, t_k)\, l_{\xi_i}\} | F_{t_k}] \\ &= E[\sum_{i \geqslant k} \Phi'(t_i, t)\, l_{\xi_i} | F_t]. \end{aligned}$$

Thus $\mathfrak{e}_t = \hat{\mathfrak{s}}_t$ for $t \in [t_{k-1}, t_k]$, and hence, by induction, for all $t \in [0, 1]$. This completes the proof and establishes Theorem (2·5) for the special case of $L(x) = l(x_{t_0} \dots x_{t_n})$ satisfying (3·1).

4. *Approximation of Fréchet differentiable functions.*

Suppose L is a Fréchet-differentiable function satisfying condition (2·4). We wish to approximate $L(x)$ by functions $L_n(x)$ depending on only a finite number of coordinates of x. For this purpose, introduce the linear interpolation $\beta_n : C \to C$ defined as follows: with $h = 2^{-n}$ and $x \in C$,

$$\beta_n(x)(t) = \frac{1}{h}[x(k+1)\,h) - x(kh)]\,t - kx((k+1)\,h) + (k+1)\,x(kh), \quad t \in [kh, (k+1)\,h].$$

Note in particular that $\|\beta_n(x)\| \leqslant \|x\|$ and $\|\beta_n(x) - x\| \to 0$ as $n \to \infty$. Now define

$$L_n(x) = L(\beta_n(x)).$$

Then clearly

$$L_n(x) = l_n(x_0, x_h, x_{2h}, \dots, x_1) \tag{4·1}$$

for some function $l_n : R^{(2^n+1)d} \to R$. The properties of L_n are summarized in the following theorem.

THEOREM 4·2.

(i) *l_n is continuously differentiable for each n.*

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press

(ii) *Let ν_x^n denote the measure corresponding to the Fréchet derivative of L_n. Then ν_x^n is continuous in the weak topology and*

$$|L_n(x)| + TV(\nu_x^n) \leqslant k(1+\|x\|)^q$$

(i.e. L_n satisfies (2·3)).

(iii) *$\nu_x^n \to \mu_x$ weakly as $n \to \infty$, for each $x \in C$.*

Proof. We give the proof for the case $d = 1$. For $d > 1$ the argument is the same but notationally 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.

$$\frac{\partial}{\partial \xi_k} l_n(\xi_0, \xi_1 \dots \xi_{2n}) = \begin{cases} \frac{1}{h}\int_0^h \nu(t)\,dt, & k = 0. \\ -\frac{1}{h}\left(\int_{(k-1)h}^{kh} \nu(t)\,dt - \int_{kh}^{(k+1)h} \nu(t)\,dt\right), & k = 1, 2, \dots 2^n - 1. \\ -\frac{1}{h}\int_{1-h}^{1} \nu(t)\,dt, & k = 2^n. \end{cases} \quad (4\cdot3)$$

Here $\nu(t) = \mu_{\beta_n(x)}([0,t])$ and $\xi_k = x_{kh}$, $k = 0, 1, \dots, 2^n$. Indeed, for $x, y \in C$ we have

$$\begin{aligned} L_n(x+y) &= L(\beta_n(x+y)) \\ &= L_n(x) + \int_{[0,1]} \beta_n(y)\,\mu_{\beta_n(x)}(ds) + o(\|y\|). \end{aligned} \quad (4\cdot4)$$

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

$$L_n(x+y) = l_n(x_0 \dots x_{(k-1)h}, x_{kh} + \alpha, x_{(k+1)h}, \dots x_1).$$

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

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

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

$$\nu_x^n(ds) = \sum_{i=0}^{2^n} \frac{\partial}{\partial \xi_i} l_n(x_0, x_h, \dots, x_{2^n h})\,\delta_{ih}(ds).$$

Using (4·3) shows that

$$\nu_x^n(t) = \nu_x^n([0,t]) = \frac{1}{h}\int_{kh}^{(k+1)h} \mu_{\beta_n(x)}(s)\,ds, t \in [kh, (k+1)h[. \quad (4\cdot5)$$

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press

By splitting up $\mu_{\beta_n(x)}$ into its positive and negative parts, it is easy to see from (4·5) that

$$TV(\nu_x^n) \leqslant TV(\mu_{\beta_n(x)}).$$

Thus
$$\begin{aligned} |L_n(x)| + TD(\nu_x^n) &\leqslant |L(\beta_n(x))| + TV(\mu_{\beta_n(x)}) \\ &\leqslant K(1 + \|\beta_n(x)\|)^q \\ &\leqslant K(1 + \|x\|)^q. \end{aligned}$$

This gives (4·2)(ii). For part (iii), note that $\nu_x^n(t)$ is a right-continuous function of t and according to the Helly–Bray Lemma (5, § 11·3*a*), weak convergence of ν_x^n to μ_x is equivalent to convergence of $\nu_x^n(t)$ to $\mu_x(t)$ at all continuity points of $\mu_x(.)$. If t is a continuity point of μ_x then $\mu_{\beta_n(x)}(s) \to \mu_x(s)$ as $n \to \infty$ for all s in a neighbourhood of t and it follows from (4·5) that $\nu_x^n(t) \to \mu_x(t)$. This completes the proof.

5. *Proof of Theorem* (2·5).

Let L be a functional satisfying (2·4) and define $L_n(x) = L(\beta_n(x))$ as in the previous section. In view of Proposition (3·8) and Theorem (4·2), $L_n(\chi)$ has the representation

$$L_n(\mathfrak{x}) = EL_n(\mathfrak{x}) + \int_0^1 E\left[\int_{]t,1]} \nu_{\mathfrak{x}}^n(ds)\, \Phi(s,t) \middle| F_t\right] \sigma(t, \mathfrak{x}_t)\, d\mathfrak{w}_t. \tag{5·1}$$

Now $\beta_n(x) \to x$ as $n \to \infty$ for any $x \in C$, so that $L_n(\mathfrak{x}) \to L(\mathfrak{x})$ a.s. It remains to show that the two terms on the right of (5·1) converge to the corresponding quantities in (2·6), (2·7) as $n \to \infty$.

Consider (1·2) and (2·8) as a *joint* system of stochastic differential equations, defining processes $\mathfrak{x}_s$ for $s \in [0,1]$ and $\Phi(s,t)$ for $s \in [t,1]$. In view of (2·2) the standard Ito conditions are satisfied, and we conclude from Theorem 5·2·3 of (2) that, for any positive integer m, there is a constant $\bar{C}(m)$ such that for any $t \in [0,1]$

$$E\{\sup_{t \leqslant s \leqslant 1} [|\mathfrak{x}_s|^m + |\Phi(s,t)|^m]\} \leqslant \bar{C}(m). \tag{5·2}$$

In particular, $E\|\mathfrak{x}\|^m < \bar{C}(m)$. Now from (2·4)(i),

$$\begin{aligned} |L_n(x)| &= |L(\beta_n(x))| \\ &\leqslant K(1 + \|\beta_n(x)\|)^q \\ &\leqslant K(1 + \|x\|)^q \end{aligned}$$

and hence applying the 'c_r inequality' (5, § 9·3*a*)

$$\begin{aligned} E|L_n(\mathfrak{x})|^2 &\leqslant EK^2(1 + \|x\|)^{2q} \\ &\leqslant K^2 c_{2q}(1 + \bar{C}(2q)). \end{aligned}$$

This shows that $\{L_n(\mathfrak{x}),\ n = 1, 2, \ldots\}$ is uniformly integrable and hence that

$$EL_n(\mathfrak{x}) \to EL(\mathfrak{x}) \quad \text{as} \quad n \to \infty.$$

Since almost all sample functions of $\Phi(.,t)$ are continuous and $\nu_x^n \to \mu_x$ weakly for each $x \in C$,

$$\int_t^1 \nu_{\mathfrak{x}}^n(ds)\, \Phi(s,t) \to \int_t^1 \mu_{\mathfrak{x}}(ds)\, \Phi(s,t) \quad \text{a.s.} \tag{5·3}$$

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press

Now in view of conditions (2·2) and part (ii) of Theorem 4·2

$$\left|\int_t^1 \nu_{\mathfrak{x}}^n(ds)\,\Phi(s,t)\,\sigma(t,\mathfrak{x}_t)\right| \leqslant \sup_{t\leqslant s\leqslant 1}|\Phi(s,t)|\,|\sigma(t,\mathfrak{x}_t)|\,TV(\nu_{\mathfrak{x}}^n)$$
$$\leqslant \sup_{t\leqslant s\leqslant 1}|\Phi(s,t)|\,K^2(1+\|\mathfrak{x}\|)^q.$$

Hence from (5·2) and the Schwarz inequality

$$E\left|\int_t^1 \nu_{\mathfrak{x}}^n(ds)\,\Phi(s,t)\,\sigma(t,\mathfrak{x}_t)\right|^3 \leqslant K_1$$

for some constant K_1. This shows that the set $\left\{\left|\int_t^1 \nu_\chi^n(ds)\,\Phi(s,t)\,\sigma(t,x_t)\right|^2,\ n=1,2\ldots\right\}$ is a uniformly integrable subset of $L_1(\Omega\times[0,1],\ dP^*dt)$, and therefore in view of (5·3) and Jensen's inequality we have

$$E\int_0^1\left|E\left[\int_t^1 \nu_{\mathfrak{x}}^n(ds)\,\Phi(s,t)-\int_t^1 \mu_{\mathfrak{x}}(ds)\,\Phi(s,t)\,\middle|\,F_t\right]\sigma(t,\mathfrak{x}_t)\right|^2 dt$$
$$\leqslant E\int_0^1\left|\left(\int_t^1 \nu_{\mathfrak{x}}^n(ds)\,\Phi(s,t)-\int_t^1 \mu_{\mathfrak{x}}(ds)\Phi(s,t)\right)\sigma(t,\mathfrak{x}_t)\right|^2 dt\to 0 \quad \text{as} \quad n\to\infty.$$

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 Thoerem (2·5) for a process $\mathfrak{x}$ satisfying

$$d\mathfrak{x}_t = f(t,\mathfrak{x})\,dt+\sigma(t,\mathfrak{x})\,d\mathfrak{w},$$

where f^i, $\sigma^{ij}:[0,1]\times C\to R$ 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, σ (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

(1) CLARK, J. M. C. The representation of functionals of Brownian motion by stochastic integrals. *Annals of Math Stat.* **41** (1970), 1282–1295.
(2) FRIEDMAN, A. *Stochastic differential equations*, vol. 1, Academic Press (New York, 1975).
(3) HAUSSMANN, U. G. Functionals of Ito processes as stochastic integrals. *SIAM J. Control Optimization* **16** (1978), 252–269.
(4) HAUSSMANN, U. G. On the integral representation of functionals of Ito processes. *Stochastics* **3** (1979), 17–28.
(5) LOÈVE, M. *Probability theory*, 3rd ed., D. van Nostrand (Princeton, 1963).

https://doi.org/10.1017/S0305004100056590 Published online by Cambridge University Press