The Annals of Probability
1997, Vol. 25, No. 2, 997–1010

SEMIMARTINGALE INTEGRAL REPRESENTATION

By Hyungsok Ahn

University of California, Santa Barbara

We provide an integral representation for smooth functionals of continuous semimartingales. The representation is related to an infinite-dimensional nonautonomous parabolic equation. Semimartingale integral representations, including a martingale representation, are given in terms of a solution to this equation.

1. Introduction.

Let X be a solution to a stochastic differential equation

$$X_t = H_t + \int_0^t \sigma(X_s)\, dW_s, \tag{1}$$

where W is a Brownian motion and H is an adapted continuous process with paths of finite variation. For a smooth functional f defined on $L^2[0, T]$, we provide a representation of $f(X)$ in terms of a stochastic integral with respect to X. This work is motivated by the Black–Scholes framework for option replication (see Duffie [3]).

A seminal study on explicit descriptions of the integrand in the martingale integral representation was initiated by Clark [2] in 1970. Fréchet differentiable functionals of a Brownian motion were considered in his work, and the integrand was the predictable projection of a process generated by the Fréchet derivative of the functional. Substantial generalizations including functionals of Itô processes have been made by many authors, and most recently, Karatzas, Ocone and Li [7] established a relevant formula for a broader class of functionals using the Malliavin derivative of the functional. More references on the subject are contained in their work.

In some cases, the relationship between the integrand and a solution of a partial differential equation can be obtained from the Feynman–Kac representation. A sophisticated example of the use of partial differential equations is found in Ma, Protter and Yong [9], in which an explicit solution of a forward–backward stochastic differential equation was obtained. However, the study in this direction has been restricted to functionals with a certain structure. In this paper, we establish such relationship in the case of smooth functionals defined on $L^2[0, T]$, with a unified treatment. In other words, this study links Clark's formula to infinite-dimensional parabolic equations.

In the next two sections, we prove an extended version of Itô's formula. The subject has been revisited numerous times, and extensions have been made in many directions; for example, see Kunita [8] for random functions such as flows of stochastic differential equations, see Föllmer, Protter and Shiryaev

Received July 1995; revised July 1996.
AMS 1991 *subject classifications.* Primary 60G44, 60H05, 60H30.
Key words and phrases. Martingale representation, semimartingale integral representation.

[4] for nonsmooth functions and see Metivier [10] for Hilbert space-valued semimartingales. Here, we study a formula for the level processes $t \to X_{\bullet \wedge t}$. We designate $\mathscr{X}$ to be the Banach space $C[0, T]$ or the Hilbert space $L^2[0, T]$. Suppose that $f\colon \mathscr{X} \to \mathbb{R}$ is twice continuously Fréchet differentiable with respect to the corresponding norm. For a real-valued continuous semimartingale X, we show that $t \to f(X^t)$, where $X_s^t = X_{s\wedge t}$, is also a continuous semimartingale. Moreover, it has a representation

$$f(X^t) = f(X^0) + \int_0^t \langle \eta_s, \triangledown f(X^s)\rangle \, dX_s + \tfrac{1}{2}\int_0^t \langle \eta_s \otimes \eta_s, \triangledown^2 f(X^s)\rangle d[X, X]_s.$$

Here, $\eta_s = \mathbb{I}_{[s, T]}$ is an element of $\mathscr{X}^{**}$, the bidual of $\mathscr{X}$, and the bracket $\langle \cdot, \cdot\rangle$ is used for dual pairs. We will call this *Itô's representation for functional f*; if $f(x) = g(x(T))$, where $g \in C^2(\mathbb{R}, \mathbb{R})$, the above formula agrees with Itô's formula. In the case of $L^2[0, T]$, the representation compares with the transformation formula for Hilbert space-valued semimartingales (see Métivier [10]), where the Hilbert space-valued stochastic integral was used. Regarding $s \to X^s$ as a Hilbert space-valued semimartingale, one has $\langle \triangledown f(X^s), dX^s\rangle = \langle \eta_s, \triangledown f(X^s)\rangle \, dX_s$. We will also provide a relevant formula for the case of $\mathbb{R}^p$-valued semimartingales.

In Section 4, we start with an infinite-dimensional nonautonomous parabolic equation

$$\frac{\partial u}{\partial t}(t, x) + A(t)u(t, x) = 0 \tag{2}$$

with $u(T, x) = f(x)$, where the differential operator $A(t)$ is defined by

$$A(t)\phi(x) = \tfrac{1}{2}\langle \eta_t \otimes \eta_t, \triangledown^2 \phi(x)\rangle \sigma(x(t))^2 + \langle \eta_t, \triangledown \phi(x)\rangle b(x(t)).$$

We prove the existence of a solution using a probabilistic method. The idea is contained in the following finite-dimensional example. Let $f\colon \mathbb{R}^3 \to \mathbb{R}$ and let W be a standard Brownian motion. For $0 \le t_1 < t_2 < t_3$, one retains a formula $E(f(W_{t_1}, W_{t_2}, W_{t_3})|\mathscr{F}_{t_2}) = u(t_2, W_{t_1}, W_{t_2}, W_{t_2})$ where u solves a heat equation with the first two space variables as parameters:

$$\frac{\partial u}{\partial t}(t, x_1, x_2, x_3) + \frac{1}{2}\frac{\partial^2 u}{\partial x_3^2}(t, x_1, x_2, x_3) = 0 \tag{3}$$

with $u(t_3, x_1, x_2, x_3) = f(x_1, x_2, x_3)$. Next, we solve another heat equation

$$\frac{\partial v}{\partial t}(t, x_1, x_2, x_3) + \frac{1}{2}\sum_{i=2}^{3}\sum_{j=2}^{3}\frac{\partial^2 v}{\partial x_i \partial x_j}(t, x_1, x_2, x_3) = 0 \tag{4}$$

with $v(t_2, x_1, x_2, x_3) = u(t_2, x_1, x_2, x_3)$. This yields

$$E(f(W_{t_1}, W_{t_2}, W_{t_3})|\mathscr{F}_{t_1}) = v(t_1, W_{t_1}, W_{t_1}, W_{t_1}).$$

Virtually, (2) is the aggregation of these heat equations defined piecewise. Finally, we describe the integrand in Clark's formula in terms of (2). Let $p(t, x, |T, y)$ be the transition density for a standard Brownian motion, which is the Green function for $u_t + u_{xx}/2 = 0$. Then the integrand in Clark's formula for $f(W_{t_1}, W_{t_2}, W_{t_3})$ can be expressed by

$$\int p_x(t, W_t | t_3, y_3) f(W_{t_1}, W_{t_2}, y_3) dy_3 \quad \text{on } [t_2, t_3], \tag{5}$$

$$\int p_x(t, W_t | t_2, y_2) \int p(t_2, y_2 | t_3, y_3) f(W_{t_1}, y_2, y_3)\, dy_3\, dy_2 \quad \text{on } [t_1, t_2] \tag{6}$$

and one more iteration yields the expression on $[0, t_1]$. Note that (5) and (6) can also be written as $(\partial u/\partial x_3)(t, W_{t_1}, W_{t_2}, W_t)$ and $\sum_{i=2}^{3} \partial u/\partial x_i(t, W_{t_1}, W_t, W_t)$, respectively, where u and v are obtained from (3) and (4). For $f\colon L^2[0, T] \to \mathbb{R}$, we exploit (2) to describe the integrand of Clark's formula. We also describe $f(X)$, where X is defined as in (1), as a stochastic integral with respect to X.

2. Itô's representation for functionals. Respectively, $\mathscr{X}^*$ and $\mathscr{X}^{**}$ will denote the dual and the bidual of $\mathscr{X}$. Suppose that $f\colon \mathscr{X} \to \mathbb{R}$ is twice continuously Fréchet differentiable. That is, $\nabla f(\cdot)\colon \mathscr{X} \to \mathscr{X}^*$ and $\nabla^2 f(\cdot)\colon \mathscr{X} \to L(\mathscr{X}, \mathscr{X}^*)$ are continuous with respect to the corresponding norm. It is known that $L(\mathscr{X}, \mathscr{X}^*)$ is also isometrically isomorphic to the dual of the cross-space $\mathscr{X} \otimes_\gamma \mathscr{X}$, where γ is the greatest crossnorm; see, Schatten [14], page 47. We will only consider the greatest crossnorm γ and its associates, and we will not specify it every time. Thus, $\otimes$ denotes $\otimes_\gamma$, $\otimes_{\gamma'}$ and $\otimes_{\gamma''}$ when it applies to $\mathscr{X}$, $\mathscr{X}^*$ and $\mathscr{X}^{**}$, respectively.

THEOREM 2.1. *Let $\mathscr{X}$ be either $C[0, T]$ or $L^2[0, T]$. Suppose that $f\colon \mathscr{X} \to \mathbb{R}$ is twice continuously Fréchet differentiable at each $x \in C[0, T]$, with respect to the corresponding norm. Then, for a continuous semimartingale X and $t \in [0, T]$, we have*

$$f(X^t) = f(X^0) + \int_0^t \langle \eta_s, \nabla f(X^s)\rangle\, dX_s + \tfrac{1}{2} \int_0^t \langle \eta_s \otimes \eta_s, \nabla^2 f(X^s)\rangle\, d[X, X]_s, \tag{7}$$

where $\eta_s = \mathbb{1}_{[s,T]}$ and $X_t^s = X_t \mathbb{1}\{t \le s\} + X_s \mathbb{1}\{t > s\}$.

One could hold that L^2 topology is luxurious in this theorem, since the uniform norm is stronger than the L^2 norm and since we are interested in continuous semimartingales only. Indeed, this is true. The L^2 topology is appended for the purpose of studying partial differential equations in Section 4. For each $t \in [0, T]$, η_t is in $\mathscr{X}^{**}$ and $\eta_t \otimes \eta_t$ in $(\mathscr{X} \otimes \mathscr{X})^{**}$. Therefore $\langle \eta_t, \nabla f(X^t)\rangle$ and $\langle \eta_t \otimes \eta_t, \nabla^2 f(X^t)\rangle$ are well defined as dual pairs, bounded linear operators on $\mathscr{X}^*$ and $\mathscr{X}^* \otimes \mathscr{X}^*$ acting on elements of $\mathscr{X}^*$ and $\mathscr{X}^* \otimes \mathscr{X}^*$, respectively. Note that this representation depends more on the path of X than on the underlying filtration. For instance, consider two different filtrations for which X remains a semimartingale. Since $\langle \eta_s, \nabla f(X^s)\rangle$ and $\langle \eta_s \otimes \eta_s, \nabla^2 f(X^s)\rangle$ are defined path-by-path, the representation will remain the same.

The proof of the above theorem is lengthy and is presented in the next section. For now, we illustrate how this formula works using simple examples.

Then we prove the regularities of the integrands to assure that the stochastic integral in (7) is well defined. A multivariate version of (7) will be stated at the end of this section.

EXAMPLE 2.1. Let $g\colon \mathbb{R}^k \to \mathbb{R}$ be a C^2 function, and let $f(x) = g(x(t_1), \dots, x(t_k))$ for each x in $\mathscr{X}$. In this case, we can derive (7) by using the usual Itô formula. Since $f(X^t) = g(X_{t\wedge t_1}, \dots, X_{t\wedge t_k})$, one has

$$
\begin{aligned}
f(X^t) - f(X^0)
&= \sum_{i=1}^k \int_0^t g_i(X_{s\wedge t_1}, \dots, X_{s\wedge t_k})\, dX_s^{t_i} \\
&\quad + \tfrac{1}{2} \sum_{i=1}^k \sum_{j=1}^k \int_0^t g_{ij}(X_{s\wedge t_1}, \dots, X_{s\wedge t_k}) d[X^{t_i}, X^{t_j}]_s \\
&= \int_0^t \sum_{i=1}^k g_i(X_{s\wedge t_1}, \dots, X_{s\wedge t_k})\, \mathbb{1}_{[0,\,t_i]}(s)\, dX_s \\
&\quad + \tfrac{1}{2} \int_0^t \sum_{i=1}^k \sum_{j=1}^k g_{ij}(X_{s\wedge t_1}, \dots, X_{s\wedge t_k})\, \mathbb{1}_{[0,\,t_i]}(s) \mathbb{1}_{[0,\,t_j]}(s)\, d[X, X]_s.
\end{aligned}
$$

Here, the subscripts on g denote the partial derivatives. Replacing $\mathbb{1}_{[0,\,t_i]}(s)$ by $\mathbb{1}_{[s,\,T]}(t_i)$, we obtain (7).

EXAMPLE 2.2. Unlike the Malliavin differential operator, $f(W) \to \langle \eta_\bullet, \nabla f(W^\bullet)\rangle$ is not closable. Let π_n be a refining sequence of partitions of $[0,1]$, and let $g\colon \mathbb{R} \to \mathbb{R}$ be a C^3 function. Consider

$$f^n(W) = \sum_{\pi_n} g'(W_{t_k^n})(W_{t_{k+1}^n} - W_{t_k^n}).$$

Then, $f^n(W)$ converges to

$$f(W) = g(W_1) - g(W_0) - \tfrac{1}{2} \int_0^1 g''(W_s)\, ds.$$

Since

$$\nabla f^n[W](h) = \sum_{\pi_n} g''(W_{t_k^n}) h(t_k^n)(W_{t_{k+1}^n} - W_{t_k^n}) + \sum_{\pi_n} \{h(t_{k+1}^n) - h(t_k^n)\} g'(W_{t_k^n}),$$

we have $\langle \eta_t, \nabla f^n(W^t)\rangle = g'(W_{t_k^n})$ for $t \in (t_k^n, t_{k+1}^n]$. However, we have

$$\nabla g[W](h) = g'(W_1)h(1) - g'(W_0)h(0) - \tfrac{1}{2} \int_0^1 g^{(3)}(W_s)h(s)\, ds,$$

and hence

$$\langle \eta_t, \nabla f(W^t)\rangle = g'(W_t) - \tfrac{1}{2}(1-t) g^{(3)}(W_t).$$

This differs from $g'(W_t)$, the limit of $\langle \eta_t, \nabla f^n(W^t)\rangle$.

EXAMPLE 2.3. Let μ be a finite signed Borel measure, and let

$$f(x) = \int_0^T g(x(s), s)\, \mu(ds),$$

where $g: \mathbb{R} \times [0, T] \to \mathbb{R}$ satisfy the following conditions.

(i) For each t, $g(\cdot, t): \mathbb{R} \to \mathbb{R}$ is C^2.
(ii) For each u, $g(u, \cdot): [0, T] \to \mathbb{R}$ is μ-measurable.

Let g_u and g_{uu} be the first and the second partial derivatives with respect to the first argument. Then

$$\nabla f[x](y) = \int_0^T g_u(x(s), s) y(s)\, \mu(ds),$$

$$\nabla^2 f[x](y, z) = \int_0^T g_{uu}(x(s), s) y(s) z(s)\, \mu(ds).$$

Thus, substituting η_t for y and z, we have

$$\langle \eta_t, \nabla f(x^t)\rangle = \int_t^T g_u(x(t), s)\, \mu(ds),$$

$$\langle \eta_t \otimes \eta_t, \nabla^2 f(x^t)\rangle = \int_t^T g_{uu}(x(t), s)\, \mu(ds).$$

Therefore, for a continuous semimartingale X, we have

$$f(X) - f(X^0) = \int_0^T \int_t^T g_u(X_t, s)\, \mu(ds)\, dX_t + \tfrac{1}{2} \int_0^T \int_t^T g_{uu}(X_t, s)\, \mu(ds) d[X, X]_s.$$

This identity can be also obtained from the Fubini theorem for stochastic integrals; see Protter [13].

We now investigate the measurability of $\langle \eta_t, \nabla f(X^t)\rangle$ and $\langle \eta_t \otimes \eta_t, \nabla^2 f(X^t)\rangle$ as processes in t. Properties of $\nabla f(x)$ are well known as a member of $\mathscr{X}^*$. Although $L(\mathscr{X}, \mathscr{X}^*)$ is not as easy to access as $\mathscr{X}^*$, especially when $\mathscr{X}$ is not reflexive, there is a topological resemblance.

LEMMA 2.1. *Suppose x_n^{**} converges weakly to x^{**} in $\mathscr{X}^{**}$; that is $\langle x_n^{**}, x^*\rangle$ converges to $\langle x^{**}, x^*\rangle$ for all $x^* \in \mathscr{X}^*$. Then, for all $A \in L(\mathscr{X}, \mathscr{X}^*)$, $\langle x_n^{**} \otimes x_n^{**}, A\rangle$ converges to $\langle x^{**} \otimes x^{**}, A\rangle$.*

This is a consequence of Theorem 3.4 of Schatten [14], which says that every operator $A \in L(\mathscr{X}, \mathscr{X}^*)$ can be approximated by operators of finite rank in the norm topology induced by the greatest crossnorm.

PROPOSITION 2.1. *Let $f: \mathscr{X} \to \mathbb{R}$ be twice continuously Fréchet differentiable at $x \in C[0, T]$. Define $x^t(s) = x(s)\mathbb{1}\{s \le t\} + x(t)\mathbb{1}\{s > t\}$. Then both*

$\langle \eta_t, \nabla f(x^t)\rangle$ and $\langle \eta_t \otimes \eta_t, \nabla^2 f(x^t)\rangle$ are cadlag in t; that is, they are left continuous and have right limits. Especially if $\mathscr{X} = L^2[0, T]$, then both are continuous in t.

PROOF. First, let $\mathscr{X} = C[0, T]$. For each x^* in $\mathscr{X}^*$, $\langle \eta_{t-\varepsilon}, x^*\rangle$ converges to $\langle \eta_t, x^*\rangle$ as $\varepsilon \downarrow 0$. This is due to the monotone convergence theorem of a finite signed Borel measure x^*. Also note that $\nabla f(x^{t-\varepsilon})$ converges to $\nabla f(x^t)$ in the norm topology as $\varepsilon \to 0$. Then the left continuity of $\langle \eta_t, \nabla f(x^t)\rangle$ follows. Next, let $\bar{\eta}_t = \mathbb{I}_{(t,T]}$. Then, for each x^* in $\mathscr{X}^*$, $\langle \eta_{t+\varepsilon}, x^*\rangle$ converges to $\langle \bar{\eta}_t, x^*\rangle$ as $\varepsilon \downarrow 0$. Again, this is due to the monotone convergence theorem of a finite signed Borel measure x^*. Since $\nabla f(x^{t+\varepsilon})$ converges to $\nabla f(x^t)$ in the norm topology, $\langle \eta_{t+\varepsilon}, \nabla f(x^{t+\varepsilon})\rangle$ has a limit as $\varepsilon \downarrow 0$. Therefore $\langle \eta_t, \nabla f(x^t)\rangle$ is cadlag in t. Similarly $\langle \eta_t \otimes \eta_t, \nabla^2 f(x^t)\rangle$ is cadlag; we apply the previous lemma.

If $\mathscr{X} = L^2[0, T]$, then $t \to \eta_t$ is continuous in the strong operator topology. Since $t \to \nabla f(x^t)$ is continuous in the uniform topology, $t \to \langle \eta_t, \nabla f(x^t)\rangle$ is continuous in t. The continuity of $\langle \eta_t \otimes \eta_t, \nabla^2 f(x^t)\rangle$ follows from the previous lemma. □

Next we discuss the multivariate case. Let p be a positive integer and

$$\mathscr{X}^p = \underbrace{\mathscr{X} \oplus \cdots \oplus \mathscr{X}}_{p \text{ times}},$$

where $\oplus$ indicates the usual direct sum of vector spaces. That is, $\mathscr{X}^p$ is a Banach space of $\mathbb{R}^p$-valued continuous functions (or L^2 functions) defined on $[0, T]$. Let $x = (x_1, \ldots, x_p)$ be an element of $\mathscr{X}_p$. Then $\|x\|_{\mathscr{X}^p} = \|\,(\|x_1\|_{\mathscr{X}}, \ldots, \|x_p\|_{\mathscr{X}})\,\|_{\mathbb{R}^p}$. Since norms in a finite-dimensional space are all equivalent, any Euclidean norm will serve as $\|\cdot\|_{\mathbb{R}^p}$. Note that

$$(\mathscr{X}^p)^* = \underbrace{\mathscr{X}^* \oplus \cdots \oplus \mathscr{X}^*}_{p \text{ times}}$$

and this is also the case for $(\mathscr{X}^p)^{**}$. Thus, we may write $\nabla f(x) = (\nabla_1 f(x), \ldots, \nabla_p f(x))$ where each $\nabla_i f(x)$ is an element of $\mathscr{X}^*$. These $\nabla_i f(x)$, $i = 1, \ldots, p$ are essentially partial Fréchet derivatives; see Ambrosetti and Prodi [1]. Similarly, if f is twice Fréchet differentiable at $x \in \mathscr{X}^p$, then $\nabla^2 f(x)$ is a $p \times p$ matrix which consists of elements in $L(\mathscr{X}, \mathscr{X}^*)$. We designate $\nabla^2_{ij} f(x)$ to be the (i, j) entry of $\nabla^2 f(x)$. The proof of the following theorem is parallel to that of Theorem 2.1, which will be given in the next section.

THEOREM 2.2. *Let $f\colon \mathscr{X}^p \to \mathbb{R}$ be twice continuously Fréchet differentiable at each $x \in C[0, T]$. Then, for an $\mathbb{R}^p$-valued continuous semimartingale $X = (X_1, \ldots, X_p)$ and $t \in [0, T]$, we have*

$$\begin{aligned} f(X^t) = f(X^0) &+ \sum_{i=1}^{p} \int_0^t \langle \eta_s, \nabla_i f(X^s)\rangle\, dX_s^i \\ &+ \frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{p} \int_0^t \langle \eta_s \otimes \eta_s, \nabla^2_{ij} f(X^s)\rangle d[X^i, X^j]_s. \end{aligned}$$

In the case of FV processes (i.e., processes with paths of finite variation), we are able to relax the smoothness of f. Again, the proof of Theorem 2.1 can be adapted for the following variations.

THEOREM 2.3. *Let* $f\colon \mathscr{X}^p \to \mathbb{R}$ *be continuously Fréchet differentiable at each* $x \in C[0,T]$. *Then, for an* $\mathbb{R}^p$*-valued continuous FV process* $X = (X_1, \dots, X_p)$ *and* $t \in [0,T]$, *we have*

$$f(X^t) = f(X^0) + \sum_{i=1}^{p} \int_0^t \langle \eta_s, \triangledown_i f(X^s)\rangle \, dX_s^i.$$

THEOREM 2.4. *Let* $f\colon \mathbb{R} \times \mathscr{X} \to \mathbb{R}$ *be differentiable with respect to the first argument and twice Fréchet differentiable at each* $x \in C[0,T]$. *Furthermore, let us assume that all the derivatives are continuous in both directions. Then, for a continuous semimartingale* X, *we have*

$$\begin{aligned} f(t, X^t) = f(0, X^0) &+ \int_0^t \frac{\partial f}{\partial s}(s, X^s)\, ds + \int_0^t \langle \eta_s, \triangledown f(s, X^s)\rangle \, dX_s \\ &+ \int_0^t \langle \eta_s \otimes \eta_s, \triangledown^2 f(s, X^s)\rangle d[X, X]_s. \end{aligned} \tag{8}$$

3. Proof of Theorem 2.1. It suffices to prove (7) for $t = T$. Let $\{\sigma_n\}_{n=1}^{\infty}$ be a refining sequence of nonrandom partitions of $[0, T]$ with $\|\sigma_n\| \downarrow 0$. Let X^n be a piecewise linear approximation of X with respect to σ_n. More precisely, if τ_k^n is the kth smallest member of σ_n, and if $t \in [\tau_k^n, \tau_{k+1}^n]$, we have

$$X_t^n = \frac{X_{\tau_{k+1}^n} - X_{\tau_k^n}}{\tau_{k+1}^n - \tau_k^n}(t - \tau_k^n) + X_{\tau_k^n}.$$

Since f is continuous, $f(X^n)$ converges to $f(X)$ almost surely. Using a telescoping sum, we have

$$f(X^n) - f(X^{n,0}) = \sum_{\sigma_n} \{ f(X^{n, \tau_{k+1}^n}) - f(X^{n, \tau_k^n})\}. \tag{9}$$

When f is defined on $L^2[0, T]$, it is easier to work with a piecewise constant approximation: that is, an approximation by simple predictable processes. Nevertheless we insist on using a piecewise linear approximation, which works for either case, $C[0, T]$ or $L^2[0, T]$. The following result is standard and works for a general Banach space.

LEMMA 3.1. *Let* $f\colon \mathscr{X} \to \mathbb{R}$ *be twice continuously Fréchet differentiable, and let* $R(x, y)$ *be the remainder of the second order Taylor expansion of* f; *that is*

$$f(x+y) - f(x) = \langle \triangledown f(x), y\rangle + \tfrac{1}{2}\langle \triangledown^2 f(x), y \otimes y\rangle + R(x, y).$$

Then, for a compact set B, $\sup_{x, y \in B} |R(x, y)| \le r_B(\|y\|)\, \|y\|^2$ *where* $r_B\colon \mathbb{R} \to \mathbb{R}$ *is nonnegative, continuous at* 0 *and* $r_B(0) = 0$.

Thus, by the Taylor expansion, (9) can be rewritten as $S_n^1 + S_n^2 + S_n^3$ where

$$S_n^1 = \sum_{\sigma_n} \langle \nabla f(X^{n,\tau_k^n}), (X^{n,\tau_{k+1}^n} - X^{n,\tau_k^n})\rangle,$$

$$S_n^2 = \tfrac{1}{2} \sum_{\sigma_n} \langle \nabla^2 f(X^{n,\tau_k^n}), (X^{n,\tau_{k+1}^n} - X^{n,\tau_k^n}) \otimes (X^{n,\tau_{k+1}^n} - X^{n,\tau_k^n})\rangle,$$

and $S_n^3 = \sum_{\sigma_n} R(X^{n,\tau_k^n}, X^{n,\tau_{k+1}^n} - X^{n,\tau_k^n})$. Next we define

$$\eta^n[\tau_k^n](t) = \frac{(t \vee \tau_{k+1}^n) - \tau_k^n}{\tau_{k+1}^n - \tau_k^n} \mathbb{1}\{t \geq \tau_k^n\}$$

so that $X^{n,\tau_{k+1}^n} - X^{n,\tau_k^n} = (X_{\tau_{k+1}^n} - X_{\tau_k^n})\, \eta^n[\tau_k^n]$. Since $\langle\ ,\ \rangle$ is bilinear, we have

$$S_n^1 = \sum_{\sigma_n} \langle \nabla f(X^{n,\tau_k^n}), \eta^n[\tau_k^n]\rangle\, (X_{\tau_{k+1}^n} - X_{\tau_k^n}),$$

$$S_n^2 = \tfrac{1}{2} \sum_{\sigma_n} \langle \nabla^2 f(X^{n,\tau_k^n}), \eta^n[\tau_k^n] \otimes \eta^n[\tau_k^n]\rangle\, (X_{\tau_{k+1}^n} - X_{\tau_k^n})^2.$$

The following lemma will be used in proving the convergence of these sums.

LEMMA 3.2. *Let $f\colon \mathscr{X} \to \mathbb{R}$ be twice continuously Fréchet differentiable. If $x_n \to x$ in the uniform topology, then*

$$\langle \eta_t, \nabla f(x_n^t)\rangle \to \langle \eta_t, \nabla f(x^t)\rangle$$

$$\langle \eta_t \otimes \eta_t, \nabla^2 f(x_n^t)\rangle \to \langle \eta_t \otimes \eta_t, \nabla^2 f(x^t)\rangle$$

uniformly in t.

PROOF. Let $B^o = \{x_n^t\colon t \in [0, T] \text{ and } n \geq 1\}$. Since x_n converges to x in the uniform topology, B^o is an equicontinuous family, and hence, it is relatively compact in $C[0, T]$. It is also relatively compact in $L^2[0, T]$; see Friedman [5], page 115. Let B be the closure of B^o with respect to the corresponding norm. Then, since $\nabla f(\cdot)$ and $\nabla^2 f(\cdot)$ are continuous in the norm topology, they are uniformly continuous on B in the norm topology. Now,

$$\sup_{t \leq T} |\langle \eta_t, \nabla f(x_n^t) - \nabla f(x^t)\rangle| \leq \sup_{t \leq T} \| \nabla f(x_n^t) - \nabla f(x^t)\|.$$

This is due to $\|\eta_t\| = 1$ for all t. The uniform convergence of $\langle \eta_t, \nabla f(x_n^t)\rangle$ to $\langle \eta_t, \nabla f(x^t)\rangle$ follows from the uniform continuity of $\nabla f(\cdot)$ on B. Similarly we prove the uniform convergence of $\langle \eta_t \otimes \eta_t, \nabla^2 f(x_n^t)\rangle$. $\square$

Now we prove the convergence of S_n^1, S_n^2, and S_n^3. Then Theorem 2.1 will follow.

CLAIM 1. The random variable $S_n^1 \to \int_0^T \langle \eta_t, \nabla f(X^t)\rangle\, dX_t$ in probability.

PROOF. Note that if H^n is a sequence of cadlag processes converging to H uniformly on compacts in probability, then a sequence of stochastic integrals $\int H^n_s\,dX_s$ also converges to $\int H_s\,dX_s$ uniformly on compacts in probability; see Protter [13], page 51. Then, by Lemma 3.2, we have

$$\sum_{\sigma_n}\langle \eta_{\tau^n_k}, \nabla f(X^{n,\tau^n_k})\rangle (X_{\tau^n_{k+1}} - X_{\tau^n_k}) \to \int_0^T \langle \eta_t, \nabla f(X^t)\rangle\,dX_t \tag{10}$$

in probability. For $s \in [0, T]$, define $\lambda_n(s) = \max\{\tau \in \sigma_n : \tau \le s\}$. Then

$$\begin{aligned}\sup_{s\le T} &|\langle \eta^n[\lambda_n(s)], \nabla f(X^{n,\lambda_n(s)})\rangle - \langle \eta_{\lambda_n(s)}, \nabla f(X^{n,\lambda_n(s)})\rangle| \\ &\le 2\sup_{s\le T}\|\nabla f(X^{n,\lambda_n(s)})\|\end{aligned}$$

is stochastically bounded. Also $\langle \eta^n[\lambda_n(t)], \nabla f(X^{n,\lambda_n(t)})\rangle - \langle \eta_{\lambda_n(t)}, \nabla f(X^{n,\lambda_n(t)})\rangle$ converges to 0 for each t almost surely. This follows from the weak convergence of $\eta^n[\lambda_n(t)] - \eta_{\lambda_n(t)}$ for each t. Therefore, by the dominated convergence theorem (see Protter [13], page 145),

$$\sum_{\sigma_n}\{\langle \eta^n[\lambda_n(s)], \nabla f(X^{n,\lambda_n(s)})\rangle - \langle \eta_{\lambda_n(s)}, \nabla f(X^{n,\lambda_n(s)})\rangle\}(X_{\tau^n_{k+1}} - X_{\tau^n_k})$$

converges to 0 in probability. This together with (10) proves Claim 1. □

CLAIM 2. The random variable $S^2_n \to \frac{1}{2}\int_0^T \langle \eta_t \otimes \eta_t, \nabla^2 f(X^t)\rangle d[X,X]_t$ in probability.

PROOF. As in Claim 1,

$$\sum_{\sigma_n}\langle \eta^n[\tau^n_k] \otimes \eta^n[\tau^n_k], \nabla^2 f(X^{n,\tau^n_k})\rangle\{[X,X]_{\tau^n_{k+1}} - [X,X]_{\tau^n_k}\}$$

converges to $\int_0^T \langle \eta_t \otimes \eta_t, \nabla^2 f(X^t)\rangle d[X,X]_t$ in probability. We need to verify that

$$\sum_{\sigma_n}\langle \eta^n[\tau^n_k] \otimes \eta^n[\tau^n_k], \nabla^2 f(X^{\tau^n_k})\rangle((X_{\tau^n_{k+1}} - X_{\tau^n_k})^2 - ([X,X]_{\tau^n_{k+1}} - [X,X]_{\tau^n_k}))$$

converges to 0 in probability. Note that the above can be rewritten as

$$\int_0^T \langle \eta^n[\lambda_n(t)] \otimes \eta^n[\lambda_n(t)], \nabla^2 f(X^{\lambda_n(t)})\rangle (X_t - X_{\lambda_n(t)})\,dX_t.$$

This converges to 0 in probability since $(X_t - X_{\lambda_n(t)})$ converges to 0 uniformly in t almost surely and $\sup_{t\le T}\langle \eta^n[\lambda_n(t)] \otimes \eta^n[\lambda_n(t)], \nabla^2 f(X^{\lambda_n(t)})\rangle$ is stochastically bounded. □

CLAIM 3. The random variable $S^3_n \to 0$ in probability.

PROOF. For each sample path w, let $B(w)$ be the closure of $\{X^{n,t}(w): t \in [0, T]$ and $n \ge 1\}$ with respect to the corresponding norm. Then B is compact with probability 1, since it is a closure of an equicontinuous family. By Lemma 3.1, we have

$$|S_n^3| \le \max_{\sigma_n} r_B(|X_{\tau_{k+1}^n} - X_{\tau_k^n}|) \sum_{\sigma_n} (X_{\tau_{k+1}^n} - X_{\tau_k^n})^2.$$

The result follows from the fact that $\max_{\sigma_n} r_B(|X_{\tau_{k+1}^n} - X_{\tau_k^n}|)$ converges to 0 almost surely and $\sum_{\sigma_n}(X_{\tau_{k+1}^n} - X_{\tau_k^n})^2$ converges in probability. □

4. PDE and integral representation. Suppose that $f\colon \mathscr{X} \to \mathbb{R}$, where $\mathscr{X}$ is either $L^2[0,T]$ or $C[0,T]$, is twice continuously Fréchet differentiable with respect to the corresponding norm. Let W be a standard one-dimensional Brownian motion, and let $x \in \mathscr{X}$. Then, by Theorem 2.1, one has

$$f(W^t + x) = f(x) + \int_0^t \langle \eta_s, \bigtriangledown f(W^s + x)\rangle \, dW_s + \tfrac{1}{2}\int_0^t \langle \eta_s \otimes \eta_s, \bigtriangledown^2 f(W^s + x)\rangle \, ds.$$

Under assumptions of integrability, one obtains

$$Ef(W^t + x) = f(x) + \int_0^t EA(s) f(W^s + x)\, ds,$$

where $A(t)f(x) = 1/2\langle \eta_t \otimes \eta_t, \bigtriangledown^2 f(x)\rangle$. Especially if the expectation and $A(\cdot)$ are interchangeable, one obtains a weak form of a partial differential equation with initial data f. Recall that $\langle \eta_t \otimes \eta_t, \bigtriangledown^2 f(w^t + x)\rangle$ as well as $\langle \eta_t, \bigtriangledown f(w^t + x)\rangle$ are continuous in t, if $w \in C[0,T]$ and if $f\colon L^2[0,T] \to \mathbb{R}$ is twice continuously Fréchet differentiable at w; see Proposition 2.1. This property enables us to differentiate $Ef(W^t + x)$ with respect to t, and for this reason, $L^2[0,T]$ is preferable to $C[0,T]$. For example, if $f(x) = g(x(t_1), \ldots, x(t_k))$ where $g \in C^2(\mathbb{R}^k, \mathbb{R})$, then $\langle \eta_t, \bigtriangledown f(w^t + x)\rangle$ has jumps at every t_i, and so does the second derivative. Thus, in this case, $Ef(W^t + x)$ has a piecewise differentiable trajectory with singularities on every t_i. Despite this inconvenience, these types of equations are indispensable; eventually one will have to deal with a finite-dimensional approximation.

In this section, we focus on the backward equation with terminal data f, which agrees with (2). Using this backward equation, we provide integral representations of a smooth functional of a diffusion process.

THEOREM 4.1. *Let $f\colon L^2[0,T] \to \mathbb{R}$ be twice continuously Fréchet differentiable at each $x \in C[0,T]$, and*

$$|f(x)| + \| \bigtriangledown f(x)\| + \| \bigtriangledown^2 f(x)\| \le K(1 + \|x\|^p) \tag{11}$$

for some positive numbers K and p. Furthermore, suppose that σ and b have bounded first derivatives. Then, for each $x \in C[0,T]$, there exists a real valued function u satisfying

$$\frac{\partial u}{\partial t}(t,x) + \frac{1}{2}\langle \eta_t \otimes \eta_t, \bigtriangledown^2 u(t,x)\rangle \sigma(x(t))^2 + \langle \eta_t, \bigtriangledown u(t,x)\rangle b(x(t)) = 0 \tag{12}$$

with $u(T,x) = f(x)$.

PROOF. Let $Y_t(s,x)$ be a solution of

$$Y_t = \int_s^t \sigma(Y_r + x(r))\,dW_r + \int_s^t b(Y_r + x(r))\,dr \tag{13}$$

for $t > s$, and 0 otherwise. It is well known that the process $Y(s,x)$ is unique, and $E\sup_{t\le T}|Y_t(s,x)|^n$ is finite for each positive integer n; see Ikeda and Watanabe [6], page 240. Again, $Y^t(s,x)$ will denote the corresponding stopped process. Note that $Y_t(r,x) - Y_s(r,x) = Y_t(s, Y^s(r,x)+x)$ holds for $0 \le r \le s \le t \le T$. Then, by the Markov property, we have

$$E\{f(Y(s,x)+x)|\mathscr{F}_t\} = [Ef(Y(t,z)+z)]_{z=Y^t(s,x)+x}, \tag{14}$$

for each $s < t$. Define a two parameter family U by $U(s,t)f(x) = Ef(Y^t(s,x)+x)$ for $0 \le s \le t \le T$. Then (14) implies that $U(s,T)f = U(s,t)U(t,T)f$ for $0 \le s < t \le T$. We will show that U is an evolution system (see Pazy [11]) corresponding to our equation. Now, for each $h > 0$, we have

$$\frac{1}{h}\{U(t,t+h) - I\}f(x) = \frac{1}{h}\int_t^{t+h} EA(s)f(Y^s(t,x)+x)\,ds,$$

where A is a differential operator defined by

$$A(t)f(x) = \tfrac{1}{2}\langle \eta_t \otimes \eta_t, \triangledown^2 f(x)\rangle \sigma(x(t))^2 + \langle \eta_t, \triangledown f(x)\rangle b(x(t)). \tag{15}$$

This follows from Itô's representation of $f(Y^{t+h}(t,x)+x) - f(x)$. Note that $A(s)f(Y^s(t,x)+x)$ converges almost surely to $A(t)f(x)$ as s tends to t from above; if $C[0,T]$ were adopted (instead of $L^2[0,T]$), it would have been $A(t+)f(x)$. Thus, by the uniform integrability, we obtain

$$\begin{aligned}\lim_{h\downarrow 0}\frac{1}{h}\{U(t,T) - U(t+h,T)\}f &= \lim_{h\downarrow 0}\frac{1}{h}\{U(t,t+h) - I\}U(t+h,T)f \\ &= A(t)U(t,T)f\end{aligned}$$

for each $t \le T$. Similarly, $(1/h)\{U(t-h,T) - U(t,T)\}f$ also converges to $A(t)U(t,T)f$ as $h \downarrow 0$. Therefore the derivative of $U(t,T)f$ with respect to t is $-A(t)U(t,T)f$, and hence $U(t,T)f(x) = Ef(Y(t,x)+x)$ satisfies the equation. □

The solution defined by $Ef(Y(t,x)+x)$ is called the *canonical solution* of the equation. Note that $Y(t,x) = W - W^t$, if $\sigma = 1$ and $b = 0$ identically. The operation in $W - W^t + x$ is a coordinate-wise shift, and one can interpret $Y(t,x)$ in the same manner. When x has jumps, neither $\sigma(x(t))$ nor $b(x(t))$ are continuous except when they are constant, and $Ef(Y(t,x)+x)$ is not differentiable with respect to t at those jump times. As we have discussed earlier, the nondifferentiability issue arises more casually when the smoothness of f is considered in the $C[0,T]$ sense only. Yet a similar result is conceivable in the weak sense:

$$u(t,x) = f(x) + \int_t^T A(s)u(s,x)\,ds$$

where A is defined as in (15). Next we discuss the regularity of the canonical solution.

LEMMA 4.1. *Let f: $L^2[0,T] \to \mathbb{R}$ be twice continuously Fréchet differentiable at each $x \in C[0,T]$, and satisfy (11). Suppose that σ and b have bounded continuous derivatives up to order 2. Then the derivatives of the canonical solution, $\partial u/\partial t$, $\triangledown u$, and $\triangledown^2 u$, are continuous in both directions (i.e., time and space).*

PROOF. First, we evaluate $\triangledown u$ and $\triangledown^2 u$ using stochastic flows. Define $M_r = M_r(t,x,h)$ by

$$M_r = \int_t^r (h(s) + M_s)\, d\xi_s(t,x),$$

where

$$\xi_r(t,x) = \int_t^r \sigma'(Y_s(t,x) + x(s))\, dW_s + \int_t^r b'(Y_s(t,x) + x(s))\, ds.$$

Then M is linear in h (see Protter [13], page 266). Since

$$E\{f(Y(t,x+h) + x + h) - f(Y(t,x) + x) - \langle \triangledown f(Y(t,x) + x), M(t,x,h) + h\rangle\}$$

is $o(\|h\|)$, by the uniqueness of the Riesz representor, we have

$$\langle \triangledown u(t,x), h\rangle = E\langle \triangledown f(Y(t,x) + x), M(t,x,h) + h\rangle. \tag{16}$$

Next, we define $N_r = N_r(t,x,h,g)$ by

$$N_r = \int_t^r N_s\, d\xi_s(t,x) + \int_t^r (M_s(t,x,h) + h(s))(M_s(t,x,g) + g(s))\, d\zeta_s,$$

where

$$\zeta_r(t,x) = \int_t^r \sigma''(Y_s(t,x) + x(s))\, dW_s + \int_t^r b''(Y_s(t,x) + x(s))\, ds.$$

Then, $\langle \triangledown^2 u(t,x), h \otimes g\rangle$ can be represented by

$$\begin{aligned} &E\langle \triangledown^2 f(Y(t,x) + x), (M(t,x,h) + h) \otimes (M(t,x,g) + g)\rangle \\ &\qquad + E\langle \triangledown f(Y(t,x) + x), N(t,x,h,g)\rangle. \end{aligned} \tag{17}$$

It can be shown that Y, M, N, ξ, and ζ are continuous in both t and x with respect to the uniform L^2 metric on compacts. See, for instance, Protter [12]. The continuity of $\triangledown u$ and $\triangledown^2 u$, then, follows directly from (16) and (17), respectively. Replacing both h and g by η_t, one proves the continuity of $\partial u/\partial t$. $\square$

Now we present applications of Itô's representation for functionals and the canonical solution. It is known that a square integrable random variable in the canonical Wiener space has a martingale representation as a stochastic integral with respect to a Brownian motion. Particularly if the random variable is Malliavin differentiable, one retains the following formula:

$$Y = E(Y) + \int_0^T E(D_t Y|\mathscr{F}_t)\,dW_t,$$

where $D_t Y$ denotes the Malliavin derivative. This formula is known as Clark's formula; see Karatzas, Ocone and Li [7]. The following result shows that if Y depends smoothly on the path of the Brownian motion or diffusion process, then the integrand in Clark's formula as well as $E(Y)$ can be expressed in terms of a solution to a partial differential equation.

THEOREM 4.2. *Let X be a diffusion process defined by*

$$X_t = X_0 + \int_0^t \sigma(X_s)\,dW_s + \int_0^t b(X_s)\,ds,$$

where σ and b have bounded continuous derivatives up to order 2. Suppose that f: $L^2[0,T] \to \mathbb{R}$ is twice continuously Fréchet differentiable at each $x \in C[0,T]$, and satisfies (11). Then we have

$$f(X) = u(0, X_0 \mathbb{1}_{[0,T]}) + \int_0^T \langle \eta_s, \nabla u(s, X^s)\rangle \sigma(X_s)\,dW_s \tag{18}$$

where u is the canonical solution of (12) with $u(T,x) = f(x)$.

Equation (18) is a consequence of Theorem 2.4. In fact, it is valid for each solution u of (12) which satisfies the regularities required for Theorem 2.4. Note that the martingale representation of $f(X)$ is unique. Therefore $u(0, x(0)\mathbb{1}_{[0,T]})$ and $\langle \eta_t, \nabla u(t, x^t)\rangle$ must be the same for any solution u which has continuous derivatives. Next, consider the following stochastic differential equation:

$$X_t = H_t + \int_0^t \sigma(X_s)\,dW_s \tag{19}$$

where H is a continuous FV process and σ has bounded continuous derivatives up to order 2. The uniqueness and the existence of the solution to (19) are given in Protter [13]. However the solution may not be a Markov process. The following result, the *semimartingale integral representation*, is another application of Theorem 2.4.

THEOREM 4.3. *Let X be a solution to (19) where H is a continuous FV process and σ has bounded continuous derivatives up to order 2. Suppose that f: $L^2[0,T] \to \mathbb{R}$ is twice continuously Fréchet differentiable at each $x \in C[0,T]$ and satisfies (11). Then we have*

$$f(X) = u(0, X_0 \mathbb{1}_{[0,T]}) + \int_0^T \langle \eta_s, \nabla u(s, X^s)\rangle\,dX_s,$$

where u is the canonical solution of (12) with $b = 0$ and $u(T,x) = f(x)$.

REFERENCES

[1] AMBROSETTI, A. and PRODI, G. (1995). *A Primer of Nonlinear Analysis*. *Cambridge Stud. Adv. Math.* 34.

[2] CLARK, J. M. C. (1970). The representation of functionals of Brownian motion by stochastic integrals. *Ann. Math. Statist.* 41 1282–1295.

[3] DUFFIE, D. (1992). *Dynamic Asset Pricing Theory.* Princeton Univ. Press.

[4] FÖLLMER, H., PROTTER, P. and SHIRYAEV, A. N. (1995). Quadratic covariation and an extension of Itô's formula. *Bernoulli* 1 149–169.

[5] FRIEDMAN, A. (1970). *Foundations of Modern Analysis.* Holt, Rinehart and Winston, New York.

[6] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam.

[7] KARATZAS, I., OCONE, D. L. and LI, J. (1991). An extension of Clark's formula. *Stochastics* 37 127–131.

[8] KUNITA, H. (1990). Stochastic flows and stochastic differential equations. *Cambridge Stud. Adv. Math.* 24.

[9] MA, J., PROTTER, P. and YONG, J. (1994). Solving forward–backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields* 98 339–359.

[10] MÉTIVIER, M. (1982). *Semimartingales, a Course on Stochastic Processes.* de Gruyter, Berlin.

[11] PAZY, A. (1983). Semigroups of linear operators and applications to partial differential equations. *Appl. Math. Sci.* 44.

[12] PROTTER, P. (1978). H^p stability of solutions of stochastic differential equations. *Z. Wahrsch. Verw. Gebiete* 44 337–352.

[13] PROTTER, P. (1990). Stochastic integration and differential equations, a new approach. *Appl. Math.* 21.

[14] SCHATTEN, R. (1950). A theory of cross-spaces. *Ann. of Math. Stud.* 26.

DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CALIFORNIA 93106