ILLINOIS JOURNAL OF MATHEMATICS
Volume 36, Number 3, Fall 1992

ON SEMIMARTINGALE DECOMPOSITIONS OF CONVEX FUNCTIONS OF SEMIMARTINGALES

BY

ERIC CARLEN[1] AND PHILIP PROTTER[2]

Let X be a semimartingale with values in $\mathbf{R}^d$, and let $X_t = X_0 + M_t + A_t$ be a decomposition of X into a local martingale M and a càdlàg, adapted, finite variation process A, with $M_0 = A_0 = 0$. Let $f\colon \mathbf{R}^d \to \mathbf{R}$ be convex. P.A. Meyer showed in 1976 [6] that $f(X)$ is again a semimartingale. We will give a new proof of this result which moreover gives the semimartingale decomposition of $f(X)$ in terms of uniform limits of explicitly identified processes.

The case where $d = 1$ is already well understood. Indeed, the Meyer-Tanaka formula allows us to give an explicit decomposition of $f(X)$:

$$(1) \qquad f(X_t) = f(X_0) + \int_0^t f'(X_{s-})\, dM_s + \left\{ \int_0^t f'(X_{s-})\, dA_s + \frac{1}{2}\int_{\mathbf{R}} L_t^a \mu(da) + \sum_{0<s\le t} \left(f(X_s) - f(X_{s-}) - f'(X_s)\,\Delta X_s\right) \right\},$$

where f' is the left continuous version of the derivative of f, L_t^a is the local time of X at the level a, the measure μ is the second derivative of f in the generalized function sense, and the term in brackets $\{\cdots\}$ is the finite variation term in a decomposition of $f(X)$. See [8] for details on this formula. Moreover in the case $d = 1$ if B is a standard Brownian motion and $f(B)$ is a semimartingale, then f must be the difference of two convex functions (see [3]), hence convex functions are the most general functions taking semimartingales into semimartingales.

We now turn to the case $d \ge 2$, where $f\colon \mathbf{R}^d \to \mathbf{R}$ is convex. Except in very special cases (see [2], [4], [5], [7], [9], [10]) no formula such as (1) is known to exist, except of course when f is $\mathscr{C}^2$, and then the Meyer-Itô formula gives

Received October 23, 1990.
1991 Mathematics Subject Classification. Primary 60H05; Secondary 60G44, 60G48.
[1]NSF Postdoctoral Fellow.
[2]Supported in part by a grant from the National Science Foundation.

an explicit decomposition of $f(X)$:

(2)

$$f(X_t) = f(X_0) + \sum_{j=1}^{d} \int_0^t \frac{\partial f}{\partial x_j}(X_{s-})\, dM_s^j$$
$$+ \left\{ \sum_{j=1}^{d} \int_0^t \frac{\partial f}{\partial x_j}(X_{s-})\, dA_s^j + \frac{1}{2} \sum_{i,j} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s-})\, d[X^i, X^j]_s^c \right.$$
$$\left. + \sum_{0<s\le t} \left(f(X_s) - f(X_{s-}) - \sum_{j=1}^{d} \frac{\partial f}{\partial x_j}(X_{s-})\, \Delta X_s^j \right) \right\},$$

where $X_t^j = X_0^j + M_t^j + A_t^j$ denotes the semimartingale decomposition of the jth component of the vector X of d semimartingales.

Let Γ denote the set of convex functions on $\mathbf{R}^d$, and recall that convex functions are always continuous. We equip Γ with the topology of uniform convergence on compacts. A standard metric ρ for this topology is given by $\rho(f, g) = \sum_{n=1}^{\infty} 2^{-n} \rho_n(f, g)$ where

$$\rho_n(f, g) = \frac{\sup_{|x| \le n} |f(x) - g(x)|}{1 + \sup_{|x| \le n} |f(x) - g(x)|}.$$

By an obvious convolution argument, $\mathscr{C}^2$ convex functions are dense in (Γ, ρ).

We show here that if $\{f_n\}$ is a sequence of $\mathscr{C}^2$ convex functions converging to f in (Γ, ρ), and if $f_n(X_t) = f_n(X_0) + N_t^n + S_t^n$ is an appropriately chosen decomposition of $f_n(X_t)$, then the corresponding local martingale terms N^n and finite variation terms S^n converge respectively to N and S, where $f(X_t) = f(X_0) + N_t + S_t$, a decomposition of $f(X)$. This gives a decomposition of $f(X)$ in terms of limits of explicitly identified processes. The proof consists essentially of verifying the hypotheses of a recent theorem of Barlow and Protter [1].

To do this, we require the following lemma:

LEMMA. *Let $\{f_n\}$ be a sequence of $\mathscr{C}^2$ convex functions on $\mathbf{R}^d$, f convex on $\mathbf{R}^d$, and $\lim_{n\to\infty} \rho(f_n, f) = 0$. Then for each $\alpha > 0$,*

$$\sup_n \sup_{|x| \le \alpha} |\nabla f_n(x)| \le C(\alpha) < \infty,$$

where $C(\alpha)$ depends only on α and f.

Proof. Since $\rho(f_n, f)$ tends to 0, the variation of f_n on $\{|x| \le \alpha + 1\}$ is uniformly bounded in n by, say, $V(\alpha)$. Let x_n be some point in $\{|x| \le \alpha\}$ such that

$$|\nabla f_n(x_n)| = \sup_{|x| \le \alpha} |\nabla f_n(x)|.$$

Let u_n denote $\nabla f_n(x_n)/|\nabla f_n(x_n)|$. Define φ_n by $\varphi_n(t) = f_n(x_n + tu_n)$. Then φ_n is a $\mathscr{C}^2$ convex function on $\mathbf{R}$. Therefore, for $t \ge 0$, $\varphi_n'(t) \ge \varphi_n'(0) = \nabla f_n(x) \cdot u_n = |\nabla f_n(x_n)|$. Since φ_n is convex, $\varphi_n'(t) \ge |\nabla f_n(x_n)|$ for all positive t. Thus

$$f_n(x_n + u_n) - f_n(x_n) = \int_0^1 \varphi_n'(t)\, dt \ge |\nabla f_n(x_n)|.$$

Since $|x_n + u_n| \le \alpha + 1$ we have $|f_n(x_n + u_n) - f_n(x_n)| \le V(\alpha)$, and therefore $|\nabla f_n(x_n)| \le V(\alpha)$. □

The next theorem is our principal theorem. Because we wish to use the result of [1], and also because of the simplifications entailed in the existence of canonical decompositions, we consider in Theorem 1 the case where the semimartingale X is in $\mathscr{H}^1$; (that is, X has a decomposition $X_t = X_0 + M_t + A_t$ where X_0, $[M, M]_\infty^{1/2}$ and $\int_0^\infty |dA_s|$ are all in L^1.) In Theorem 2 we consider the general case where X is locally in $\mathscr{H}^1$; that is there exists a sequence $(T^n)_{n \ge 1}$ of stopping times increasing to ∞ a.s. such that $X_{t \wedge T^n} 1_{\{T_n > 0\}}$ is in $\mathscr{H}^1$ for each n. Note that if X is a *continuous* semimartingale, the X is automatically at least locally in $\mathscr{H}^1$. We let $\|\cdot\|_{\mathscr{H}^1}$ denote the H^1 norm (see [8]), and $A_t^* = \sup_{s \le t} |A_s|$.

THEOREM 1. *Let X be an $\mathbf{R}^d$-valued semimartingale in $\mathscr{H}^1$. Let $X_0 = 0$ and $X_t = N_t + S_t$ be its canonical decomposition. For $\alpha > 0$, let*

$$T_\alpha = \inf\{t > 0: |X_t| > \alpha\}.$$

Let f be a convex function, and let $\{f_n\}$ be a sequence of $\mathscr{C}^2$ convex functions with $\lim_{n \to \infty} \rho(f_n, f) = 0$. Then $f(X)$ is a semimartingale with canonical decomposition $f(X_t) = f(X_0) + M_t + A_t$, and moreover, for each $\alpha > 0$, we have,

$$\lim_{n \to \infty} \|(M^n - M)^{T_\alpha}\|_{\mathscr{H}^1} = 0,$$

$$\lim_{n \to \infty} E\{(A^n - A)_{T_\alpha}^*\} = 0,$$

where

$$M_t^n = \int_0^t \nabla f_n(X_{s-})\, dN_s$$

and

$$(3) \qquad A_t^n = \int_0^t \nabla f_n(X_{s-})\, dS_s + \frac{1}{2} \sum_{i,j} \int_0^t \frac{\partial^2 f_n}{\partial x_i \partial x_j}(X_{s-})\, d[X^i, X^j]_s^c$$
$$+ \sum_{0<s\le t} \left\{ f_n(X_s) - f_n(X_{s-}) - \sum_i \frac{\partial f}{\partial x_i}(X_{s-})\, \Delta X_s^i \right\}.$$

Proof. We need to verify only that the hypotheses of Theorem 1 of Barlow and Protter [1] are satisfied; specifically we must show that for each $\alpha > 0$,

$$(4) \qquad \lim_{n\to\infty} E\left\{ \sup_{t\le T_\alpha} |f_n(X_t) - f(X_t)| \right\} = 0,$$

and that there is a $K_\alpha < \infty$ such that

$$(5) \qquad \sup_n E\left\{ \int_0^{T_\alpha} |dA_s^n| \right\} \le K_\alpha,$$

$$(6) \qquad \sup_n E\left\{ \sup_{t\le T_\alpha} |M_t^n| \right\} \le K_\alpha.$$

First observe that (4) is a trivial consequence of $\lim_{n\to\infty} \rho(f_n, f) = 0$. Also, note that using the lemma together with the Davis inequality,

$$E\left\{ \sup_{t\le T_\alpha} \left| \int_0^t \nabla f_n(X_{s-})\, dN_s \right| \right\} \le cE\left\{ \left(\int_0^{T_\alpha} |\nabla f_n(X_{s-})|^2 d[N,N]_s \right)^{1/2} \right\}$$
$$\le cC(\alpha) E\left\{ [N,N]_{T_\alpha}^{1/2} \right\},$$

since $|X_-|$ is bounded by α on $[0, T_\alpha]$. The above holds for each n and since the bound is independent of n, we have (6).

We next turn to (5). We treat separately the three terms in (3). First, again using the lemma,

$$\text{Variation}\left(\int_0^t \nabla f_n(X_{s-})\, dS_s \right) \le C(\alpha) \int_0^{T_\alpha} |dS_s|,$$

which is independent of n. Second, let B^n denote the process

$$B_t^n = \frac{1}{2}\sum_{i,j}\int_0^t \frac{\partial^2 f_n}{\partial x_i\,\partial x_j}(X_{s-})\,d[X^i, X^j]_s^c.$$

Since f_n is convex,

$$\left(\frac{\partial^2 f_n}{\partial x_i\,\partial x_j}\right)$$

is a positive matrix, and also $d[X^i, X^j]^c$ is positive in the sense that for any constants $a_i, \dots, a_d$, $\sum_{i,j=1}^d a_i a_j [X^i, X^j]^c$ is an increasing process. Thus B^n is an increasing process. Next, let D^n denote the third term in (3); that is,

$$\begin{aligned} D_t^n &= \sum_{0<s\le t}\{f_n(X_s) - f_n(X_{s-}) - \nabla f_n(X_{s-})\,\Delta X_s\} \\ &= \sum_{0<s\le t}\frac{\partial^2 f}{\partial x_i\,\partial x_j}(X_{s-} + \mathcal{O}_s)\,\Delta X_s^i\,\Delta X_s^j \end{aligned}$$

where $\mathcal{O}_s = \lambda_s\,\Delta X_s$ for some $\lambda_s \in [0,1]$ by Taylor's theorem. The convexity of f_n yields that D^n is also an increasing process.

Next observe that, letting V_α denote total variation on $[0, T_\alpha]$:

$$\begin{aligned} V_\alpha(A_t^n) &= V_\alpha\left(\int_0^t \Delta f_n(X_{s-})\,dS_s + B_t^n + D_t^n\right) \\ &\le C(\alpha)|S|_{T_\alpha} + B_{T_\alpha}^n + D_{T_\alpha}^n. \end{aligned} \tag{7}$$

However by the Meyer-Itô formula (2) and since the expectation of the (true) martingale term is zero,

$$E\{B_{T_\alpha}^n + D_{T_\alpha}^n\} = E\{f_n(X_{T_\alpha}) - f_n(X_0)\} + E\left\{\int_0^{T_\alpha} \nabla f_n(X_{s-})\,dS_s\right\}. \tag{8}$$

Since f_n tends uniformly to f, and since $E\{\int_0^{T_\alpha} \nabla f_n(X_{s-})\,dS_s\}$ is bounded by $C(\alpha)E\{|S|_{T_\alpha}\}$ independently of n, the right side of (8) is bounded by a K_α for n sufficiently large, and hence for all n. Combining this with (7) and taking expectations yields (5) and completes the proof. □

We next turn to the general case which is handled by "prelocal" stopping: Suppose X is a semimartingale with $X_0 = 0$. Then as is well known (see, e.g. [8, p. 192]) there exist stopping times T^k increasing to ∞ a.s. such that X^{T^k-}

is in $\mathscr{H}^1$, each k, where

$$X_t^{T^k-} = X_t 1_{(t<T^k)} + X_{T^k-} 1_{(t \geq T^k)}.$$

Therefore, by taking $T^{k,\alpha}$ to be $T_\alpha \wedge T^k$, we can further assume without loss that $|X^{T^{k,\alpha}-}| \leq \alpha$, for a sequence T_α as given in Theorem 1. We combine the sequences to get T_α increasing to ∞ a.s. such that $|X^{T_\alpha-}| \leq \alpha$ *and* $X^{T_\alpha-} \in \mathscr{H}^1$, each α. We then have:

THEOREM 2. *Let X be an $\mathbf{R}^d$-valued semimartingale with $X_0 = 0$. Let T^α be stopping times increasing to ∞ such that $|X^{T_\alpha-}| \leq \alpha$ and $X^{T_\alpha-} \in \mathscr{H}^1$. Let $X^{T_\alpha-} = N^\alpha + S^\alpha$ be the canonical decomposition, f be a convex function, and f_n a sequence of $\mathscr{C}^2$ convex functions with $\lim_{n\to\infty} \rho(f_n, f) = 0$. Then $f(X)$ is a semimartingale with prelocal canonical decompositions*

$$f(X)^{T_\alpha-} = f(X_0) + M^\alpha + A^\alpha;$$

moreover

$$\lim_{n\to\infty} \|M^{n,\alpha} - M^\alpha\|_{\mathscr{H}^1} = 0$$

$$\lim_{n\to\infty} E\{(A^{n,\alpha} - A^\alpha)^*\} = 0$$

where

$$M_t^{n,\alpha} = \int_0^t \nabla f_n(X_{s-})\, dN_s^\alpha,$$

$$A_t^{n,\alpha} = \int_0^t \nabla f_n(X_{s-})\, dS_s^\alpha + \frac{1}{2}\sum_{i,j}\int_0^t \frac{\partial^2 f_n}{\partial x_i \partial x_j}(X_{s-})\, d[X^i, X^j]_s^{c,T_\alpha-} + \sum_{0<s\leq t}\left\{f_n(X_s)^{T_\alpha-} - f_n(X_{s-})^{T_\alpha-} - \sum_i \frac{\partial f}{\partial x_i}(X_{s-})(\Delta X_s^i)^{T_\alpha-}\right\}.$$

Proof. This is merely a localization of Theorem 1; since f is continuous $f(X)^{T-} = f(X^{T-})$. $\square$

Remarks (i) Note that in case X is *continuous* the situation is much simpler:

$$A_t^n = \int_0^t \nabla f_n(X_s)\, dS_s,$$

since there are no jump terms; decompositions are unique, hence there is no need to invoke "canonical" decompositions; there is no need of "pre-local" stopping, since stopping at $T-$ is the same as stopping at T.

(ii) The general case where X_0 need not be zero is easily handled: take $\hat{f}(X) = f(X) - f(0)$, so that without loss of generality we can assume $f(0) = 0$. Since $X_0 \neq 0$, one cannot assume that $|X^{T_\alpha -}| \leq \alpha$, however one can construct T_α tending to ∞ a.s. such that $|X^{T_\alpha -}1_{\{T_\alpha > 0\}}| \leq \alpha$. Since $f(0) = 0$ and f is continuous, $f(X^{T_\alpha -}1_{\{T_\alpha > 0\}}) = f(X)^{T_\alpha -}1_{\{T_\alpha > 0\}}$, and the proof now proceeds analogously.

(iii) "Knowing" M^α and A^α in the decomposition $f(X)^{T_\alpha -} = f(X_0) + M^\alpha + A^\alpha$ also means we "know" a decomposition for $f(X)^{T_\alpha}$: namely, we can take

$$f(X_t)^{T_\alpha} = f(X_0) + M_t^\alpha + \left\{A_t^\alpha + \left(f(X_{T_\alpha}) - f(X_{T_\alpha -})\right)1_{\{t \geq T_\alpha\}}\right\}. \tag{9}$$

Note however that we cannot in general combine these decompositions (9) to obtain only one, because of the lack of a canonical way to choose them. (Of course, in the continuous case this is not a problem.)

(iv) Finally we would like to point out that we have used the convexity of f in two ways in the proofs of Theorems 1 and 2: first through the lemma to control the size of $\int \nabla f_n(X_{s-})\, dS_s$; second, to establish that $A^n - \int \nabla f_n(X_{s-})\, dS_s$ is an increasing process—this gave us the estimate (7) which in turn allowed us to take expectations in the Meyer-Itô formula.

Acknowledgement. The first author would like to thank Purdue University for its hospitality during a visit on which discussions leading to this paper took place. The second author wishes to acknowledge preliminary discussions on related problems with Tom Kurtz, Vigirdas Mackevičius, and Jaime San Martin.

References

1. M.T. Barlow and P. Protter, "On Convergence of Semimartingales" in *Séminaire de Probabilités XXIV*, Lecture Notes in Math., vol. 1426, Springer-Verlag, New York, 1990, pp. 188–193.
2. G.A. Brosamler, *Quadratic variation of potentials and harmonic functions*, Trans. Amer. Math. Soc., 149 (1970), pp. 243–257.
3. E. Cinlar, J. Jacod, P. Protter and M. Sharpe, *Semimartingales and Markov processes*, Z. Wahrsch. Verw. Geberte, vol. 54 (1980), pp. 161–220.
4. N.V. Krylov, *Controlled diffusion processes*, Springer-Verlag, New York, 1980.
5. H.H. Kuo and N.R. Shieh, *A generalized Itô's formula for multidimensional Brownian motion and Its applications*, Chinese J. Math., vol. 15 (1987), pp. 163–174.
6. P.A. Meyer, "Un Cours sur les Intégrales Stochastiques" in *Séminaire de Probabilités X*, Lecture Notes in Math., vol. 511, Springer-Verlag, New York, 1976, pp. 246–400.

7. ______, "La Formule d'Itô pour le Mouvement Brownien d'après G. Brosamler" in *Séminaire de Probabilités XII*, Lecture Notes in Math., vol. 649, Springer-Verlag, New York, 1978, pp. 763–769.
8. P. Protter, *Stochastic integration and differential equations*, Springer-Verlag, New York, 1990.
9. J. Rosen, *A representation for the intersection local time of Brownian motion in space*, Ann. Probability, vol. 13 (1985), pp. 145–153.
10. M. Yor, "Complements aux Formulas de Tanaka-Rosen" in *Séminaire de Probabilités XIX*, Lecture Notes in Math., vol. 1123 (1985), pp. 332–349. Springer-Verlag, New York, 1985, pp. 332–349.

Princeton University
Princeton, New Jersey

Purdue University
West Lafayette, Indiana