On Itô's formula for multidimensional Brownian motion

by

*Hans Föllmer and Philip Protter**

Institut für Mathematik
Humboldt-Universität
D-10099 Berlin

Mathematics and Statistics Department
Purdue University
West Lafayette IN 47907-1395, USA

Abstract. Consider a d-dimensional Brownian motion $X = (X^1, \ldots, X^d)$ and a function F which belongs locally to the Sobolev space $W^{1,2}$. We prove an extension of Itô's formula where the usual second order terms are replaced by the quadratic covariations $[f_k(X), X^k]$ involving the weak first partial derivatives f_k of F. In particular we show that for any locally square-integrable function f the quadratic covariations $[f(X), X^k]$ exist as limits in probability for any starting point, except for some polar set. The proof is based on new approximation results for forward and backward stochastic integrals.

Key words: Itô's formula, Brownian motion, stochastic integrals, quadratic covariation, Dirichlet spaces, polar sets.

**Supported in part by ONR grant # N00014-96-1-0262 and NSF grant # 9401109-INT*

1. Introduction

The behavior of a smooth function F on R^d along the paths of d-dimensional Brownian motion is described as follows by Itô's formula. Let P_x be the distribution of Brownian motion with initial point x, and let $X = (X^1, \ldots, X^d)$ denote the coordinate process on the canonical path space $\Omega = C([0,\infty), R^d)$. Consider the process A defined by

$$A_t = F(X_t) - F(X_0) - \sum_{k=1}^{d} \int_0^t f_k(X_s) dX_s^k, \tag{1.1}$$

where we denote by $f_k = \frac{\partial F}{\partial x_k}$ the partial derivatives of F. Itô's formula provides an alternative description of the process A:

$$A_t = \frac{1}{2} \int_0^t \Delta F(X_s) ds \qquad P_x - a.s. \tag{1.2}$$

for any $t \geq 0$, and for any starting point $x \in R^d$.

Note, however, that the description (1.2) in terms of the Laplace operator Δ involves second order differentiability of F, while definition (1.1) requires only differentiability of first order. In fact, the process in (1.1) is well defined whenever F belongs to the Sobolev space $W^{1,2}$, at least locally. In this case, we choose an appropriate version of F and use the weak first derivatives f_k in order to define (1.1) P_x- almost surely for all $x \notin E$, where E is some polar set. Thus the question arises how to formulate an analogue to (1.2) for a general function $F \in W^{1,2}_{\text{loc}}$. Of course we can always approximate F by smooth functions $F^{(n)}$ in such a way that the terms in (1.1) converge to the corresponding terms for F, and then we get the description

$$A_t = \lim_{n \to \infty} \frac{1}{2} \int_0^t \Delta F^{(n)}(X_s) ds. \tag{1.3}$$

But rather we are interested in an intrinsic description which directly involves the function F itself.

It turns out that such an intrinsic description can be given in terms of quadratic covariation. We show that for any initial point $x \in R^d$, except for some polar set, the quadratic covariations $[f_k(X), X^k]$ exist as limits in probability of the usual sums under the measure P_x. Our extension of Itô's formula consists in identifying the process A defined by (1.1) as

$$A_t = \frac{1}{2} \sum_{k=1}^{d} [f_k(X), X^k]_t \qquad P_x - \text{a.s.} \tag{1.4}$$

for all x except for some polar set.

If F is the difference of two positive superharmonic functions so that the distribution $\frac{1}{2}\Delta F$ is given by a signed measure μ, then (1.5) provides an explicit description of the additive functional associated to μ which appears in the extended Itô formula of Brosamler (1970) and Meyer (1978). In the general case $F \in W^{1,2}$, we can view F as a function in the Dirichlet space associated to d-dimensional Brownian motion. From this point of view, A is the process of zero energy appearing in Fukushima's decomposition of the process $F(X_t)$ $(t \geq 0)$; cf. Fukushima (1980). Thus, our formula (1.5) provides an explicit construction of the process of zero energy in terms of quadratic covariation.

In the one-dimensional case, the extension (1.5) of Itô's formula was shown in Föllmer, Protter and Shiryaev (1995). In this paper we consider the case $d \geq 2$. The basic idea is the same: The existence of the quadratic covariations in (1.4) is shown by proving that the forward and the backward stochastic integrals of $f_k(X)$ can be approximated by the corresponding sums. But in contrast to the one-dimensional case, these approximation results hold only for all initial points x outside some exceptional set of capacity zero, and the proofs are more subtle. In section 2 we fix a starting point $x_0 \in R^d$ and a measurable function f on R^d. We formulate two integrability conditions on f in terms of x_0 which guarantee that both the forward and the backward stochastic integral can be constructed in a straightforward manner as limits in probability

$$\int_0^t f(X_s)dX_s^k = \lim_{n\to\infty} \sum_{\substack{t_i \in D_n \\ 0<t_i<t}} f(X_{t_i})(X_{t_{i+1}}^k - X_{t_i}^k) \tag{1.5}$$

and

$$\int_0^t f(X_s)d^*X_s^k = \lim_{n\to\infty} \sum_{\substack{t_i \in D_n \\ 0<t_i<t}} f(X_{t_{i+1}})(X_{t_{i+1}}^k - X_{t_i}^k) \tag{1.6}$$

under the measure P_{x_0}. This implies the existence of the quadratic covariations

$$[f(X), X^k]_t = \lim_{n\to\infty} \sum_{\substack{t_i \in D_n \\ 0<t_i<t}} \{f(X_{t_{i+1}}) - f(X_{t_i})\}(X_{t_{i+1}}^k - X_{t_i}^k) \tag{1.7}$$

as limits in probability under the measure P_{x_0}, and their identification as differences

$$[f(X), X^k]_t = \int_0^t f(X_s)d^*X_s^k - \int_0^t f(X_s)dX_s^k \tag{1.8}$$

of backward and forward stochastic integrals. In section 3 we consider a measurable function f such that

$$P_{x_0}[\int_0^t f^2(X_s)ds < \infty] = 1. \tag{1.9}$$

at least for some x_0 and for some t. Note that condition (1.9) is clearly a minimal requirement if we want to talk about stochastic integrals of $f(X)$. Using results of Höhnle and Sturm (1993) on multidimensional analogues of the Engelbert-Schmidt 0 - 1 law, we show that condition (1.9) implies that our integrability conditions in section 2 for the existence of the quadratic covariations $[f(X), X^k]$ are satisfied for all starting points except for some polar set. In section 4 we apply these results to the weak derivatives f_k of a function $F \in W^{1,2}$. This leads us to our characterization (1.5) of the process A defined by (1.1).

Using the identification (1.8) of the quadratic covariations $[f_k(X), X^k]$, our version (1.5) of Itô's formula can also be written in the form

$$F(X_t) - F(X_0) = \sum_{k=1}^{d} \int_0^t f_k(X_s) \circ dX_s^k, \tag{1.10}$$

where for a function $f \in \mathcal{L}^2_{\mathrm{loc}}(R^d)$ we define the Stratonovich integral as

$$\int_0^t f(X_s) \circ dX_s^k = \frac{1}{2}\left(\int_0^t f(X_s) dX_s^k + \int_0^t f(X_s) d^* X_s^k\right). \tag{1.11}$$

The idea of deriving an extended Itô formula in terms of quadratic covariations defined by (1.8) or in terms of Stratonovich integrals defined as in (1.10) has appeared independently in Russo and Vallois (1996) in a general semimartingale context, and in Lyons and Zhang (1994) in the context of Dirichlet spaces. Note that it makes sense to use both (1.8) and (1.10) as a *definition* of the quantities appearing on the left hand side whenever the processes X^k are semimartingales after time reversal. However, the explicit approximation of the stochastic integrals in (1.5) and (1.6) and the resulting identification of the quadratic covariations as limits in probability of the sums in (1.7) is another matter. Such an approximation is of course straightforward if f is continuous. Russo and Vallois (1996) consider a different approximation where they first smoothe the right hand side of (1.7) by taking integrals over time instead of the usual sums. In Lyons and Zhang (1994), the identification (1.7) of the quadratic covariations $[f(X), X^k]$ is shown under the regularity assumption that the function f belongs to the Dirichlet space, and convergence in probability is formulated with respect to a reversible reference measure.

In this paper, we concentrate on the classical case of Brownian motion. But here we insist on two improvements. First, the approximations (1.5), (1.6) and (1.7) are established with respect to a given starting point $x_0 \in R^d$ under explicit integrability conditions involving f and x_0. The second point is that we remove any smoothing and any regularity assumptions on the measurable function f. We require only the minimal integrability conditions which are needed in order to guarantee existence of the forward stochastic integral in (1.11). Thus, the existence of the quadratic covariations in (1.4) is established on exactly the same level of generality which is appropriate for defining the stochastic integrals in (1.1).

2. Existence of Quadratic Covariation

Let f be a measurable function on R^d where $d \geq 2$. Our purpose in this section is to establish the existence of the quadratic covariations $[f(X), X^k]$ under appropriate integrability hypotheses on f, but without assuming any regularity conditions. Consider the sums

$$\sum_{\substack{t_i \in D_n \\ 0<t_i<t}} \{f(X_{t_{i+1}}) - f(X_{t_i})\}(X^k_{t_i+1} - X^k_{t_i}) \tag{2.1}$$

along a sequence of partitions D_n of R^+. As in Föllmer, Protter and Shiryaev (1995), the idea is to decompose (2.1) and to show that the two sums

$$\sum_{\substack{t_i \in D_n \\ 0<t_i<t}} f(X_{t_i})(X^k_{t_{i+1}} - X^k_{t_i})$$

and

$$\sum_{\substack{t_i \in D_n \\ 0<t_i<t}} f(X_{t_i+1})(X^k_{t_i+1} - X^k_{t_i})$$

converge separately, to respectively a forward and a backward stochastic integral. To this end we assume that the sequence of partitions satisfies the following conditions:

$$\lim_{n\to\infty} \sup_{t_i \in D_n} (t_{i+1} - t_i) = 0, \quad M := \sup_n \sup_{t_i \in D_n} \frac{t_{i+1}}{t_i} < \infty; \tag{2.2}$$

note that the second condition is satisfied whenever the partitions are equidistant.

For a given point $x_0 \in R^d$ we define two norms for f:

$$||f||_1(x_0) = \int |f(y)| \; ||x_0 - y||^{1-d} dy \tag{2.3}$$

and

$$||f||_2(x_0) = (\int f(y)^2 v(||x_0 - y||) dy)^{\frac{1}{2}}, \tag{2.4}$$

where

$$v(r) = \begin{cases} (-\log r) \vee 1 & \text{if } d = 2 \\ r^{2-d} & \text{if } d \geq 3 \end{cases} \tag{2.5}$$

(2.6) **Remark.** Suppose that f has compact support. If f is also bounded then both norms $||f||_i(x_0)$ $(i = 1, 2)$ are clearly finite for every point $x_0 \in R^d$. This is still true if f is in $\mathcal{L}^p$ for some $p > d$; see remark (3.24). In section 3 we will see that, in view of a general result on the existence of quadratic covariation, it is natural to assume finiteness of both norms for all points $x_0 \notin E$, where E is an exceptional set which is not hit by Brownian motion.

(2.7) **Proposition.** *Let f be a measurable function on R^d with compact support, and let $x_0 \in R^d$ be such that $||f||_2(x_0) < \infty$. Then the forward stochastic integral satisfies*

$$\int_0^t f(X_s)dX_s^k = \lim_{n\to\infty} \sum_{\substack{t_i\in D_n \\ 0<t_i<t}} f(X_{t_i})(X_{t_{i+1}}^k - X_{t_i}^k) \quad \text{in } \mathcal{L}^2(P_{x_0}) \tag{2.8}$$

for each $k \in \{1, \dots, d\}$.

Proof. It suffices to consider only the case $t = 1$.
1) Define the processes ϕ and ϕ_n by

$$\phi(\omega, s) = f(X_s(\omega)), \tag{2.9}$$

$$\phi_n(\omega, s) = \sum_{t_i\in D_n} f(X_{t_i}(\omega))I_{(t_i,t_{i+1}]}(s). \tag{2.10}$$

The convergence in (2.8) is equivalent to

$$\lim_{n\to\infty} ||\phi - \phi_n||_2 = 0, \tag{2.11}$$

where we use the norm

$$||\psi||_2 = E_{x_0}[\int_0^1 \psi(\omega, s)^2 ds]^{\frac{1}{2}} \tag{2.12}$$

for any measurable function ψ on $\Omega \times [0, 1]$. Observe that if $f \in \mathcal{C}_b(R^d)$, then (2.11) holds by Lebesgue's dominated convergence theorem. The general case will follow by approximating f by continuous functions in the norm $|| \cdot ||_2(x_0)$.
2) Note that the Gaussian density

$$p_s(z) = (2\pi s)^{-\frac{d}{2}} \exp(-||z||^2/2s)$$

satisfies the inequality

$$\int_0^1 p_s(z)ds \le c(R)v(||z||) \tag{2.13}$$

for any $z \in R^d$ with $||z|| \leq R$, where $c(R)$ is some constant depending on R; see, e.g., Dynkin (1965, VIII, 8.16). Denoting by K the compact support of f and choosing $R \geq \sup_{y \in K} ||y - x_0||$, we obtain the estimate

$$\begin{aligned} ||\phi||_2^2 &= \int_0^1 \int f^2(y) p_s(y - x_0) dy ds \\ &\leq c(R) \int f^2(y) v(||y - x_0||) dy, \end{aligned} \tag{2.14}$$

hence

$$||\phi||_2 \leq a_2 ||f||_2(x_0), \tag{2.15}$$

where $a_2 = \sqrt{c(R)}$.

3) In order to obtain a similar estimate for the approximating process ϕ_n, note that

$$p_{t_i}(z) \leq (2\pi t_i)^{-\frac{d}{2}} \exp(-||z||^2/2s) \leq M^{\frac{d}{2}} p_s(z), \tag{2.16}$$

for $t_i \leq s \leq t_{i+1}$, due to our assumption (2.2). Again using (2.13) we get

$$\begin{aligned} ||\phi_n||_2^2 &= \int f^2(y) \sum_{\substack{t_i \in D_n \\ 0 < t_i < t}} p_{t_i}(y - x_0)((t_{i+1} \wedge 1) - t_i) dy \\ &\leq M^{\frac{d}{2}} \int f^2(y) \int_0^1 p_s(y - x_0) ds dy \\ &\leq M^{\frac{d}{2}} c(R) ||f||_2^2(x_0), \end{aligned} \tag{2.17}$$

hence

$$||\phi_n||_2 \leq b_2 ||f||_2(x_0) \tag{2.18}$$

where $b_2 = \sqrt{c(R)} M^{\frac{d}{4}}$.

4) Next we choose a continuous function g with compact support such that $||g - f||_2(x_0) \leq \varepsilon$, and denote by ψ and ψ_n the processes associated to g as in (2.9) and (2.10). We have

$$||\phi - \phi_n||_2 \leq ||\phi - \psi||_2 + ||\psi - \psi_n||_2 + ||\psi_n - \phi_n||_2. \tag{2.19}$$

But

$$\lim_n ||\psi - \psi_n||_2 = 0 \tag{2.20}$$

since g is bounded and continuous, and also

$$||\psi_n - \phi_n||_2 = ||(\psi - \phi)_n||_2 \leq b_2 ||f - g||_2(x_0), \tag{2.21}$$

due to our estimate (2.18) applied to the function $f-g$. This together with (2.15) implies

$$\begin{aligned}\limsup_{n\to\infty}||\phi-\phi_n||_2 &\le ||\phi-\psi||_2+\limsup_{n\to\infty}||\psi_n-\phi_n||_2\\ &\le a_2||f-g||_2(x_0)+b_2||f-g||_2(x_0)\\ &\le (a_2+b_2)\,\varepsilon.\end{aligned} \tag{2.22}$$

Since $\varepsilon>0$ was arbitrary we have shown (2.11) and hence (2.8).

(2.23) **Proposition.** *Let f be a measurable function on R^d with compact support, and let $x_0\in R^d$ be such that $||f||_i(x_0)<\infty$ for $i=1,2$. Then the backward stochastic integral satisfies*

$$\int_0^t f(X_s)d^*X_s^k=\lim_{n\to\infty}\sum_{\substack{t_i\in D_n\\ 0<t_i<t}} f(X_{t_{i+1}})(X^k_{t_{i+1}}-X^k_{t_i})\quad\text{in }\mathcal{L}^1(P_{x_0}) \tag{2.24}$$

for each $k\in\{1,\dots,d\}$.

Proof. It suffices to consider the case $t=1$.

1) Let $P^*_{x_0}$ be the distribution of the time reversed process $X\circ R$ under P_{x_0}, where $(RX)_t=X_{1-t}$. The time reversed process $X\circ R$ is a d-dimensional Brownian bridge tied down to $0\in R^d$ and starting with initial distribution $N(0,I)$, where I is the identity matrix. Under $P^*_{x_0}$, each component X^k is a semimartingale with decomposition

$$X_t^k=X_0^k+W_t^k+\int_0^t\frac{x_0^k-X_s^k}{1-s}ds, \tag{2.25}$$

where W^k $(k=1,\dots,d)$ are independent Wiener processes. The convergence in (2.24) for $t=1$ is equivalent to the convergence

$$\int_0^1 f(X_s)dX_s^k=\lim_{n\to\infty}\sum_{\substack{s_i\in D_n^*\\ 0<s_i<1}} f(X_{s_i})(X^k_{s_{i+1}}-X^k_{s_i})\quad\text{in }\mathcal{L}^1(P^*_{x_0}), \tag{2.26}$$

where $D_n^*=\{1-t_i|t_i\in D_n\}$. Let us now use the decomposition (2.25) of X^k under $P^*_{x_0}$, and let us first show that condition $||f||_2(x_0)<\infty$ implies

$$\lim_{n\to\infty}\sum_{\substack{s_i\in D_n^*\\ 0<s_i<1}} f(X_{s_i})(W^k_{s_{i+1}}-W^k_{s_i})=\int_0^1 f(X_s)dW_s^k\quad\text{in }\mathcal{L}^2(P^*_{x_0}). \tag{2.27}$$

This follows as in the proof of proposition (2.7). We have only to check that the estimates (2.15) and (2.18) have analogues in terms of the norm $||\psi||_2^*$ defined by

$$||\psi||_2^*=E^*_{x_0}[\int_0^1\psi(\omega,s)ds] \tag{2.28}$$

for any measurable function ψ on $\Omega \times [0,1]$. This is clear for (2.15) since

$$
\begin{aligned}
||\phi||_2^{*2} &= E^*_{x_0}[\int_0^1 f^2(X_s)ds] \\
&= E_{x_0}[\int_0^1 f^2(X_1 - s)ds] \qquad (2.29)\\
&= ||\phi||_2^2 \le a_2 ||f||_2^2(x_0).
\end{aligned}
$$

In order to obtain an analogue to (2.18), consider the term

$$
\begin{aligned}
||\phi_n||_2^{*2} &= E^*_{x_0}[\sum_{\substack{s_i \in D_n^* \\ 0 < s_i < 1}} f^2(X_{s_i})(s_{i+1} - s_i)] \\
&= E_{x_0}[\sum_{\substack{t_i \in D_n \\ 0 < t_i < 1}} f^2(X_{t_{i+1}})(t_{i+1} - t_i)] \qquad (2.30)\\
&= \int f^2(y) \sum_{\substack{t_i \in D_n \\ 0 < t_i < 1}} p_{t_{i+1}}(y - x_0)(t_{i+1} - t_i)dy.
\end{aligned}
$$

For $t_i \le s \le t_{i+1}$ and for any $z \in R^d$ we have

$$
\begin{aligned}
p_{t_{i+1}}(z) &\le (2\pi s)^{-\frac{d}{2}} \exp(-\frac{||z||^2}{2t_{i+1}}) \qquad (2.31)\\
&\le (2\pi s)^{-\frac{d}{2}} \exp(-\frac{||M^{-\frac{1}{2}} z||^2}{2s}) = p_s(M^{-\frac{1}{2}} z)
\end{aligned}
$$

due to (2.2). Using again the estimate (2.13), and observing that $v(M^{-\frac{1}{2}} r) \le \alpha v(r)$ for some constant α which only depends on M and d, we get

$$
\begin{aligned}
\sum_{\substack{t_i \in D_n \\ 0 < t_i < 1}} p_{t_{i+1}}(z)(t_{i+1} - t_i) &\le \int_0^1 p_s(M^{-\frac{1}{2}} z)ds \le c(R)v(||M^{-\frac{1}{2}} z||) \qquad (2.32)\\
&\le \alpha c(R) v(||z||).
\end{aligned}
$$

Returning to (2.30) we see that

$$
||\phi_n||_2^* \le b_2^* ||f||_2(x_0) \qquad (2.33)
$$

for some constant b_2^*. Using the estimates (2.29) and (2.33), we can now conclude as in part 3) of the proof of proposition (2.7) that (2.27) holds.

2) It remains to show

$$
\lim_{n \to \infty} \sum_{\substack{s_i \in D_n^* \\ 0 < s_i < 1}} f(X_{s_i}) \int_{s_i}^{s_{i+1}} \frac{X_s^k - x_0^k}{1 - s} ds = \int_0^1 f(X_s) \frac{X_s^k - x_0^k}{1 - s} ds \quad \text{in } \mathcal{L}^1(P^*_{x_0}) \qquad (2.34)
$$

or, equivalently,

$$\lim_{n\to\infty} \sum_{\substack{t_i\in D_n\\ 0<t_i<1}} f(X_{t_{i+1}}) \int_{t_i}^{t_{i+1}} \frac{X_s^k - x_0^k}{1-s} ds = \int_0^1 f(X_s) \frac{X_s^k - x_0^k}{1-s} ds \quad \text{in } \mathcal{L}^1(P_{x_0}). \tag{2.35}$$

Let us define the norm

$$||\psi||_1 = E_{x_0}[\int_0^1 |\psi(\omega,s)| \frac{|X_s^k - x_0^k|}{s} ds] \tag{2.36}$$

for any measurable function ψ on $\Omega \times [0,1]$. For the process ϕ defined in (2.9) we have

$$\begin{aligned} ||\phi||_1 &= \mathrm{E}_{x_0}[\int_0^1 |f(X_s)| \frac{|X_s^k - x_0^k|}{s} ds] \\ &= \int |f(y)||y^k - x_0^k| \int_0^1 \frac{p_s(y-x_0)}{s} ds dy. \end{aligned} \tag{2.37}$$

However for the Gaussian density p_s there is a constant a_1 such that

$$||z|| \int_0^1 \frac{p_s(z)}{s} ds \leq a_1 ||z||^{1-d} \tag{2.38}$$

for any $z \in R^d$; see, e.g., Dynkin (1980, VIII, 8.45). Combining (2.37) and (2.38) yields

$$||\phi||_1 \leq a_1 ||f||_1(x_0). \tag{2.39}$$

3) We also need an estimate of the form (2.39) for the approximating processes

$$\phi_n^*(\omega,s) = \sum_{t_i\in D_n} f(X_{t_{i+1}}(\omega)) I_{(t_i,t_{i+1}]}(s). \tag{2.40}$$

We will write $\sum_i$ for $\sum_{t_i\in D_n, 0<t_i<1}$. Then

$$\begin{aligned} ||\phi_n^*||_1 &= \sum_i \mathrm{E}_{x_0}[|f(X_{t_{i+1}})| \int_{t_i}^{t_{i+1}} \frac{|X_s^k - x_0|}{s} ds] \\ &= \sum_i \mathrm{E}_{x_0}[\int_{t_i}^{t_{i+1}} \frac{|X_s^k - x_0^k|}{s} ds |f(X_{t_{i+1}})|] \\ &= \sum_i \int_{t_i}^{t_{i+1}} \frac{1}{s} \mathrm{E}_{x_0}[\mathrm{E}_{x_0}[|X_s^k - x_0^k||X_{t_{i+1}}]|f(X_{t_{i+1}})|] ds. \end{aligned} \tag{2.41}$$

Recall that a normal random variable Z with law $N(m,\sigma^2)$ satisfies

$$\mathrm{E}[|Z|] = m(2\Phi(\frac{m}{\sigma}) - 1) + \sqrt{\frac{2}{\pi}} \sigma \exp(\frac{-m^2}{2\sigma^2}) \leq |m| + \sqrt{\frac{2}{\pi}} \sigma \tag{2.42}$$

where Φ denotes the distribution function of $N(0,1)$. In the conditional expectation appearing in (2.41), the normal random variable $Z = X_s^k - x_0^k$ has conditional mean

$$m = \frac{s}{t_{i+1}}(X_{t_{i+1}}^k - x_0^k) \tag{2.43}$$

and conditional variance

$$\sigma^2 = \frac{s}{t_{i+1}}(t_{i+1} - s) \le (M-1)s, \tag{2.44}$$

where we use our assumption (2.2). Thus, (2.42) implies

$$E_{x_0}[|X_s^k - x_0^k||X_{t_{i+1}}] \le A_i(s) + B_i(s) \tag{2.45}$$

where we put

$$A_i(s) = \frac{s}{t_{i+1}}||X_{t_{i+1}} - x_0||, \quad B_i(s) = \sqrt{\frac{2Ms}{\pi}} \tag{2.46}$$

for $t_i \le s < t_{i+1}$. Returning to (2.41) we obtain the estimate

$$||\phi_n^*||_1 \le \mathrm{E}_{\mathrm{x}_0}[\sum_i \int_{t_i}^{t_i+1} \frac{|f(X_{t_{i+1}})|}{s}(A_i(s) + B_i(s))ds]. \tag{2.47}$$

4) Let us first consider the term

$$\mathrm{E}_{x_0}[\sum_i \int_{t_i}^{t_{i+1}} \frac{|f(X_{t_{i+1}})|}{s} A_s^i ds] = \int |f(y)| \; ||y - x_0|| \sum_i \frac{t_{i+1} - t_i}{t_{i+1}} p_{t_{i+1}}(y - x_0)dy. \tag{2.48}$$

Due to (2.31) and (2.38) we obtain the estimate

$$\begin{aligned} ||z|| \sum_i \frac{s_{i+1} - s_i}{s_{i+1}} p_{s_{i+1}}(z) &\le ||z|| \int_0^1 \frac{1}{s} p_s(M^{-\frac{1}{2}} z) ds \\ &\le cM^{\frac{d}{2}} ||z||^{1-d}. \end{aligned} \tag{2.49}$$

Returning to (2.48) we see that

$$\begin{aligned} \mathrm{E}_{\mathrm{x}_0}[\sum_i \int_{t_i}^{t_{i+1}} \frac{1}{s} |f(X_{t_{i+1}})| A_s^i ds] &\le k_1 \int |f(y)|||y - x_0||^{1-d} dy \\ &= k_1 ||f||_1(x_0) \end{aligned} \tag{2.50}$$

where $k_1 = c_1 M^{\frac{d}{2}}$.

5) As to the second term on the right hand side of (2.47), we again use (2.31) to obtain the estimate

(2.51)
$$\begin{aligned} \mathrm{E}_{x_0}[\sum_i \int_{t_i}^{t_{i+1}} \frac{|f(X_{t_{i+1}})|}{s} B_i(s) ds] &= \sqrt{\frac{2M}{\pi}} \int |f(y)| \sum_i \int_{t_i}^{t_{i+1}} \frac{1}{\sqrt{s}} p_{t_{i+1}}(y-x_0) ds dy \\ &\leq \sqrt{\frac{2M}{\pi}} \int |f(y)| \int_0^1 \frac{1}{\sqrt{s}} p_s(M^{-\frac{1}{2}}(y-x_0) ds dy. \end{aligned}$$

But for any $z \in R^d$ we have

(2.52)
$$\int_0^1 \frac{1}{\sqrt{s}} p_s(z) ds \leq c_2 ||z||^{1-d}$$

for some constant c_2, and so (2.51) implies

(2.53)
$$\begin{aligned} \mathrm{E}_{x_0}[\sum_i \int_{t_i}^{t_{i+1}} \frac{|f(X_{t_{i+1}})|}{s} B_i(s) ds] &\leq l_1 \int |f(y)| \; ||y-x_0||^{1-d} dy \\ &= l_1 ||f||_1(x_0) \end{aligned}$$

where $l_1 = \sqrt{\frac{2}{\pi}} c_2 M^{\frac{d}{2}}$.

6) Combining (2.47) with (2.50) and (2.53) we see that

(2.54)
$$||\phi_n^*||_1 \leq b_1 ||f||_1(x_0),$$

where $b_1 = k_1 + l_1$. Using the estimates (2.54) and (2.39), we can now conclude as in part 3) of the proof of proposition (2.7) that $||\phi - \phi_n^*||_1$ tends to 0. This implies the convergence in (2.35) and (2.34), and so the proposition is proved.

We combine propositions (2.7) and (2.23) to obtain:

(2.55) **Corollary.** *Let f be a measurable function on R^d, let $x_0 \in R^d$, and let K_m $(m \geq 1)$ be a sequence of compact sets such that*

(2.56)
$$\lim_{m\to\infty} P_{x_0}[X_t \in K_m \quad \forall\, t \in [0,T]] = 1 \quad \forall T > 0.$$

Assume that for any $m \geq 1$ the restriction f_m of f to K_m satisfies

(2.57)
$$||f_m||_i(x_0) < \infty \quad (i = 1, 2).$$

Then the quadratic covariation

(2.58)
$$\begin{aligned} [f(X), X^k]_t &\equiv \lim_{n\to\infty} \sum_{\substack{t_i \in D_n \\ 0<t_i<t}} \{f(X_{t_{i+1}}) - f(X_{t_i})\}(X_{t_{i+1}}^k - X_{t_i}^k) \\ &= \int_0^t f(X_s) d^* X_s^k - \int_0^t f(X_s) dX_s^k \end{aligned}$$

exists in probability under P_{x_0} and satisfies

$$[f(X), X^k]_t = \int_0^t f(X_s) d^* X_s^k - \int_0^t f(X_s) dX_s^k, \tag{2.59}$$

for each $k \in \{1, \ldots, d\}$.

Proof. Let t be fixed and $\varepsilon > 0$, and let $T_m = \inf\{t > 0 \,|X_t \notin K_m\}$ be the exit time from K_m. We denote by S_n the n-th sum in (2.58), by S the difference of the forward and backward stochastic integrals, and by S_n^m and S^m the corresponding terms if the function f is replaced by f_m. Since $S_n = S_n^m$ and $S = S^m$ P_{x_0}-a.s. on $\{T_m > t\}$, we have

$$P_{x_0}[|S_n - S| \geq \varepsilon] \leq P_{x_0}[T_m \leq t] + P_{x_0}[|S_n^m - S^m| \geq \varepsilon] \tag{2.60}$$

for any $m \geq 1$. Applying propositions (2.7) and (2.23) to the function f_m we see that the last term converges to 0 as n tends to ∞. Thus,

$$\limsup_{n \to \infty} P_{x_0}[|S_n - S| \geq \varepsilon] \leq P_{x_0}[T_m \leq t], \tag{2.61}$$

and due to (2.56) the result follows by letting m tend to ∞.

3. Exceptional sets

Let f be a measurable function on R^d. In our approximation (2.55) of the quadratic covariation

$$[f(X), X^k]_t = \int_0^t f(X_s) d^* X_s^k - \int_0^t f(X_s) dX_s^k. \tag{3.1}$$

as a limit in probability under P_{x_0}, we have assumed integrability conditions on f which are formulated in terms of the initial point x_0. In this section we show that it is no loss of generality to make these assumptions for all x_0 outside some exceptional set which is not hit by Brownian motion.

(3.2) **Definition.** A measurable set $E \subseteq R^d$ is called *polar* if

$$P_x[X_t \in E \quad \text{for some } t > 0] = 0 \quad \forall x \in R^d. \tag{3.3}$$

(3.4) **Remark.** This probabilistic notion of an exceptional set is equivalent to the potential theoretic notion of a set of *capacity zero*; see, e.g., Fukushima (1980, Th. 4.3.1 and Example 4.3.1). Equivalently, we can define these exceptional sets in terms of the Bessel capacity of order (1,2) as in Ziemer (1989, 2.6); see, e.g., Fukushima (1993, p.25).

Note first that in order to introduce the forward stochastic integral in (3.1) with respect to the measure P_{x_0}, at least for some $x_0 \in R^d$ and some $t > 0$, we clearly need the condition

$$(3.5) \qquad P_{x_0}[\int_0^t f^2(X_s)ds < \infty] = 1.$$

But the results in Höhnle and Sturm (1993, Th.1.1) show that the validity of (3.2) for some $x_0 \in R^d$ and some $t_0 > 0$ implies that the function f satisfies condition (3.5) for all $t > 0$ and for all initial points $x \notin E$ where E is some polar set. Moreover, it follows that all the assumptions we used in theorem (2.55) hold for any starting point lying outside some polar set:

(3.6) **Proposition.** *Let f be a measurable function on R^d such that condition (3.5) holds for some $x_0 \in R^d$ and some $t < \infty$. Then there exist a sequence of compact sets $K_m \subseteq R^d$ $(m \geq 1)$ and a polar set E such that the conditions*

$$(3.7) \qquad \lim_{m\to\infty} P_{x_0}[X_t \in K_m \quad \forall\, t \in [0,T]] = 1 \quad \forall T > 0$$

and

$$(3.8) \qquad \int f_m^2(x)dx < \infty, \quad ||f_m||_i(x_0) < \infty \qquad (i = 1, 2)$$

are satisfied for all $x_0 \notin E$ and for all $m \geq 1$, where f_m denotes the restriction $f \cdot I_{K_m}$ of f to the set K_m.

Proof. 1) It is shown in Höhnle and Sturm (1993, Th. 3.5 and 3.7) that the validity of (3.2) for some $x_0 \in R^d$ and some $t < \infty$ is equivalent to the fact that there exists a polar set E_2 and a sequence of compact sets K_m $(m \geq 1)$ with

$$(3.9) \qquad \int_{K_m} f^2(x)dx < \infty \quad (m \geq 1)$$

such that the conditions (3.7) and

$$(3.10) \qquad \int_{K_m} v(\| x_0 - y \|) f^2(y)dy < \infty \quad (m \geq 1)$$

are satisfied for any $x_0 \notin E_2$. Thus, the restrictions f_m of f to K_m satisfy

$$(3.11) \qquad f_m \in \mathcal{L}^2(R^d), \quad ||f_m||_2(x_0) < \infty$$

for all $x_0 \notin E_2$.

2) It remains to verify the integrability condition

$$||f_m||_1(x_0) < \infty \tag{3.12}$$

for all x_0 outside some polar set. In view of 1) we may assume that f has compact support K and is square integrable. Consider the Bessel potential $g_1 * |f|$ of order 1 defined by

$$(g_1 * |f|)(x) = \int g_1(x-y)|f|(y)dy, \tag{3.13}$$

where g_α is defined as that function whose Fourier transform is

$$\hat{g}_\alpha(x) = (2\pi)^{-\frac{\alpha}{2}}(1 + ||x||^2)^{-\frac{\alpha}{2}}. \tag{3.14}$$

Note that

$$g_\alpha(z) = \frac{1}{\gamma(\alpha)}||z||^{\alpha-d} + o(||z||^{\alpha-d}) \tag{3.15}$$

with some constant $\gamma(\alpha)$ as $||z|| \to 0$; see, e.g., Ziemer (1989, p. 65). Thus, there is a constant c such that

$$||y - x_0||^{1-d} \le c \cdot g_1(x_0 - y) \quad \text{on } K \tag{3.16}$$

and so we have

$$\begin{aligned} ||f||_1(x_0) &= \int_K |f|(y)||y-x_0||^{1-d}dy \\ &\le c \cdot (g_1 * |f|)(x_0). \end{aligned} \tag{3.17}$$

But the function $u = g_1 * f$, being the Bessel potential with index $\alpha = 1$ associated to the square integrable function f, belongs to the Sobolev space $\mathcal{W}^{1,2}$; see Ziemer (1989, Th. 2.6.1). This implies that the version $\tilde{u}$ defined by

$$\tilde{u}(x) := \lim_{\delta \downarrow 0} \frac{1}{\text{vol}(B_\delta(x))} \int_{B_\delta(x)} u(z)dz \tag{3.18}$$

satisfies $\tilde{u}(x) < \infty$ outside some polar set E_1; see, e.g., Ziemer (1989, Th.3.1.4 and 3.10.2) or Fukushima (1993, Th. 2.1). Since

$$\lim_{\delta \downarrow 0} \frac{1}{\text{vol}(B_\delta(x))} \int_{B_\delta(x)} g_1(y-z)dz = g_1(y-x) \tag{3.19}$$

for any $y \neq x$ by Lebesgue's theorem, we obtain

$$\begin{aligned} u(x) &= \int |f|(y) \lim_{\delta \downarrow 0} \frac{1}{\text{vol}(B_\delta(x))} \int_{B_\delta(x)} g_1(y-x)dzdy \\ &\le \liminf_{\delta \downarrow 0} \frac{1}{\text{vol}(B_\delta(x))} \int_{B_\delta(x)} u(z)dz, \end{aligned} \tag{3.20}$$

using Fatou's lemma and Fubini's theorem. For $x_0 \notin E_1$, we have thus shown $u(x_0) < \infty$, hence $||f||_1(x_0) < \infty$ due to (3.17).

Combining proposition (3.6) with corollary (2.55) we obtain:

(3.21) **Theorem.** *Let f be a measurable function on R^d such that condition (3.5) holds for some $x_0 \in R^d$ and some $t < \infty$. Then there exists a polar set E such that for any $x_0 \notin E$ the quadratic covariation*

$$\begin{aligned}[f(X), X^k]_t &\equiv \lim_{n\to\infty} \sum_{\substack{t_i \in D_n \\ 0<t_i<t}} \{f(X_{t_{i+1}}) - f(X_{t_i})\}(X^k_{t_{i+1}} - X^k_{t_i}) \\ &= \int_0^t f(X_s)d^*X^k_s - \int_0^t f(X_s)dX^k_s \end{aligned} \tag{3.22}$$

exists in probability under P_{x_0} and satisfies

$$[f(X), X^k]_t = \int_0^t f(X_s)d^*X^k_s - \int_0^t f(X_s)dX^k_s, \tag{3.23}$$

for each $k \in \{1, \ldots, d\}$.

(3.24) **Remark.** If $f \in \mathcal{L}^p_{loc}(R^d)$ for some $p > d$ then the conlusion of the theorem holds for *every* starting point $x_0 \in R^d$, without exception. To see this we may assume that f has compact support. In this case, the assumption that $f \in \mathcal{L}^p(R^d)$ for some $p > d$ implies

$$||f||_i(x_0) < \infty \qquad (i = 1, 2) \tag{3.25}$$

for *every* $x_0 \in R^d$. Thus, our assumptions in corollary (2.55) for the existence of quadratic covariation are satisfied *everywhere.* Conditions (3.25) can be verified directly, using Hölder's inequality. They also follow from the Sobolev embedding theorem. Note that the function u_2 defined by

$$u_2(x) = \int f^2(y)v(\| x - y \|)dy \tag{3.26}$$

belongs to $W^{2,p/2}$; see Gilbarg and Trudinger (1983, Th. 9.9). This implies $u_2 \in C(R^d)$ for $p > d$, hence $||f||_2(x) < \infty$ for any $x \in R^d$; see Ziemer (1989, Th.2.4.2). Note also that for $f \in \mathcal{L}^p$ the Bessel potential $u = g_1 * |f|$ belongs to $W^{1,p}$; see Ziemer (1989, Th. 2.6.1). This implies $u \in C(R^d)$ for $p > d$, again by Ziemer (1989, Th. 2.4.2), hence $||f||_1(x) < \infty$ for any $x \in R^d$, due to (3.17).

4. Itô's formula

Let $\mathcal{W}^{1,2}$ denote the Sobolev space of functions in $\mathcal{L}^2(R^d)$ such that the weak first partial derivatives belong to $\mathcal{L}^2(R^d)$. Recall that a function in $\mathcal{W}^{1,2}$ can be defined everywhere, except for a polar set, in terms of its integral averages, i.e., we can choose a version F and a polar set E_0 such that

$$F(x) = \lim_{\delta \downarrow 0} \frac{1}{\mathrm{vol}(B_\delta(x))} \int_{B_\delta(x)} F(y)dy \tag{4.1}$$

for all $x \in E_0$; see, e.g., Fukushima (1993, Th. 2.1 and p.25) or Ziemer (1989, Th. 3.1.4).

Let us now consider a function in $\mathcal{W}^{1,2}_{\mathrm{loc}}$, i.e., a measurable function on R^d which coincides on each compact set with a function in $\mathcal{W}^{1,2}$. We fix a version F such that (4.1) holds outside some polar set E_0, and we denote by

$$f_k = \frac{\partial F}{\partial x_k} \in \mathcal{L}^2_{loc}(R^d) \tag{4.2}$$

the k-th weak partial derivative of F. With probability 1, Brownian motion does not enter a given polar set after time 0, and so the values $F(X_t(\omega))$ of the function F along a Brownian path are well defined P_x-almost surely for any starting point $x \notin E_0$.

(4.3) **Theorem.** *Let $F \in \mathcal{W}^{1,2}_{\mathrm{loc}}$ be given as above. For all $x_0 \in R^d$ except for some polar set, the quadratic covariation*

$$[f_k(X), X^k]_t = \lim_n \sum_{\substack{t_i \in D_n \\ t_i \le t}} (f_k(X_{t_{i+1}}) - f_k(X_{t_i}))(X^k_{t_{i+1}} - X^k_{t_i}) \tag{4.4}$$

exists as a limit in probability under P_{x_0} for each $k \in \{1, \ldots, d\}$, and Itô's formula holds in the form

$$F(X_t) = F(X_0) + \sum_{k=1}^d \int_0^t f_k(X_s)dX^k_s + \frac{1}{2}\sum_{k=1}^d [f_k(X), X^k]_t \qquad P_{x_0} - \text{a.s.} \tag{4.5}$$

for all $t \geq 0$.

Proof. 1) By a localization argument as in the proof of corollary (2.55), we can assume that F has compact support and belongs to $W^{1,2}$. Since $f_k \in \mathcal{L}^2(R^d)$, we have

$$P_{x_0}[\int_0^t f_k^2(X_s)ds < \infty] = 1 \tag{4.6}$$

for all x_0 except for some polar set; see Fukushima (1980, (5.4.23)). Due to (3.21) we can conclude that the quadratic covariations $[f_k(X), X^k]_t$ are well defined as limits in probability under P_{x_0} for all x_0, except for some polar set. Alternatively we can apply propositions (2.7) and (2.23), using the estimates $||f_k||_i(x_0) < \infty \quad (i = 1, 2)$ for $x_0 \notin E_1 \cup E_2$ which are implied by (4.20) and (4.25) below.

2) Let $x_0 \in R^d$. Suppose that we can approximate F by functions $F^{(n)} \in C^2(R^d)$ with compact support in such a way that

$$F(x) = \lim_{n\to\infty} F^{(n)}(x) \tag{4.7}$$

for all x outside some polar set, and that the partial derivatives $f_k^{(n)} = \frac{\partial}{\partial x_k} F^{(n)}$ satisfy

$$\lim_{n\to\infty} ||f_k^{(n)} - f_k||_i(x_0) = 0 \qquad (i = 1, 2) \tag{4.8}$$

for the two norms introduced in section 2. Due to our estimates (2.15), (2.29), (2.39) we can conclude, as in the proof of (2.7) and (2.23), that

$$\lim_{n\to\infty} \int_0^t f_k^{(n)}(X_s)dX_s^k = \int_0^t f_k(X_s)dX_s^k \quad \text{in } \mathcal{L}^2(P_{x_0}) \tag{4.9}$$

and

$$\lim_{n\to\infty} \int_0^t f_k^{(n)}(X_s)d^*X_s^k = \int_0^t f_k(X_s)d^*X_s^k \quad \text{in } \mathcal{L}^1(P_{x_0}). \tag{4.10}$$

In particular,

$$\begin{aligned} [f_k(X), X^k]_t &= \int_0^t f_k(X_s)d^*X_s^k - \int_0^t f_k(X_s)dX_s^k \\ &= \lim_{n\to\infty} [f_k^{(n)}(X), X^k]_t \qquad \text{in } \mathcal{L}^1(P_{x_0}). \end{aligned} \tag{4.11}$$

Applying Itô's formula to the functions $F^{(n)}$, we obtain

$$\begin{aligned} \frac{1}{2}\sum_{k=1}^d [f_k(X), X^k]_t &= \lim_{n\to\infty} \frac{1}{2}\sum_{k=1}^d [f_k^{(n)}(X), X^k]_t \\ &= \lim_{n\to\infty} (F^{(n)}(X_t) - F^{(n)}(X_0) - \sum_{k=1}^d \int_0^t f_k^{(n)}(X_s)dX_s^k) \\ &= F(X_t) - F(X_0) - \sum_{k=1}^d \int_0^t f_k(X_s)dX_s^k \quad P_{x_0} - \text{a.s.}, \end{aligned} \tag{4.12}$$

where we have applied once more (4.9) in the last step.

3) In order to construct such an approximation, let us take functions $F^{(n)} \in C^2(R^d)$ with compact support such that

$$\sum_{n=1}^{\infty} ||F^{(n)} - F||_{1,2} < \infty \tag{4.13}$$

where $||\cdot||_{1,2}$ denotes the Sobolev norm in $\mathcal{W}^{1,2}$. In particular, the functions

$$h_k := \sum_{n=1}^{\infty} |f_k^{(n)} - f_k| \tag{4.14}$$

belong to $\mathcal{L}^2(R^d)$ since

$$\begin{aligned}(\int h_k^2(x)dx)^{\frac{1}{2}} &\leq \sum_{n=1}^{\infty}(\int |f_k^{(n)} - f_k|^2 dx)^{\frac{1}{2}} \\ &\leq \sum_{n=1}^{\infty} ||F^{(n)} - F||_{1,2} < \infty.\end{aligned} \tag{4.15}$$

It follows as in part 2) of the proof of proposition (3.6) that the Bessel potential

$$(g_1 * h_k)(x) = \int g_1(x-y)h_k(y)dy, \tag{4.16}$$

is finite for all x_0 outside some polar set E_1. But due to (3.17) we have

$$||h_k||_1(x_0) \leq c \cdot \ (g_1 * h_k)(x_0), \tag{4.18}$$

for some constant c, and so we have shown that $||h_k||_1(x_0) < \infty$ for all $x_0 \notin E_1$. This implies

$$\sum_{n=1}^{\infty} ||f_k^{(n)} - f_k||_2(x_0) < \infty, \tag{4.19}$$

by monotone integration, and so we get the desired approximation

$$\lim_{n\to\infty} ||f_k^{(n)} - f_k||_1(x_0) = 0 \tag{4.20}$$

for all $x_0 \notin E_1$.

4) Let us define

$$\tilde{h}_k := (\sum_{n=1}^{\infty}(f_k^{(n)} - f_k)^2)^{\frac{1}{2}}. \tag{4.21}$$

Since

$$\int \tilde{h}_k^2 dx = \sum_{n=1}^{\infty} \int (f_k^{(n)} - f_2)^2 dx < \infty, \tag{4.22}$$

part 1) of the proof of (3.6) shows that

$$\int_{K_m} v(||x_0 - y||)\tilde{h}_k^2(y)dy < \infty \tag{4.23}$$

for all $m \geq 1$ and for all x_0 except for some polar set E_2, where (K_m) is a sequence of compact sets satisfying (3.7). In view of the localization argument in the proof of (2.55), we can assume without loss of generality that F vanishes outside K_{m_0} for some $m_0 \geq 1$. Then we get

$$\sum_{n=1}^{\infty} \int v(||x_0 - y||)(f_k^{(n)} - f_k)^2(y)dy < \infty \tag{4.24}$$

for all $x_0 \notin E_2$, and this implies the desired approximation

$$\lim_{n\to\infty} ||f_k^{(n)} - f_k||_2(x_0) = 0 \tag{4.25}$$

for all $x_0 \notin E_2$.

5) Let us also check that (4.7) holds for all points x_0 except for some polar set. We have

$$(F - F^{(n)})(x_0) = \lim_{\delta\downarrow 0} \frac{1}{\mathrm{vol}(B_\delta(x_0))} \int_{B_\delta(x_0)} (F - F^{(n)})(x)dx \tag{4.26}$$

for all points x_0 except for the polar set E_0 involved in the choice of the version F. Since

$$\sum_{n=1}^{\infty} |F - F^{(n)}| \in \mathcal{W}^{1,2} \tag{4.27}$$

due to (4.13), we get the existence of

$$\lim_{\delta\downarrow 0} \frac{1}{\mathrm{vol}(B_\delta(x_0))} \int_{B_\delta(x_0)} \sum_{n=1}^{\infty} |F - F^{(n)}|(x)dx < \infty \tag{4.28}$$

for all x_0 except for a polar set E_3. For $x_0 \notin E_0 \cup E_3$, we can conclude that

$$\begin{aligned} \sum_{n=1}^{\infty} |F - F^{(n)}|(x_0) &\leq \sum_{n=1}^{\infty} \liminf_{\delta\downarrow 0} \frac{1}{\mathrm{vol}(B_\delta(x_0))} \int_{B_\delta(x_0)} |F - F^{(n)}|(x)dx \\ &\leq \liminf_{\delta\downarrow 0} \frac{1}{\mathrm{vol}(B_\delta(x_0))} \int_{B_\delta(x_0)} \sum_{n=1}^{\infty} |F - F^{(n)}|(x)dx < \infty, \end{aligned} \tag{4.29}$$

and this implies (4.7). Thus, all properties of the approximation which were used in part 2) of the proof are satisfied for any $x_0 \notin E_0 \cup E_1 \cup E_2 \cup E_3$.

References

1] Brosamler, G. (1970): Quadratic variation of potentials and harmonic functions. *Trans. Am. Math. Soc.* 149, p.243 - 257

[2] Dynkin, E.B. (1980): *Markov Processes* vol. I . Springer-Verlag, Berlin-Göttingen-Heidelberg

[3] Föllmer, H., Protter, Ph. and Shiryaev, A.N. (1995): Quadratic covariation and an extension of Itô 's formula. *Bernoulli* 1 (1/2), 149-169

[4] Fukushima, M. (1980): *Dirichlet Forms and Markov Processes.* Amsterdam: North-Holland

[5 Fukushima, M. (1993): Two topics related to Dirichlet forms: quasi everywhere convergences and additive functionals. In: *Dirichlet Forms.* Springer Lecture Notes in Mathematics 1563, 21 - 53

[6] Gilbarg, D. and Trudinger, N.S. (1983): *Elliptic Partial Differential Equations of Second Order.* Springer-Verlag, New York-Berlin-Heidelberg

[7] Höhnle, R. and Sturm, K.-Th. (1993): A Multidimensional Analogue to the 0-1-Law of Engelbert and Schmidt. *Stochastics and Stochastics Reports* 44, 27-41

[8] Itô, K. (1944): Stochastic integral. *Proc. Imperial Acad. Tokyo*, 20, 519-524

[9] Lyons, T.J. and Zhang,T.S. (1994): Decomposition of Dirichlet processes and its application. *Ann. Probability* 22, 494-524

[10] Meyer, P.A. (1978): La formule d'Itô pour le mouvement Brownian d'après G. Brosamler. In: *Séminaire de Probabilités XII.* Springer Lecture Notes in Mathematics 649

[11] Russo, F. and Vallois, P. (1996): Itô formula for C^1-functions of semimartingales.*Prob. Theory Relat. Fields* 104, 27-41

[12] Ziemer, W.P. (1989): *Weakly Differentiable Functions.* Graduate Texts in Mathematics 120, Springer-Verlag, New York Berlin Heidelberg