The Annals of Probability
1989, Vol. 17, No. 1, 208–238

INTEGRATION BY PARTS AND TIME REVERSAL FOR DIFFUSION PROCESSES

BY A. MILLET,[1] D. NUALART AND M. SANZ

Université d'Angers and Universitat de Barcelona

In this paper we obtain necessary and sufficient conditions for the reversibility of the diffusion property, assuming the existence of a density at every time t. The proofs are based on techniques of the stochastic calculus of variations.

Introduction. Suppose that $\{X_t, 0 \le t \le 1\}$ is a d-dimensional diffusion process solution of the stochastic differential equation

$$dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dW_t,$$

where $\{W_t, 0 \le t \le 1\}$ is a Brownian motion in $\mathbb{R}^l$. It is well known that the reversed process $\overline{X}_t = X_{1-t}$ is Markovian, and we may ask if the diffusion property is preserved too. This is not true in general (see [18]). Then an interesting problem is to find minimal conditions that guarantee the reversibility of the diffusion property and to compute the diffusion and drift coefficients of the reversed process.

Different methods have been used to solve this problem. The approach of Anderson [1] and Haussmann and Pardoux [6] is based on the study of the solution of forward and backward Kolmogorov equations. The technique of the enlargement of a filtration has been used by Pardoux [14]. Föllmer [4] presents an approach based on the notion of entropy which, when $\sigma = Id$, allows one to deal with the non-Markov case.

Following the ideas contained in the paper of Föllmer [4], we have used the integration by parts formula to attack the time reversal problem. This formula establishes the duality between the derivative operator D on the Wiener space and the Skorohod stochastic integral, and it is a fundamental tool in the applications of Malliavin calculus.

Assuming that X_t has a density for any $t > 0$, we obtain a necessary and sufficient condition for the reversibility of the diffusion property. The condition is as follows.

(C) The sums of the distributional derivatives $\sum_{j=1}^d \nabla_j(a^{ij}(t,x)p_t(x))$, $i = 1, \dots, d$ (with $a = \sigma\sigma^*$), are locally integrable functions.

This result includes the sufficient conditions given by Haussmann and Pardoux [6], and we want to point out that our work has been inspired by this paper, although our proofs are essentially different.

Received May 1987; revised December 1987.

[1]This work was carried out while the first author was visiting the "Centre de Recerca Matemàtica" at Barcelona.

AMS 1980 *subject classification*. 60J60.

Key words and phrases. Time reversal, diffusions, integration by parts formula.

The organization of the paper is as follows. In Section 1 we present some preliminary and well-known results about the derivation on the Wiener space that will be needed in the sequel. In Section 2 we prove the main result (necessary and sufficient conditions for the reversibility of the diffusion property), assuming that the coefficients $b(t, x)$ and $\sigma(t, x)$ are globally Lipschitz functions of x, uniformly in t. Section 3 is devoted to extending this result to the case of locally Lipschitz and bounded coefficients. The necessary part follows easily from an approximation argument. The proof of the sufficiency is more sophisticated than the corresponding one in the case of globally Lipschitz coefficients. An integration by parts formula for particular elements which are locally in $\mathbb{D}_{2,1}$ has to be proved. We also need a detailed analysis of the dependence of the diffusion process on the initial condition. In this section some additional assumptions are introduced: the conservative character of the diffusion and the existence of exponential moments for the gradients of the coefficients b and σ.

In Section 4 we present some sufficient conditions on the coefficients that ensure the absolute continuity of the law of X_t and condition (C).

Finally, we include in the Appendix some technical results on Sobolev spaces that have been used in the paper.

Some remarks on the notation: All constants will be denoted by the same letter, although they may vary from one expression to another one. We also use the usual convention on summation of repeated indexes.

1. Basic results about derivation on the Wiener space. In this section we recall the main properties of the derivative operator on the Wiener space. For a more detailed exposition of this topic we refer to Malliavin [10], Ikeda and Watanabe [8], Watanabe [17], Zakai [19] and Nualart and Pardoux [12].

Let $\{W_t,\ 0 \le t \le 1\}$ be an l-dimensional standard Wiener process defined on the canonical probability space $(\Omega, \mathscr{F}, P)$. That is, $\Omega = C([0,1], \mathbb{R}^l)$, P is the Wiener measure, $\mathscr{F}$ is the completion of the Borel σ-algebra of Ω with respect to P and $W_t(\omega) = \omega(t)$. We denote by H the Hilbert space $L^2([0,1], \mathbb{R}^l)$.

Let E be a real separable Hilbert space. An E-valued random variable F: $\Omega \to E$ will be called smooth if

$$F = \sum_{i=1}^{M} f_i(W_{t_1}, \dots, W_{t_n}) v_i,$$

where $f_i \in \mathscr{C}_b^\infty(\mathbb{R}^{ln})$, $t_1, \dots, t_n \in [0,1]$ and $v_1, \dots, v_M \in E$.

Here $\mathscr{C}_b^\infty(\mathbb{R}^{ln})$ denotes the set of $\mathscr{C}^\infty$ functions f: $\mathbb{R}^{ln} \to \mathbb{R}$ which are bounded, together with all their derivatives.

The derivative of a smooth random variable is the random variable taking values in the Hilbert space $H \otimes E = L^2([0,1], \mathbb{R}^l \otimes E)$, given by

$$D_t^j F = \sum_{i=1}^{M} \sum_{k=1}^{n} \frac{\partial f_i}{\partial x^{jk}}(W_{t_1}, \dots, W_{t_n}) 1_{[0, t_k]}(t) v_i,$$

for $t \in [0,1]$ and $j = 1, \dots, l$.

Notice that for any $h \in H$, $\sum_{i=1}^{l} \int_0^1 D_t^i F h^i(t)\,dt$ can be interpreted as the directional derivative

$$\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0}\left(F\left(\omega_\cdot + \varepsilon\int_0^\cdot h(s)\,ds\right)\right).$$

The Nth derivative of F will be the $H^{\otimes N} \otimes E$-valued random variable defined by

$$(D^{(N)}F)_{t_1,\dots,t_N}^{j_1,\dots,j_N} = D_{t_1}^{j_1} D_{t_2}^{j_2} \cdots D_{t_N}^{j_N} F.$$

For any integer $N \geq 1$ and any real number $p > 1$, $\mathbb{D}_{p,N}(E)$ will denote the Banach space which is the completion of the set of smooth random variables with respect to the norm

$$\|F\|_{L^p(\Omega, E)} + \sum_{M=1}^{N} \|D^{(M)}F\|_{L^p(\Omega, H^{\otimes M} \otimes E)}.$$

Set $\mathbb{D}_{\infty, N}(E) = \bigcap_{p>1} \mathbb{D}_{p,N}(E)$ and $\mathbb{D}_\infty(E) = \bigcap_{N \geq 1} \mathbb{D}_{\infty, N}(E)$. The set $\mathbb{D}_\infty(E)$ is the space of test functionals which plays a basic role in Malliavin calculus. If $E = \mathbb{R}$, we write $\mathbb{D}_{p,N}$ for $\mathbb{D}_{p,N}(E)$.

We will basically deal with the space $\mathbb{D}_{2,1}(E)$. Consider the orthogonal Wiener-chaos decomposition $L^2(\Omega, E) = \oplus_{n=0}^{\infty} \mathscr{H}_n$ and denote by J_n the projection on $\mathscr{H}_n$. The space $\mathbb{D}_{2,1}(E)$ coincides with the set of random variables $F \in L^2(\Omega, E)$ such that

$$E\left(\|DF\|_{H\otimes E}^2\right) = \sum_{n=1}^{\infty} nE\left(\|J_n F\|_E^2\right) < \infty.$$

We recall the following two basic properties of the derivative operator.

PROPOSITION 1.1 (Chain rule). *Let $g\colon \mathbb{R}^d \to \mathbb{R}$ be a continuously differentiable function with bounded partial derivatives. Suppose that $F = (F^1, \dots, F^d)$ is a random vector whose components belong to $\mathbb{D}_{p,1}$ for some $p \geq 2$. Then $g(F) \in \mathbb{D}_{p,1}$ and*

$$D(g(F)) = (\nabla_i g)(F) DF^i. \tag{1.1}$$

PROPOSITION 1.2. *Let $F \in L^2(\Omega, \mathscr{F}_A, P)$, where A is a Borel subset of $[0,1]$ and $\mathscr{F}_A$ is the σ-algebra generated by the random vectors $\{W(G),\ G \subset A\}$. Assume $F \in \mathbb{D}_{2,1}$. Then $D_t F = 0$ for all (ω, t) in $\Omega \times A^c$ a.e.*

We will also need the next result, which is a refinement of the chain rule. In the sequel $\alpha_n(x)$ will represent a sequence of regularization kernels of the form

$$\alpha_n(x) = n^d \alpha(nx), \tag{1.2}$$

where $\alpha \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ is a nonnegative function whose compact support contains 0 and such that $\int_{\mathbb{R}^d} \alpha(x)\,dx = 1$.

PROPOSITION 1.3. *Let $g: \mathbb{R}^d \to \mathbb{R}$ be a globally Lipschitz function. Suppose that $F = (F^1, \dots, F^d)$ is an absolutely continuous random vector whose components are in $\mathbb{D}_{p,1}$ for some $p \geq 2$. Then $g(F) \in \mathbb{D}_{p,1}$ and* (1.1) *holds.*

PROOF. Set $g_n = g * \alpha_n$. Then we know that $\lim_n g_n(x) = g(x)$ and $\lim_n(\nabla g_n)(x) = (\nabla g)(x)$ a.e., by choosing a suitable subsequence denoted again by g_n. Then the Lebesgue dominated convergence implies that

$$g_n(F) \xrightarrow{L^p(\Omega)} g(F)$$

and

$$(\nabla_i g_n)(F) DF^i \xrightarrow{L^p(\Omega \times [0,1], \mathbb{R}^l)} (\nabla_i g)(F) DF^i$$

as $n \to \infty$ and the proposition follows. □

Denote by δ the dual of the operator D defined on $\mathbb{D}_{2,1}(E)$. The domain of δ [denoted by $\operatorname{dom} \delta(E)$] is the set of square integrable processes $u \in L^2(\Omega \times [0,1], \mathbb{R}^l \otimes E) = L^2(\Omega, H \otimes E)$ such that

$$|E(\langle u, DF \rangle_{H \otimes E})| \leq C \|F\|_{L^2(\Omega, E)},$$

for any $F \in \mathbb{D}_{2,1}(E)$, where C is some constant. In this case we have

$$E(\langle u, DF \rangle_{H \otimes E}) = E(\langle \delta(u), F \rangle_E). \tag{1.3}$$

Expression (1.3) is known as the *integration by parts formula*. Another important relation between the operators D and δ is given by the following proposition.

PROPOSITION 1.4. *Let $u \in \mathbb{D}_{2,1}(H \otimes E)$ be such that $D_t^i u \in \operatorname{dom} \delta(H \otimes E)$ for any $i = 1, \dots, l$ and $t \in [0,1]$ a.e. and suppose that $\sum_{i=1}^l \int_0^1 E(|\delta(D_t^i u)|^2)\, dt < \infty$. Then $\delta(u) \in \mathbb{D}_{2,1}(H \otimes E)$ and*

$$D_t^i(\delta(u)) = \delta(D_t^i u) + u_t^i. \tag{1.4}$$

Consider the l-dimensional Wiener process W as a Gaussian orthogonal measure on $T = [0,1] \times \{1, \dots, 1\}$ and let $u \in L^2(\Omega \times [0,1], \mathbb{R}^l \otimes E)$. By means of the Wiener-chaos decomposition of $L^2(\Omega, \mathbb{R}^l \otimes E)$ we can write u_t as an orthogonal series

$$u_t = \sum_{m=0}^{\infty} I_m(f_m(\cdot, t)),$$

where for any m, $f_m^{j_1, \dots, j_m, j}(s_1, \dots, s_m, t)$ is a symmetric function of the m variables $(s_1, j_1), \dots, (s_m, j_m)$ for each fixed (t, j) and belongs to $L^2([0,1]^{m+1}, \mathbb{R}^{l(m+1)} \otimes E)$.

The Skorohod integral of the process u is then defined by

$$\int_0^1 u_t\, dW_t = \sum_{m=0}^{\infty} I_{m+1}(\tilde{f}_m),$$

where $\tilde{f}_m$ denotes the symmetrization of f_m in the $(m+1)$ variables (s_i, j_i), $1 \le i \le m$, (t, j), provided that the series converges in $L^2(\Omega)$ (see Skorohod [15], Nualart and Zakai [13] and Nualart and Pardoux [12]). Notice that this integral allows one to integrate nonadapted processes and is in fact an extension of the Itô integral (see [13]).

Gaveau and Trauber have proved in [5] that the operator δ coincides with the Skorohod integral. Using the Wiener–chaos expansion, it can be proved, as in [13], that $\mathbb{D}_{2,1}(H \otimes E) \subset \operatorname{dom} \delta(E)$.

We suppose that our reference probability space is the product of the Wiener space $(\Omega, \mathscr{F}, P)$ with some separable probability space $(\Omega_0, \mathscr{F}_0, P_0)$. Then we can take $E = L^2(\Omega_0, \mathscr{F}_0, P_0)$ and identify $L^2(\Omega, E)$ with $L^2(\Omega \times \Omega_0, \mathscr{F} \otimes \mathscr{F}_0, P \times P_0)$. In this way we can define the operator D on square integrable random variables defined on the product space $\Omega \times \Omega_0$. The preceding results concerning the operators D and δ can be properly translated into this new framework.

2. Time reversal for diffusions: Globally Lipschitz coefficients. Consider the d-dimensional diffusion process $\{X_t = (X_t^1, \dots, X_t^d),\ 0 \le t \le 1\}$, solution of the stochastic differential equation

$$X_t = X_0 + \int_0^t \sigma(s, X_s)\, dW_s + \int_0^t b(s, X_s)\, ds, \tag{2.1}$$

where $X_0 = (X_0^1, \dots, X_0^d)$ is a random vector, and $\sigma\colon [0,1] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^l$ and $b\colon [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$ are Borel measurable functions satisfying the hypothesis:

(H1) There exists a constant K such that

$$\sup_t \left[|\sigma(t,x) - \sigma(t,y)| + |b(t,x) - b(t,y)|\right] \le K|x-y|,$$

$$\sup_t \left[|\sigma(t,x)| + |b(t,x)|\right] \le K(1 + |x|),$$

for every $x, y \in \mathbb{R}^d$.

In order to study the reversibility of the diffusion property we need a result which is partly contained in [2].

THEOREM 2.1. *The process X given by* (2.1) *belongs to* $\mathbb{D}_{\infty,1}(L^2([0,1], \mathbb{R}^d))$. *Moreover, assuming*

(H2) *for any $t > 0$, X_t has a density $p_t(x)$,*

the derivative of X is the solution of the stochastic differential system

$$\begin{aligned} D_r^\beta X_t^i &= 0, \qquad \text{if } t < r, \\ D_r^\beta X_t^i &= \sigma_\beta^i(r, X_r) + \int_r^t \nabla_k \sigma_\alpha^i(s, X_s) D_r^\beta X_s^k\, dW_s^\alpha \\ &\quad + \int_r^t \nabla_k b^i(s, X_s) D_r^\beta X_s^k\, ds, \qquad r \le t, \end{aligned} \tag{2.2}$$

$$i = 1, \dots, d,\ \beta = 1, \dots, l.$$

PROOF. 1. If σ and b are $\mathscr{C}^1$ functions in the x variable the result holds without assumption (H2). This can be proved by a slight modification of the approximation argument used by Ikeda and Watanabe in [8].

2. In the general case, we consider the sequence of regularization kernels defined by (1.2) and take new coefficients $\sigma_n(t,\cdot) = \sigma(t,\cdot) * \alpha_n(\cdot)$, $b_n(t,\cdot) = b(t,\cdot) * \alpha_n(\cdot)$. σ_n and b_n are $\mathscr{C}^\infty$ functions in the second variable, and it is easy to check that they satisfy hypothesis (H1) uniformly in n.

Moreover,

$$\lim_{n\to\infty} \sup_{x,t} \{|\sigma_n(t,x) - \sigma(t,x)| + |b_n(t,x) - b(t,x)|\} = 0. \tag{2.3}$$

Now consider the stochastic differential equation

$$X_t^n = X_0 + \int_0^t \sigma_n(s, X_s^n)\,dW_s + \int_0^t b_n(s, X_s^n)\,ds.$$

By the first part of the proof, $X^n \in \mathbb{D}_{\infty,1}(L^2([0,1],\mathbb{R}^d))$ and

$$\begin{aligned} D_r^\beta X_t^{n,i} &= 0, &&\text{if } t < r,\\ D_r^\beta X_t^{n,i} &= \sigma_{n,\beta}^i(r, X_r^n) + \int_r^t \nabla_k \sigma_{n,\alpha}^i(s, X_s^n) D_r^\beta X_s^{n,k}\,dW_s^\alpha \\ &\quad + \int_r^t \nabla_k b_n^i(s, X_s^n) D_r^\beta X_s^{n,k}\,ds, &&\text{otherwise.}\end{aligned}$$

We have

$$\lim_{n\to\infty} E\Big\{\sup_{0\le t\le 1} |X_t^n - X_t|^p\Big\} = 0, \qquad \forall\, p \ge 2. \tag{2.4}$$

In fact,

$$\begin{aligned} &E\Big\{\sup_{0\le t\le 1} |X_t^n - X_t|^p\Big\}\\ &\quad\le CE\Big\{\int_0^1 [|\sigma_n(s, X_s^n) - \sigma(s, X_s)|^p + |b_n(s, X_s^n) - b(s, X_s)|^p]\,ds\Big\}\\ &\quad\le CE\Big\{\int_0^1 \sup_{0\le u\le s} |X_u^n - X_u|^p\,ds\\ &\qquad\quad + \int_0^1 [|\sigma_n(s, X_s) - \sigma(s, X_s)|^p + |b_n(s, X_s) - b(s, X_s)|^p]\,ds\Big\}.\end{aligned}$$

Therefore, using (2.3) and Gronwall's lemma, (2.4) follows.

On the other hand, it is easy to prove that

$$\sup_n \sup_{0\le r\le 1} E\Big\{\sup_{r\le t\le 1} |D_r X_t^n|^p\Big\} < \infty,$$

for all $p \ge 2$. Hence the sequence $\{DX^n,\ n \ge 1\}$ is bounded in $L^p(\Omega \times [0,1]^2, \mathbb{R}^d \otimes \mathbb{R}^l)$. Denote again by $\{DX^n,\ n \ge 1\}$ a subsequence converging in the weak topology $\sigma(L^p, L^q)$ to some $Y \in L^p(\Omega \times [0,1]^2, \mathbb{R}^d \otimes \mathbb{R}^l)$.

The next step is to identify Y with DX, by proving that they both have the same projection on every Wiener chaos.

To this end, consider $Z \in L^2(\Omega \times [0,1]^2, \mathbb{R}^d \otimes \mathbb{R}^l)$ having a finite Wiener-chaos representation. By the integration by parts formula and (2.4),

$$\langle DX^n, Z\rangle = \langle X^n, \delta Z\rangle \underset{n\to\infty}{\longrightarrow} \langle X, \delta Z\rangle$$

and

$$\langle DX^n, Z\rangle \underset{n\to\infty}{\longrightarrow} \langle Y, Z\rangle.$$

Consequently, $Y = DX$, where DX is formally given by a series expansion like $\sum_{n\geq 1} nJ_n X$. This fact, together with (2.4), implies that $X \in \mathbb{D}_{\infty,1}(L^2([0,1], \mathbb{R}^d))$.

Suppose that (H2) holds. Once we know that $X \in \mathbb{D}_{\infty,1}(L^2([0,1], \mathbb{R}^d))$ we can write

$$\begin{aligned} D_r^\beta X_t^i &= D_r^\beta\left[\int_0^t \sigma_\alpha^i(s, X_s)\, dW_s^\alpha + \int_0^t b^i(s, X_s)\, ds\right] \\ &= \int_0^t D_r^\beta\left[\sigma_\alpha^i(s, X_s)\right] dW_s^\alpha + \sigma_\beta^i(r, X_r) + \int_0^t D_r^\beta\left(b^i(s, X_s)\right) ds, \quad \text{if } t \geq r, \end{aligned}$$

by the derivative rules of D. By Proposition 1.3 we obtain

$$D_r^\beta X_t^i = \sigma_\beta^i(r, X_r) + \int_r^t \nabla_k \sigma_\alpha^i(s, X_s) D_r^\beta X_s^k\, dW_s^\alpha + \int_r^t \nabla_k b^i(s, X_s) D_r^\beta X_s^k\, ds, \qquad t \geq r.$$

Notice that $\nabla_k \sigma_\alpha^i(s, X_s)$ and $\nabla_k b^i(s, X_s)$ are well-defined, measurable and adapted processes, and by the usual methods it can be proved that (2.2) has a unique continuous solution. Therefore we are done. □

We next state necessary and sufficient conditions for the Markov process $\{\overline{X}_t,\ 0 \leq t < 1\}$ to be a solution of the martingale problem associated with a second-order operator

$$\overline{L}_t f(x) = \tfrac{1}{2}\bar{a}^{ij}(t,x)\nabla_{ij} f(x) + \bar{b}^i(t,x)\nabla_i f(x), \tag{2.5}$$

for $f \in \mathscr{C}_0^\infty(\mathbb{R}^d)$.

THEOREM 2.2. *Assume that* σ *and* b *satisfy hypotheses* (H1) *and* (H2). *Consider measurable functions* $\bar{a}^{ij}$: $[0,1] \times \mathbb{R}^d \to \mathbb{R}$, $\bar{b}^i$: $[0,1] \times \mathbb{R}^d \to \mathbb{R}$, $i, j = 1, \ldots, d$, *such that*

$$\int_{t_0}^1 \int_D \left[|\bar{a}^{ij}(1-t, x)| + |\bar{b}^i(1-t, x)|\right] p_t(x)\, dx\, dt < \infty, \tag{2.6}$$

for any bounded open set $D \subset \mathbb{R}^d$ *and any* $t_0 > 0$.

Let $\overline{L}_t$ *be defined by* (2.5) *and assume that* $\{\overline{X}_t, 0 \leq t < 1\}$ *is a solution of the martingale problem associated with* $\overline{L}_t$.

Then the sums of distributional derivatives $\nabla_j(a^{ij}(t,x)p_t(x))$, $i = 1, \ldots, d$ $[a^{ij} = (\sigma\sigma^*)_{ij}]$, *are locally integrable functions, i.e.,*

$$\int_{t_0}^1 \int_D \left|\nabla_j\left[a^{ij}(t,x)p_t(x)\right]\right| dx\, dt < \infty, \tag{2.7}$$

for any bounded open set $D \subset \mathbb{R}^d$ *and any* $t_0 > 0$.

In addition,

$$\bar{a}^{ij}(t,x) = a^{ij}(1-t,x)$$

and

$$p_t(x)\left[\bar{b}^i(1-t,x) + b^i(t,x)\right] = \nabla_j\left(a^{ij}(t,x)p_t(x)\right),$$

for all t, x a.e.

PROOF. Since $\bar{X}$ is Markovian,

$$E\left[f(\bar{X}_t) - f(\bar{X}_s) - \int_s^t (\bar{L}_u f)(\bar{X}_u)\,du \middle| \bar{X}_s\right] = 0,$$

for any $f \in \mathscr{C}_0^\infty(\mathbb{R}^d)$, $0 \le s \le t < 1$. Equivalently,

$$E\left[f(X_t) - f(X_s) - \int_s^t (\tilde{L}_u f)(X_u)\,du \middle| X_t\right] = 0,$$

where $\tilde{L}_t = -\bar{L}_{1-t}$. That is,

$$E\{[f(X_t) - f(X_s)]g(X_t)\} = E\left\{\left(\int_s^t (\tilde{L}_u f)(X_u)\,du\right) g(X_t)\right\},$$

where g is any function in $\mathscr{C}_0^\infty(\mathbb{R}^d)$.

By Lebesgue's differentiation theorem

$$\lim_{h\downarrow 0}\frac{1}{h}E\left\{\left[\int_{t-h}^t (\tilde{L}_u f)(X_u)\,du\right]g(X_t)\right\}$$
$$= E\left[\left(-\frac{1}{2}\bar{a}^{ij}(1-t,X_t)\nabla_{ij}f(X_t) - \bar{b}^i(1-t,X_t)\nabla_i f(X_t)\right)g(X_t)\right],$$

for almost every $t \in [0,1]$.

By Itô's formula

$$\lim_{h\downarrow 0}\frac{1}{h}E\left[(f(X_t) - f(X_{t-h}))g(X_t)\right]$$
$$= \lim_{h\downarrow 0}\frac{1}{h}E\left\{\left(\int_{t-h}^t \left[\nabla_i f(X_s)\sigma_\alpha^i(s,X_s)\,dW_s^\alpha + \nabla_i f(X_s)b^i(s,X_s)\,ds\right.\right.\right.$$
$$\left.\left.\left. + \tfrac{1}{2}\nabla_{ij}f(X_s)a^{ij}(s,X_s)\,ds\right]\right)g(X_t)\right\}.$$

The martingale property of stochastic integrals together with Itô's formula applied to g yield

$$\lim_{h\downarrow 0}\frac{1}{h}E\left[\left(\int_{t-h}^t \nabla_i f(X_s)\sigma_\alpha^i(s,X_s)\,dW_s^\alpha\right)g(X_t)\right]$$
$$= \lim_{h\downarrow 0}\frac{1}{h}E\left[\left(\int_{t-h}^t \nabla_i f(X_s)\sigma_\alpha^i(s,X_s)\,dW_s^\alpha\right)\right.$$
$$\left.\times\left(\int_{t-h}^t \nabla_j g(X_s)\sigma_\alpha^j(s,X_s)\,dW_s^\alpha + \int_{t-h}^t G(s)\,ds\right)\right],$$

where

$$G(s) = \nabla_j g(X_s) b^j(s, X_s) + \tfrac{1}{2} \nabla_{jk} g(X_s) a^{jk}(s, X_s)$$

is bounded. Since $\nabla_i f(X_s)\sigma_\alpha^i(s, X_s)$ is also bounded, Schwarz's inequality and standard estimations imply that

$$\lim_{h \downarrow 0} \frac{1}{h} E\left[\left(\int_{t-h}^{t} \nabla_i f(X_s)\sigma_\alpha^i(s, X_s)\, dW_s^\alpha\right)\left(\int_{t-h}^{t} G(s)\, ds\right)\right] \le C \lim_{h \downarrow 0} h^{1/2} = 0.$$

Hence by the Lebesgue differentiation theorem

$$\begin{aligned}
\lim_{h \downarrow 0} \frac{1}{h} & E[(f(X_t) - f(X_{t-h}))g(X_t)] \\
&= \lim_{h \downarrow 0} \frac{1}{h} E\left\{\left(\int_{t-h}^{t} \left[\nabla_i f(X_s) b^i(s, X_s) + \frac{1}{2}\nabla_{ij} f(X_s) a^{ij}(s, X_s)\right] ds\right) g(X_t)\right. \\
&\qquad\qquad \left. + \int_{t-h}^{t} \nabla_i f(X_s) a^{ij}(s, X_s) \nabla_j g(X_s)\, ds\right\} \\
&= E\left\{\left[\nabla_i f(X_t) b^i(t, X_t) + \frac{1}{2} \nabla_{ij} f(X_t) a^{ij}(t, X_t)\right] g(X_t)\right. \\
&\qquad\qquad \left. + \nabla_i f(X_t) a^{ij}(t, X_t) \nabla_j g(X_t)\right\},
\end{aligned}$$

for all $t \in [0,1]$ a.e.

We remark that this result could also be obtained using the integration by parts formula (1.3). Therefore for all $t \in [0,1]$ a.e. we have

$$\begin{aligned}
\int_{\mathbb{R}^d} & \left(-\tfrac{1}{2}\bar{a}^{ij}(1-t, x)\nabla_{ij} f(x) - \bar{b}^i(1-t, x)\nabla_i f(x)\right) g(x) p_t(x)\, dx \\
&= \int_{\mathbb{R}^d} \nabla_i f(x) a^{ij}(t, x) \nabla_j g(x) p_t(x)\, dx \\
&\quad + \int_{\mathbb{R}^d} \left(\nabla_i f(x) b^i(t, x) + \tfrac{1}{2} \nabla_{ij} f(x) a^{ij}(t, x)\right) g(x) p_t(x) \qquad (2.8) \\
&= -\int_{\mathbb{R}^d} \nabla_j\left(\nabla_i f(x) a^{ij}(t, x) p_t(x)\right) g(x)\, dx \\
&\quad + \int_{\mathbb{R}^d} \left(\nabla_i f(x) b^i(t, x) + \tfrac{1}{2} \nabla_{ij} f(x) a^{ij}(t, x)\right) g(x) p_t(x).
\end{aligned}$$

In particular, for $f(x) \approx x_i$ we obtain

$$\int_{\mathbb{R}^d} \bar{b}^i(1-t, x) p_t(x) g(x)\, dx = \int_{\mathbb{R}^d} \left[\nabla_j\left(a^{ij}(t, x) p_t(x)\right) - b^i(t, x) p_t(x)\right] g(x)\, dx,$$

for any $t \in [0,1]$ a.e. and every $g \in \mathscr{C}_0^\infty(\mathbb{R}^d)$.

Consequently,

$$\nabla_j\big(a^{ij}(t,x)p_t(x)\big) = \big[\bar{b}^i(1-t,x) + b^i(t,x)\big]\,p_t(x),$$

for all t, x a.e., and (2.7) is satisfied.

Looking back to (2.8) and taking $f(x) \approx x_i x_j$, we finally obtain $\bar{a}^{ij}(t,x) = a^{ij}(1-t,x)$ for every t, x a.e. □

THEOREM 2.3. *Assume that σ and b satisfy hypotheses* (H1) *and* (H2). *Suppose also that the sums of distributional derivatives $\nabla_j(a^{ij}(t,x)p_t(x))$, $i = 1,\dots,d$, are locally integrable functions, i.e., satisfy* (2.7). *Then $\{\bar{X}_t, 0 \le t < 1\}$ is a diffusion process with generator $\bar{L}_t$ given by* (2.5), *where*

$$\begin{aligned} &\bar{a}^{ij}(1-t,x) = a^{ij}(t,x),\\ (2.9)\qquad &\bar{b}^i(1-t,x) = -b^i(t,x) + \frac{1}{p_t(x)}\nabla_j\big(a^{ij}(t,x)p_t(x)\big), \end{aligned}$$

with the convention that the term involving $(p_t(x))^{-1}$ is 0 *if $p_t(x) = 0$.*

Before giving the proof of this theorem we will state a lemma.

LEMMA 2.4. *Let $\{X_{s,t}(x),\ s \le t \le 1\}$ be the process solution of*

$$(2.10)\qquad X_{s,t}(x) = x + \int_s^t \sigma\big(u, X_{s,u}(x)\big)\,dW_u + \int_s^t b\big(u, X_{s,u}(x)\big)\,du$$

and $g \in \mathscr{C}_0^\infty(\mathbb{R}^d)$. Define

$$\varphi(s,x) = E\big[g(X_t)|X_s = x\big] = E\big[g\big(X_{s,t}(x)\big)\big],$$

where $t \in [0,1]$ is fixed. Then $\varphi(s,x)$ is globally Lipschitz in x, uniformly in s.

PROOF. We will use the following fact. If $g \in \mathscr{C}_0^1(\mathbb{R}^d)$ and $a\colon [0,1]\times\mathbb{R}^d \to \mathbb{R}$ is locally Lipschitz in x, uniformly in t, then $x \to \nabla g(x)a(t,x)$ is also globally Lipschitz in x, uniformly in t.

For any $x, y \in \mathbb{R}^d$ we have, using Itô's formula,

$$\begin{aligned} |\varphi(s,x) - \varphi(s,y)| \le \Bigg| & g(x) - g(y)\\ &+ E\bigg[\int_s^t \Big(\nabla_i g\big(X_{s,u}(x)\big)b^i\big(u, X_{s,u}(x)\big)\\ &\qquad - \nabla_i g\big(X_{s,u}(y)\big)b^i\big(u, X_{s,u}(y)\big)\Big)\,du\bigg]\\ &+ \tfrac{1}{2}E\bigg[\int_s^t \Big(\nabla_{ij} g\big(X_{s,u}(x)\big)a^{ij}\big(u, X_{s,u}(x)\big)\\ &\qquad - \nabla_{ij} g\big(X_{s,u}(y)\big)a^{ij}\big(u, X_{s,u}(y)\big)\Big)\,du\bigg]\Bigg|. \end{aligned}$$

But

$$E\left[|\nabla_i g(X_{s,u}(x))b^i(u, X_{s,u}(x)) - \nabla_i g(X_{s,u}(y))b^i(u, X_{s,u}(y))|\right]$$
$$\leq CE\left[|X_{s,u}(x) - X_{s,u}(y)|\right],$$

by the remark at the beginning of the proof, and the same majorization also holds for the last term. Therefore

$$|\varphi(s,x) - \varphi(s,y)| \leq C\Big\{|x-y| + \sup_{s\leq u\leq t} E\left[|X_{s,u}(x) - X_{s,u}(y)|\right]\Big\}$$
$$\leq C|x-y|,$$

as can be easily seen from Gronwall's lemma. □

Proof of Theorem 2.3. We need only to show that $\{\overline{X}_t, 0 \leq t < 1\}$ is a solution of the martingale problem associated with $\overline{L}_t$, i.e., since $\overline{X}_t$ is Markov

$$E\{[f(X_t) - f(X_s)]g(X_t)\} = E\left\{\left(\int_s^t (\tilde{L}_u f)(X_u)\,du\right)g(X_t)\right\}, \qquad s < t. \tag{2.11}$$

First we remark that the left-hand side of (2.11) is absolutely continuous as a function of $s \in [0, t]$. In fact, using Itô's formula and integration by parts, we obtain

$$\begin{aligned}
&E\{[f(X_t) - f(X_s)]g(X_t)\}\\
&\quad= E\left\{\left[\int_s^t \nabla_i f(X_u)\sigma_\alpha^i(u, X_u)\,dW_u^\alpha\right.\right.\\
&\qquad\qquad \left.\left. + \nabla_i f(X_u)b^i(u, X_u)\,du + \tfrac{1}{2}\nabla_{ij} f(X_u)a^{ij}(u, X_u)\,du\right]g(X_t)\right\}\\
&\quad= E\left[\int_s^t \nabla_i f(X_u)\sigma_\alpha^i(u, X_u)\nabla_k g(X_t)D_u^\alpha X_t^k\,du\right.\\
&\qquad\qquad + \nabla_i f(X_u)b^i(u, X_u)g(X_t)\,du\\
&\qquad\qquad \left. + \tfrac{1}{2}\nabla_{ij} f(X_u)a^{ij}(u, X_u)g(X_t)\,du\right].
\end{aligned}$$

Hence in order to establish (2.11) it will suffice to prove that for all $s \in [0, t]$ a.e.

$$\lim_{h\downarrow 0}\frac{1}{h}E[(f(X_s) - f(X_{s-h}))g(X_t)] = E[\tilde{L}_s f(X_s)g(X_t)]. \tag{2.12}$$

We have

$$\begin{aligned}
&\lim_{h\downarrow 0}\frac{1}{h}E\left[\int_{s-h}^s\left(\nabla_i f(X_u)b^i(u, X_u)g(X_t)\right.\right.\\
&\qquad\qquad \left.\left. + \frac{1}{2}\nabla_{ij} f(X_u)a^{ij}(u, X_u)g(X_t)\right)du\right]\\
&\quad= E\left[\nabla_i f(X_s)b^i(s, X_s)g(X_t) + \frac{1}{2}\nabla_{ij} f(X_s)a^{ij}(s, X_s)g(X_t)\right],
\end{aligned} \tag{2.13}$$

for any $s \leq t$ a.e.

Furthermore,

$$\lim_{h\downarrow 0}\frac{1}{h}E\left[\left(\int_{s-h}^{s}\nabla_i f(X_u)\sigma_\alpha^i(u,X_u)\,dW_u^\alpha\right)g(X_t)\right]$$
$$(2.14)\qquad = \lim_{h\downarrow 0}\frac{1}{h}E\left[\left(\int_{s-h}^{s}\nabla_i f(X_u)\sigma_\alpha^i(u,X_u)\,dW_u^\alpha\right)\varphi(s,X_s)\right],$$

where φ is defined in Lemma 2.4.

The solution of (2.2) can be written as $D_r^\beta X_t^i = \sigma_\beta^k(r,X_r)Y_k^i(t,r)$, where $Y_k^i(t,r)$, $t\ge r$, is given by the stochastic differential system

$$Y_k^i(t,r) = \delta_k^i + \int_r^t \nabla_j\sigma_\alpha^i(s,X_s)Y_k^j(s,r)\,dW_s^\alpha + \int_r^t \nabla_j b^i(s,X_s)Y_k^j(s,r)\,ds,\qquad i,k=1,\dots,d.$$

$\{Y_k^i(t,r),\, t\ge r\}$ has a continuous version in (t,r), as follows from Kolmogorov's continuity criterion.

By the integration by parts formula, Lemma 2.4 and Proposition 1.3, we have

$$(2.14) = \lim_{h\downarrow 0}\frac{1}{h}E\left[\int_{s-h}^{s}\nabla_i f(X_u)\sigma_\alpha^i(u,X_u)(\nabla_k\varphi)(s,X_s)D_u^\alpha X_s^k\,du\right]$$
$$(2.15)\qquad = \lim_{h\downarrow 0}\frac{1}{h}E\left[\int_{s-h}^{s}\sum_{\alpha=1}^{l}\nabla_i f(X_u)\sigma_\alpha^i(u,X_u)(\nabla_k\varphi)(s,X_s)\times\sigma_\alpha^j(u,X_u)Y_j^k(s,u)\,du\right]$$
$$= E\left[\nabla_i f(X_s)a^{ij}(s,X_s)(\nabla_j\varphi)(s,X_s)\right],\qquad \forall\, s\le t \text{ a.e.}$$

By Lemma A.1 we obtain

$$(2.15) = \int_{\mathbb{R}^d}\nabla_i f(x)a^{ij}(s,x)(\nabla_j\varphi)(s,x)p_s(x)\,dx$$
$$= -\int_{\mathbb{R}^d}\nabla_j\big(\nabla_i f(x)a^{ij}(s,x)p_s(x)\big)\varphi(s,x)\,dx$$
$$= -\int_{\mathbb{R}^d}\nabla_{ij}f(x)a^{ij}(s,x)p_s(x)\varphi(s,x)\,dx - \int_{\mathbb{R}^d}\nabla_i f(x)\nabla_j\big(a^{ij}(s,x)p_s(x)\big)\varphi(s,x)\,dx,$$

and finally by Lemma A.2

$$(2.15) = -E\left[\nabla_{ij}f(X_s)a^{ij}(s,X_s)g(X_t)\right]$$
$$(2.16)\qquad -E\left[\nabla_i f(X_s)\frac{1}{p_s(X_s)}\nabla_j\big(a^{ij}(s,X_s)p_s(X_s)g(X_t)\big)\right].$$

The results obtained in (2.13) and (2.16) prove the equality stated in (2.12). □

3. Locally Lipschitz coefficients. In this section we suppose that the coefficients σ and b of (2.1) satisfy the following conditions.

($\mathrm{H1}_{\mathrm{loc}}$)(a) Given any constant A, there exists a constant $K(A)$ such that

$$\sup_t \left[|\sigma(t,x) - \sigma(t,y)| + |b(t,x) - b(t,y)|\right] \le K(A)|x-y|,$$

$$\sup_t \left[|\sigma(t,x)| + |b(t,x)|\right] \le K(A)(1+|x|),$$

for each $|x|, |y| \le A$.

(b) The solution X_t does not explode on $[0,1]$.

($\mathrm{H2}_{\mathrm{loc}}$) For all $t > 0$, X_t has a density $p_t(x)$.

Our goal is to extend in this situation the results proved in Section 2. The proof of Theorem 2.2 only uses the local Lipschitz and boundedness properties of b, σ and a; hence it remains true when b and σ satisfy ($\mathrm{H1}_{\mathrm{loc}}$) and ($\mathrm{H2}_{\mathrm{loc}}$). In order to prove the analogous of Theorem 2.3 we need some preliminaries.

3.1. *Preliminary results.* We approximate the solution X_t of (2.1) by a sequence of process (X_t^n) which coincide with X_t up to a stopping time T_n and have a density; this is done by a slight modification of the classical argument (see Kunita [9]).

Let $\tilde{W}$ denote a d-dimensional Brownian motion independent of W. For each $n \ge 1$ let $\varphi_n\colon \mathbb{R}_+ \to \mathbb{R}$ be defined by $\varphi_n(x) = [(2 - n^{-1}x) \vee 0]$ and let $\psi_n\colon \mathbb{R} \to [0,1]$ be a $\mathscr{C}^\infty$ function such that $\psi_n(x) = 1$ if $|x| \ge 2n$, $\psi_n(x) = 0$ if $|x| \le n$ and $\psi_n(x) > 0$ for $n < x$.

Define $\sigma_n\colon [0,1] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^{l+d}$ and $b_n\colon [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$ by

$$\begin{aligned} (\sigma_n)^i_\beta(t,x) &= \sigma^i_\beta(t,x)\varphi_n(|x|), && \text{if } 1 \le \beta \le l, 1 \le i \le d, \\ (\sigma_n)^i_{l+\beta}(t,x) &= \delta^i_\beta \psi_n(|x|), && \text{if } 1 \le \beta \le d, 1 \le i \le d, \\ b_n(t,x) &= b(t,x)\varphi_n(|x|). \end{aligned} \tag{3.1}$$

An easy computation shows the existence of a constant K_n, depending on the constant $K(2n)$ given in ($\mathrm{H1}_{\mathrm{loc}}$) and of $\|\psi_n'\|_\infty$, such that

$$\sup_t \left[|\sigma_n(t,x) - \sigma_n(t,y)| + |b_n(t,x) - b_n(t,y)|\right] \le K_n|x-y|$$

and

$$\sup_t \left[|\sigma_n(t,x)| + |b_n(t,x)|\right] \le K_n(1+|x|).$$

Let X^n denote the solution of the stochastic differential equation (on the enlarged probability space)

$$\begin{aligned} dX_t^n &= \sigma_n(t, X_t^n)\,d(W_t, \tilde{W}_t) + b_n(t, X_t^n)\,dt, \\ X_0^n &= X_0. \end{aligned} \tag{3.2}$$

Set $T_n = \inf\{t, |X_t^n| \geq n\} \wedge 1$. Then if $0 \leq s \leq t$ on the set $\{t \leq T_n\}$ one has that $(\sigma_n)^i_\beta(s, X_s^n) = \sigma^i_\beta(s, X_s^n)$ for $1 \leq \beta \leq l$, $(\sigma_n)^i_{l+\beta}(s, X_s^n) = 0$ for $1 \leq \beta \leq d$ and $b_n(s, X_s^n) = b(s, X_s^n)$. Thus on the set $\{t \leq T_n\}$, $X_t^n = X_t^{n+k}$ for each $k \geq 0$. Therefore setting $X_t = X_t^n$ on $\{t \leq T_n\}$, we define a solution of the given stochastic differential system on $[0, 1]$ by assumption $(\mathrm{H1}_{\mathrm{loc}})$(b).

Each random variable X_t^n belongs to $\mathbb{D}_{2,1}$ and the sequence $\{(\{T_n \geq t\}, X_t^n); n \geq 1\}$ localizes the random variable X_t. Hence X_t belongs to $\mathbb{D}_{2,1}^{\mathrm{loc}}$ (see [12]), and we can define, up to a $dP \times dt$ equivalence class, the derivative DX_t by setting

$$D_r^\alpha X_t|_{\{T_n \geq t\}} := D_r^\alpha X_t^n|_{\{T_n \geq t\}}.$$

However, the derivative is not square integrable and we do not have an integration by parts formula for random variables in $\mathbb{D}_{2,1}^{\mathrm{loc}}$.

We prove that assumption $(\mathrm{H1}_{\mathrm{loc}})$(b) implies that the approximating processes (X_t^n) have a density.

LEMMA 3.1. *For each $t > 0$, $n \geq 1$, the random variable X_t^n has a density.*

PROOF. Set $S_n = \inf\{t, |X_t^n| > n\}$. Then for $0 \leq s \leq t$ on $\{S_n \geq t\}$, $(\sigma_n)^i_\beta(s, X_s^n) = \sigma^i_\beta(s, X_s)$ for $1 \leq \beta \leq l$, $(\sigma_n)^i_{\beta+l}(s, X_s^n) = 0$ for $1 \leq \beta \leq d$ and $b_n(s, X_s^n) = b(s, X_s)$.

Let B be a Borel subset of $\mathbb{R}^d$ with Lebesgue measure $\lambda(B) = 0$. Then if P denotes also the probability on the enlarged space,

$$\begin{aligned} P\{X_t^n \in B\} &= P\{X_t^n \in B, S_n \geq t\} + P\{X_t^n \in B, S_n < t\} \\ &= P\{X_t \in B, S_n \geq t\} + P\{X_t^n \in B, S_n < t\}. \end{aligned}$$

Since X_t has a density, the first term in the sum is 0. Notice that $X_t^n \in \mathbb{D}_{2,1}$. Therefore by Theorem 9 of [2],

$$E\Big\{1_{\{X_t^n \in B\}} \det\big(\langle DX_t^{n,i}, DX_t^{n,j}\rangle_{1 \leq i,j \leq d}\big)\Big\} = 0.$$

The determinant is nonnegative. Consequently,

$$E\Big\{1_{\{X_t^n \in B, S_n < t\}} \det\big(\langle DX_t^{n,i}, DX_t^{n,j}\rangle_{1 \leq i,j \leq d}\big)\Big\} = 0.$$

Therefore, in order to conclude the proof, it suffices to check that $\det(\langle DX_t^{n,i}, DX_t^{n,j}\rangle_{1 \leq i,j \leq d}) > 0$ on $\{X_t^n \in B, S_n < t\}$.

Let $A = \{(s, x) \in \mathbb{R}_+ \times \mathbb{R}^d, \operatorname{rank} \sigma_n(s, x) = d\}$ and suppose that $S_n(\omega, \tilde{\omega}) < t$. The path continuity of the process X^n implies $(\omega, \tilde{\omega})$ a.s. the existence of a (random) interval $[s - h, s + h] \subset [0, t]$ such that $|X_u^n(\omega, \tilde{\omega})| > n$ for $u \in [s - h, s + h]$. By definition of σ_n, the rank of $\sigma_n(u, X_u^n)$ is larger than that of $\psi_n(|X_u^n|) Id_d$ and is equal to d for $u \in [s - h, s + h]$. Hence if

$$S = \inf\Big\{t, \int_0^t 1_A(s, X_s^n)\, ds > 0\Big\},$$

one has that $S(\omega, \tilde{\omega}) < t$ on $\{S_n(\omega, \tilde{\omega}) < t\}$. The proof of Theorem 19 in [2] shows that

$$\det\left(\langle DX_t^{n,i}, DX_t^{n,j}\rangle_{1\le i,j\le d}\right) > 0 \quad \text{on } \{S < t\} \supset \{S_n < t\}. \qquad \square$$

In the sequel, when no confusion arises, we still denote by P (resp. W) the probability [resp. the $(d+l)$-Brownian motion $(W(\omega), \tilde{W}(\tilde{\omega}))$] on the extended space.

3.2. *Direct part: Preliminary results on the dependence on the initial condition.* Our purpose is to state an analogue of Theorem 2.3 in the case of locally Lipschitz coefficients. The proof is more complicated than the corresponding one in the case of globally Lipschitz coefficients. Indeed, we do not know if the functions $\varphi(s, x) = E[g(X_t)|X_s = x]$ is Lipschitz, because the expected value "mixes" the various Lipschitz constants corresponding to the various ω. In order to avoid this problem we study the dependence of the solution of the stochastic differential equation upon the initial condition and then prove a result which replaces the integration by parts formula.

Fix $s \ge 0$, $x \in \mathbb{R}^d$ and let $X_{s,t}(x)$ denote the solution of the stochastic differential equation (2.10). When $s = 0$ simply set $X_t(x) = X_{0,t}(x)$.

Let $T(s,x)$ denote the explosion time of $X_{s,\cdot}(x)$ and set $D_{s,t}(\omega) = \{x, T(s,x) > t\}$ for $s < t$. We suppose that the solution $X_{s,t}(x)$ is strictly conservative, i.e., $P\{T(s,x) > 1 \text{ for all } (s,x)\} = 1$. Then almost every set $D_{s,t}(\omega)$ is equal to $\mathbb{R}^d$, and we have $X_t(\omega) = X_{s,t}(X_s(\omega))$ a.e. for $s < t$ (see, e.g., Kunita [9], Section 5).

Given a bounded domain $D \subset \mathbb{R}^m$, $m \ge 1$, let $W^{1,2}(D)$ denote the Sobolev space of functions $f\colon D \to \mathbb{R}^d$ such that f and its distributional derivatives are square integrable functions, i.e.,

$$\|f\|_{1,2}^2 = \int_D \left[|f(x)|^2 + \sum_i |\nabla_i f(x)|^2\right] dx < \infty.$$

Let $W^{1,2}_{\mathrm{loc}}$ denote the intersection of $W^{1,2}(D)$ over bounded domains D of $\mathbb{R}^d$.

If a (real- or vector-valued) process $X_t(x)$ belongs to $L^2(\Omega \times [0,1], W^{1,2}(D))$ set

$$\|X\|^2 = E\left[\int_0^1 \|X_t(x)\|_{1,2}^2\, dt\right].$$

First we prove that the process $X_{s,t}(x)$ belongs to $W^{1,2}_{\mathrm{loc}}$, and, under additional assumptions, that the gradient $\nabla_x X_{s,t}(x)$ is the solution of a linear stochastic differential equation and is square integrable. All the results needed are summarized in the following lemma, which is stated with $s = 0$ for the sake of simplicity.

LEMMA 3.2. *Suppose that the coefficients b and σ satisfy the condition* ($\mathrm{H1}_{\mathrm{loc}}$)(a) *and that the solution $X_t(x)$ is strictly conservative. Then:*

1. *For any t, $X_t(\cdot)(\omega)$ belongs to $W^{1,2}_{\mathrm{loc}}$ a.s.*
2. *Suppose furthermore* ($\mathrm{H2}_{\mathrm{loc}}$) *for $X_0 \equiv x$, i.e., $X_t(x)$ has a density for $t > 0$.*

Then the distributional gradient $\nabla X_t(x)$ *is solution of the stochastic differential equation*

$$\nabla_i X_t^k(x) = \delta_i^k + \int_0^t \nabla_j \sigma_\beta^k(u, X_u(x)) \nabla_i X_u^j(x)\, dW_u^\beta + \int_0^t \nabla_j b^k(u, X_u(x)) \nabla_i X_u^j(x)\, du. \tag{3.3}$$

3. *Suppose that each random variable* $X_t(x)$, $t > 0$, *has a density, and that for*

$$B_s(x) = \left[\sum_{i,k} \nabla_i b^k(s, X_s(x))^2\right]^{1/2}, \qquad A_s^\beta(x) = \left[\sum_{i,k} \nabla_i \sigma_\beta^k(s, X_s(x))^2\right]^{1/2},$$
$$\sup_{x\in D} E\left\{\exp\left(\int_0^1 \left[4B_s(x) + 8\sum_\beta \left(A_s^\beta(x)\right)^2\right] ds\right)\right\} < \infty, \tag{3.4}$$

where D *is a bounded open subset of* $\mathbb{R}^d$. *Then* $\|\nabla X_t(x)\|^2$ *is integrable for each* $t > 0$, *and*

$$\sup_{x\in D} E\left[\sup_t \sum_i \sum_k \left(\nabla_i X_t^k(x)\right)^2\right] < \infty.$$

PROOF. (A) We at first suppose that the coefficients b and σ satisfy assumption (H1) of Section 2. Define recursively the sequence $\xi_t^n(x)$ of processes by

$$\xi_t^0(x) = x,$$
$$\xi_t^{n+1}(x) = x + \int_0^t \sigma(s, \xi_s^n(x))\, dW_s + \int_0^t b(s, \xi_s^n(x))\, ds, \qquad n \geq 0.$$

Standard arguments show that

$$E\left[\sum_n \sup_{0\leq t\leq 1} |\xi_t^n(x) - \xi_t^{n+1}(x)|^2\right] < \infty,$$

and hence that $X_t(x) = \lim_n \xi_t^n(x)$ in $L^2(\Omega \times [0,1])$.

1. Fix a bounded domain $D \subset \mathbb{R}^d$. We prove by induction on n that for any t, $\xi_t^n(x)$ belongs to $W^{1,2}(D)$ a.s. and that

$$\sup_{0\leq t\leq 1} \sup_n E\left(\|\xi_t^n(x)\|_{1,2}^2\right) < \infty. \tag{3.5}$$

Then for any fixed t (3.5) yields the existence of a subsequence $\xi_t^{n'}$ converging weakly to $\xi_t \in W^{1,2}(D)$ in $L^2(\Omega, W^{1,2}(D))$. Thus $X_t = \xi_t$ and $X_t(x) \in W^{1,2}(D)$ a.s.

To prove (3.5), we show that $\sigma_\beta^i(s, \xi_s^n(x))$ and $b^i(s, \xi_s^n(x))$ belong to $L^2(\Omega \times [0,1], W^{1,2}(D))$ and that

$$E\left[\int_0^1 \sum_i \|\sigma_\beta^i(s, \xi_s^n(x))\|_{1,2}^2\, ds\right] \leq c_1 + c_2\|\xi^n\|^2,$$
$$E\left[\int_0^1 \sum_i \|b^i(s, \xi_s^n(x))\|_{1,2}^2\, ds\right] \leq c_1 + c_2\|\xi^n\|^2, \tag{3.6}$$

where c_1 and c_2 do not depend on n.

These inequalities imply (see Lemma A.3) that the integrals $\int_0^t \sigma_\beta(s, \xi_s^n(x))\, dW_s^\beta$ and $\int_0^t b(s, \xi_s^n(x))\, ds$ belong to $L^2(\Omega, W^{1,2}(D))$ and

$$\nabla_i\left(\int_0^t \sigma_\beta^k(s, \xi_s^n(x))\, dW_s^\beta\right) = \int_0^t \nabla_i\left[\sigma_\beta^k(s, \xi_s^n(x))\right] dW_s^\beta,$$

$$\nabla_i\left(\int_0^t b^k(s, \xi_s^n(x))\, ds\right) = \int_0^t \nabla_i\left[b^k(s, \xi_s^n(x))\right] ds. \tag{3.7}$$

Then (3.6) and (3.7) imply that there exist constants A_1 and A_2 such that

$$\sup_{0 \le s \le 1} E\left(\|\xi_s^{n+1}(x)\|_{1,2}^2\right) \le A_1 + A_2 \int_0^1 \sup_{0 \le u \le s} E\left(\|\xi_u^n(x)\|_{1,2}^2\right) ds,$$

and (3.5) follows.

The inequalities (3.6) are proved by using the sequence of regularization kernels defined by (1.2). Suppose that $S_m(t, x) = \sigma(t, x) * \alpha_m(x)$, $B_m(t, x) = b(t, x) * \alpha_m(x)$. Then S_m and B_m satisfy hypothesis (H1) with a constant K independent of m. Moreover,

$$B_m^k(s, \xi_s^n(x)) \to b^k(s, \xi_s^n(x)) \quad \text{in } L^2(\Omega \times [0,1]) \text{ a.s. } m \to \infty,$$

and

$$|\nabla_i[B_m^k(s, \xi_s^n(x))]| = |(\nabla_j B_m^k)(s, \xi_s^n(x)) \nabla_i \xi_s^{n,j}(x)| \le K \sum_j |\nabla_i \xi_s^{n,j}(x)|.$$

Thus the sequence $\{B_m^k(s, \xi_s^n(x)),\ m \ge 1\}$ is bounded in $L^2(\Omega \times [0,1], W^{1,2}(D))$. Hence it has a weakly convergent subsequence to some element which must coincide with $b^k(s, \xi_s^n(x))$ and verifies the second inequality of (3.6). The proof of the first one is similar.

2. Assume furthermore that $X_t(x)$ has a density for each $t > 0$. Then by Lemma A.4

$$\nabla_i\left(\int_0^t \sigma_\beta^k(s, X_s(x))\, dW_s^\beta\right) = \int_0^t \nabla_j \sigma_\beta^k(s, X_s(x)) \nabla_i X_s^j(x)\, dW_s^\beta$$

and similarly

$$\nabla_i\left(\int_0^t b^k(s, X_s(x))\, ds\right) = \int_0^t \nabla_j b^k(s, X_s(x)) \nabla_i X_s^j(x)\, ds.$$

Therefore, taking the distributional derivative on both sides of the stochastic differential equation defining $X_t(x)$, we obtain (3.3).

(B) Now suppose that the coefficients b and σ are locally Lipschitz and bounded and let $X^n(t)$ denote the solution of (3.2) for $X_0 = x$, on the enlarged probability space described in Section 3.1.

1. Set $T_n(x) = \inf\{s,\ |X_s^n(x)| > n\}$ and for any fixed $t \in [0,1]$ let $D_t^n(\omega) = \{x, T_n(x)(\omega) > t\}$. Then $X_t(x) = X_t^n(x)$ on $\{t \le T_n\}$ and $X_t^n(\cdot) \in W^{1,2}(D_t^n(\omega))$ a.s. The strict conservativeness of $X_t(x)$ shows that for every t and almost every ω, $D_t^n(\omega) \nearrow \mathbb{R}^d$. Thus $X_t(\cdot)(\omega) \in W_{\text{loc}}^{1,2}$ a.s.

2. Suppose furthermore that each random variable $X_t(x)$, $t > 0$, has a density. Then by Lemma 3.1 the approximating processes $X_t^n(x)$, $t > 0$, also have a

density. Therefore (A) shows that the distributional derivatives $\nabla_i X_t^{n,\,t}(x)$ satisfy the stochastic differential equation

$$\nabla_i X_t^{n,\,k}(x) = \delta_i^k + \int_0^t \nabla_j(\sigma_n)_\beta^k(s, X_s^n(x))\nabla_i X_s^{n,\,j}(x)\,dW_s^\beta$$
$$+ \int_0^t \nabla_j(\psi_n(|X_s^n(x)|)\nabla_i X^{n,\,j}(x)\,d\tilde{W}_s^k$$
$$+ \int_0^t \nabla_j b_n^k(s, X_s^n(x))\nabla_i X_s^{n,\,j}(x)\,ds.$$

Since $\{t \le T_n\} \uparrow \Omega$ a.s. we have that $\nabla_i X_t^k(x)$ satisfies (3.3).

3. For the sake of simplicity, we drop x from the notation. The argument involves an arbitrary vector $x \in D$.

Set $Y_i^k(t) = \nabla_i X_t^k(x)$ and apply Itô's formula to each $|Y_i^k(t)|^2$. Then for $Y_t = \sum_{i,k}|Y_i^k(t)|^2$, one has that $Y_t > 0$ a.s. and

$$Y_t = d + \int_0^t U_\beta(s)Y_s\,dW_s^\beta + \int_0^t (V'(s) + V''(s))Y_s\,ds, \tag{3.8}$$

with

$$U_\beta(s) = \left[2\sum_{i,j,k} Y_i^k(s)\nabla_j\sigma_\beta^k(s, X_s(x))Y_i^j(s)\right]Y_s^{-1},$$

$$V'(s) = \left[2\sum_{i,j,k} Y_i^k(s)\nabla_j b^k(s, X_s(x))Y_i^j(s)\right]Y_s^{-1}$$

and

$$V''(s) = \left[\sum_{i,j,k,l,\beta} \nabla_j\sigma_\beta^k(s, X_s(x))Y_i^j(s)\nabla_l\sigma_\beta^k(s, X_s(x))Y_i^l(s)\right]Y_s^{-1}.$$

Apply Schwarz's inequality to $\sum_i$ and then $\sum_k$; we obtain that

$$|U_\beta(s)| \le 2\left[\sum_{i,k}\nabla_i\sigma_\beta^k(s, X_s(x))^2\right]^{1/2} = 2A_s^\beta,$$

$$|V'(s)| \le 2\left[\sum_{i,k}\nabla_i b^k(s, X_s(x))^2\right]^{1/2} = 2B_s.$$

Schwarz's inequality applied to $\sum_i$, $\sum_j$ and then $\sum_l$ yields

$$|V''(s)| \le \sum_{i,k,\beta}\nabla_i\sigma_\beta^k(s, X_s(x))^2 = \sum_\beta (A_s^\beta)^2.$$

The integrability assumption (3.4) implies the integrability of

$$Z_t = d\exp\left\{\int_0^t U_\beta(s)\,dW_s^\beta + \int_0^t [V'(s) + V''(s)]\,ds - \tfrac{1}{2}\int_0^t \sum_\beta U_\beta(s)^2\,ds\right\},$$

which is the solution of the linear stochastic differential equation (3.8) and thus equal to Y_t. Indeed,

$$Z_t = d\exp\left\{\int_0^t U_\beta(s)\,dW_s^\beta - \int_0^t \sum_\beta U_\beta(s)^2\,ds\right\}$$

$$\times\exp\left\{\int_0^t\left[V'(s) + V''(s) + \tfrac{1}{2}\sum_\beta U_\beta(s)^2\right]ds\right\}.$$

The first term in the product is square integrable, since by assumption

$$E\left\{\exp\left[\int_0^1 2\sum_\beta U_\beta(s)^2\,ds\right]\right\} \le E\left\{\exp\left[\int_0^1 8\sum_\beta \left(A_s^\beta\right)^2 ds\right]\right\} < \infty.$$

The second term also is square integrable. Indeed,

$$E\left\{\exp\left[\int_0^t\left[2|V'(s)| + 2|V''(s)| + \sum_\beta U_\beta(s)^2\right]ds\right]\right\}$$

$$\le E\left\{\exp\left[\int_0^1\left[4B_s + 6\sum_\beta\left(A_s^\beta\right)^2\right]ds\right]\right\} < \infty.$$

Therefore Hölder's inequality yields that

$$\sup_{x\in D} E\left[\sup_t Z_t\right] \le \sup_{x\in D} d\left\{E\left\{\exp\left[\int_0^1\left[4B_s + 6\sum_\beta\left(A_s^\beta\right)^2\right]ds\right]\right\}\right\}^{1/2} < \infty.$$

This completes the proof of the lemma. □

REMARKS.

1. The lemma remains valid if the initial condition is $X_0 = (X_0', x)$, where $X_0' \in \mathbb{R}^{d-d'}$ is fixed, $x \in \mathbb{R}^{d'}$, $d > d'$, and the function $X_t(\cdot)(\omega)$ is studied on $\mathbb{R}^{d'}$. In this case the property of strict conservativeness with respect to $x \in \mathbb{R}^{d'}$ is sufficient.
2. The existence of the density of $X_t(x)$ for $t > 0$ is only required to ensure the definition of $\nabla_j b^k(u, X_u(x))$ and $\nabla_j \sigma_\beta^k(u, X_u(x))$ as P equivalence classes. It is therefore not necessary if $b(t,\cdot)$ and $\sigma(t,\cdot)$ are of class $\mathscr{C}^1$ with derivatives locally bounded uniformly in t.
3. If for some components $1 \le i \le d'$, the coefficients b^i and σ_β^i are null (and hence X_t^i, $t \ge 0$, remain constant), then the existence of the density (with respect to Lebesgue's measure on $\mathbb{R}^{d-d'}$) of $\{X_t^i,\ d' < i \le d\}$ is sufficient to obtain (3.3), and the integrability condition (3.4) is expressed in terms of

$$B_s(x) = \left[\sum_{d'<i}\sum_{k\le d} \nabla_i b^k(s, X_s(x))^2\right]^{1/2}$$

and

$$A_s^\beta(x) = \left[\sum_{d'<i}\sum_{k\le d} \nabla_i \sigma_\beta^k(s, X_s(x))^2\right]^{1/2}.$$

3.3. *Direct part: The main theorem.* We now prove the direct part by a technique similar to that of Theorem 2.3. Since $X_t \in \mathbb{D}_{2,1}^{\mathrm{loc}}$, we have to establish a formula of integration by parts for X_t. This is done in Lemma 3.4 by means of Girsanov's theorem and Lemma 3.2. For each $x \in \mathbb{R}^d$, $s < t$, let $X_{s,t}(x)$ be the solution of (2.10) and set

$$B_{s,t}(x) = \left[\sum_{i,k} \nabla_i b^k\big(t, X_{s,t}(x)\big)^2\right]^{1/2}, \qquad A^\beta_{s,t}(x) = \left[\sum_{i,k} \nabla_i \sigma^k_\beta\big(t, X_{s,t}(x)\big)^2\right]^{1/2}.$$

THEOREM 3.3. *Suppose that the coefficients b and σ of (2.1) satisfy conditions* ($\mathrm{H1}_{\mathrm{loc}}$) *and* ($\mathrm{H2}_{\mathrm{loc}}$), *that the system $X_{s,t}(x)$ of solutions of the stochastic differential equation* (2.10) *is strictly conservative and that each random variable $X_{s,t}(x)$ has a density when $s < t$. Suppose finally that the coefficients satisfy the integrability condition* (2.7) *and also*

$$\sup_{x \in D} \sup_{s} E\left\{\exp\left[\int_s^1 \left[4B_{s,t}(x) + 8\sum_\beta |A^\beta_{s,t}(x)|^2\right] dt\right]\right\} < \infty, \tag{3.9}$$

for each bounded domain $D \subset \mathbb{R}^d$.

Then the reversed Markov process $\overline{X}_t = X_{1-t}$ is a diffusion with generator $\overline{L}$ defined in (2.5), *with coefficients $\bar{a}$ and $\bar{b}$ given by* (2.9).

Before proving this theorem, we first establish an integration by parts formula.

LEMMA 3.4. *Suppose that the assumptions of Theorem* 3.3 *are satisfied and let $\{\phi_\beta(u, x), 1 \le \beta \le l\}$ be bounded functions with compact supports, Lipschitz in x, uniformly in u. Then for $g \in \mathscr{C}_0^\infty$,*

$$\begin{aligned} &E\left[g(X_t)\int_0^t \phi_\beta(u, X_u)\, dW_u^\beta\right] \\ &\quad = E\left[\int_0^t \phi_\beta(u, X_u)\nabla_j g(X_t)\sigma_\beta^k(u, X_u) Y_k^j(t, u)\, du\right], \end{aligned} \tag{3.10}$$

where $Y_k^j(t, u)$ is the solution of the stochastic differential system

$$\begin{aligned} Y_k^j(t, u) &= 0, \qquad \text{if } u > t, \\ Y_k^j(t, u) &= \delta_k^j + \int_u^t \nabla_i \sigma_\beta^j(r, X_r) Y_k^i(r, u)\, dW_r^\beta \\ &\quad + \int_u^t \nabla_i b^j(r, X_r) Y_k^i(r, u)\, dr, \quad \text{if } u \le t. \end{aligned} \tag{3.11}$$

PROOF. For $\varepsilon \in \mathbb{R}$ let X^ε denote the solution of the stochastic differential equation

$$X_t^\varepsilon = X_0 + \int_0^t \sigma_\beta(u, X_u^\varepsilon)\, dW_u^\beta + \int_0^t \left[b(u, X_u^\varepsilon) - \varepsilon \sum_\beta \phi_\beta(u, X_u)\sigma_\beta(u, X_u^\varepsilon)\right] du.$$

Then if $\tilde{W}^{\beta}_{\varepsilon, t} = W^{\beta}_t - \varepsilon\int_0^t \phi_\beta(u, X_u)\, du$,

$$\frac{dP_\varepsilon}{dP}(\omega) = \exp\left[\varepsilon\int_0^1 \phi_\beta(u, X_u)\, dW^\beta_u - \frac{1}{2}\varepsilon^2\int_0^1 \sum_\beta \phi_\beta(u, X_u)^2\, du\right],$$

and Girsanov's theorem shows that $\tilde{W}_\varepsilon$ is an l-dimensional Brownian motion under P_ε and that the law of X^ε under P is the same as the law of X under P_ε (see, e.g., Ikeda and Watanabe [8]). Hence

$$\varepsilon^{-1}E\left[g(X^\varepsilon_t) - g(X_t)\right] = \varepsilon^{-1}E\left[g(X_t)\left(\frac{dP_\varepsilon}{dP} - 1\right)\right].$$

The derivative of the function

$$\varepsilon \to g(X_t)\exp\left[\varepsilon\int_0^1 \phi_\beta(u, X_u)\, dW^\beta_u - \tfrac{1}{2}\varepsilon^2\int_0^1 \sum_\beta \phi_\beta(u, X_u)^2\, du\right]$$

is dominated by a P-integrable random variable for $|\varepsilon| \le 1$. Thus

$$\begin{aligned}\lim_{\varepsilon\to 0}\varepsilon^{-1}E\left[g(X^\varepsilon_t) - g(X_t)\right] &= E\left[g(X_t)\frac{d}{d_\varepsilon}\left(\frac{dP_\varepsilon}{dP}\right)\bigg|_{\varepsilon=0}\right]\\ &= E\left[g(X_t)\int_0^t \phi_\beta(u, X_u)\, dW^\beta_u\right].\end{aligned}$$

On the other hand, for $u \in [0,1]$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, $z \in \mathbb{R}$, $1 \le \beta \le l$, $1 \le i \le d$, set

$$\begin{aligned}\tilde{\sigma}^i_\beta(u,(x,y,z)) &= \sigma^i_\beta(u,x),\\ \tilde{\sigma}^{d+i}_\beta(u,(x,y,z)) &= \sigma^i_\beta(u,y),\\ \tilde{\sigma}^{2d+1}_\beta(u,(x,y,z)) &= 0,\\ \tilde{b}^i(u,(x,y,z)) &= b^i(u,x),\\ \tilde{b}^{d+i}(u,(x,y,z)) &= b^i(u,y) - z\sum_{\beta=1}^{l}\sigma^i_\beta(u,y)\phi_\beta(u,x),\\ \tilde{b}^{2d+1}(u,(x,y,z)) &= 0.\end{aligned} \tag{3.12}$$

Then $\tilde{\sigma}$ and $\tilde{b}$ are locally Lipschitz and locally bounded in (x, y, z), uniformly in u. Clearly, $\tilde{X}^\varepsilon_t = (X_t, X^\varepsilon_t, \varepsilon)$ is the solution of the stochastic differential equation

$$\tilde{X}^\varepsilon_t = (X_0, X_0, \varepsilon) + \int_0^t \tilde{\sigma}_\beta(u, \tilde{X}^\varepsilon_u)\, dW^\beta_u + \int_0^t \tilde{b}(u, \tilde{X}^\varepsilon_u)\, du. \tag{3.13}$$

Since P_ε is absolutely continuous with respect to P, X is strictly conservative under P_ε and hence $\tilde{X}^\varepsilon$ is conservative with respect to P.

Fix a bounded open set containing 0. Let $T(\varepsilon)$ be the explosion time of X^ε. We know that $P\{T(\varepsilon) = \infty\} = 1$. For any fixed $t > 0$ set $D_t = \{\varepsilon \in I, T(\varepsilon) > t\}$. It holds that (cf. [9]) D_t is an open set dense in I a.s.

Consider the sequence $\tilde{X}^{\varepsilon, n}$ of approximations of $\tilde{X}^{\varepsilon}$ described in Section 3.1 and define $T_n(\varepsilon) = \inf\{s, |X_s^{\varepsilon, n}| \geq n\}$ and $D_t^n = \{\varepsilon \in I, T_n(\varepsilon) > t\}$. Then $D_t^n \uparrow D_t$ a.s. as $n \to \infty$. On the set D_t^n the processes X^{ε} and $X^{\varepsilon, n}$ coincide. Using the first part of the proof of Lemma 3.2 and Remarks 1 and 3 after this lemma, we obtain that $\varepsilon \to X_t^{\varepsilon, n}$ belongs to $W^{1,2}(D_t^n)$ for any n. We want to show that $D_t = I$ and that $\varepsilon \to X_t^{\varepsilon}$ belongs to $W^{1,2}(I)$ a.s.

To this end let us consider the process Z_t^{ε} given by

$$Z_t^{\varepsilon} = \int_0^t \nabla_k \sigma_\beta(u, X_u^{\varepsilon}) Z_u^{\varepsilon, k}\, dW_u^{\beta} - \varepsilon \int_0^t \sum_\beta \nabla_k \sigma_\beta(u, X_u^{\varepsilon}) Z_u^{\varepsilon, k} \phi_\beta(u, X_u)\, du$$
$$+ \int_0^t \left[\nabla_k b(u, X_u^{\varepsilon}) Z_u^{\varepsilon, k} - \sum_\beta \sigma_\beta(u, X_u^{\varepsilon}) \phi_\beta(u, X_u) \right] du,$$

and the process

$$Z_t^{\varepsilon, n} = \frac{dX_t^{\varepsilon, n}}{d\varepsilon}.$$

For almost all ω and for all $\varepsilon \in D_t^n(\omega)$ a.e. we have

$$\frac{dX_t^{\varepsilon}}{d\varepsilon} = \frac{dX_t^{\varepsilon, n}}{d\varepsilon} = Z_t^{\varepsilon, n} = Z_t^{\varepsilon}.$$

Consider the process Z_t, which is the solution of the stochastic differential equation,

$$Z_t = \int_0^t \nabla_k \sigma_\beta(u, X_u) Z_u^k\, dW_u^{\beta} + \int_0^t \left[\nabla_k b(u, X_u) Z_u^k - \sum_\beta \sigma_\beta(u, X_u) \phi_\beta(u, X_u) \right] du. \tag{3.14}$$

Girsanov's theorem implies that the law of the pair (X_t^{η}, Z_t^{η}) under P is the same as the law of (X_t, Z_t) under P_η.

Furthermore, an easy computation shows that

$$Z_t^k = \int_0^t \sum_{\beta=1}^{l} \sigma_\beta^j(u, X_u) \phi_\beta(u, X_u) Y_j^k(t, u)\, du, \tag{3.15}$$

where $Y(t, u)$ is the solution of (3.11).

Lemma 3.2 shows that under the integrability assumption (3.9) we have

$$\sup_u E\left[\sup_t \sum_{j,k} |Y_k^j(t, u)|^2 \right] < \infty.$$

Since $\phi_\beta(u, x)$ and $\sum_{\beta=1}^{l} \sigma_\beta^j(u, x)\phi_\beta(u, x)$ are bounded, we have $\sup_t E[|Z_t|^2] < \infty$ and $E[(dP_\eta/dP)^2] < \infty$. Therefore

$$E\left[\int_I |Z_t^{\eta}|\, d\eta \right] \leq \sup_{\eta \in I} E\left(\frac{dP_\eta}{dP} |Z_t| \right) |I|$$
$$\leq \sup_{\eta \in I} |I| \left\{ E\left(\frac{dP_\eta}{dP} \right)^2 E|Z_t|^2 \right\}^{1/2} < \infty.$$

As a consequence the process $\{J_t^\varepsilon,\ \varepsilon \in I\}$ given by

$$J_t^\varepsilon = X_t + \int_0^\varepsilon Z_t^\eta\, d\eta$$

is well defined and has a.s. continuous paths.

Denote by $(\varepsilon_1(\omega), \varepsilon_2(\omega))$ the connex component of $D_t(\omega)$ containing 0. Notice that we may assume $0 \in D_t(\omega)$ a.s. The preceding arguments show that the variables X_t^ε and J_t^ε coincide on $(\varepsilon_1(\omega), \varepsilon_2(\omega))$ a.s. On the other hand, X_t^ε explodes at times $\varepsilon = \varepsilon_1(\omega)$ and $\varepsilon = \varepsilon_2(\omega)$. Hence we must have $D_t = I$ and $\varepsilon \to X_t^\varepsilon$ belongs to $W^{1,2}(I)$ a.s.

Therefore for almost all ω,

$$\frac{d}{d\varepsilon}[g(X_t^\varepsilon)] = \nabla_j g(X_t^\varepsilon)\frac{d}{d\varepsilon}X_t^{\varepsilon, j}$$

and, as a consequence,

$$g(X_t^\varepsilon) - g(X_t) = \int_0^\varepsilon \nabla_j g(X_t^\eta)\frac{d}{d\eta}X_t^{\eta, j}\, d\eta.$$

The Fubini theorem implies

$$\begin{aligned}\varepsilon^{-1}E[g(X_t^\varepsilon) - g(X_t)] &= \varepsilon^{-1}\int_0^\varepsilon E\left[\nabla_j g(X_t^\eta)\frac{d}{d\eta}X_t^{\eta, j}\right] d\eta\\ &= \varepsilon^{-1}\int_0^\varepsilon E\left[\nabla_j g(X_t)Z_t^j\frac{dP_\eta}{dP}\right] d\eta.\end{aligned}$$

Consequently,

$$\begin{aligned}\Delta_\varepsilon &= \left|\varepsilon^{-1}E[g(X_t^\varepsilon) - g(X_t)] - E\left[\nabla_j g(X_t)Z_t^j\right]\right|\\ &\le \varepsilon^{-1}\int_0^\varepsilon E\left[|\nabla_j g(X_t)Z_t^j|\left|\frac{dP_\eta}{dP} - 1\right|\right] d\eta\\ &\le K\varepsilon^{-1}\int_0^\varepsilon E\big(|Z_t|^2\big)^{1/2}E\left(\left|\frac{dP_\eta}{dP} - 1\right|^2\right)^{1/2} d\eta\\ &\le K\varepsilon^{-1}\int_0^\varepsilon E\left(\left|\frac{dP_\eta}{dP} - 1\right|^2\right)^{1/2} d\eta.\end{aligned}$$

The map $\eta \to E[|dP_\eta/dP - 1|^2]^{1/2}$ is continuous and null for $\eta = 0$. Therefore $\lim_{\varepsilon \to 0}\Delta_\varepsilon = 0$.

Replacing Z_t^j by the integral in (3.15), we obtain (3.10). □

Proof of Theorem 3.3. As in the proof of Theorem 2.3, let $0 < h < s < t$, $f, g \in \mathscr{C}_0^\infty$. By Itô's formula

$$\frac{1}{h}E\{[f(X_s) - f(X_{s-h})]g(X_t)\} = I_1 + I_2,$$

with

$$\lim_{h \to 0} I_1 = E\left\{g(X_t)\left[\nabla_i f(X_s)b^i(s, X_s) + \tfrac{1}{2}\nabla_{ij}f(X_s)a^{ij}(s, X_s)\right]\right\},$$

for all s a.e., and

$$I_2 = \frac{1}{h} E\left[g(X_t) \int_{s-h}^{s} \nabla_i f(X_u) \sigma_\beta^i(u, X_u)\, dW_u^\beta\right].$$

For $\beta = 1, \dots, l$ set

$$\phi_\beta(u, x) = \frac{1}{h} \nabla_i f(x) \sigma_\beta^i(u, x) 1_{[s-h, s]}(u).$$

ϕ_β is bounded with compact support and is Lipschitz in x, uniformly in u. Therefore by Lemma 3.4 we have

$$I_2 = \frac{1}{h} \int_{s-h}^{s} E\left[\sum_{\beta=1}^{l} \nabla_i f(X_u) \sigma_\beta^i(u, X_u) \times \nabla_j g(X_t) \sigma_\beta^k(u, X_u) Y_k^j(t, u)\right] du, \tag{3.16}$$

where $Y_k^j(t, u)$ is the solution of (3.11).

As in the proof of Theorem 2.3 we have that $s \to E[f(X_s) g(X_t)]$ is absolutely continuous. Furthermore, Lebesgue's differentiation theorem implies that

$$\lim_{h \to 0} I_2 = E\left[\nabla_i f(X_s) a^{ik}(s, X_s) \nabla_j g(X_t) Y_k^j(t, s)\right],$$

for any $s \in [0, t]$ a.e.

Let $X_{s,t}(x)$ be the solution of (2.10) and apply Lemma 3.2. Then for almost every (ω, t) the map $x \to X_{s,t}(x)$ belongs to $W_{\text{loc}}^{1,2}$. The integrability assumption (3.9) shows that, given any bounded domain D of $\mathbb{R}^d$,

$$\sup_{x \in D} \sup_{s} E\left[\sup_{t} |\nabla X_{s,t}(x)|^2\right] < \infty.$$

Set $\varphi(s, x) = E[g(X_{s,t}(x))]$. Lemma A.4 implies that

$$\nabla_k \varphi(s, x) = E\left[\nabla_j g(X_{s,t}(x)) \nabla_k X_{s,t}^j(x)\right]. \tag{3.17}$$

Since the solution $X_{s,t}(x)$ of (2.10) is strictly conservative, we have that $X_t(\omega) = X_{s,t}(X_s(\omega))$, ω a.s. (see, e.g., Kunita [9]). We next prove that $\nabla_k X_{s,t}^j(X_s(\omega))$ is well defined a.s. and that

$$\nabla_k X_{s,t}^j(X_s) = Y_k^j(t, s) \quad \text{a.s.} \tag{3.18}$$

Indeed, $\nabla_k X_{s,t}^j(x)$ is defined for almost every x. Let $\xi_1(s, t, \omega, x) = \xi_2(s, t, \omega, x)$ for almost every (ω, x) be elements in the equivalence class $\nabla X_{s,t}(x)(\omega)$. Then

$$\begin{aligned} \Delta &= E\left[|\xi_1(s, t, \omega, X_s(\omega)) - \xi_2(s, t, \omega, X_s(\omega))|^2\right] \\ &= E\left[E\left[|\xi_1(s, t, \omega, X_s(\omega)) - \xi_2(s, t, \omega, X_s(\omega))|^2 | X_s\right]\right]. \end{aligned}$$

Since $\xi_i(s, t, \omega, x)$ only depends on increments of W after time s, it is independent of X_s. Hence

$$\Delta = \int p_s(x) E\left[|\xi_1(s, t, \omega, x) - \xi_2(s, t, \omega, x)|^2\right] dx = 0.$$

Therefore $\nabla X_{s,t}(X_s)$ is well defined, ω a.s.

Furthermore, Lemma 3.2 shows that $\nabla_k X_{s,t}(x)$ is the solution of the stochastic differential equation

$$\begin{aligned}\nabla_k X^j_{s,t}(x) &= \delta^j_k + \int_s^t \nabla_i \sigma^j_\beta\big(u, X_{s,u}(x)\big) \nabla_k X^i_{s,u}(x)\, dW^\beta_u \\ &\quad + \int_s^t \nabla_i b^j\big(u, X_{s,u}(x)\big) \nabla_k X^i_{s,u}(x)\, du\end{aligned} \tag{3.19}$$

(except on a null set of ω independent of x).

In order to establish (3.18) it suffices to check that $\nabla_k X^j_{s,t}(X_s(\omega))$ is the solution of the stochastic differential equation (3.11).

The composition of stochastic flows $X_u = X_{s,u}(X_s)$ a.s. shows that

$$\left(\int_s^t \nabla_i b^j\big(u, X_{s,u}(x)\big) \nabla_k X^i_{s,u}(x)\, du\right) \circ X_s = \int_s^t \nabla_i b^j\big(u, X_s\big) \nabla_k X^i_{s,u}(X_s)\, du \quad \text{a.s.}$$

Fix j, k; for $\beta = 1, \dots, l$ set $\chi_\beta(u, x) = \nabla_i \sigma^j_\beta(u, X_{s,u}(x)) \nabla_k X^i_{s,u}(x)$. Notice that $\chi_\beta(u, x)$ is $\mathscr{F}_u$-adapted and only depends on increments of W after time s. We will prove that

$$\left(\int_s^t \chi_\beta(u, x)\, dW^\beta_u\right) \circ X_s = \int_s^t \chi_\beta(u, X_s)\, dW^\beta_u. \tag{3.20}$$

Indeed, let $\psi \in \mathscr{C}^\infty_0(\mathbb{R}^d)$ and let $R(u)$ be a bounded l-dimensional adapted process. Then

$$\begin{aligned}J &= E\left\{\psi(X_s)\left[\int_s^t R_\beta(u)\, dW^\beta_u\right]\left[\int_s^t \chi_\beta(u, X_s)\, dW^\beta_u\right]\right\} \\ &= E\left\{\psi(X_s) E\left[\left(\int_s^t \chi_\beta(u, X_s)\, dW^\beta_u\right)\left(\int_s^t R_\beta(u)\, dW^\beta_u\right)\Big| \mathscr{F}_s\right]\right\} \\ &= E\left\{\psi(X_s) E\left[\int_s^t \sum_\beta \chi_\beta(u, X_s) R_\beta(u)\, du \Big| \mathscr{F}_s\right]\right\}.\end{aligned}$$

Since the stochastic integrals $\int_s^t \chi_\beta(u, x)\, dW^\beta_u$ and $\int_s^t R_\beta(u)\, dW^\beta_u$ are independent of $\mathscr{F}_s$, we obtain

$$\begin{aligned}J &= \int_s^t \int_{\mathbb{R}^d} \psi(x) p_s(x) E\left[\sum_\beta R_\beta(u) \chi_\beta(u, x)\right] dx\, du \\ &= \int_{\mathbb{R}^d} \psi(x) p_s(x) E\left[\left(\int_s^t R_\beta(u)\, dW^\beta_u\right)\left(\int_s^t \chi_\beta(u, x)\, dW^\beta_u\right)\right] dx \\ &= \int_{\mathbb{R}^d} \psi(x) p_s(x) E\left[\left(\int_s^t R_\beta(u)\, dW^\beta_u\right)\left(\int_s^t \chi_\beta(u, x)\, dW^\beta_u\right)\Big| X_s = x\right] dx \\ &= E\left\{\psi(X_s)\left(\int_s^t R_\beta(u)\, dW^\beta_u\right)\left[\left(\int_s^t \chi_\beta(u, x)\, dW^\beta_u\right) \circ X_s\right]\right\}.\end{aligned}$$

The function ψ and the process R are arbitrary, therefore (3.18) holds. Equations (3.17) and (3.18) show that

$$(3.21)\qquad E\left[\nabla_j g(X_t) Y_k^j(t, s) | X_s\right] = \nabla_k \varphi(s, X_s) \quad \text{a.s.}$$

Then (3.21) implies that

$$\lim_{h\to 0} I_2 = E\left\{\nabla_i f(X_s) a^{ik}(s, X_s)\, E\left[\nabla_j g(X_t) Y_k^j(t, s) | X_s\right]\right\}$$
$$= E\left[\nabla_i f(X_s) a^{ik}(s, X_s) \nabla_k \varphi(s, X_s)\right],$$

for any $s \in [0, t]$ a.e.

Now we proceed as in the proof of Theorem 2.3 to conclude the proof. □

4. Sufficient conditions for absolute continuity. Assume that σ and b satisfy hypothesis (H1) of Section 2. We will now discuss sufficient conditions that guarantee the following properties:

(Ai) X_t has a density $p_t(x)$ for all $t \in (0, 1]$.
(Aii) The sums of distributional derivatives, $\nabla_j(a^{ij}(t, x)p_t(x))$, $i = 1, \dots, d$, are locally integrable functions.

From Section 2 we know that if (Ai) holds, then (Aii) is equivalent to the reversibility of the diffusion property. We first observe that Hörmander's conditions (see [7] and [3]) imply the smoothness of the density $p_t(x)$ for $t > 0$, and then properties (Ai) and (Aii) are true. These conditions include the assumptions that b and σ are $\mathscr{C}^\infty$ functions of x. On the other hand, under conditions (Ci) and (Cii) of Theorem 3.1 in [6], the process $\overline{X}_t = X_{1-t}$ is a Markov diffusion process, and therefore these conditions imply (Ai) and (Aii) due to Theorem 2.2. We remark that conditions (Ci) and (Cii) of [6] require very little regularity on the coefficients and the existence of an initial density p_0 satisfying some growth assumptions.

In relation to the existence of a density for X_t, we can state the following result.

PROPOSITION 4.1. *One of the next conditions implies that X_t has a density for any $t > 0$:*

(Hi) *The initial value is a fixed point $x_0 \in \mathbb{R}^d$, and* $\operatorname{rank} \sigma(s, x_0) = d$ *for all s in some neighborhood of* 0.
(Hii) *X_0 has a density, and the coefficients σ and b are of class $\mathscr{C}^{1,\alpha}$ for some $\alpha > 0$, i.e., with α-Hölder continuous derivatives.*

PROOF. Suppose first that (Hi) is true and define

$$T = \inf\{t \geq 0, |\{s \in [0, t], \operatorname{rank} \sigma(s, X_s) = d\}| > 0\}.$$

Then by (Hi) we have $P\{T = 0\} = 1$, and Theorem 19 of [2] implies that (Ai) holds.

Now assume (Hii). We know (cf. [9]) that for almost all ω, $x \to X_t(x)$ is a diffeomorphism of $\mathbb{R}^d$ onto $\mathbb{R}^d$ of class $\mathscr{C}^{1,\beta}$ for any $0 < \beta < \alpha$ and $0 \le t \le 1$. We can write for any Borel subset B of $\mathbb{R}^d$,

$$P\{X_t \in B\} = \int_{\mathbb{R}^d} P\{X_t(x) \in B\} p_0(x)\, dx$$

$$= E\int_{\mathbb{R}^d} 1_{\{x,\, X_t(x) \in B\}} p_0(x)\, dx.$$

Then $|B| = 0$ implies $|\{x, X_t(x) \in B\}| = 0$ and so $P\{X_t \in B\} = 0$. □

Finally, the next result provides sufficient conditions for (Aii) to hold.

PROPOSITION 4.2. *The following hypothesis implies* (Aii) [*and also* (Ai)]:

(Hiii) *The coefficients σ and b are of class $\mathscr{C}^2$ in x, and there exists $\varepsilon > 0$ such that $a(t, x) \ge \varepsilon I$.*

PROOF. Denote by Γ_t the Malliavin matrix of X_t,

$$\Gamma_t = \left(\langle DX_t^i, DX_t^j \rangle_{1 \le i, j \le d}\right).$$

Assume that $\det \Gamma_t > 0$ a.s. [which is true under hypothesis (Hiii) and implies (Ai)]. Let $\varphi \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ be a test function and compute

$$\int_{\mathbb{R}^d} \varphi(x) \nabla_j\left(a^{ij}(t, x) p_t(x)\right) dx$$

$$\begin{aligned} &= -\int_{\mathbb{R}^d} \nabla_j \varphi(x) a^{ij}(t, x) p_t(x)\, dx \\ &= -E\left[\nabla_j \varphi(X_t) a^{ij}(t, X_t)\right] \\ &= -E\left[\nabla_j(\varphi a^{ij})(t, X_t)\right] + E\left[\varphi(X_t) \nabla_j a^{ij}(t, X_t)\right]. \end{aligned} \tag{4.1}$$

Then

$$D\left(\varphi(X_t) a^{ij}(t, X_t)\right) = \nabla_k(\varphi a^{ij})(t, X_t)\, DX_t^k.$$

Therefore

$$\langle D\left(\varphi(X_t) a^{ij}(t, X_t)\right), DX_t^k \rangle = \sum_{\theta=1}^{d} \nabla_\theta(\varphi a^{ij})(t, X_t)(\Gamma_t)_{k\theta}$$

and

$$\langle D\left(\varphi(X_t) a^{ij}(t, X_t)\right), DX_t^k \rangle \left(\Gamma_t^{-1}\right)_{jk} = \nabla_j(\varphi a^{ij})(t, X_t).$$

As a consequence,

$$
\text{(4.2)}\qquad
\begin{aligned}
E\left[\nabla_j(\varphi a^{ij})(t, X_t)\right] &= E\left[\langle D\left(\varphi(X_t)a^{ij}(t, X_t)\right), \left(\Gamma_t^{-1}\right)_{jk} DX_t^k\rangle\right] \\
&= E\left(\varphi(X_t)a^{ij}(t, X_t)\delta\left[\left(\Gamma_t^{-1}\right)_{jk} DX_t^k\right]\right),
\end{aligned}
$$

if the stochastic processes

$$
\text{(4.3)}\qquad u_t^{j\beta}(s) = \left(\Gamma_t^{-1}\right)_{jk} D_s^\beta X_t^k, \qquad s \in [0,1],\ \beta = 1,\dots,l,
$$

are Skorohod integrable, for any $j = 1,\dots,d$.

Now, from (4.1) and (4.2) we deduce the following fact.

If $\det\Gamma_t > 0$ a.s. and the processes u_t^j given by (4.3) are Skorohod integrable, then (Ai) and (Aii) hold for this value of t, and the coefficient $c^i(t,x) = (1/p_t(x))\nabla_j(a^{ij}(t,x)p_t(x))$ verifies $E(|c(t, X_t)|) < \infty$.

In fact, we have

$$
c^i(t, X_t) = \nabla_j a^{ij}(t, X_t) - a^{ij}(t, X_t)E\left\{\delta\left[\left(\Gamma_t^{-1}\right)_{jk} DX_t^k\right] | X_t\right\}.
$$

Under hypothesis (Hiii) (and assuming that the first partial derivatives of σ and b are bounded) we can show that $X \in \mathbb{D}_{\infty,2}(L^2([0,1],\mathbb{R}^d))$, and $\sup_t E[(\det_t)^{-p}] < \infty$ for any $p \geq 2$ (see Stroock [16]). From these properties we deduce that u_t^j is Skorohod integrable. This completes the proof. □

APPENDIX

In this last section we state some technical lemmas which have been used throughout the paper.

Lemma A.1. *Let $h \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ and $f^j, g\colon \mathbb{R}^d \to \mathbb{R}$, $j = 1,\dots,d$, be functions such that g and its distributional derivatives $\nabla_j g$ are locally bounded functions, f^j and $\nabla_j f^j$ are locally integrable functions. Then*

$$
\int_{\mathbb{R}^d} hf^j\, \nabla_j g\, dx = -\int_{\mathbb{R}^d} \nabla_j\left(hf^j\right) g\, dx.
$$

Proof. Let $\{\alpha_n(x),\ n \geq 1\}$ be a sequence of regularization kernels. Then

$$
\begin{aligned}
\int_{\mathbb{R}^d} \nabla_j\left(hf^j\right) g\, dx &= \lim_n \int_{\mathbb{R}^d} \nabla_j\left(hf^j\right)\left(g * \alpha_n\right) dx \\
&= -\lim_n \int_{\mathbb{R}^d} hf^j\left(\nabla_j g * \alpha_n\right) dx = -\int_{\mathbb{R}^d} hf^j\, \nabla_j g\, dx.
\end{aligned}
$$

In fact, the first equality holds by dominated convergence because $\nabla_j(hf^j) = h\nabla_j f^j + (\nabla_j h)f^j$ and $|\nabla_j(hf^j)(g * \alpha_n)| \leq |\nabla_j(hf^j)|.\ \sup_{x\in K}|g(x)|$, where K is a suitable bounded set containing the support of h. The second equality follows from the relation $\nabla_j(g * \alpha_n) = \nabla_j g * \alpha_n$. Finally, the last convergence is also immediate. □

LEMMA A.2. *Let $p(x)$ be a probability density on $\mathbb{R}^d$ and let $a\colon \mathbb{R}^d \to \mathbb{R}^d$ be a locally Lipschitz function. Assume that $a^j \nabla_j p$ is a locally integrable function, then $a^j \nabla_j p = 0$ a.e. on $\{x, p(x) = 0\}$.*

PROOF. It follows the same lines as that of Lemma A.2 of [6]. We recall the main points. Let $\{\alpha_n, n \geq 1\}$ be a sequence of regularization kernels of the form (1.2). Clearly, the sequence $p_n = p * \alpha_n$ converges in $L^1(\mathbb{R}^d)$ and a.s. (by taking a suitable subsequence) to p. Furthermore, if we fix a bounded open set $D \subset \mathbb{R}^d$, it can be proved that $a^j \nabla_j p_n$ converges to $a^j \nabla_j p$ in the weak topology $\sigma(L^1(D), L^\infty(D))$ by choosing a suitable subsequence. This result uses the fact that a is locally Lipschitz, as in Lemma A.1 of [6].

It suffices to show that $a^j \nabla_j p = 0$ on the set $\{x, p(x) = 0\} \cap \{x, |a(x)| \neq 0\}$, or more precisely, on $\{x, p(x) = 0\} \cap \{x, k_1 < a^1(x) < k_2\}$. So, dividing by $a^1(x)$, we may assume that $a^1(x) = 1$. Fix x_0 on the open set $\{x, k_1 < a^1(x) < k_2\}$ and consider the transformation $x \to y(x)$ given by

$$y^1 = x^1,$$

$$\text{(A.1)} \qquad \frac{dy^i}{dx^1} = a^i(y(x)), \qquad y^i(x_0^1, x^2, \dots, x^d) = x^i, \qquad i = 2, \dots, d,$$

for x in some neighborhood V of x_0. Call U the image of V under this transformation. Let μ be the image of the Lebesgue measure by (A.1); then μ and Lebesgue's measure are equivalent.

Replacing p by ψp with a $\mathscr{C}^\infty$-function ψ with support contained in U, we may assume that p has compact support contained in U. Then we have

$$\text{(A.2)} \qquad p(y(x)) = \int_{-\infty}^{x^1} (a^j \nabla_j p)(y(\theta, x^2, \dots, x^d))\, d\theta,$$

for all x in V a.e. In fact, this is obviously true for p_n, and we can pass to the limit using the preceding convergences.

Finally, as in [6] this shows the desired property, taking into account that p is minimum on $\{x, p(x) = 0\}$. □

LEMMA A.3. *Let $\phi_s(x)$ be an adapted process of $L^2(\Omega \times [0,1], W^{1,2}(D))$. (We assume for simplicity that $l = 1$.) Then for any t the integrals $\int_0^t \phi_s(x)\, ds$ and $\int_0^t \phi_s(x)\, dW_s$ belong to $L^2(\Omega, W^{1,2}(D))$ and*

$$\nabla_i\left(\int_0^t \phi_s(x)\, ds\right) = \int_0^t \nabla_i \phi_s(x)\, ds,$$

$$\nabla_i\left(\int_0^t \phi_s(x)\, dW_s\right) = \int_0^t \nabla_i \phi_s(x)\, dW_s.$$

PROOF. We only show the second equality. First note that $\nabla_i \phi_s(x)$ is measurable in (x, s, ω) and adapted and $E\int_0^1\int_D |\nabla_i \phi_s(x)|^2\, dx\, ds < \infty$. So the preceding stochastic integral is well defined. Let $\varphi \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ be a test function. The stochastic Fubini theorem yields

$$\int_D \varphi(x)\left(\int_0^t \nabla_i \phi_s(x)\,dW_s\right)dx = \int_0^t\left(\int_D \varphi(x)\nabla_i\phi_s(x)\,dx\right)dW_s$$
$$= -\int_0^t\left(\int_D (\nabla_i\varphi)(x)\phi_s(x)\,dx\right)dW_s$$
$$= -\int_D (\nabla_i\varphi)(x)\left(\int_0^t \phi_s(x)\,dW_s\right)dx,$$

which proves the result. □

LEMMA A.4. *Let $\{Y^i(x),\ i = 1,\dots,d\}$ be random variables belonging to $L^2(\Omega, W^{1,2}(D))$ and let $f\colon \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz function. Assume that for almost all x, $Y(x)$ has a density. Then $f(Y(x)) \in W^{1,2}(D)$ and*

$$\nabla_i(f(Y(x))) = (\nabla_j f)(Y(x))\nabla_i Y^j(x),$$

for all (x, ω) a.e.

PROOF. Let $\{\alpha_k,\ k \geq 1\}$ be a sequence of regularization kernels and set $f_k = f * \alpha_k$. It holds that

$$\nabla_i(f_k(Y(x))) = (\nabla_j f_k)(Y(x))\nabla_i Y^j(x).$$

In fact, for any $\varphi \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ we can write

$$\int_{\mathbb{R}^d} f_k(Y(x))\nabla_i\varphi(x)\,dx = \lim_n \int_{\mathbb{R}^d} f_k(Y * \alpha_n)\nabla_i\varphi\,dx$$
$$= -\lim_n \int_{\mathbb{R}^d} (\nabla_j f_k)(Y * \alpha_n)(\nabla_i Y^j * \alpha_n)\varphi\,dx$$
$$= -\int_{\mathbb{R}^d} (\nabla_j f_k)(Y)(\nabla_i Y^j)\varphi\,dx.$$

By choosing a suitable subsequence we have

$$\nabla_j f_k \xrightarrow{\text{a.e.}} \nabla_j f.$$

Hence

$$\nabla_j f_k(Y(x)) \to (\nabla_j f)(Y(x)),$$

for all (x, ω) a.e., and the result follows easily. □

REFERENCES

[1] ANDERSON, B. D. O. (1982). Reverse time diffusion equation models. *Stochastic Process. Appl.* **12** 313–326.

[2] BOULEAU, M. and HIRSCH, F. (1986). Propietés d'absolue continuité dans les espaces de Dirichlet et applications aux équations différentielles stochastiques. *Séminaire de Probabilités XX. Lecture Notes in Math.* **1204** 131–161. Springer, New York.

[3] CHALEYAT-MAUREL, M. and MICHEL, D. (1984). Hypoellipticity theorems and conditional laws. *Z. Wahrsch. verw. Gebiete* **65** 573–597.

[4] FÖLLMER, H. (1986). Time reversal on Wiener space. *Stochastic Processes in Mathematics and Physics. Lecture Notes in Math.* **1158** 119–129. Springer, New York.

[5] GAVEAU, B. and TRAUBER, P. (1982). L'intégrale stochastique comme opérateur de divergence dans l'espace fonctionnel. *J. Funct. Anal.* **46** 230–238.

[6] HAUSSMANN, U. and PARDOUX, E. (1986). Time reversal of diffusions. *Ann. Probab.* **14** 1188–1205.

[7] HÖRMANDER, L. (1967). Hypoelliptic second order differential equations. *Acta Math.* **119** 147–171.

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

[9] KUNITA, H. (1984). Stochastic differential equations and stochastic flow of diffeomorphisms. *Ecole d'Été de Probabilités de Saint-Flour XII–1982. Lecture Notes in Math.* **1097**. Springer, New York.

[10] MALLIAVIN, P. (1978). Stochastic calculus of variations and hypoelliptic operators. *Proc. of the International Symposium on Stochastic Differential Equations, Kyoto, 1976* (K. Itô, ed.) 195–263. Wiley, New York.

[11] MAZ'JA, V. G. (1985). *Sobolev Spaces.* Springer, New York.

[12] NUALART, D. and PARDOUX, E. (1988). Stochastic calculus with anticipating integrands. *Probab. Theory Related Fields* **78** 535–581.

[13] NUALART, D. and ZAKAI, M. (1986). Generalized stochastic integrals and the Malliavin calculus. *Probab. Theory Related Fields* **73** 255–280.

[14] PARDOUX, E. (1986). Grossissement d'une filtration et retournement du temps d'une diffusion. *Séminaire de Probabilités XX. Lecture Notes in Math.* **1204**. Springer, New York.

[15] SKOROHOD, A. V. (1975). On a generalization of a stochastic integral. *Theory Probab. Appl.* **20** 219–233.

[16] STROOCK, D. (1981). The Malliavin calculus and its applications. *Stochastic Integrals. Lecture Notes in Math.* **851** 394–432. Springer, New York.

[17] WATANABE, S. (1984). *Lectures on Stochastic Differential Equations and Malliavin Calculus.* Springer, New York.

[18] WILLIAMS, D. (1979). *Diffusions, Markov Processes, and Martingales.* Wiley, New York.

[19] ZAKAI, M. (1985). The Malliavin calculus. *Acta Appl. Math.* **3-2** 175–207.

A. MILLET
FACULTÉ DES SCIENCES
UNIVERSITÉ D'ANGERS
2, BOULEVARD LAVOISIER
49045 ANGERS CEDEX
FRANCE

D. NUALART AND M. SANZ
FACULTAT DE MATEMÀTIQUES
UNIVERSITAT DE BARCELONA
GRAN VIA 585
08007 BARCELONA
SPAIN