The Annals of Probability
1986, Vol. 14, No. 4, 1188–1205

TIME REVERSAL OF DIFFUSIONS[1]

BY U. G. HAUSSMANN AND E. PARDOUX

University of British Columbia and Université de Provence

It is shown that if a diffusion process, $\{X_t\colon 0 \le t \le 1\}$, on R^d satisfies

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

then the reversed process, $\{\overline{X}_t\colon 0 \le t \le 1\}$ where $\overline{X}_t = X_{1-t}$, is again a diffusion with drift $\bar{b}$ and diffusion coefficient $\bar{\sigma}$, provided some mild conditions on b, σ, and p_0, the density of the law of X_0, hold. Moreover $\bar{b}$ and $\bar{\sigma}$ are identified.

1. Introduction. It is well known that a Markov process remains a Markov process under a time reversal. On the other hand, the strong Markov property is not necessarily preserved under time reversal [18] so it is of interest to see whether the diffusion property is preserved. Specifically if $\{X_t\colon 0 \le t \le 1\}$ is a diffusion process in $\mathbb{R}^d$ (hence a Markov process), solution of

$$dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dw_t, \tag{1.1}$$

where $\{w_t\colon 0 \le t \le 1\}$ is a standard Brownian motion in $\mathbb{R}^l$, and if $\overline{X}_t = X_{1-t}$ is the reversed process, we ask whether there exist $\bar{b}$, $\bar{\sigma}$, and a Brownian motion $\overline{w}$ such that

$$d\overline{X}_t = \bar{b}(t, \overline{X}_t)\,dt + \bar{\sigma}(t, \overline{X}_t)\,d\overline{w}_t \tag{1.2}$$

and we seek to identify $\bar{b}, \bar{\sigma}, \overline{w}$.

The problem has been of interest to physicists, most notably Nelson [15], who uses formally the reversibility of the diffusion property, as well as to control theorists, [1], [13], [16]. In [1] and [16] rather unverifiable conditions on the solution of the Fokker–Planck equation were given which guarantee the reversibility of the diffusion property, and $\bar{b}, \bar{\sigma}, \overline{w}$ were identified. Another approach, related to the problem of the enlargement of a filtration (grossissement d'une filtration) is used in [4] and [7], but again with unverifiable hypotheses or with incomplete proofs. Föllmer [8] has an interesting approach to the problem in the non-Markov case (but with $\sigma = I$), and in an infinite dimensional case. A related problem is treated in [19]. Finally, Azéma (private communication) has pointed out that it seems likely that the reversibility of the diffusion property follows under similar hypotheses as ours from the general theory of time reversal of Markov processes [2]; see also the work of Kunita and Watanabe referenced in [2].

In Section 2 we give conditions which insure that $\overline{X}$ is again a diffusion. The hypotheses are rather mild, but are still implicit to the extent that they require a certain integrability of the density of X_t. The method of proof uses weak (i.e., H^1-valued) solutions of the forward and backward Kolmogorov equations for X_t

Received November 1984.

[1] This work was carried out while the first author was visiting the Université de Provence.

AMS 1980 *subject classifications.* Primary 60J60; secondary 35K15.

Key words and phrases. Time reversal, diffusion process, Markov process, martingale problem, Kolmogorov equation.

to show that $\bar{X}_t$ solves the martingale problem corresponding to the generator of the solution of (1.2), i.e., the martingale problem $(\bar{\sigma}\bar{\sigma}^*, \bar{b})$, (* denotes transpose) with a given fixed initial law.

In Section 3 we give conditions only on b, σ, and p_0, the density of the law of X_0, which imply the hypotheses made in Section 2, and in the appendix we establish two technical lemmata.

These results were announced in [11]; one can find related results concerning the boundedness of $\bar{b}$ in [10].

2. Time reversal. We are given a diffusion process $\{X_t: 0 \le t \le 1\}$ on $\mathbb{R}^d$ satisfying the differential equation

$$dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dw_t. \tag{2.1}$$

We make the following hypotheses:

(A)(i) b: $[0,1] \times \mathbb{R}^d \to \mathbb{R}^d$, σ: $[0,1] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^l$ are Borel measurable and satisfy

$$\begin{aligned} |b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| &\le K|x-y|, \\ |b(t,x)| + |\sigma(t,x)| &\le K(1+|x|), \end{aligned} \tag{2.2}$$

for some constant K.

(ii) For almost all $t > 0$, $X(t)$ has a density $p(t,x)$ such that for all $t_0 > 0$,

$$p \in L^2(t_0, 1; \tilde{H}^1_{\mathrm{loc}}).$$

The notation here is that $p \in L^2(t_0,1; \tilde{H}^1_{\mathrm{loc}})$ if for any open bounded set $\mathcal{O}$

$$\int_{t_0}^1 \int_{\mathcal{O}} |p(t,x)|^2 + \sum_i \left|\sigma^{ji}(t,x) p(t,x)_{x_j}\right|^2 dx\,dt < \infty,$$

where $p(t,x)_{x_i}$ denotes the partial derivative of $p(t,\cdot)$ in the distribution sense, and where we use the convention that repeated indices are summed. The condition (A)(i) implies that the unique strong solution, $\{X_t\}$, of (2.1) is a Markov process with generator

$$\begin{aligned} L_t v(x) &= \tfrac{1}{2} a^{ij}(t,x) v_{x_i x_j} + b^i(t,x) v_{x_i} \\ &= \tfrac{1}{2}\left[a^{ij}(t,x) v_{x_j}\right]_{x_i} + \tilde{b}^i(t,x) v_{x_i} \end{aligned}$$

if $\tilde{b}^i(t,x) = b^i(t,x) - [a^{ij}(t,x)]_{x_j}/2$. We are denoting components by superscripts. $\{w_t: 0 \le t \le 1\}$ is a standard Brownian motion on $\mathbb{R}^l$ and $a(t,x) = \sigma(t,x)\sigma(t,x)^*$. Later we shall give some hypotheses involving only b, σ, and p_0 which imply (A)(ii).

We define

$$\begin{aligned} \bar{b}^i(t,x) &= -b^i(1-t,x) \\ &\quad + p(1-t,x)^{-1}\left[a^{ij}(1-t,x)p(1-t,x)\right]_{x_j}, \\ \bar{a}^{ij}(t,x) &= a^{ij}(1-t,x), \qquad \bar{\sigma}^{ij}(t,x) = \sigma^{ij}(1-t,x), \\ \bar{L}_t f(x) &= \tfrac{1}{2}\bar{a}^{ij}(t,x) f_{x_i x_j}(x) + \bar{b}^i(t,x) f_{x_i}(x), \end{aligned} \tag{2.3}$$

with the convention that any term involving $p^{-1}(t,x)$ is taken to be zero if $p(t,x)=0$. We have the following result.

THEOREM 2.1. *Assume* (A). *Then* $\{\bar{X}_t\colon 0\le t<1\}$ *is a Markov diffusion process with generator* $\bar{L}_t$.

Before proceeding with the proof, we remark that if in (A)(ii) we replace $t_0>0$ by 0 then $\bar{X}$ is a diffusion up to and including $t=1$.

PROOF. Since $\bar{L}_t$ is a second-order partial differential operator and since it is already known that $\bar{X}_t$ is a Markov process, we need only show that $\bar{L}_t$ is its generator, or that $\{\bar{X}_t\colon 0\le t<1\}$ is a solution of the martingale problem associated with $\bar{L}_t$, i.e., with $(\bar{a},\bar{b})$. If f, g are two arbitrary functions in $C_c^\infty(\mathbb{R}^d)$, i.e., $\mathbb{R}^d\mapsto\mathbb{R}$ infinitely differentiable with compact support, then we need to show that for $1>t>s\ge 0$

$$E\left\{f(\bar{X}_t)-f(\bar{X}_s)-\int_s^t\bar{L}_\theta f(\bar{X}_\theta)\,d\theta\,\middle|\,\bar{X}_r\colon 0\le r\le s\right\}=0$$

or, since $\bar{X}$ is Markovian,

$$E\left\{f(\bar{X}_t)-f(\bar{X}_s)-\int_s^t\bar{L}_\theta f(\bar{X}_\theta)\,d\theta\,\middle|\,\bar{X}_s\right\}=0$$

or again, since g is arbitrary in $C_c^\infty(\mathbb{R}^d)$,

$$E\left\{\left[f(\bar{X}_t)-f(\bar{X}_s)-\int_s^t\bar{L}_\theta f(\bar{X}_\theta)\,d\theta\right]g(\bar{X}_s)\right\}=0$$

or, with the change of variable $1-s\to t$, $1-t\to s$ (so $1\ge t>s>0$),

$$E\left\{\left[f(X_t)-f(X_s)-\int_s^t\tilde{L}_\theta f(X_\theta)\,d\theta\right]g(X_t)\right\}=0 \tag{2.4}$$

if

$$\begin{aligned}\tilde{L}_\theta f(x)&=-\bar{L}_{1-\theta}f(x)\\&=-\tfrac{1}{2}a^{ij}(\theta,x)f_{x_ix_j}(x)\\&\quad+\left\{b^i(\theta,x)-p(\theta,x)^{-1}\left[a^{ij}(\theta,x)p(\theta,x)\right]_{x_j}\right\}f_{x_i}(x).\end{aligned}$$

Recall the convention regarding the case $p(\theta,x)=0$, and observe that

$$E\int_s^t|\bar{L}_\theta f(\bar{X}_\theta)|\,d\theta=E\int_s^t|\tilde{L}_\theta f(X_\theta)|\,d\theta<\infty,$$

so that despite the fact that $\bar{b}$ need not be locally bounded, $f(\bar{X}_t)-\int_s^t\bar{L}_\theta f(\bar{X}_\theta)\,d\theta$ is integrable. Note also that it suffices to establish (2.4) for almost all t, s.

Let us write

$$(f,g)=\int_{\mathbb{R}^d}f(x)\cdot g(x)\,dx,$$

where we do not distinguish between the cases of vector and scalar valued f, g. We continue with this convention throughout.

We shall now establish (2.4). Since the density p is assumed to exist, then local boundedness of b, σ [implied by (A)(i)] and Itô's lemma imply that p satisfies the Kolmogorov forward equation

$$\frac{\partial p}{\partial t} = L_t^* p, \qquad t > 0,$$

in a weak sense, i.e., for $\phi \in C_c^\infty(\mathbb{R}^d)$

$$\frac{d}{dt}(p(t), \phi) = (p(t), L_t \phi), \tag{2.5}$$

where $p(t)$ is the function in $L^2_{\mathrm{loc}}(\mathbb{R}^d)$ with values $p(t,x)$. Here L_t^* is the formal adjoint of L_t. For $0 < s \le t \le 1$ define

$$v(s,x) = E\{g(X_t)|X_s = x\} \equiv E_{sx} g(X_t)$$

so that

$$Ef(X_s)g(X_t) = Ef(X_s)v(s, X_s) = (fp(s), v(s)). \tag{2.6}$$

Formally v satisfies the Kolmogorov backward equation

$$\frac{\partial v}{\partial s} + L_s v = 0, \qquad 0 < s < t, \qquad v(t) = g, \tag{2.7}$$

so that

$$\begin{aligned}(fp(s), v(s)) &= (fp(t), v(t)) - \int_s^t \left(\frac{d}{d\theta} fp(\theta), v(\theta)\right) + \left(fp(\theta), \frac{dv}{d\theta}(\theta)\right) d\theta \\ &= Ef(X_t)g(X_t) - \int_s^t (fL_\theta^* p, v(\theta)) - (fp(\theta), L_\theta v)\, d\theta \\ &= Ef(X_t)g(X_t) - \int_s^t (p(\theta)\tilde{L}_\theta f, v(\theta)) + (L_\theta^*(fp), v(\theta)) \\ &\quad - (fp(\theta), L_\theta v(\theta))\, d\theta \\ &= Ef(X_t)g(X_t) - \int_s^t E\tilde{L}_\theta f(X_\theta) v(\theta, X_\theta)\, d\theta \\ &= E\left\{\left[f(X_t) - \int_s^t \tilde{L}_\theta f(X_\theta)\, d\theta\right] g(X_t)\right\}\end{aligned} \tag{2.8}$$

and (2.4) is established provided (2.8) is justified. We proceed to do this.

Since $f \in C_c^\infty(\mathbb{R}^d)$ is fixed, we can choose $f_1, f_2 \in C_c^\infty(\mathbb{R}^d)$ such that $f(x) = f_1(x)f_2(x)$. We let $H = L^2(\mathbb{R}^d)$, and define the usual Sobolev spaces H^1, H^{-1}, cf. [3], Chapter 2. Observe that if $F \in L^2_{\mathrm{loc}}(\mathbb{R}^d)$ then $Ff_i \in H$, $i = 1,2$. Recall that s, t are fixed. Now we define the following spaces:

$$\begin{aligned}\mathscr{H} &= L^2(s,t;H) = L^2((s,t) \times \mathbb{R}^d), \\ \mathscr{H}^1 &= L^2(s,t;H^1), \\ \tilde{\mathscr{H}}^1 &= \overline{C_c^\infty((s,t) \times \mathbb{R}^d)},\end{aligned}$$

where the overbar denotes closure with respect to the norm

$$\|u\|_{\sim} = \left\{ \int_s^t \int_{\mathbb{R}^d} |u(\theta, x)|^2 + \sum_i \left| \chi(x) \sigma^{ji}(\theta, x) u(\theta, x)_{x_j} \right|^2 dx\, d\theta \right\}^{1/2},$$

where χ is a fixed but arbitrarily chosen function in $C_c^\infty(\mathbb{R}^d)$ such that $\chi(x) = 1$ for x in $\operatorname{supp} f_1 \cup \operatorname{supp} f_2$.

On $\mathscr{H}^1$ we have the norm

$$\|u\|_1 = \left\{ \int_s^t \int_{\mathbb{R}^d} |u(\theta, x)|^2 + \sum_i |u(\theta, x)_{x_i}|^2 dx\, d\theta \right\}^{1/2}.$$

We define $\tilde{\mathscr{H}}^{-1}$ as the dual of $\tilde{\mathscr{H}}^1$ and $\mathscr{H}^{-1}$ as the dual of $\mathscr{H}^1$, so that $\mathscr{H}^{-1} = L^2(s, t; H^{-1})$. Observe that we can identify $\mathscr{H}$ with it dual and hence we have

$$\mathscr{H}^1 \subset \tilde{\mathscr{H}}^1 \subset \mathscr{H} \simeq \mathscr{H}' \subset \tilde{\mathscr{H}}^{-1} \subset \mathscr{H}^{-1} \tag{2.9}$$

with continuous injections. We use here the fact that $\chi\sigma$ is bounded. Let us now define a distribution on $(s, t) \times \mathbb{R}^d$, i.e., a linear functional on $C_c^\infty((s, t) \times \mathbb{R}^d)$, by

$$\begin{aligned} A_{f_1}(p, \phi) = \int_s^t \bigl[-\tfrac{1}{2}(\sigma^*(\theta)\nabla p(\theta), \sigma^*(\theta)\nabla(f_1\phi(\theta))) \\ + (p(\theta), \tilde{b}(\theta) \cdot \nabla\{f_1\phi(\theta)\}) \bigr] d\theta, \end{aligned} \tag{2.10}$$

where $\phi \in C_c^\infty((s, t) \times \mathbb{R}^d)$, ∇ represents gradient with respect to x, and $\phi(\theta)$ is the function $x \to \phi(\theta, x)$ (similarly for σ^*, $\tilde{b}$, etc.). Then

$$|A_{f_1}(p, \phi)| \le K_{f_1} \|\chi p\|_{\sim} \|\phi\|_1, \tag{2.11}$$

where the constant K_{f_1} depends on the essential supremum over $[s, t] \times \operatorname{supp} f_1$ of $|\sigma^*|$, $|b|$, $|a^{ij}_{x_j}|$, $|f_1|$, $|\nabla f_1|$, all of which are finite.

Since (A)(ii) implies that $\|\chi p\|_{\sim} < \infty$, then $A_{f_1}(p, \cdot)$ can be extended to be an element of $\mathscr{H}^{-1}$.

Next we define

$$\begin{aligned} B_{f_2}(v, \phi) = \int_s^t -\tfrac{1}{2}(\sigma^*(\theta)\nabla v(\theta), \sigma^*(\theta)\nabla(f_2\phi(\theta))) \\ + (\tilde{b}(\theta) \cdot \nabla v(\theta), f_2\phi(\theta))\, d\theta \end{aligned} \tag{2.12}$$

so again there exists a constant K_{f_2} such that

$$|B_{f_2}(v, \phi)| \le K_{f_2} \|\chi v\|_1 \|\phi\|_{\sim}. \tag{2.13}$$

According to Lemma 2.1 below $\chi v \in \mathscr{H}^1$ so that $B_{f_2}(v, \cdot)$ can be extended to be in $\tilde{\mathscr{H}}^{-1}$. The same lemma also gives

$$
\begin{aligned}
(fp(t), v(t)) - (fp(s), v(s)) &= A_{f_1}(p, f_2 v) - B_{f_2}(v, f_1 p) \\
&= \int_s^t -\tfrac{1}{2}(\sigma^*\nabla p, v\sigma^*\nabla f) + \tfrac{1}{2}(\sigma^*\nabla v, p\sigma^*\nabla f) \\
&\quad + (\tilde{b}\cdot\nabla f, pv)\,d\theta \\
&= \int_s^t -(\sigma^*\nabla p, v\sigma^*\nabla f) - \tfrac{1}{2}\left(a^{ij}f_{x_i x_j}, pv\right) \\
&\quad + \left(\left(b^i - a^{ij}_{x_j}\right)f_{x_1}, pv\right)d\theta \\
&= \int_s^t (\tilde{L}_\theta f, p(\theta)v(\theta))\,d\theta,
\end{aligned}
\tag{2.14}
$$

where for the last equality we have used Lemma A.2 and for the next to last we used

$$(\sigma^*\nabla f, v\sigma^*\nabla p) + (\sigma^*\nabla f, p\sigma^*\nabla v) = -\left(a^{ij}_{x_j}f_{x_i}, pv\right) - \left(a^{ij}f_{x_i x_j}, pv\right),$$

which formula can be established with the aid of the smooth approximations p_m introduced in the appendix, Lemma A.1. We point out that the convention of taking $p^{-1}[a^{ij}p]_{x_j} = 0$ on the set where p vanishes is completely arbitrary. Whatever convention is used, the $dx\,d\theta$ integral over $A = \{(x, \theta): p(x, \theta) = 0\}$ of $(\tilde{L}fpv)(x, \theta)$ is zero. We have thus established (2.8) and hence the theorem. □

We have made use of the following

LEMMA 2.1. *Assume* (A). *Then*

(a) $\chi v \in \mathscr{H}^1$

(b) $(fp(t), v(t)) - (fp(s), v(s)) = A_{f_1}(p, f_2 v) - B_{f_2}(v, f_1 p)$.

REMARK 2.1. The right side of (b) makes sense because $f_2 v = f_2\chi v \in \mathscr{H}^1$ according to (a); $f_1 p = f_1 \chi p \in \tilde{\mathscr{H}}^1$, so the right side is well defined. The fact that the left side makes sense is established in the proof.

PROOF. We wish to show that $\chi v \in \mathscr{H}^1$, or equivalently that $v \in L^2(s, t; H^1_{\mathrm{loc}})$. We begin by considering the case where b and σ are C^1 in x. Observe that if ξ^i is the solution of

$$
\begin{aligned}
d\xi_\theta &= \nabla b(\theta, X_\theta)\cdot\xi_\theta\,d\theta + \nabla\sigma^j(\theta, X_\theta)\cdot\xi_\theta\,dw_j, \\
\xi_r &= e^i,
\end{aligned}
\tag{2.15}
$$

where σ^j is the jth column of σ and where e^i is the ith column of I, $0 < r \le t$, then the global Lipschitz condition on b, σ implies

$$E_{rx}|\xi^i_t|^2 \le K_0^2, \qquad 0 < r \le t \le 1, \qquad x \in \mathbb{R}^d, \qquad i = 1, \dots, d, \tag{2.16}$$

where the constant K_0 depends only on the Lipschitz constant for b, σ. Now

$$|v_{x_i}(r,x)| = |E_{rx}\{\nabla g(X_t)\cdot \xi_t^i\}| \le \|\nabla g\|_{L^\infty} K_0. \tag{2.17}$$

It follows that $v_{x_i} \in L^\infty((s,t)\times \mathbb{R}^d)$, and hence $\chi v \in \mathscr{H}^1$.

The case when b, σ are not C^1 but only globally Lipschitz (in x) is treated by regularization. We shall at the same time show that v satisfies the Kolmogorov backwards equation (in a weak sense) since this is needed in the proof of (b). Let $\alpha^n(x), \beta^n(t)$ be regularization kernels, i.e., nonnegative C_c^∞ functions whose support converges to $\{0\}$, with L^1 norm equal to 1. Extend b, σ to be zero outside $0 \le t \le 1$. Let $b_n = b * (\alpha^n\beta^n)$, $\sigma_n = \sigma * (\alpha^n\beta^n)$ where $*$ denotes convolution. Then b_n^i, σ_n^{ij} are in C^∞ with respect to t, x and satisfy (A)(i) uniformly in n. Moreover for any $N < \infty$, any i, j

$$\int_0^1 \sup_{|x|\le N}\left\{|b_n^i(r,x) - b^i(r,x)|^2 + |\sigma_n^{ij}(r,x) - \sigma^{ij}(r,x)|^2\right\} dr \to 0.$$

This last claim is established as follows:

$$\begin{aligned}|b_n(r,x) - b(r,x)| &\le \int\int |b(r-\tau, x-\xi) - b(r,x)|\alpha^n(\xi)\beta^n(\tau)\,d\tau\,d\xi \\ &\le I_1 + I_2\end{aligned}$$

with

$$I_1 = \int\int |b(r-\tau, x-\xi) - b(r-\tau,x)|\alpha^n(\xi)\beta^n(\tau)\,d\tau\,d\xi \le K\int |\xi|\alpha^n(\xi)\,d\xi$$

by the Lipschitz continuity of b which is uniform in the time variable, and with

$$\begin{aligned}I_2 &= \int\int |b(r-\tau,x) - b(r,x)|\alpha^n(\xi)\beta^n(\tau)\,d\tau\,d\xi \\ &= \int |b(r-\tau,x) - b(r,x)|\beta^n(\tau)\,d\tau.\end{aligned}$$

The (uniform in t) Lipschitz continuity of b implies that on $\{|x| \le N\}$, $b(\theta, x)$ can be approximated uniformly in θ by $b(\theta, y)$ for some $y \in \{y^1, y^2, \dots, y^M\}$, a fixed finite set depending on the Lipschitz constant and the degree of approximation desired. But for each y^i

$$\int |b(r-\tau, y^i) - b(r, y^i)|\beta^n(\tau)\,d\tau \to 0$$

in $L^2(0,1)$, and also $\int|\xi|\alpha^n(\xi)\,d\xi \to 0$ so that the claim is established for b. σ is treated in the same manner.

Next let $\tilde{w}$ be a standard Brownian motion on $\mathbb{R}^d$ independent of w (we may have to enlarge the underlying probability space) and let X^n be the solution of

$$\begin{aligned}dX_\theta^n &= b_n(\theta, X_\theta^n)\,d\theta + \sigma_n(\theta, X_\theta^n)\,dw_\theta + n^{-1}\,d\tilde{w}_\theta, \\ X_0^n &= X_0.\end{aligned}$$

We set

$$v^n(r,x) = E_{rx}g(X_t^n).$$

According to a slight variation of [9], Chapter 2, Section 7, Theorem 2, for each r, x

$$\lim_{n\to\infty} E_{rx}\{|X_t^n - X_t|^2\} = 0$$

so that $v^n(r,x) \to v(r,x)$. Since we already know that $|v^n(r,x)| \le \|g\|_{L^\infty}$, then $v^n \to v$ in $L^2(s,t; H_{\text{loc}})$. On the other hand, (2.17) implies that ∇v^n lies in a bounded set of $L^2(s,t; L^2(D)^d)$ for any bounded domain D, i.e., ∇v^n lies in a weakly compact set. Hence a subsequence, again denoted by $\{\nabla v^n\}$, converges weakly to a limit in $L^2(s,t; L^2(D)^d)$. Since we already know that $\nabla v^n \to \nabla v$ in the distribution sense, then we can identify the L^2 limit as ∇v. This procedure can be done for a sequence of domains $D_k \uparrow \mathbb{R}^d$, so that a diagonalization argument yields a subsequence v^n such that $v^n_{x_i} \to v_{x_i}$ weakly in $L^2(s,t; H_{\text{loc}})$, and hence $v \in L^2(s,t; H^1_{\text{loc}})$. This establishes (a).

We will now show that $d\bar{v}/dt \in \mathscr{H}^{-1}$ where $\bar{v} = f_2 v$. This amounts to showing that for $\phi \in C_c^\infty(s,t)$, the map

$$\phi \to \int_s^t \phi'(\theta)\bar{v}(\theta)\,d\theta \in H$$

(ϕ' is the derivative of ϕ) can be extended as a bounded linear map of $L^2(s,t) \to H^{-1}$, c.f. [3]. We begin by observing that Itô's lemma and the uniqueness of the classical solution of

$$\frac{dv}{d\theta} + L_\theta^n v = 0, \qquad \theta \le t,$$

$$v(t) = g,$$

where L^n is the generator of X^n (hence a uniformly elliptic operator with smooth coefficients), imply that v^n is this solution. Hence for $\psi \in C_c^\infty(\mathbb{R}^d)$, $\bar{v}^n \equiv f_2 v^n$, and $B^n_{f_2}$ defined by (2.12) with $\sigma^{ij}, \tilde{b}^i$ replaced by σ_n^{ij}, $b_n^i - (\sigma_n\sigma_n^*)^{ij}_{x_j}/2$, we have

$$\begin{aligned}\left(\int_s^t \phi'(\theta)\bar{v}^n(\theta)\,d\theta, \psi\right) &= \int_s^t \phi'(\theta)(v^n(\theta), f_2\psi)\,d\theta \\ &= \int_s^t \phi(\theta)(L_\theta^n v^n, f_2\psi)\,d\theta \\ &= B^n_{f_2}(v^n, \psi\phi) - \frac{1}{2n^2}\int_s^t (\nabla v^n, \nabla(f_2\psi\phi))\,d\theta.\end{aligned}$$

Now (2.13), the strong convergence in $L^2(s,t; H_{\text{loc}})$ of $(\sigma_n\sigma_n^*)^{ij}$ and of $\tilde{b}_n^i$, and the weak convergence of ∇v^n imply

$$\left(\int_s^t \phi'(\theta)\bar{v}(\theta)\,d\theta, \psi\right) = B_{f_2}(v, \psi\phi), \tag{2.18}$$

$$|B_{f_2}(v, \psi\phi)| \le K'_{f_2}\|\chi v\|_1 \|\psi\|_{H^1}\|\phi\|_{L^2},$$

where $\|\psi\|_{H^1} = \{(\psi,\psi) + (\psi_{x_i}, \psi_{x_i})\}^{1/2}$. Hence indeed $\bar{v} \in \mathscr{H}^1$, $d\bar{v}/dt \in \mathscr{H}^{-1}$, or $\bar{v} \in W(s,t)$ in the notation of [3], Chapter 2.

Let us set $\bar{p} = f_1 p$. Then (A)(ii) implies that $\bar{p} \in \tilde{\mathscr{H}}^1$. Next we show that $d\bar{p}/dt \in \mathscr{H}^{-1}$. Using (2.5) and ϕ, ψ as above we find

$$
\begin{aligned}
\left(\int_s^t \phi'(\theta)\bar{p}(\theta)\,d\theta, \psi\right) &= \int_s^t \phi'(\theta)(p(\theta), f_1\psi)\,d\theta \\
&= -\int_s^t \phi(\theta)(p(\theta), L_\theta(f_1\psi))\,d\theta \qquad (2.19)\\
&= -A_{f_1}(p, \psi\phi)
\end{aligned}
$$

and

$$|A_{f_1}(p, \psi\phi)| \le K_{f_1}\|\chi p\|_{\sim}\|\psi\|_{H^1}\|\phi\|_{L^2},$$

so that indeed $d\bar{p}/dt \in \mathscr{H}^{-1}$.

To establish (b) we wish to apply an integration by parts formula which is valid in $W(s,t)$. This requires us to approximate $\bar{p}$ by something in $W(s,t)$. Let $p_m = \bar{\beta}_m * \bar{p}$ be as given by Lemma A.1 with p replaced by $\bar{p}$. Thus $p_m \in L^2(s, t; C^1(\operatorname{supp} f))$ and $p_m \to \bar{p}$ in $\tilde{\mathscr{H}}^1$. Clearly $p_m \in \mathscr{H}^1$, and for $\phi \in C_c^\infty(s,t)$, $\psi \in H^1$,

$$\left(\int_s^t \phi'(\theta)p_m(\theta)\,d\theta, \psi\right) = \int_s^t \phi'(\theta)(\bar{p}(\theta), \psi_m)\,d\theta,$$

where $\psi_m = \bar{\beta}_m * \psi$ so that $\|\psi_m\|_{L^2} \le \|\bar{\beta}_m\|_{L^1}\|\psi\|_{L^2} = \|\psi\|_{L^2}$. But

$$
\begin{aligned}
\left|\left(\int_s^t \phi'(\theta)\bar{p}(\theta)\,d\theta, \psi_m\right)\right| &\le \left\|\int_s^t \phi'(\theta)\bar{p}(\theta)\,d\theta\right\|_{H^{-1}}\|\psi_m\|_{H^1} \\
&\le \|\phi\|_{L^2}\left\|\frac{d\bar{p}}{dt}\right\|_{\mathscr{H}^{-1}}\|\psi_m\|_{H^1}
\end{aligned}
$$

since $\phi \to -\int_s^t \phi'(\theta)\bar{p}(\theta)\,d\theta$ is a bounded linear map of $L^2(s,t) \to H^{-1}$ because $d\bar{p}/dt \in \mathscr{H}^{-1}$. It follows that $dp_m/dt \in \mathscr{H}^{-1}$ and $\|dp_m/dt\|_{\mathscr{H}^{-1}} \le \|d\bar{p}/dt\|_{\mathscr{H}^{-1}}$. Thus $p_m \in W(s,t)$. We proceed to show that $dp_m/dt \to d\bar{p}/dt$ weak $*$ in $\mathscr{H}^{-1}$.

$$\left|\left(\int_s^t\left(\frac{dp_m}{dt} - \frac{d\bar{p}}{dt}\right)\phi\,d\theta, \psi\right)\right| \le \|\phi\|_{L^2}\left\|\frac{dp}{dt}\right\|_{\mathscr{H}^{-1}}\|\psi_m - \psi\|_{H^1}. \qquad (2.20)$$

But $\psi \in H^1$ so $\psi_m = \bar{\beta}_m * \psi \to \psi$ in H^1. Finally since any $\nu \in \mathscr{H}^1$ can be written as the limit in $\mathscr{H}^1$ of ν_n of the form $\nu_n = \sum_1^N 1_{A_i}(t)\psi^i$ with $\psi^i \in H^1$, and since $\|dp_m/dt - dp/dt\|_{\mathscr{H}^{-1}} \le 2\|dp/dt\|_{\mathscr{H}^{-1}} < \infty$, then (2.20) suffices to establish the claim.

Now

$$
\begin{aligned}
&(p_m(t), \bar{v}(t)) - (p_m(s), \bar{v}(s)) \\
&\quad = \int_s^t \left(\frac{dp_m}{dt}(\theta), \bar{v}(\theta)\right) + \left(p_m(\theta), \frac{d\bar{v}}{dt}(\theta)\right)d\theta \qquad (2.21)
\end{aligned}
$$

by the product rule in $W(s,t)$. Since (see Lemma A.1) $p_m(\theta) \to \bar{p}(\theta)$ in $L^1(\mathbb{R}^d)$ for all θ, and since $\bar{v}(\theta) \in L^\infty(\mathbb{R}^d)$ for all θ, then the left side of (2.21) converges, as $m \to \infty$, to

$$(\bar{p}(t), \bar{v}(t)) - (\bar{p}(s), \bar{v}(s)).$$

On the other side, the weak $*$ convergence of dp_m/dt gives

$$\int_s^t \left(\frac{dp_m}{dt}, \bar{v} \right) d\theta \to \int_s^t \left(\frac{d\bar{p}}{dt}, \bar{v} \right) d\theta$$

and

$$\int_s^t \left(p_m, \frac{d\bar{v}}{dt} \right) d\theta \to \int_s^t \left(\bar{p}, \frac{d\bar{v}}{dt} \right) d\theta \tag{2.22}$$

because on the left side of (2.22) the duality $\mathscr{H}^1, \mathscr{H}^{-1}$ can be replaced by the duality $\tilde{\mathscr{H}}^1, \tilde{\mathscr{H}}^{-1}$, c.f. (2.9), since by (2.18) $d\bar{v}/dt = -B_{f_2}(v, \cdot) \in \tilde{\mathscr{H}}^{-1}$. But (2.19) implies that $d\bar{p}/dt = A_{f_1}(p, \cdot)$, so that (b) now follows by passing to the limit in (2.21).

This completes the proof of Lemma 2.1 and hence of the theorem. □

We conclude this section with some remarks.

REMARK 2.2. Condition (A)(i) was given in global form because it seems to be most useful as such, cf. Section 3, but it can be relaxed slightly to local form. We can replace (2.2) by:

(2.23) b, σ are locally bounded, and are locally Lipschitz in x uniformly in t.

However, this does not guarantee that the process $\{X_t\}$ does not explode so we must add

$$X_t \text{ does not explode on } [0, 1]. \tag{2.24}$$

In addition, we no longer have the bound (2.16) which can, however, be established as follows, at least locally in x, which suffices.

$$E|\xi_\theta|^2 \le k \left\{ 1 + \left[\int_r^\theta \sum_i E|\nabla b^i|^2 \, d\rho + \int_r^\theta \sum_{ij} |\nabla \sigma^{ij}|^2 \, d\rho \right] \int_r^\theta E|\xi_\rho|^2 \, d\rho \right\}$$

since

$$E\left| \int \nabla \sigma^{ij} \cdot \xi \, dw^j \right|^2 \le l \sum_j E \int |\nabla \sigma^{ij} \cdot \xi|^2 \, dt.$$

By Gronwall's lemma $E_{rx}|\xi_\theta|^2 \le K_0^2$ provided for each compact set B there exists a constant c such that

$$\int_r^t E_{rx} \left\{ \sum_i \left| \nabla b^i(\rho, X_\rho) \right|^2 + \sum_{ij} \left| \nabla \sigma^{ij}(\rho, X_\rho) \right|^2 \right\} d\rho \le c < \infty, \tag{2.25}$$

where c is independent of i and of $r \in [s, t]$, $x \in B$. Thus (2.2) can be replaced by (2.23), (2.24), (2.25).

There is one other point in the proof of Lemma 2.1 which must be altered. Since b, σ hence b_n, σ_n may not satisfy a linear bound then X^n may not exist.

Instead first alter b, σ outside $\{|X| \le m\}$ to be bounded (call them b^m, σ^m), and proceed with the proof. This can be done so that (2.25) holds uniformly in m. The nonexplosion of X implies that $v^m(\theta, x) \to v(\theta, x)$ pointwise, hence in $\mathscr{H}$, if v^m is defined in the obvious manner. Since (2.25) holds uniformly in m we have weak compactness of ∇v^m and so for a subsequence we can go to the limit to obtain again Lemma 2.1.

REMARK 2.3. We observe that we had a tradeoff in the foregoing derivation: $\bar{p} \in \tilde{\mathscr{H}}^1$, $\bar{v} \in \mathscr{H}^1$. This seems to be the natural way to proceed because we can give simple conditions to imply $\bar{p} \in \tilde{\mathscr{H}}^1$ (or we might say $p \in \tilde{\mathscr{H}}^1_{\mathrm{loc}}$), cf. Section 3, but if we know by some other means that the stronger condition $\bar{p} \in \mathscr{H}^1$, holds, then we can work with the case where $\bar{v} \in \tilde{\mathscr{H}}^1$ only. The definitions of A_{f_1}, B_{f_2} have to be changed slightly so as to eliminate $b \cdot \nabla v$ in B_{f_2} by introducing $\operatorname{div}(\tilde{b}p)$ in A_{f_1}. The advantage here is that we can prove by p.d.e. methods that $\bar{v} \in \tilde{\mathscr{H}}^1$ rather than introduce the ξ^i's; hence, the global Lipschitz condition on b, σ is avoided. This is the result announced in [11]. The precise hypotheses are:

(A′) b, σ are Borel measurable; (2.23) holds; (2.24) holds; the distributional derivatives $\sigma^{ij}_{x_k x_l}$ exist as locally bounded functions on $[0,1] \times \mathbb{R}^d$ for all i, j, k, l; and for almost all $t > 0$, $X(t)$ has a density $p(t)$ such that for all $t_0 > 0$

$$p \in L^2\big(t_0, 1; H^1_{\mathrm{loc}}\big).$$

REMARK 2.4. We have now found that for $\overline{X}$ the drift is $\bar{b}$ and the diffusion is $\bar{\sigma}$, but since $\bar{b}$ need not be locally bounded we should check that $\int_0^t \bar{b}(s, \overline{X}_s)\, ds$ makes sense, i.e., that $\bar{b}(s, \overline{X}_s) \in L^1(0, t)$ for any $t < 1$. This point is related to the integrability of $\overline{L}f$ established after (2.4). Let $1_n(x)$ be the characteristic function of $\{|x| \le n\}$ and let $\tau_n = \inf\{s \ge 0: |X_s| \ge n\}$. The nonexplosion of $\{X_s\}$ implies that $\tau_n \uparrow 1$ w.p.1. Now with $t_0 = 1 - t$, $a\nabla p \in L^2_{\mathrm{loc}}([t_0, 1] \times \mathbb{R}^d)$, cf. (A)(i), implies that

$$E\int_{t_0}^1 \big| p(s, X_s)^{-1}\big[a^{ij}(s, X_s)p(s, X_s)\big]_{x_j}\big| 1_n(X_s)\, ds < \infty,$$

so that

$$\int_{\min\{t_0, \tau_n\}}^{\min\{1, \tau_n\}} |\bar{b}(1 - s, X_s)|\, ds < \infty \quad \text{w.p.1},$$

or after $n \to \infty$,

$$\int_0^t |\bar{b}(s, \overline{X}_s)|\, ds < \infty \quad \text{w.p.1}.$$

REMARK 2.5. To identify $\bar{w}$, we set $\hat{w}_t = \bar{w}_{1-t}$. Rewriting the equations for X and $\overline{X}$ in Stratonovich form it follows as in [16], Section 3, if σ is continuous in t, that

$$\sigma(t, X_t) \circ d\hat{w}_t = \sigma(t, X_t) \circ dw_t + p(t, X_t)^{-1}\sigma(t, X_t)\big[p(t, X_t)\sigma^{*(j)}(t, X_t)\big]_{x_j}\, dt,$$

where $\circ$ denotes the Stratonovich integral and $\sigma^{*(j)}$ denotes the jth column of σ^*. If for each (t, x), $\sigma(t, x)$ is 1–1, i.e., rank$\sigma(t, x) = l \le d$, then

$$\hat{w}_t^i = w_t^i - w_1^i - \int_t^1 p(s, X_s)^{-1}\left[p(s, X_s)\sigma^{ji}(s, X_s)\right]_{x_j} ds.$$

Using a different approach one can show that this equality holds even when σ is not 1–1, but (C) of Theorem 3.1 below holds, cf. [17].

3. Locally integrable densities. We will now give explicit conditions which guarantee the implicit condition (A)(ii). First observe that smoothness of p implies local boundedness, so that regardless of the law of X_0, if

(B) $\dfrac{\partial}{\partial t} + L_t$ is hypoelliptic

then (A)(ii) holds. The conditions of Hörmander's theorem, which imply (B), can be found in [12]. We point out that these conditions include the assumptions that b and σ be C^∞ in (t, x). A version of Hörmander's theorem which does not require smoothness in t can be found in [5].

Let us now turn to cases where b, σ have much less regularity, but where an initial density p_0 satisfying some growth conditions is assumed.

THEOREM 3.1. *Assume* (A)(i) *and*

(C)(i) *the law of* X_0 *has a density* p_0 *such that for some* $\lambda < 0$,

$$p_0 \in L^2\left(\mathbb{R}^d, (1 + |x|^2)^\lambda\, dx\right),$$

(C)(ii) *either*

(a) *there exists* $\alpha > 0$ *such that* $a(t, x) \ge \alpha I$, *or*

(b) $a^{ij}_{x_i x_j} \in L^\infty((0,1) \times \mathbb{R}^d)$.

Then (A)(ii) *holds.*

PROOF. We follow Menaldi [14] in introducing the following Sobolev spaces with weights. Let

$$\beta_0(x) = (1 + |x|^2)^{\lambda/2}, \qquad \beta_1(x) = (1 + |x|^2)^{(\lambda+1)/2}, \qquad \gamma(x) = x(1 + |x|^2)^{-1/2},$$

$$\hat{\sigma}(t, x) = \sigma(t, x)(1 + |x|^2)^{-1/2}, \qquad \hat{b}(t, x) = \tilde{b}(t, x)(1 + |x|^2)^{-1/2},$$

$$\hat{H} = \left\{v \colon \beta_0 v \in L^2(\mathbb{R}^d)\right\} = L^2\left(\mathbb{R}^d, (1 + |x|^2)^\lambda\, dx\right).$$

Observe that $\hat{\sigma}$, $\hat{a} \equiv \hat{\sigma}\hat{\sigma}^*$, $\hat{b}$ are bounded due to (A)(i). On $\hat{H}$ introduce the inner product

$$(u, v)_0 = (\beta_0 u, \beta_0 v)$$

with corresponding norm $|u|_0$. We also define $\hat{\mathscr{H}} = L^2(0,1; \hat{H})$, $\hat{\mathscr{H}}^1 = L^2(0,1; \hat{H}^1)$, where $v \in \hat{H}^1$ if $\|v\|_\wedge = \int_0^1 |v|_0^2 + (\beta_1 v_{x_i}, \beta_1 v_{x_i})\, dt^{1/2} < \infty$, and finally we let $\hat{\hat{\mathscr{H}}}^1$ be the closure of $C_c^\infty((0,1) \times \mathbb{R}^d)$ under

$$\|u\|_\sim = \left[\int_0^1 |u|_0^2 + (\beta_1 \hat{\sigma}^* \nabla u, \beta_1 \hat{\sigma}^* \nabla u)\, dt\right]^{1/2}.$$

Let $\hat{\mathscr{H}}^{-1}$, $\hat{\hat{\mathscr{H}}}^{-1}$ be the dual spaces. Then

$$\hat{\mathscr{H}}^1 \subset \hat{\hat{\mathscr{H}}}^1 \subset \hat{\mathscr{H}} \simeq \hat{\mathscr{H}}' \subset \hat{\hat{\mathscr{H}}}^{-1} \subset \hat{\mathscr{H}}^{-1}.$$

Let us define

$$\hat{A}(t,u,v) = \int_0^t \tfrac{1}{2}(\beta_1\hat{\sigma}^*\nabla u, \beta_1\hat{\sigma}^*\nabla v) - (\beta_0\hat{b}u, \beta_1\nabla v)$$
$$+\lambda(\beta_1\hat{\sigma}^*\nabla u, \beta_0\hat{\sigma}^*\gamma v) + 2\lambda(\hat{b}u, \gamma v)_0\, d\theta$$

so that

$$|\hat{A}(t,u,v)| \le k_0\|u\|_{\sim}\|v\|_{\wedge}.$$

Moreover for $v \in C_c^\infty((0,1)\times\mathbb{R}^d)$ we have

$$\int_0^t (-L_s^* u, v)_0\, ds = \hat{A}(t,u,v).$$

Hence $\hat{A}(t,u,\cdot)$ is an extension of $\int_0^t(-L_s^*u,\cdot)_0\, ds$ to $\hat{\mathscr{H}}^{-1}$.

Let us consider the case (a). Now $\hat{\hat{\mathscr{H}}}^1 = \hat{\mathscr{H}}^1$ and the norms are equivalent, so we work only on $\hat{\mathscr{H}}^1$. Choose b_n, σ_n such that $b_n = b, \sigma_n = \sigma$ on $\{|x| \le n\}$, $b_n, \sigma_n, (\sigma_n)_x$ are in $L^\infty([0,1]\times\mathbb{R}^d)$, $a_n = \sigma_n\sigma_n^* \ge \alpha I$, and $\hat{b}_n, \hat{\sigma}_n$ bounded uniformly in n (since $\hat{b}$ and $\hat{\sigma}$ are). Let p_n be the unique solution in $L^2(0,1;H^1)$ of

$$\frac{dp}{dt} = {}_nL_t^*p, \qquad p(0) = q_n,$$

where $q_n \in L^2(\mathbb{R}^d)$, $q_n \to p_0$ in $\hat{H}$ and $q_n = p_0$ on $\{|x| \le n\}$. N.b. ${}_nL_t$ is the generator corresponding to b_n, σ_n. Since $\beta_0, \beta_0\beta_1^{-1}, \nabla\beta_0 \in L^\infty$ then $\beta_0 p_n \in L^2(0,1;H^1)$, $d(\beta_0 p_n)/dt \in L^2(0,1;H^{-1})$ so that

$$\tfrac{1}{2}|p_n(t)|_0^2 - \tfrac{1}{2}|q_n|_0^2 = \int_0^t (\beta_{0n}L_s^*p_n(s), \beta_0 p_n(s))\, ds$$
$$= -\hat{A}_n(t, p_n, p_n)$$
$$\le \int_0^t \left[-\frac{\alpha}{4}(\beta_1\nabla p_n, \beta_1\nabla p_n) + k_0|p_n|_0^2\right] ds$$
$$\le k_0\int_0^t |p_n(s)|_0^2\, ds.$$

Then

$$|p_n(t)|_0^2 \le K|q_n|_0^2 \le K_0,$$
$$\int_0^1 |\beta_1\nabla p_n(t)|^2\, dt \le \hat{K}|q_n|_0^2 \le \hat{K}_0,$$

so that for a subsequence $p_n \to p$ weakly in $L^2(0,1;\hat{H}^1)$ for some p.

It remains only to show that $p(t)$ is the requisite density. Let $\phi \in C_c^\infty(\mathbb{R}^d)$ and let X^n satisfy

$$dX_t = b_n(t,X_t)\,dt + \sigma_n(t,X_t)\,dw, \qquad X_0 \sim q_n\,dx.$$

Then by regularizing b_n, σ_n it follows that

$$(p_n(t), \phi) = E\phi(X_t^n).$$

But we can now let $n \to \infty$ in this last equation to observe that for almost all t,

$p(t)$ is the density of X_t. Thus X_t has a density $p(t)$ which lies in $L^2(0,1;\hat{H}^1) \subset L^2(0,1;H^1_{\mathrm{loc}})$.

Now consider the case (b). With $a_n = \sigma_n\sigma_n^* + 1/nI$, we obtain

$$\begin{aligned}|p_n(t)|_0^2 - |q_n|_0^2 &= -\hat{A}_n(t, p_n, p_n)\\ &\le -\tfrac{1}{4}\int_0^t\int|\beta_1\hat{\sigma}^*\nabla p_n|^2\,dx\,ds + k\int_0^t|p_n|_0^2\,ds,\end{aligned}$$

where we have used the fact that

$$\begin{aligned}(\hat{b}\beta_1, \beta_0 p\nabla p) &= \left(\hat{b}\sqrt{1+|x|^2}, \beta_0^2 p\nabla p\right)\\ &= \left(\tilde{b}, \tfrac{1}{2}\nabla(\beta_0 p)^2 - \beta_0\nabla\beta_0 p^2\right)\\ &= \left(\tilde{b}, \tfrac{1}{2}\nabla(\beta_0 p)^2 - (\beta_0 p)^2\gamma\lambda(1+|x|^2)^{-1/2}\right)\\ &= -\tfrac{1}{2}\left(\nabla\cdot\tilde{b}, (\beta_0 p)^2\right) - \lambda\left(\hat{b}\cdot\gamma, (\beta_0 p)^2\right)\end{aligned}$$

and by hypothesis $\operatorname{div}\tilde{b} \in L^\infty((0,1)\times\mathbb{R}^d)$. Now the result follows as in the case (a). □

APPENDIX

We give here two lemmata required above. The first concerns the approximation of p, or more precisely of p multiplied by a $C_c^\infty(\mathbb{R}^d)$ function. Let us then in fact take p to have support in the open set $D^0 \subset D$ with D compact.

LEMMA A.1. *Assume σ is Borel measurable, locally Lipschitz with respect to x uniformly in t, and locally bounded. If $p \in \tilde{\mathscr{H}}^1$ and $\operatorname{supp} p \subset D^0 \subset D$ compact then there exists a sequence $\{p_m\} \subset L^2(s,t;C^1(D))$ such that $p_m \to p$ in $\tilde{H}^1$ and $p_m(\theta) \to p(\theta)$ in $L^1(\mathbb{R}^d)$ if $p(\theta) \in L^1(\mathbb{R})$.*

PROOF. Let $\gamma \in C^1(\mathbb{R})$, $\int\gamma(x)\,dx = 1$, $\operatorname{supp}\gamma \in [-1,1]$, $\gamma(x) \ge 0$. Let

$$\beta_m(y) = \prod_{i=1}^d m\gamma(my_i), \qquad y \in \mathbb{R}^d,$$
$$\tilde{p}_m = \beta_m * p.$$

Then

$$\nabla\tilde{p}_m(t,x) = \nabla\beta_m * p$$

and $\tilde{p}_m \to p$ in $L^2(s,t;H)$ since p lies in this space.

We shall require a weak compactness argument so we begin by showing that $\{\sigma^*\nabla\tilde{p}_m\}$ is bounded in $L^2(s,t;L^2(D))$. In fact

$$\begin{aligned}\sigma^*(t,x)\nabla\tilde{p}_m(t,x) &= \int\sigma^*(t,x)\nabla\beta_m(x-y)p(t,y)\,dy\\ &= \int\sigma^*(t,y)\nabla\beta_m(x-y)p(t,y)\,dy + R_m(t,x)\\ &= \beta_m * \nabla(\sigma^* p)(t,x) + R_m(t,x).\end{aligned}$$

But

$$\|\beta_m * \nabla(\sigma^* p)\|_{L^2(s,t;H)} \le \|\beta_m\|_{L^1(\mathbb{R}^d)} \|\nabla(\sigma^* p)\|_{L^2(s,t;H)} < \infty$$

since $\nabla(\sigma^* p) = \sigma^* \nabla p + (\nabla \cdot \sigma^*) p \in L^2(s, t; L^2(D))$ for each component. Note that in the definition of $\|\cdot\|_{\sim}$ we take $\chi(x) = 1$ on D^0, and that we have used the local Lipschitz property of σ^*. As for R_m

$$(1) \qquad |R_m(t,x)| \le K \int |x-y| |\nabla \beta_m(x-y)| p(y)\, dy.$$

If we set $\alpha_m(x) = |x|\,|\nabla\beta_m(x)|$, $\tilde{\alpha}_m(x) = \alpha_m(x)/\|\alpha_m\|_{L^1(\mathbb{R}^d)}$, then

$$0 \le \alpha_m(x) \le |x| \sqrt{d}\, 2m^{d+1}$$

and

$$\operatorname{supp} \alpha_m = \{x: |x_i| \le m^{-1}\ \forall\, i\} \subset \{x: |x| \le \sqrt{d}/m\}$$

so that

$$(2) \qquad \|\alpha_m\|_{L^1(\mathbb{R}^d)} \le \sqrt{d}\, 2m^{d+1} K_d \int_0^{\sqrt{d}/m} r^d\, dr \equiv \tilde{K}_d.$$

It follows that $\tilde{\alpha}_m$ is again a regularization kernel and by (1)

$$(3) \qquad \frac{|R_m(t,x)|}{\|\alpha_m\|_{L^1(\mathbb{R}^d)}} \le K \tilde{\alpha}_m * p(x) \to K p(x)$$

in $L^2(s, t; H)$. (2) and (3) imply that $\{R_m\}$ is bounded in $L^2(s, t; D)$, and hence that $\{\sigma^* \nabla \tilde{p}_m\}$ is bounded.

Since $\sigma^* \nabla \tilde{p}_m \to \sigma^* \nabla p$ in the sense of distribution, then the weak compactness (i.e., boundedness) implies that the convergence is weak in $L^2(s, t; H)$. Since this space is reflexive then for each m there exist constants $\delta_i^m \ge 0$, $i = 1, \ldots, I_m < \infty$, such that

$$\sum_{i=m}^{I_m} \delta_i^m = 1,$$

$$\sum_{i=m}^{I_m} \delta_i^m \sigma^* \nabla \tilde{p}_i \to \sigma^* \nabla p$$

strongly in $L^2(s, t; H)$, cf. [6], page 439, Section 43. If

$$p_m = \sum_{i=m}^{I_m} \delta_i^m \tilde{p}_i$$

then $\sigma^* \nabla p_m \to \sigma^* \nabla p$ strongly and

$$\begin{aligned} \|p_m - p\|_{L^2(s,t;H)} &\le \sum_{i=m}^{I_m} \delta_i^m \|\tilde{p}_i - p\|_{L^2(s,t;H)} \\ &\le \sup_{i \ge m} \|\tilde{p}_i - p\|_{L^2(s,t;H)} \\ &\to 0. \end{aligned}$$

Thus $p_m \to p$ in $\tilde{\mathscr{H}}^1$. Clearly $p_m \in L^2(s, t; C^1(D))$.

It remains to show that $p_m(\theta) \to p(\theta)$ in $L^1(\mathbb{R}^d)$. But if $p(\theta, \cdot)$ is in $L^1(\mathbb{R}^d)$, then $\bar{\beta}_m * p(\theta) \to p(\theta)$ in $L^1(\mathbb{R}^d)$, where $\bar{\beta}_m = \sum_{i=m}^{I_m} \delta_i^m \beta_i$. □

Let us now turn to the second result. We wish to show that $\sigma^* \nabla p = 0$ a.e. on the set $A = \{(t, x): p(t, x) = 0\}$. This follows because on A, p is a minimum so that any (directional) derivative which may exist must be zero a.e. But the columns of $\sigma^* \nabla p$ are directional derivatives, and they exist since $p \in \tilde{\mathscr{H}}^1$—except we do not know that these derivatives in the distributional sense are necessarily derivatives in the absolutely continuous sense. We show this in the next lemma. To minimize the regularity assumptions we use a certain localization.

LEMMA A.2. *Assume σ is as in Lemma* A.1 *and assume* (A)(ii). *Then* $\sigma^*(t, x) \nabla p(t, x) = 0$ *a.e. on* $A = \{(t, x): p(t, x) = 0\}$.

PROOF. We need only give the proof for the case when σ consists of only one column, and then it suffices to show that for each fixed $t \in (0, 1]$, each compact set $K \subset \mathbb{R}^d$, and each n, $\sigma(t, x) \cdot \nabla p(t, x) = 0$ a.e. (x) on

$$K_n^t = K \cap \{x: |\sigma(x)| > n^{-1},\ p(t, x) = 0\}.$$

Let us suppress the argument t.

We let $\xi_s(x)$ be the unique solution of

$$\frac{d\xi_s}{ds}(x) = \sigma(\xi_s(x)), \qquad \xi_0(x) = x, \tag{4}$$

which exists locally by the Lipschitz continuity. From the compactness of K_n and the fact that $|\sigma(x)| \ge n^{-1}$ on K_n, it follows that there exists a finite covering of K_n by domains D_n^i with the properties that for each D_n^i

(i) there exists a domain $\tilde{D}_n^i \supset D_n^i$ such that the distance between D_n^i and the complement of D_n^i is positive,

(ii) there exists a unit vector v and a constant $\alpha > 0$ such that

$$v \cdot \sigma(x) \ge \alpha, \qquad \forall\, x \in \tilde{D}_n^i,$$

(iii) there exists a hyperplane Δ orthogonal to v such that

$$\tilde{D}_n^i = \tilde{D}_n^i \cap \left\{ \bigcup_{x \in \tilde{D}_n^i \cap \Delta} \ \bigcup_{s \in R} \xi_s(x) \right\}.$$

It now suffices to prove that $\sigma \cdot \nabla p = 0$ a.e. on $A_D = \{x \in D: p(x) = 0\}$ with $D = D_n^i$ for any i, n. By (ii) and the local boundedness of σ we can renormalize $\sigma(x)$ such that $v \cdot \sigma(x) = 1$ on $\tilde{D} \equiv \tilde{D}_n^i$. This does not change the curve ξ, only its parameterization, so that (iii) is preserved under the renormalization (as are the properties of σ).

On the Borel sets of $\tilde{D}$, we define a measure

$$\mu(B) = \int_{\Delta'} dy \int_{-\infty}^{\infty} 1_B(\xi_s(0, y))\, ds$$

after changing variables so that $\Delta = \{0\} \times \Delta'$, $\Delta' \subset \mathbb{R}^{d-1}$. We shall show that μ is equivalent to Lebesgue measure. If Φ_s is the flow of the differential equation in

(4), then for any $B' \subset \Delta'$ define

$$\bar{\mu}_s(B') = \int_{\pi\Phi_s^{-1}(\{s\}\times B')} dy,$$

where π is the projection of Δ on Δ'. The Lipschitz continuity of σ and the boundedness of $\tilde{D}$ imply that there exists k such that

$$0 < k^{-1} \leq \left| \frac{\partial}{\partial y} \xi_s(0, y) \right| \leq k$$

for all s, y such that $\xi_s(0, y) \in \tilde{D}$; hence $\bar{\mu}_s$ is equivalent to Lebesgue measure, uniformly in such s.

Since

$$\mu(B) = \int_a^b \bar{\mu}_s(B')\, ds$$

for B of the form $[a, b] \times B'$, then the equivalence of μ and Lebesgue measure follows.

Let $\psi \in C_c^\infty(\mathbb{R}^d)$, $\psi(x) = 1$ on D, $\psi(x) = 0$ off $\tilde{D}$. Replacing p by ψp we obtain a function (again called p) of compact support equal to the original p on D. We shall now establish

$$p(\xi_s(0, y)) = \int_{-\infty}^{s} (\sigma \cdot \nabla p)(\xi_\theta(0, y))\, d\theta \tag{5}$$

a.e. (s, y). Since $\sigma \cdot \nabla p$ is in $L^2(\tilde{D}, \mu)$, the right side of (5) is well defined a.e. y. Since (5) is true for p_m as given by Lemma A.1 then the result follows on passing to the limit.

Finally a.e. y, $s \to p(\xi_s(0, y))$ is absolutely continuous, hence a.e. differentiable, and

$$\frac{d}{ds} p(\xi_s(0, y)) = (\sigma \cdot \nabla p)(\xi_s(0, y)), \qquad (s, y)\mu \quad \text{a.e.} \tag{6}$$

But on A_D, p is a minimum, so that the left side of (6) is zero a.e. The conclusion follows by (6) and the equivalence of μ and Lebesgue measure. □

REFERENCES

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

[2] AZEMA, J. (1973). Théorie générale des processus et retournement du temps. *Ann. Sci. École Norm. Sup.* (4) **6** 459–519.

[3] BENSOUSSAN, A. and LIONS, J. L. (1982). *Applications of Variational Inequalities in Stochastic Control.* North Holland, Amsterdam.

[4] CASTANON, D. A. (1982). Reverse time diffusion processes. *IEEE Trans. Inform. Theor.* **IT-L8** 953–956.

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

[6] DUNFORD, N. and SCHWARTZ, J. T. (1966). *Linear Operators* **1**. Wiley, New York.

[7] ELLIOTT, R. J. and ANDERSON, B. D. O. (1985). Reverse time diffusions. *Stochastic Process. Appl.* **19** 327–339.

[8] FÖLLMER, H. (1985). An entropy approach to the time reversal of diffusion processes. *Lecture Notes in Control Inform. Sci.* **69** 156–163. Springer, New York.

[9] GIHMAN, I. I. and SKOROHOD, A. V. (1972). *Stochastic Differential Equations*. Springer, New York.

[10] HAUSSMANN, U. G. (1985). On the drift of a reversed diffusion. *Lecture Notes in Control Inform. Sci.* **69** 164–175. Springer, New York.

[11] HAUSSMANN, U. G. and PARDOUX, E. (1985). Time reversal of diffusion processes. *Lecture Notes in Control Inform. Sci.* **69** 176–182. Springer, New York.

[12] HÖRMANDER, L. (1967). Hypoelliptic second order differential equations. *Acta Math.* **117** 147–171.

[13] LINDQUIST, A. and PICCI, G. (1979). On the stochastic realization problem. *SIAM J. Control Optim.* **17** 365–389.

[14] MENALDI, J. L. (1980). On the optimal impulse control problem for degenerate diffusions. *SIAM J. Control Optim.* **18** 722–739.

[15] NELSON, E. (1967). *Dynamical Theories of Brownian Motion*. Princeton Univ. Press.

[16] PARDOUX, E. (1983). Smoothing of a diffusion process conditioned at final time. *Lecture Notes in Control Inform. Sci.* **43** 187–196. Springer, New York.

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

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

[19] ZHENG, W. A. (1985). Tightness results for laws of diffusion processes, application to stochastic mechanics. Preprint. Univ. Louis Pasteur, Strasbourg.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
121-1984 MATHEMATICS ROAD
VANCOUVER, BRITISH COLUMBIA V6T 1Y4
CANADA

UNIVERSITÉ DE PROVENCE
U.E.R. DE MATHÉMATIQUES
3, PLACE VICTOR-HUGO
13331 MARSEILLE, CEDEX 3
FRANCE