SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

JAMES R. NORRIS

Simplified Malliavin calculus

Séminaire de probabilités (Strasbourg), tome 20 (1986), p. 101-130
<http://www.numdam.org/item?id=SPS_1986__20__101_0>

SIMPLIFIED MALLIAVIN CALCULUS

by James Norris

We aim to show, as economically as possible, using the Malliavin Calculus that the solution x_t of a certain stochastic differential equation:

$$dx_t = X_o(x_t)dt + X_i(x_t)\partial w_t^i$$
$$x_o = x \in \mathbb{R}^d$$

has a smooth probability density function on $\mathbb{R}^d$, whenever the following hypothesis is satisfied at the starting point x :

$$H_1 : \quad X_1,\ldots,X_m; \quad [X_i,X_j]_{i,j=0}^m; \quad [X_i[X_j,X_k]]_{i,j,k=0}^m; \ \ldots \text{ etc.,}$$

evaluated at x, span $\mathbb{R}^d$.

We assume above that:

- $X_0,X_1,\ldots,X_m$ are C^∞ vector fields on $\mathbb{R}^d$ satisfying certain boundedness conditions,

- $w_t \equiv (w_t^1,\ldots,w_t^m)$ is an $(\mathcal{F}_t, \mathbb{P})$-Brownian motion on $\mathbb{R}^m$.

We use ∂w_t to denote the Stratonovich differential, the symbol dw_t being reserved for the Itô differential. We sum the index i from 1 to m whenever it is repeated. Of course

$$[X_i,X_j] \equiv DX_j\cdot X_i - DX_i\cdot X_j .$$

Programmes for establishing this result have been given by Malliavin, Stroock [12], [13], Bismut [3] and others, though only Stroock [13] has obtained the full result. All are agreed that the proof splits naturally into two parts: namely, for a certain $d\times d$ random matrix C_t, associated with x_t, known as the Malliavin Covariance Matrix,

$$C_t^{-1} \in L^p(\mathbb{P}) \text{ for all } p<\infty \Rightarrow x_t \text{ has } C^\infty \text{ density}$$

and

$$H_1 \text{ holds at } x \Rightarrow C_t^{-1} \in L^p(\mathbb{P}) \text{ for all } t>0 \text{ and } p<\infty .$$

Our proof of the first implication, given in Sections 1 to 3, uses Bismut's approach, which seems the most efficient in this context. We have made some simplifications of Bismut's work and been more explicit in iterating the integration by parts formula. Simplified versions of Bismut [3] have also been given by Bichteler and Fonken [1] and Fonken [6].

Our proof of the second implication, given in Section 4, follows for the most part Meyer's [10] presentation of Stroock's [13] argument. But by the application of a new semimartingale inequality (Lemma 4.1) we are able to shorten the argument considerably.

Before we start on the probabilistic arguments we give a well known result from Fourier analysis which explains how we set about obtaining smooth density.

Theorem 0.1

Let X be an $\mathbb{R}^d$-valued random variable with law μ. Let $n \geq d+1$. Suppose there exists a constant $C_n < \infty$ such that for all multi-indices α with $|\alpha| \leq n$,

$$\mathbb{E}[D^\alpha f(X)] \leq C_n \|f\|_\infty, \quad \text{for all} \quad f \in C_b^n(\mathbb{R}^d) .$$

Then there exists $g \in C^{n-d-1}(\mathbb{R}^d)$ such that

$$\mu(dy) = g(y)dy .$$

Proof

Let

$$\hat{\mu}(u) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i<u|x>}\mu(dx) , \qquad u \in \mathbb{R}^d .$$

Then for $|\alpha| \leq n$, and $f_u(x) \equiv e^{-i<u|x>}$,

$$\begin{aligned} |u^\alpha||\hat{\mu}(u)| &= \frac{1}{(2\pi)^{d/2}} \left| \int_{\mathbb{R}^d} D^\alpha f_u(x)\mu(dx)\right| \\ &= \frac{1}{(2\pi)^{d/2}} \left|\mathbb{E}[D^\alpha f_u(X)]\right| \\ &\leq C_n / (2\pi)^{d/2} . \end{aligned}$$

So for $|\alpha| \leq n-d-1$, and $|\beta| \leq d+1$,

$$|\widehat{D^{\alpha}\mu}(u)| = |u^{\alpha}||\hat{\mu}(u)|$$

$$\leq |u^{\beta}|^{-1}C_n \ / \ (2\pi)^{d/2} .$$

Hence

$$\widehat{D^{\alpha}\mu} \in L^1(\mathbb{R}^d), \quad |\alpha| \leq n-d-1 .$$

So, inverting the Fourier transform,

$$D^{\alpha}\mu \in C_b(\mathbb{R}^d), \quad |\alpha| \leq n-d-1 . \qquad \square$$

Acknowledgement

This work was supported by the Science and Engineering Research Council.

1. L^p-Estimates and Differentiability in Initial Data for S.D.E's.

In this section we first state two well known results for reference. Then in Proposition 1.3 we deduce our main technical result on s.d.e's. This result enables us to deal easily with certain s.d.e's arising in Sections 2 and 3 whose coefficients are not globally Lipschitz; so that the technical difficulties they present do not become confused with the ideas of Malliavin Calculus. Sections 2 and 3 should perhaps be read before the proof of Proposition 1.3 - for motivation.

We use systematically the symbol $C(p_1,...,p_n)$ to denote a finite constant depending only on $p_1,...,p_n$.

Proposition 1.1 (Existence and L^p-Estimates for Solutions of S.D.E's)

For $i = 0,1,...,m$, let $X_i : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be previsible, and differentiable as a function of $x \in \mathbb{R}^d$. Fix $p < \infty$. Suppose there exists a constant $B < \infty$ such that, for $i = 0,1,...,m$,

$$\mathbb{E}[\sup_{t \leq T} |X_i(\omega,t,0)|^p] \leq B ,$$

and $|DX_i| \le B$ on $\Omega \times [0,T] \times \mathbb{R}^d$.

Then, for each $x \in \mathbb{R}^d$, the s.d.e.

$$\left.\begin{aligned} dx_t &= X_o(t,x_t)dt + X_i(t,x_t)dw_t^i \\ x_o &= x \end{aligned}\right\} \tag{1.1}$$

has a unique strong solution with

$$\sup_{|x|\le R} \mathbb{E}[\sup_{s\le t}|x_s - x|^p] \le C(p,T,d,R,B)t^{p/2} \tag{1.2}$$

for all $t \in [0,T]$.

Proof:

For the existence of the solution x_t see (for example) Bichteler and Jacod [2], Theorem (A.6). The L^p-bound is a straightforward exercise in Burkholder-Davis-Gundy inequalities and Gronwall's Lemma. □

Proposition 1.2 (Differentiability Theorem for S.D.E's.)

Let $X_0, X_1, \dots, X_m$ be C^∞ vector fields on $\mathbb{R}^d$, with bounded derivatives of all orders. Then there exists a map $\phi : \Omega \times [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ such that

(i) For each $x \in \mathbb{R}^d$, $x_t(\omega) \equiv \phi(\omega,t,x)$ is the unique solution of the s.d.e.

$$\left.\begin{aligned} dx_t &= X_0(x_t)dt + X_i(x_t)dw_t^i \\ x_o &= x \end{aligned}\right\} \tag{1.3}$$

(ii) For each ω and t the map $\phi(\omega,t,\cdot)$ is C^∞ on $\mathbb{R}^d$ with derivatives of all orders satisfying the s.d.e's obtained from (1.3) by successive formal differentiation. (So, for example, $U_t(\omega) \equiv D\phi(\omega,t,x)$ and $W_t(\omega) \equiv D^2\phi(\omega,t,x)$ satisfy the s.d.e's

$$\left.\begin{aligned} dU_t &= DX_o(x_t)U_t dt + DX_i(x_t)U_t dw_t^i \\ U_0 &= I \in \mathbb{R}^d \otimes \mathbb{R}^d \end{aligned}\right\} \tag{1.4}$$

and

$$\left.\begin{aligned} dW_t &= DX_0(x_t)W_t dt + DX_i(x_t)W_t dw_t^i + D^2X_0(x_t)(U_t,U_t)dt \\ &\quad + D^2X_i(x_t)(U_t,U_t)dw_t^i \\ W_0 &= 0 \in (\mathbb{R}^d \otimes \mathbb{R}^d) \otimes \mathbb{R}^d \end{aligned}\right\} \qquad (1.5)$$

respectively.)

Proof:

This result is well known. See for example Carverhill and Elworthy [4]: the s.d.e's for the derivatives are obtained using Itô's Formula from the associated s.d.e. on the diffeomorphism group.

In Section 2 we will require an extension of Proposition 1.2 in which the hypothesis is weakened to allow a wider class of vector fields, which we now define. The extension is made in Proposition 1.3.

Definition of $\mathcal{S}(d_1,\ldots,d_k)$

For $d_1,\ldots,d_k,d \in \mathbb{N}\setminus\{0\}$, with $d_1+\ldots+d_k = d$, and $\alpha \in \mathbb{N}$, we denote by $\mathcal{S}_\alpha(d_1,\ldots,d_k)$ the set of $X \in C^\infty(\mathbb{R}^d,\mathbb{R}^d)$ of the form

$$X(x) = \begin{pmatrix} X^{(1)}(x^1) \\ \vdots \\ X^{(j)}(x^1,\ldots,x^j) \\ \vdots \\ X^{(k)}(x^1,\ldots,x^k) \end{pmatrix} \quad \text{for} \quad x = \begin{pmatrix} x^1 \\ \vdots \\ x^k \end{pmatrix}, \qquad (1.6)$$

where $\mathbb{R}^d$ is identified with $\mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_k}$, and such that

$$\|X\|_{\mathcal{S}_{\alpha,N}} \equiv \sup_{x\in\mathbb{R}^d}\left(\sup_{0\le n\le N} \frac{|D^nX(x)|}{(1+|x|^\alpha)} \vee \sup_{1\le j\le k} |D_jX^{(j)}(x)|\right)$$

$$< \infty \quad \text{for all} \quad N \in \mathbb{N}.$$

We denote

$$\mathcal{S}(d_1,\ldots,d_k) = \bigcup_{\alpha\in\mathbb{N}} \mathcal{S}_\alpha(d_1,\ldots,d_k).$$

When manipulating $S(d_1,\ldots,d_k)$ vector fields below we will often assume without comment they are given in the form (1.6).

To provide some motivation for this definition note that equations (1.3), (1.4) and (1.5) may be considered together as a single s.d.e. for (x_t,U_t,W_t) whose coefficients are $S(d,d^2,d^3)$ but do not satisfy the hypothesis of Proposition 1.2.

A similar class of "lower triangular" coefficients is introduced by Stroock [12], §6 in his version of the Malliavin Calculus to play more or less the same role that $S(d_1,\ldots,d_k)$ will play below.

Proposition 1.3

Let $X_0,X_1,\ldots,X_m \in S_\alpha(d_1,\ldots,d_k)$. Then there exists a map $\phi : \Omega \times [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ such that:

(i) For each $x \in \mathbb{R}^d$, $x_t(\omega) \equiv \phi(\omega,t,x)$ is the unique solution of the s.d.e.

$$\left.\begin{aligned} dx_t &= X_0(x_t)dt + X_i(x_t)dw_t^i \\ x_0 &= x \end{aligned}\right\} \tag{1.7}$$

(ii) For each ω and t, the map $\phi(\omega,t,\cdot)$ is C^∞ on $\mathbb{R}^d$ with derivatives of all orders satisfying the s.d.e's obtained from (1.7) by formal differentiation.

(iii) $$\begin{aligned} &\sup_{|x|\le R} \mathbb{E}[\sup_{s\le t}|D^N\phi(\omega,s,x)|^p] \\ &\quad \le C(p,t,R,N,d_1,\ldots,d_k,\alpha, \|X_0\|_{S_{\alpha,N}},\ldots,\|X_m\|_{S_{\alpha,N}}) \end{aligned} \tag{1.8}$$

for all $p < \infty$, $t \ge 0$, $R < \infty$ and $N \in \mathbb{N}$.

Furthermore the following approximation result holds. Let $(X_{i,n})$, $i = 0,1,\ldots,m$, be sequences in $S_\alpha(d_1,\ldots,d_k)$ such that, for all n and $N \in \mathbb{N}$,

$$\begin{aligned} &X_{i,n} = X_i \quad \text{on } \{|x| \le n\}, \\ &\sup_{n\in\mathbb{N}} \|X_{i,n}\|_{S_{\alpha,N}} < \infty . \end{aligned} \tag{1.9}$$

Let ϕ_n denote the flow map associated with the s.d.e.

$$\left.\begin{aligned} dx_t &= X_{o,n}(x_t)dt + X_{i,n}(x_t)dw_t^i \\ x_o &= x \end{aligned}\right\} \qquad (1.10)$$

then

$$\sup_{|x|\le R} \mathbb{E}[\sup_{s\le t} |D^N\phi_n(\omega,s,x) - D^N\phi(\omega,s,x)|^p] \to 0 \qquad (1.11)$$

as $n \to \infty$, for all $p < \infty$, $t \ge 0$, $R < \infty$ and $N \in \mathbb{N}$.

<u>Proof</u>

(a) We show (1.7) has a unique solution with

$$\sup_{|x|\le R} \mathbb{E}[\sup_{s\le t} |x_s|^p]$$

$$\le C(p,t,R,d_1,\dots,d_k,\alpha, \|X_o\|_{S_{\alpha,o}},\dots, \|X_m\|_{S_{\alpha,o}}).$$

Write (1.7) as a system of s.d.e's $(j = 1,\dots,k)$

$$\left.\begin{aligned} dx_t^j &= X_o^{(j)}(x_t^1,\dots,x_t^j)dt + X_i^{(j)}(x_t^1,\dots,x_t^j)dw_t^i \\ x_o^j &= x^j \in \mathbb{R}^{d_j} . \end{aligned}\right\} \qquad (1.12j)$$

We show by induction on j that (1.12j) has a unique solution with

$$\sup_{|x|\le R} \mathbb{E}[\sup_{s\le t} |x_s^j|^p] \le C_j(p) \qquad (1.13j)$$

where $C_j(p)$ depends as C above. Suppose true for $1,\dots,j-1$. Let

$$\tilde{X}_i(\omega,t,x^j) = X_i^{(j)}(x_t^1(\omega),\dots,x_t^{j-1}(\omega),x^j) .$$

Then, for $i = 0,1,\dots,m$ and $p < \infty$,

$$\sup_{|x|\le R} \mathbb{E}[\sup_{s\le t} |\tilde{X}_i(\omega,s,0)|^p]$$

$$\le 2^p \|X_i\|_{S_{\alpha,o}}^p (1 + (j-1)^{\alpha p/2}(C_1(\alpha p) + \dots + C_{j-1}(\alpha p))) ,$$

$$|D\tilde{X}_i| \le \|X_i\|_{S_{\alpha,o}} .$$

So Proposition 1.1 applies to the s.d.e. (1.12j) when rewritten in the form

$$dx_t^j = \tilde{X}_o(t,x_t^j)dt + \tilde{X}_i(t,x_t^j)dw_t^i$$

$$x_o^j = x^j$$

and (1.13j) follows from (1.2).

(b) Here and in part (d) of the proof we make use of a particular choice of approximating coefficients which also satisfy the hypothesis of Proposition 1.2. For $j = 1,\dots,k$, choose a sequence (ψ_n^j) in $C^\infty(\mathbb{R}^{d_j},[0,1])$ such that for all $n \in \mathbb{N}\setminus\{0\}$:

$$1_{\{|x^j|\le n\}} \le \psi_n^j \le 1_{\{|x^j|\le 3n\}} ,$$

$$\|D\psi_n^j\|_\infty \le \frac{1}{n} ,$$

$$\sup_{n\in\mathbb{N}} \|D^N\psi_n^j\|_\infty < \infty , \quad \text{for all} \quad N \in \mathbb{N} .$$

Let

$$\psi_n^{(j)} = \psi_n^1\cdot\ldots\cdot\psi_n^j , \quad j = 1,\dots,k , \quad \text{and} \quad \psi_n^{(o)} \equiv 1 ,$$

$$X_{i,n}^{(j)} = X_i^{(j)}\cdot\psi_n^{(j-1)}\cdot\psi_{n^\alpha}^j ,$$

$$X_{i,n} = \begin{pmatrix} X_{i,n}^{(1)} \\ \vdots \\ X_{i,n}^{(k)} \end{pmatrix}$$

It is easy to check that Proposition 1.2 applies to the s.d.e. with coefficients $X_{o,n},X_{1,n},\dots,X_{m,n}$. Denote by ϕ_n the flow map thus obtained. Observe that for each $x \in \mathbb{R}^d$ the solution x_t of (1.7) obtained in (a) satisfies

$$x_t(\omega) = \phi_n(\omega,t,x) \quad \text{for all} \quad t \in [0,\tau_n(\omega,x)) , \quad \text{a.s.}$$

where $\tau_n(\omega,x) \equiv \inf\{t\ge 0 : |\phi_n(\omega,t,x)| = n\}$.

Note also that since $\phi_n(\omega,t,\cdot)$ is continuous the set $\{x : \tau_n(\omega,x) > t\}$ is open in $\mathbb{R}^d$ for all ω,t and n .

Since $\phi_n(\omega,t,x) = \phi_{n+1}(\omega,t,x)$ for all $t \in [0,\tau_n(\omega,x))$ and all $x \in \mathbb{R}^d$, a.s., we may piece together a map $\phi : \{(\omega,t,x) : t < \zeta(\omega,x)\} \to \mathbb{R}^d$, where $\zeta(\omega,x) \equiv \lim_{n\to\infty} \tau_n(\omega,x)$, such that

$$\phi(\omega,t,x) = \phi_n(\omega,t,x) \quad \text{on} \quad \{(\omega,t,x) : t < \tau_n(\omega,x)\}$$

and $\phi(\omega,t,\cdot)$ is C^∞ on the open set

$$\{x : t < \zeta(\omega,x)\} \quad \text{for all } \omega \text{ and } t .$$

For fixed $x \in \mathbb{R}^d$, (a) implies $\zeta(\omega,x) = \infty$ a.s. so $\phi(\omega,t,x)$ "is" the solution of (1.7). Moreover the derivatives of ϕ must satisfy the s.d.e's obtained from (1.7) by successive differentiation, since they agree with the derivatives of ϕ_n up to τ_n.

Thus parts (i) and (ii) of the proposition will be established as soon as it is shown that $\zeta(\omega,x) = \infty$ for all x,a.s. This is actually a rather delicate point (see for example Leandre [8] or Elworthy [5] p.91).

(c) Proof of (iii) and the approximation result.

Fix $x \in \mathbb{R}^d$. For $i = 0,1,\dots,m$, let $(X_{i,n})$ be a sequence in $S_\alpha(d_1,\dots,d_k)$ satisfying (1.9). Denote by ϕ and ϕ_n the flow maps associated with equations (1.7) and (1.10) respectively. (We have shown in part (b) that these may be defined up to explosion time $\zeta(\omega,x)$ and that, for fixed x, $\zeta(\omega,x) = \infty$ a.s.) For $N \in \mathbb{N}$, let

$$U_t^N = D^N\phi(\omega,t,x) \quad \text{and} \quad U_t^N(n) = D^N\phi_n(\omega,t,x) .$$

Now fix $N \in \mathbb{N}$. Successive differentiation of (1.7) generates a system of s.d.e's for $(U_t^0,U_t^1,\dots,U_t^N)$ with coefficients in $S_{\alpha'}(d_1,\dots,d_k;dd_1,\dots,dd_k;\dots;d^Nd_1,\dots,d^Nd_k)$ for some α' (depending on α and N). Moreover the $S_{\alpha'}$-norm of these coefficients may be bounded by a quantity depending only on the $S_{\alpha,N}$-norm of the coefficients of (1.7), and N. (It may help to recall (1.4) and (1.5) where s.d.e's for the first two derivatives are written out.)

Assertion (iii) now follows from part (a) of the proof.

Applying the above argument to ϕ_n and using (1.9) we have

$$\sup_{n\in\mathbb{N}} \sup_{|x|\le R} \mathbb{E}[\sup_{s\le t} |D^N\phi_n(\omega,s,x)|^p] < \infty \qquad (1.14)$$

for all $p<\infty$, $t\ge 0$, $R<\infty$ and $N\in\mathbb{N}$.

$$\mathbb{E}[\sup_{s\le t} |U_s^N - U_s^N(n)|^p]$$

$$= \mathbb{E}[\sup_{s\le t} |U_s^N - U_s^N(n)|^p \cdot 1_{\{\sup_{s\le t}|x_s|\ge n\}}] \quad \text{by (1.9)}$$

$$\le n^{-1}\,\mathbb{E}[\sup_{s\le t} |U_s^N - U_s^N(n)|^p \cdot \sup_{s\le t}|x_s|]$$

$$\le 2^p n^{-1}\Big(\mathbb{E}[\sup_{s\le t} |U_s^N|^{2p}]^{\frac12} + \mathbb{E}[\sup_{s\le t} |U_s^N(n)|^{2p}]^{\frac12}\Big)$$

$$\cdot\, \mathbb{E}[\sup_{s\le t}|x_s|^2]^{\frac12}\ .$$

So (1.11) follows from (1.14).

(d) We show $\zeta(\omega,x)=\infty$, for all $x\in\mathbb{R}^d$, a.s.

We recall the particular choice of approximating coefficients used in part (b). We show firstly that $(X_{i,n})$ satisfies (1.9) for $i=0,1,\ldots,m$. It suffices to observe that

$$|D^N X_{i,n}^{(j)}| = \Big|\sum_{r=0}^{N} \binom{N}{r} D^r X_i^{(j)} D^{N-r}(\psi_n^{(j-1)}\psi_{n^\alpha}^j)\Big|$$

$$\le \sum_{r=0}^{N} \binom{N}{r}(1+|x|^\alpha)\,\|X_i\|_{\mathcal{S}_{\alpha,N}}$$

$$\cdot \sup_{n\in\mathbb{N}} \sup_{r\le N} \|D^r(\psi_n^{(j-1)}\psi_{n^\alpha}^j)\|_\infty$$

and, since

$$|X_i^{(j)}\cdot\psi_n^{(j-1)}(x^1,\ldots,x^j)|$$

$$\le \sup_{x\in\mathbb{R}^d} |X_i^{(j)}\psi_n^{(j-1)}(x^1,\ldots,x^{j-1},0)| + (1+|x^j|)\,\|D_j X_i^{(j)}\|_\infty$$

$$\leq \|X_i\|_{S_{\alpha,o}} [(1+(3n)^{\alpha}) + (1+|x^j|)]$$

we have

$$|D_j X^{(j)}_{i,n}| = |D_j X^{(j)}_i \cdot \psi^{(j-1)}_n \cdot \psi^j_{n^\alpha} + X^{(j)}_i \cdot \psi^{(j-1)}_n \cdot D\psi^j_{n^\alpha}|$$

$$\leq \|X_i\|_{S_{\alpha,o}} (1 + 2(1+(3n)^{\alpha})/n^{\alpha}) .$$

Thus (1.9) holds. We deduce (1.11) :

$$\sup_{|x|\leq R} \mathbb{E}[\sup_{s\leq t} |D^N\phi_n(\omega,s,x) - D^N\phi(\omega,s,x)|^p] \to 0$$

as $n \to \infty$, for all $p < \infty$, $t \geq 0$, $R < \infty$ and $N \in \mathbb{N}$.

We turn now to a well known inequality of Sobolev.

For C^{∞} functions ψ on $\mathbb{R}^d$ define

$$\|\psi\|^R_{p,N} = \sum_{M=0}^{N} \left(\int_{|x|<R} \left| D^M\psi(x) \right|^p dx \right)^{1/p} ,$$

$$\|\psi\|^R_{\infty,N} = \sum_{M=0}^{N} \sup_{|x|\leq R} |D^M\psi(x)| .$$

Then (Sobolev, [11]), for each R and $N \geq 0$, there exist $\tilde{R} > R$, $\tilde{N} > N$, $p < \infty$ and a constant $K < \infty$ such that

$$\|\psi\|^R_{\infty,N} \leq K\|\psi\|^{\tilde{R}}_{p,\tilde{N}} \quad \text{for all } \psi \in C^{\infty}(\mathbb{R}^d) .$$

It follows from (1.11) that

$$\mathbb{E}\left(\sup_{s\leq t} \int_{|x|\leq R} \left| D^N\phi_n(\omega,s,x) - D^N\phi_m(\omega,s,x) \right|^p dx \right) \to 0$$

as $n,m \to \infty$, for all $p < \infty$, $t \geq 0$, $R < \infty$ and $N \in \mathbb{N}$.

So, extracting a subsequence if necessary, there exists a null set $\Gamma \subseteq \Omega$ such that for $\omega \notin \Gamma$

$$\sup_{s\leq t} \int_{|x|\leq R} \left| D^N\phi_n(\omega,s,x) - D^N\phi_m(\omega,s,x) \right|^p dx \to 0 \qquad (1.15)$$

for all t,R,N and $p < \infty$. By the Sobolev inequality, (1.15) then holds for all t,R,N and $p = \infty$.
In particular, for $\omega \notin \Gamma$, $\phi_n(\omega,s,x)$ converges to $\phi(\omega,s,x)$ uniformly on compact subsets of $[0,\infty) \times \mathbb{R}^d$. So $\zeta(\omega,x) = \infty$ for all x, for $\omega \notin \Gamma$.

2. The Bismut Integration by Parts Formula

For $X_0, X_1, \ldots, X_m \in S(d_1, \ldots, d_k)$ and $x \in \mathbb{R}^d$, by Proposition 1.3, the s.d.e.

$$\left.\begin{aligned} dx_t &= X_0(x_t)dt + X_i(x_t)dw_t^i \\ x_0 &= x \end{aligned}\right\} \qquad (2.1)$$

has a unique solution with $\sup_{s \leq t} |x_s| \in L^p(\mathbb{P})$ for all $t \geq 0$ and $p < \infty$.

We obtain in this section an integration by parts formula involving x_t under conditions sufficiently general for the purposes of Section 3. The formula first appeared in Bismut [3] as Theorem 2.1, but written without the helpful Dx_t notation of Bichteler and Fonken [1]. We follow in outline Meyer's simplification of Bismut's proof [9] but work in greater generality. This generality is needed for the iterations of the integration by parts formula involved in proving the smooth density result.

The integration by parts formula is obtained by viewing a perturbed solution of (2.1) in two ways. Let $u : \mathbb{R}^d \to \mathbb{R}^m \otimes \mathbb{R}^r$ be C^∞ and bounded, with all its derivatives of polynomial growth. For $h \in \mathbb{R}^r$, let

$$w_t^h = w_t + \int_0^t u(x_s).h\, ds$$

The perturbed process x_t^h is defined by

$$\left.\begin{aligned} dx_t^h &= X_0(x_t^h)dt + X_i(x_t^h)\, dw_t^{h,i} \\ x_0^h &= x \end{aligned}\right\} \qquad (2.2)$$

or equivalently (writing $(u(x_s).h)^i$ for the ith component)

$$\left.\begin{aligned} dx_t^h &= (X_0(x_t^h) + X_i(x_t^h)(u(x_t).h)^i)dt + X_i(x_t^h)dw_t^i \\ x_0^h &= x \end{aligned}\right\} \qquad (2.2)'$$

Using Girsanov's Theorem a new probability measure $\mathbb{P}^h$ may be found to make w_t^h an $\mathbb{R}^m$ - Brownian motion. Since x_t is a measurable function of the path $(w_s)_{s\leq t}$, (2.2) thus implies that the law of x_t^h under $\mathbb{P}$ is independent of $\bar{h}$, i.e.

$$\frac{\partial}{\partial h}\int_\Omega f(x_t^h)\,d\mathbb{P}^h = 0 \qquad \text{for all} \quad f \in C_b(\mathbb{R}^d).$$

Using (2.2) one can show x_t^h is differentiable in h and a differentiation under the integral sign is possible yielding an integration by parts formula.

Let

$$z_t^h = \exp\left(-\int_0^t (u(x_s).h)^i dw_s^i - \frac{1}{2}\int_0^t |u(x_s).h|^2 ds\right) \tag{2.3}$$

and let

$$\mathbb{P}^h = z_t^h \mathbb{P} \text{ on } \mathcal{F}_t .$$

Lemma 2.1

(a) For each $h \in \mathbb{R}^r$, z_t^h satisfies the s.d.e.

$$\left.\begin{aligned} dz_t^h &= -z_t^h(u(x_t).h)^i dw_t^i \\ z_0 &= 1 \end{aligned}\right\} \tag{2.4}$$

(b) For all $t \geq 0$ and $p < \infty$

$$\sup_{|h|\leq 1} \mathbb{E}\left(\sup_{s\leq t}|z_s^h|^p\right) < \infty \tag{2.5}$$

(c) Under $\mathbb{P}^h$, w_t^h is an $\mathbb{R}^m$-Brownian motion.

Proof:

(a) Use Itô's Formula

(b) Apply Proposition 1.3 to the system of s.d.e's.

$$\left.\begin{aligned} dh_t &= 0 , \quad h_0 = h \\ dx_t &= X_0(x_t)dt + X_i(x_t)dw_t^i , \quad x_0 = x \\ dz_t &= - z_t(u(x_t).\psi(h_t)h_t)^i dw_t^i , \quad z_0 = 1 \end{aligned}\right\}$$

where ψ is a C^∞ function of compact support on $\mathbb{R}^r$ with $\psi(h) = 1$ for $|h| \leq 1$.

The coefficients lie in $S(r,d,1)$!

(c) By (a) and (b), Z_t^h is the exponential associated with the martingale $-\int_0^t (u(x_s)h)^i dw_s^i$ and is itself a martingale. So by the Girsanov Theorem (see example Jacod [7], Theorem 7.24), w_t being a $\mathbb{P}$-martingale,

$$w_t^h = w_t - \langle -\int_0^{\cdot} (u(x_s)h)^i dw_s^i, w \rangle_t$$

is a $\mathbb{P}^h$-martingale. But the quadratic variation of w_t^h under $\mathbb{P}^h$ is exactly that of w_t under $\mathbb{P}$. So by Levy's characterization of Brownian motion we have (c).

In the next proposition we obtain, for each ω, differentiability in a parameter of solutions of s.d.e's by the trick of turning the parameter into a starting point.

<u>Proposition 2.2</u> <u>(Differentiability with respect to a parameter)</u>

Let $X_0, X_1, \ldots, X_m \in S(d_1, \ldots, d_k)$ and $d_1 + \ldots + d_k = d$.
Let $u : \mathbb{R}^d \to \mathbb{R}^m \otimes \mathbb{R}^r$ be C^∞ with all derivatives of polynomial growth. Then there exists a function

$$\phi : \Omega \times [0,\infty) \times \mathbb{R}^r \times \mathbb{R}^d \to \mathbb{R}^d$$

such that

(i) For each $(h,x) \in \mathbb{R}^r \times \mathbb{R}^d$, $x_t^h(\omega) \equiv \phi(\omega,t,h,x)$ is the unique solution of (2.2)':

$$dx_t^h = (X_0(x_t^h) + X_i(x_t^h)(u(x_t).h)^i)dt + X_i(x_t^h)dw_t^i$$

$$x_0^h = x .$$

(ii) For each ω and t, $\phi(\omega,t,\cdot,\cdot)$ is continuously differentiable on $\mathbb{R}^r \times \mathbb{R}^d$ with

$$\sup_{|h|\leq 1} \mathbb{E}\left\{\sup_{s\leq t} \left|\frac{\partial\phi}{\partial h}(\omega,s,h,x)\right|^p\right\} < \infty$$

for all $x \in \mathbb{R}^d$, $t \geq 0$ and $p < \infty$.

(iii) Define $Dx_t^h \equiv \frac{\partial\phi}{\partial h}(\omega,t,h,x)$. Then Dx_t^h satisfies the s.d.e. obtained by differentiating (2.2)' formally:

$$\left.\begin{aligned} dDx_t^h &= \left[DX_0(x_t^h)Dx_t^h + DX_i(x_t^h)Dx_t^h(u(x_t).h)^i\right] dt \\ &\quad + X_i(x_t^h)u(x_t)^i\, dt + DX_i(x_t^h)Dx_t^h dw_t^i \\ Dx_0^h &= 0 \in \mathbb{R}^d \otimes \mathbb{R}^r \end{aligned}\right\} \quad (2.6)$$

(where $u(x_t)^i$ denotes the ith row of $u(x_t)$).

Proof:

Apply Proposition 1.3 to the system of s.d.e's.

$$\left.\begin{aligned} dh_t &= 0, \quad h_0 = h \\ dx_t &= X_0(x_t)dt + X_i(x_t)dw_t^i \quad, \quad x_0 = x \\ dx_t^h &= (X_0(x_t^h) + X_i(x_t^h)(u(x_t).\psi(h_t)h_t)^i)dt + X_i(x_t^h)dw_t^i \\ x_0^h &= x \end{aligned}\right\} \quad (2.7)$$

(where $\psi : \mathbb{R}^r \to \mathbb{R}$ is C^∞, of compact support and $\psi(h) = 1$ for $|h| \leq 1$). This system has $S(r;d_1,\ldots,d_k;d_1,\ldots,d_k)$ coefficients. The conclusion of Proposition 1.3, parts (i)-(iii), implies (i), (ii) and (iii) above.

Theorem 2.3 (Integration by Parts Formula)

Let $X_0,X_1,\ldots,X_m \in S(d_1,\ldots,d_k)$. Let x_t be the solution of (2.1) :

$$\left.\begin{aligned} dx_t &= X_0(x_t)dt + X_i(x_t)dw_t^i \\ x_0 &= x \in \mathbb{R}^d. \end{aligned}\right\}$$

Let $u : \mathbb{R}^d \to \mathbb{R}^m \otimes \mathbb{R}^r$ be C^∞, with all derivatives of polynomial growth. Then the linear s.d.e.

$$\left.\begin{aligned} dDx_t &= DX_0(x_t)Dx_t dt + DX_i(x_t)Dx_t dw_t^i \\ &\quad + X_i(x_t)u(x_t)^i dt \\ Dx_0 &= 0 \in \mathbb{R}^d \otimes \mathbb{R}^r \end{aligned}\right\} \quad (2.8)$$

has a unique solution with $\sup_{s\le t}|Dx_t| \in L^p(\mathbb{P})$ for all $t \geq 0$ and $p < \infty$.

Furthermore, for any function $f : G \to \mathbb{R}$, where G is an open subset of $\mathbb{R}^d$ with $x_t \in G$ a.s., such that f is differentiable and $Df(x_t)$ and $f(x_t) \in L^2(\mathbb{P})$,

$$\mathbb{E}\left[Df(x_t)Dx_t\right] = \mathbb{E}\left[f(x_t)\int_0^t u(x_s)^i dw_s^i\right] . \tag{2.9}$$

Remark:

Equations (2.1) and (2.8) combine to form a system of s.d.e's with $S(d_1,\ldots,d_k;\ d.d_1,\ldots,d.d_k)$ coefficients. This is the crucial observation for iterations of the formula.

Proof:

Assume for now u and all its derivatives are bounded. Let x_t^h be the solution of (2.2)/(2.2)'. Define z_t^h by (2.3). We make three observations:

(i) By Lemma 2.1 (c)

$$\frac{\partial}{\partial h}\mathbb{E}\left[f(x_t^h)z_t^h\right] = 0 \quad \text{for all} \quad f \in C_b^1(\mathbb{R}^d). \tag{2.10}$$

(ii) By Proposition 2.2 we may assume $x_t^h(\omega)$ is differentiable in h a.s. with

$$\sup_{|h|\le 1}\mathbb{E}\left[\left|\frac{\partial x_t^h}{\partial h}\right|^p\right] < \infty \quad \text{for all} \quad t \geq 0 \text{ and } p < \infty.$$

(iii) For each ω, z_t^h is evidently differentiable in h with

$$\frac{\partial}{\partial h} z_t^h = -z_t^h\left[\int_0^t u(x_s)^i dw_s^i + \int_0^t u(x_s)^T u(x_s) h\, ds\right]$$

so by Lemma 2.1 (b)

$$\sup_{|h|\le 1}\mathbb{E}\left[\left|\frac{\partial}{\partial h} z_t^h\right|^p\right] < \infty \quad \text{for all} \quad t \geq 0 \text{ and } p < \infty.$$

It follows that we may differentiate (2.10) under the expectation sign at 0 to obtain (2.9). Equation (2.8) is just (2.6) with $h = 0$.

It remains to relax the boundedness conditions on u and f. For f as in the statement of the theorem take a sequence f_n in $C_b^1(\mathbb{R})$ with

$$|f_n| \leq |f|, |Df_n| \leq |Df| \text{ and } f_n \to f, \ Df_n \to Df$$

(pointwise) on G. Then (2.9) extends to f by the Dominated Convergence Theorem.

We extend the result to functions u with derivatives of polynomial growth by means of an approximating sequence of compactly supported functions u_n. Choose C^∞ functions $\psi_n : \mathbb{R}^d \to [0,1]$ with

$$1_{\{|x| \leq n\}} \leq \psi_n \leq 1_{\{|x| \leq n+1\}}$$

with derivatives of all orders uniformly bounded in n and on $\mathbb{R}^d$. Let $u_n = u.\psi_n$ then (2.8)/(2.9) holds for u_n. Since $\sup_{s \leq t} |x_s|^p \in L^p(\mathbb{P})$ for all $t \geq 0$ and $p < \infty$,

$$\int_0^t u_n(x_s)^i dw_s^i \to \int_0^t u(x_s)^i dw_s^i \quad \text{in} \quad L^p(\mathbb{P})$$

for all $p < \infty$ and $t \geq 0$. We may thus extend (2.8)/(2.9) to u by taking the limit as $n \to \infty$, provided that (with an obvious notation)

$$Dx_t(n) \to Dx_t \quad \text{in} \quad L^p(\mathbb{P}), \quad \text{some } p \geq 2$$

But by virtue of our assumptions on ψ_n this is a consequence of the approximation result of Proposition 1.3 applied to the system of s.d.e's

(2.1) and

(2.8) with u replaced by u_n.

<u>Alternative proof</u>:

One can avoid using the full strength of Proposition 1.3 by establishing the formula first for X_i bounded with bounded derivatives of all orders then extending to $X_i \in \mathcal{S}(d_1,...,d_k)$ by the same approximation as was used in parts (b) and (d) of the proof of Proposition 1.3. Thus the use of the Sobolev inequality in Proposition 1.3 may be avoided.

3. Application of the Integration by Parts Formula: Smooth Density and the Covariance Matrix

We fix vector fields $X_0, X_1, \ldots, X_m$ on $\mathbb{R}^d$ which are assumed C^∞ with bounded first derivatives and higher derivatives of polynomial growth. We obtain a sufficient condition for the s.d.e.

$$\left.\begin{aligned} dx_t &= X_0(x_t)dt + X_i(x_t)dw_t^i \\ x_0 &= x \end{aligned}\right\} \qquad (3.1)$$

to have a smooth density.

We will make use of two processes U_t and V_t associated to the s.d.e. (3.1), which are in fact the derivative of the flow associated to (3.1) and its inverse. However we will regard them as defined by the following s.d.e's.

$$\left.\begin{aligned} dU_t &= DX_0(x_t)U_t dt + DX_i(x_t)U_t dw_t^i \\ U_0 &= I \in \mathbb{R}^d \otimes \mathbb{R}^d \end{aligned}\right\} \qquad (3.2)$$

$$\left.\begin{aligned} dV_t &= -V_t\left\{DX_0(x_t) - \sum_{i=1}^m DX_i(x_t)^2\right\}dt - V_t DX_i(x_t)dw_t^i \\ V_0 &= I \in \mathbb{R}^d \otimes \mathbb{R}^d \end{aligned}\right\} \qquad (3.3)$$

The system $\{(3.1),(3.2),(3.3)\}$ has $S(d,d^2,d^2)$ coefficients so by Proposition 1.3

$$\sup_{s\leq t} |U_s| \text{ and } \sup_{s\leq t} |V_s| \in L^p(\mathbb{P}), \text{ for all } p < \infty \text{ and } t \geq 0.$$

Furthermore an easy application of Itô's Formula shows that for U_t and V_t so defined we indeed have $U_t^{-1} = V_t$ for all $t \geq 0$, a.s.

We now make the optimal choice of perturbation u for the process x_t. We aim by this choice to make the matrix Dx_t non-degenerate. Recall that

$$dDx_t = DX_0(x_t)\, Dx_t\, dt + DX_i(x_t)Dx_t\, dw_t^i + X_i(x_t)\, u(x_t)^i\, dt$$

So

$$\begin{aligned} d(V_t\, Dx_t) &= V_t dDx_t + dV_t Dx_t + \langle V_t, Dx_t\rangle \\ &= V_t X_i(x_t).u(x_t)^i\, dt \end{aligned}$$

Thus $\quad Dx_t = U_t\int_0^t V_s X_i(x_s).u(x_s)^i\, ds.$ So we would like to take "$u(x_s)^i = (V_s X_i(x_s))^T$". That we can allow u to depend on V_t as well as x_t follows by the technical device of applying Theorem 2.3 not to (3.1) but to {(3.1),(3.3)}. We then choose $u(x,V) = (VX_i(x))^T$ ($x \in \mathbb{R}^d$, $V \in \mathbb{R}^d \otimes \mathbb{R}^d$) so that $Dx_t = U_t C_t$, where

$$C_t = \int_0^t V_s X_i(x_s) \otimes V_s X_i(x_s)\, ds \tag{3.4}$$

- the Malliavin Covariance Matrix.

The main result of this section is that if for some $t > 0$, $C_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$, then x_t has a smooth density.

Assume for the rest of this section that for a certain $t > 0$, $C_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$. Then since $Dx_t = U_t C_t$, $Dx_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$.

The following definition will be used to provide classes of functions to which the Integration by Parts Formula applies.

Definition:

For an $\mathbb{R}^n$- valued random variable Y, denote by $\mathcal{D}[Y]$ the set of all functions $f : \mathbb{R}^n \to \mathbb{R}$ such that for some open set $W \subseteq \mathbb{R}^n$:
(i) $Y \in W$ a.s., (ii) $f|_W$ is C^∞ and (iii) $D^\alpha f(Y) \in L^2(\mathbb{P})$ for all $\alpha \geq 0$.

The point of this definition is that the inverse map on d×d matrices lies in $\mathcal{D}[Dx_t]$.

Recall the remark following Theorem 2.3: if a process y_t satisfies an s.d.e. with $\mathcal{S}(d_1,\ldots,d_k)$ coefficients, then (y_t, Dy_t) satisfies one with $\mathcal{S}(d_1,\ldots,d_k;\, dd_1,\ldots,dd_k)$ coefficients. So (for a fixed $u(y_t)$) we may define inductively

$$D^n y_t = D(D^{n-1} y_t) \ .$$

In particular let

$$y_t^{(0)} = (x_t, V_t, R_t)$$

where

$$R_t = \int_0^t (V_s X_i(x_s))^T dw_s^i$$

then $y_t^{(0)}$ satisfies an s.d.e. with $\mathcal{S}(d,d^2,1)$ coefficients so we may define for $n \geq 1$

$$y_t^{(n)} = (y_t^{(0)}, Dy_t^{(0)}, \ldots, D^n y_t^{(0)}) .$$

<u>Theorem 3.1</u>

Suppose for some $t > 0$, $C_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$. Then for each $n \geq 1$ and $k = 1,\ldots,d$ there exists a map

$$A_k^n : \mathcal{D}[y_t^{(n)}] \to \mathcal{D}[y_t^{(n+1)}]$$

such that:

$$\mathbb{E}\,[(D_k f)(x_t) g(y_t^{(n)})] = \mathbb{E}\,[f(x_t)(A_k^n g)(y_t^{(n+1)})] \tag{3.5}$$

for all $f \in C_b^1(\mathbb{R}^d)$ and $g \in \mathcal{D}[y_t^{(n)}]$ (where D_k is the kth partial derivative).

<u>Proof</u>:

Apply Theorem 2.3 to the process $y_t^{(n)}$ and the matrix function F such that

$$y_t^{(n)} \to f(x_t)\, \Psi\, (Dx_t) g(y_t^{(n)})$$

(where $\Psi(D) \equiv D^{-1}$ for $D \in \mathbb{R}^d \otimes \mathbb{R}^d$). The components of F are all $\mathcal{D}[y_t^{(n)}]$ since $Dx_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$. So we have the following equality in $(\mathbb{R}^d \otimes \mathbb{R}^d) \otimes \mathbb{R}^d$

$$\begin{aligned}
&\mathbb{E}\,[Df(x_t)Dx_t \otimes \Psi(Dx_t) g(y_t^{(n)})] \\
&\quad + \mathbb{E}\,[f(x_t) D\Psi(Dx_t) D^2 x_t g(y_t^{(n)})] \\
&\quad + \mathbb{E}\,[f(x_t)\Psi(Dx_t) \otimes Dg(y_t^{(n)}) Dy_t^{(n)}] \\
&\quad = \mathbb{E}\,[f(x_t)\Psi(Dx_t) g(y_t^{(n)}) \otimes R_t]
\end{aligned}$$

Summing the (k,j,j) component over j and rearranging we have (3.5) with

$$(A_k^n g)(y_t^{(n+1)}) \equiv \sum_{j=1}^{d} \Big\{ \Psi(Dx_t) g(y_t^{(n)}) \otimes R_t - D\Psi(Dx_t) D^2 x_t g(y_t^{(n)}) - \Psi(Dx_t) \otimes Dg(y_t^{(n)}) Dy_t^{(n)} \Big\}_{(k,j,j)}$$

The fact that $A_k^n g \in \mathcal{D}[y_t^{(n+1)}]$ follows from $g \in \mathcal{D}[y_t^{(n)}]$ and $Dx_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$.

Theorem 3.1 is ready-made for iteration, which we now perform and which leads immediately to the main result.

Theorem 3.2

Suppose for some $t > 0$, $C_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$. Then the law of x_t has a C^∞ density with respect to Lebesgue measure on $\mathbb{R}^d$.

Proof:

Let $g(y_t^{(1)}) \equiv 1$ then $g \in \mathcal{D}[y_t^{(1)}]$. By repeated application of Theorem 3.1, for each $n \geq 1$,

$$\mathbb{E}\,[(D_{k_1}\ldots D_{k_n})f(x_t)] = \mathbb{E}\,[f(x_t)(A_{k_n}^n\circ\ldots\circ A_{k_1}^1\; g)(y_t^{n+1})]$$

for all $f \in C_b^\infty(\mathbb{R}^d)$.

So $\quad |\mathbb{E}\;[(D_{k_1}\;\ldots\;D_{k_n})f(x_t)]| \leq c(k_1,\ldots,k_n)\,\|\,f\,\|_\infty$,

where $c(k_1,\ldots,k_n) = \mathbb{E}\,[|(A_{k_n}^n\circ\ldots\circ A_{k_1}^1 g)(y_t^{(n+1)})|]$.

The result follows by Theorem 0.1. □

4. Non-Degeneracy of the Covariance Matrix under the H_1 condition.

It is convenient in this section to relabel the coefficient X_0 appearing in (3.1) as $\tilde{X}_0$, whilst preserving in all other respects the set up of §3. This is because we wish to reserve the symbol X_0 for the dt coefficient $X_0 \equiv \tilde{X}_0 - \frac{1}{2} DX_i\cdot X_i$ of the associated Stratonovich s.d.e.:

$$\left.\begin{aligned} dx_t &= X_0(x_t)dt + X_i(x_t)\partial w_t^i \\ x_0 &= x \in \mathbb{R}^d \end{aligned}\right\} .$$

We show that the covariance matrix C_t, defined at (3.4), satisfies $C_t^{-1} \in L^p(\mathbb{P})$ for all $p<\infty$ and $t>0$, provided that the following local condition on the vector fields $X_0,\ldots,X_m$ is met:

H_1 : $X_1,\ldots,X_m$; $[X_i,X_j]_{i,j=0}^m$; $[X_i[X_j,X_k]]_{i,j,k=0}^m$; ... etc., evaluated at x, span $\mathbb{R}^d$.

This result, combined with Theorem 3.2, completes the task of showing that H_1 is sufficient for the smooth density of $x_t, t>0$.

The proof of the main result is given in Theorem 4.2 following Meyer [10], himself following Stroock [13]. The new contribution is the semimartingale inequality set out in Lemma 4.1.

Lemma 4.1

Let $\alpha, y \in \mathbb{R}$. Let $\beta_t, \gamma_t \equiv (\gamma_t^1, \ldots, \gamma_t^m)$ and $u_t \equiv (u_t^1, \ldots, u_t^m)$ be previsible processes. Let

$$a_t = \alpha + \int_0^t \beta_s ds + \int_0^t \gamma_s^i dw_s^i \qquad \text{and}$$

$$Y_t = y + \int_0^t a_s ds + \int_0^t u_s^i dw_s^i \quad .$$

Suppose T is a bounded stopping time ($T \le t_0$ say) such that for some constant $C < \infty$:

$$|\beta_t|, \quad |\gamma_t|, \quad |a_t| \;\&\; |u_t| \le C \text{ for all } t \le T.$$

Then for any $q > 8$ and $\nu < (q-8)/9$

$$\mathbb{P}\left\{\int_0^T Y_t^2 dt < \varepsilon^q \text{ and } \int_0^T (|a_t|^2 + |u_t|^2) dt \ge \varepsilon\right\}$$
$$\le \text{const}(C, t_0, q, \nu) e^{-1/\varepsilon^\nu} .$$

Proof

We adopt some notation. Let

$$A_t = \int_0^t a_s ds \quad ,$$

$$M_t = \int_0^t u_s^i dw_s^i \quad ,$$

$$N_t = \int_0^t Y_s u_s^i dw_s^i \quad ,$$

and $Q_t = \int_0^t A_s \gamma_s^i dw_s^i$.

Define for $\varepsilon, \delta > 0$

$$B_1(\varepsilon, \delta) = \{\langle N,N\rangle_T < \varepsilon \text{ and } \sup_{t \le T} |N_t| \ge \delta\},$$

$$B_2(\varepsilon,\delta) = \{<M,M>_T < \varepsilon \text{ and } \sup_{t\le T}|M_t| \ge \delta\}$$

and $B_3(\varepsilon,\delta) = \{<Q,Q>_T < \varepsilon \text{ and } \sup_{t\le T}|Q_t| \ge \delta\}$

By a well known exponential martingale inequality,

$$\mathbb{P}(B_i(\varepsilon,\delta)) \le 2e^{-\delta^2/2\varepsilon} \quad \text{for } i = 1,2,3.$$

Let $q_1 = \frac{1}{2}(q-\nu)$, $q_2 = \frac{1}{2}(\frac{1}{2}q_1-\nu)$ and $q_3 = \frac{1}{2}(2q_2-\nu)$.

Then $q_3 = \frac{1}{8}(q-9\nu) > 1$. For $i = 1,2,3$, let $\delta_i = \varepsilon^{q_i}$.

We will choose below in an appropriate way $\varepsilon_i>0$ such that $B_i\equiv B_i(\varepsilon_i,\delta_i)$ has probability $0(e^{-1/\varepsilon^\nu})$, $i = 1,2,3$. For our choice of ε_1, ε_2, ε_3 we will show further that

$$\left\{\int_0^T Y_t^2 dt < \varepsilon^q \text{ and } \int_0^T (|a_t|^2+|u_t|^2)dt \ge \varepsilon\right\} \subseteq B_1 \cup B_2 \cup B_3$$

for sufficiently small ε, thus completing the proof.

Suppose that $\omega \notin B_1 \cup B_2 \cup B_3$ and $\int_0^T Y^2 dt < \varepsilon^q$.

Then $<N,N>_T = \int_0^T Y_t{}^2|u_t|^2 dt < C^2\varepsilon^q$. Choose $\varepsilon_1 = C^2\varepsilon^q$.

Then since $\omega \notin B_1$, $\sup_{t\le T}\left|\int_0^t Y_s u_s^i dw_s^i\right| < \delta_1 = \varepsilon^{q_1}$.

Also $\sup_{t\le T}\left|\int_0^t Y_s a_s ds\right| \le (t_0\int_0^T Y_t^2\, a_t^2\, dt)^{\frac{1}{2}} < t_0^{\frac{1}{2}}\, C\, \varepsilon^{q/2}$

Thus $\sup_{t\le T}\left|\int_0^t Y_s dY_s\right| < (1 + t_0^{\frac{1}{2}}\, C\, \varepsilon^{\nu/2})\, \varepsilon^{q_1}$. By Itô's Formula

$$Y_t^2 = y^2 + 2\int_0^t Y_s dY_s + <M,M>_t. \text{ So}$$

$$\begin{aligned}\int_0^T <M,M>_t\, dt &= \int_0^T Y_t^2\, dt - Ty^2 - 2\int_0^T\int_0^t Y_s dY_s\, dt \\ &< \varepsilon^q + 2t_o(1 + t_0^{\frac{1}{2}} C\varepsilon^{\nu/2})\, \varepsilon^{q_1} \\ &< (2t_0+1)\varepsilon^{q_1} \quad \text{for sufficiently small } \varepsilon.\end{aligned}$$

Since $\langle M,M\rangle_t$ is an increasing process, we must have $\langle M,M\rangle_{T-\gamma} < (2t_0+1)\,\varepsilon^{q_1}/\gamma$ and hence $\langle M,M\rangle_T < (2t_0+1)\,\varepsilon^{q_1}/\gamma + C^2\gamma$, for any $\gamma > 0$. Choose $\gamma = (2t_0+1)^{1/2}\,\varepsilon^{q_1/2}$ and $\varepsilon_2 = (1+C^2)(2t_0+1)^{1/2}\,\varepsilon^{q_1/2}$. Then since $\omega \notin B_2$

$$\sup_{t\le T}|M_t| < \delta_2 = \varepsilon^{q_2}.$$

Recall that $\int_0^T Y^2 dt < \varepsilon^q$ so that

$$\text{Leb}\,\{t \in [0,T] : |Y_t| \ge \varepsilon^{q/3}\} \le \varepsilon^{q/3} \quad \text{and so}$$

$$\text{Leb}\,\{t \in [0,T] : |y+A_t| \ge \varepsilon^{q/3} + \varepsilon^{q_2}\} \le \varepsilon^{q/3}.$$

So for each $t \in [0,T]$, there exists $s \in [0,T]$ such that $|s-t| \le \varepsilon^{q/3}$ and $|y+A_s| < \varepsilon^{q/3} + \varepsilon^{q_2}$. Therefore $|y+A_t| \le |y+A_s| + \left|\int_s^t a_r dr\right| < (1+C)\varepsilon^{q/3} + \varepsilon^{q_2}$. In particular $|y| < (1+C)\varepsilon^{q/3} + \varepsilon^{q_2}$ so for all $t \in [0,T]$, $|A_t| < 2\left((1+C)\varepsilon^{q/3} + \varepsilon^{q_2}\right) \le 3\varepsilon^{q_2}$ for sufficiently small ε.

By Itô's Formula

$$\int_0^T a_t^2\,dt = \int_0^T a_t\,dA_t = a_T A_T - \int_0^T A_t(\beta_t dt + \gamma_t^i dw_t^i).$$

We have $|a_T A_T| < 3C\,\varepsilon^{q_2}$,

$$\left|\int_0^T A_T \beta_t dt\right| < 3C\,t_0\,\varepsilon^{q_2} \quad \text{and}$$

$$\langle Q,Q\rangle_t = \int_0^T A_t^2 |\gamma_t|^2 dt < 9C^2\,t_0\,\varepsilon^{2q_2}.$$

So, since $\omega \notin B_3$, choosing $\varepsilon_3 = 9C^2\,t_0\,\varepsilon^{2q_2}$:

$$|Q_T| = \left|\int_0^T A_t \gamma_t^i\,dw_t^i\right| < \delta_3 = \varepsilon^{q_3}$$

Therefore $\int_0^T a_t^2 dt < 3C(1+t_0)\,\varepsilon^{q_2} + \varepsilon^{q_3} \le 2\,\varepsilon^{q_3}$ for sufficiently small ε.

We have thus shown that for

$$\varepsilon_1 = C^2\,\varepsilon^q$$

$$\varepsilon_2 = (1+C^2)(2t_0+1)^{\frac{1}{2}}\,\varepsilon^{q_1/2} \quad \text{and}$$

$$\varepsilon_3 = 9C^2\,t_0\,\varepsilon^{2q_2},$$

for any $\omega \notin B_1 \cup B_2 \cup B_3$ such that $\int_0^T Y_t^2 dt < \varepsilon^q$ we have for sufficiently small ε (depending only, as the reader may easily check, on C, t_0, q and ν).

$$\int_0^T (|a_t|^2 + |u_t|^2)dt < 2\varepsilon^{q_3} + (1+C^2)(2t_0+1)^{\frac{1}{2}}\varepsilon^{q_1/2} < \varepsilon \text{ for}$$

sufficiently small ε.

It is furthermore clear that, for $i = 1,2,3$

$$\left(\delta_i^2 / 2\varepsilon_i\right)^{-1} = 0(\varepsilon^\nu) \quad \text{as } \varepsilon \to 0$$

with constants depending only on C, t_0, q and ν.

Remark

The above lemma is more powerful than we actually need. It suffices in Theorem 4.2 that

$$\mathbb{P}\left\{\int_0^T Y_t^2 dt < \varepsilon^q \text{ and } \int_0^T (|a_t|^2 + |u_t|^2)dt \geq \varepsilon\right\} = 0(\varepsilon^p) \quad \text{for all } p < \infty.$$

However, if it were necessary to establish that

$$\mathbb{E}\left[\exp(\nu|C_t^{-1}|)\right] < \infty \quad \text{for some } \nu > 0,$$

Lemma 4.1 would still provide good enough estimates.

Theorem 4.2

Suppose H_1 is satisfied at x, and $t > 0$.
Then $C_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$.

Proof:

In this proof $t > 0$ is fixed. Let K_ℓ be the set of vector fields appearing as brackets of length at most ℓ in H_1. Fix an integer ℓ such that K_ℓ spans $\mathbb{R}^d$ at x. Then

$$\delta \equiv \inf_{|v|=1} \left\{\sup_{K \in K_\ell} \langle K(x)|v\rangle^2\right\} > 0$$

For a given $B > 0$ define the stopping time

$$T = \inf\{s \geq 0 : |x_s - x| \geq 1/B \text{ or } |V_s - I| \geq 1/B\} \wedge t.$$

Then for $\varepsilon \in (0,t)$,

$$\{T \le \varepsilon\} = \{\sup_{s\le\varepsilon}|x_s - x| \vee \sup_{s\le\varepsilon}|V_s - I| \ge 1/B\}$$

By Proposition 1.2(b)

$$\mathbb{E}\left\{\sup_{s\le\varepsilon}|x_s - x|^p \vee \sup_{s\le\varepsilon}|V_s - I|^p\right\} = O(\varepsilon^{p/2}) \quad \text{for all } p < \infty.$$

It follows that $T^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$.

Since the coefficients X_i, $i = 0,1,\ldots,m$ and their derivatives are continuous, by choosing B sufficiently large we have

(a) $\sup_{s\le T}|V_s K(x_s)| \le B$ for all $K \in \mathcal{K}_{\ell+2}$

(b) For all $v \in S$ $(S \equiv \{u \in \mathbb{R}^d : |u| = 1\})$, there exists $K \in \mathcal{K}_\ell$ and a neighbourhood N of v in S such that

$$\inf_{s\le T,\, u\in N} \langle V_s K(x_s)|u\rangle^2 \ge \delta/2$$

We deduce immediately from (b) and the fact $T^{-1} \in L^p(\mathbb{P})$ for all $p<\infty$ that:

(c) For all $v \in S$, there exist $K \in \mathcal{K}_\ell$ and a neighbourhood N of v in S such that

$$\sup_{u\in N} \mathbb{P}\{\int_0^T \langle V_s K(x_s)|u\rangle^2 ds < \varepsilon\} \le \mathbb{P}\{\frac{\delta T}{2} < \varepsilon\} = O(\varepsilon^p) \quad \text{for all } p < \infty.$$

We divide the remainder of the proof into two parts.

<u>Claim 1</u>

(d) $\Rightarrow C_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$, where

(d) For all $v \in S$, there exist $i \in \{1,\ldots,m\}$ and a neighbourhood N of v in S with

$$\sup_{u\in N} \mathbb{P}\{\int_0^T \langle V_s X_i(x_s)|u\rangle^2 ds < \varepsilon\} = O(\varepsilon^p) \quad \text{for all } p < \infty.$$

<u>Claim 2</u>

(c) $\Rightarrow$ (d)

<u>Proof of Claim 1</u>

To show $C_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$, it suffices to show $(\det C_t)^{-1} \in L^p(\mathbb{P})$, for all $p < \infty$; so it suffices to show

$\lambda_{min}^{-1} \in L^p(\mathbb{P})$, for all $p < \infty$, (where λ_{min} is the smallest eigenvalue of C_t)

i.e. $\mathbb{P}\{\inf_{v\in S} \int_0^t \sum_{i=1}^m \langle V_s X_i(x_s)|v\rangle^2 ds < \varepsilon\} = O(\varepsilon^p)$, for all p.

So it suffices to show

$$\mathbb{P}\{\inf_{v\in S} \int_0^T \sum_{i=1}^m \langle V_s X_i(x_s)|v\rangle^2 ds < \varepsilon\} = O(\varepsilon^p), \quad \text{for all } p.$$

By our choice of T the random quadratic forms

$$v \to \int_0^T \sum_{i=1}^m \langle V_s X_i(x_s)|v\rangle^2 ds \text{ are uniformly Lipschitz on } S.$$

Denote their common Lipschitz constant by Θ and cover S with balls of radius ε/Θ, centre v_j. The number of these balls may be chosen less than $D(\varepsilon/\Theta)^{-d}$ for some fixed $D < \infty$. Note that

$$\int_0^T \sum_{i=1}^m \langle V_s X_i(x_s)|v\rangle^2 ds < \varepsilon \quad \text{for some} \quad v \in S$$

$$\Rightarrow \int_0^T \sum_{i=1}^m \langle V_s X_i(x_s)|v_j\rangle^2 ds < 2\varepsilon \quad \text{for some } j .$$

So $\mathbb{P}\{\inf_{v\in S} \int_0^T \sum_{i=1}^m \langle V_s X_i(x_s)|v\rangle^2 ds < \varepsilon\}$

$$\le D(\varepsilon/\Theta)^{-d} \sup_j \mathbb{P}\{\int_0^T \sum_{i=1}^m \langle V_s X_i(x_s)|v_j\rangle^2 ds < 2\varepsilon\} .$$

So to show $C_t^{-1} \in L^p(\mathbb{P})$ for all $p < \infty$ it suffices to show

$$\sup_{v\in S} \mathbb{P}\{\int_0^T \sum_{i=1}^m \langle V_s X_i(x_s)|v\rangle^2 ds < \varepsilon\} = O(\varepsilon^p) \text{ for all } p < \infty ,$$

which by compactness of S is equivalent to (d).

Proof of Claim 2

Let $v \in S$ and suppose (c) holds. Choose $K \in \mathcal{K}_\ell$ and a neighbourhood N of v in S with

$$\sup_{u\in N} \mathbb{P}\{\int_0^T \langle V_s K(x_s)|u\rangle^2 ds < \varepsilon\} = O(\varepsilon^p) \quad \text{for all } p < \infty.$$

We may write K in the form $\pm[X_{i_k},[\ldots,[X_{i_2},X_{i_1}]\ldots]]$ where $i_1,\ldots,i_k \in \{0,1,\ldots,m\}$, $i_1 \neq 0$ and $k \leq \ell$. Define

$$K_1 = X_{i_1}$$

$$K_j = [X_{i_j},K_{j-1}] \quad j = 2,\ldots,k$$

so $K = K_k$. We show by induction on j (decreasing) that for $j = 1,\ldots,k$,

$$\sup_{u\in N} \mathbb{P}\{\int_0^T \langle V_s K_j(x_s)|u\rangle^2 ds < \varepsilon\} = 0(\varepsilon^p) \quad \text{for all} \quad p < \infty$$

which completes the proof of (d) - with $i = i_1$.

By Itô's Formula

$$d(V_s K_{j-1}(x_s)) = V_s[X_i,K_{j-1}](x_s)dw_s^i$$

$$+ V_s([X_o, \; K_{j-1}](x_s) + \tfrac{1}{2}[X_i, [X_i,K_{j-1}]](x_s))ds.$$

Let $Y_s = \langle V_s K_{j-1}(x_s)|u\rangle$,

$$y = \langle K_{j-1}(x)|u\rangle,$$

$$a_s = \langle V_s([X_o \;,K_{j-1}](x_s) + \tfrac{1}{2}[X_i, [X_i,K_{j-1}]](x_s))|u\rangle$$

and

$$u_s^i = \langle V_s[X_i,K_{j-1}](x_s)|u\rangle.$$

It is easy to check the conditions of Lemma 4.1 hold for

$$Y_s = y + \int_0^s a_r dr + \int_0^s u_r^i dw_r^i, \quad \text{with C chosen independently of } u \in N.$$

So we have for $q > 8$,

$$\mathbb{P}\{\int_0^T \langle V_s K_{j-1}(x_s)|u\rangle^2 ds < \varepsilon^q \quad \text{and}$$

$$\int_0^T \Big(\langle V_s([X_o \;,K_{j-1}](x_s) + \tfrac{1}{2}[X_i,[X_i,K_{j-1}]](x_s))|u\rangle^2$$

$$+ \sum_{i=1}^m \langle V_s[X_i,K_{j-1}](x_s)|u\rangle^2\Big) ds \geq \varepsilon\}$$

$$= 0(\varepsilon^p) \quad \text{for all} \quad p < \infty \quad \text{uniformly in} \quad u \in N.$$

If $i_j \neq 0$ this is all that is required to complete the inductive step.

If $i_j = 0$, we apply Lemma 4.1 as above but with K_{j-1} replaced by $[X_i, K_{j-1}]$, $i = 1,\dots,m$ to deduce

$$\mathbb{P}\left\{\int_0^T \langle V_s[X_i,K_{j-1}](x_s)|u\rangle^2 ds < \varepsilon^q \text{ and } \int_0^T \langle V_s[X_i,[X_i,K_{j-1}]](x_s)|u\rangle^2 ds \geq \varepsilon\right\}$$

$$= O(\varepsilon^p) \text{ for all } p < \infty \quad (i \text{ is not summed}).$$

Hence

$$\mathbb{P}\left\{\int_0^T \langle V_s K_{j-1}(x_s)|u\rangle^2 ds < \varepsilon^{q^2} \text{ and } \int_0^T \Big(\sum_{i=1}^m \langle V_s[X_i,[X_i,K_{j-1}]](x_s)|u\rangle\Big)^2 ds > \varepsilon\right\}$$

$$= O(\varepsilon^p) \quad \text{for all} \quad p < \infty.$$

But then

$$\mathbb{P}\left\{\int_0^T \langle V_s K_{j-1}(x_s)|u\rangle^2 ds < \varepsilon^{q^2} \text{ and } \int_0^T \langle V_s[X_o \quad ,K_{j-1}](x_s)|u\rangle^2 ds \geq 3\varepsilon\right\}$$

$$= \quad O(\varepsilon^p) \quad \text{for all} \quad p < \infty$$

(using the first application of Lemma 4.1), which completes the inductive step. □

Finally, Theorems 3.2 and 4.2 combine to give:

<u>Theorem 4.3</u>

Let $X_0,\dots,X_m$ be C^∞ vector fields on $\mathbb{R}^d$. Let $\tilde{X}_0 \equiv X_0 + \frac{1}{2} DX_i \cdot X_i$. Suppose that $\tilde{X}_0, X_1,\dots,X_m$ have bounded derivatives and higher derivatives of polynomial growth. Suppose that H_1 is satisfied at some $x \in \mathbb{R}^d$. Then, for any $t > 0$, the solution x_t of the s.d.e.

$$\left.\begin{aligned} dx_t &= X_0(x_t)dt + X_i(x_t)\partial w_t^i \\ x_0 &= x \end{aligned}\right\}$$

has a C^∞ density with respect to Lebesgue measure on $\mathbb{R}^d$. □

References

1. K. Bichteler and D. Fonken, "A Simple Version of the Malliavin Calculus in Dimension One", Lecture Notes in Mathematics 939 (Springer 1982).

2. K. Bichteler and J. Jacod, "Calcul de Malliavin pour les Diffusions avec Sauts: Existence d'une Densité dans le cas Unidimensionel", Lecture Notes in Mathematics 986 (Springer 1983) 132-157.

3. J.M. Bismut, "Martingales, the Malliavin Calculus and Hypoellipticity under General Hormander's Conditions", Z. Wahrs 56 (1981) 469-505.

4. A.P. Carverhill and K.D. Elworthy, "Flows of Stochastic Dynamical Systems: The Functional Analytic Approach", Z. Wahrs 65 (1983) 245-267.

5. K.D. Elworthy, Stochastic Differential Equations on Manifolds (C.U.P. 1982).

6. D. Fonken, "A Simple Version of the Malliavin Calculus with Applications to the Filtering Equation" (Preprint).

7. J. Jacod, Calcul Stochastique et Problèmes de Martingales, Lecture Notes in Mathematics 714 (Springer 1979).

8. R. Leandre, "Un Exemple en Theorie des Flots Stochastiques", Lecture Notes in Mathematics 986 (Springer 1983) 158-161.

9. P.A. Meyer, "Variation des Solutions d'une E.D.S.", Lecture Notes in Mathematics 921 (Springer 1982) 151-164.

10. P.A. Meyer, "Malliavin Calculus, and some Pedagogy", (Preprint).

11. S.L. Sobolev, Applications of Functional Analysis in Mathematical Physics (Amer. Math. Soc., Providence 1963).

12. D. Stroock, "The Malliavin Calculus, Functional Analytic Approach", J. Funct. Anal. 44 (1981) 212-257.

13. D. Stroock, "Some Applications of Stochastic Calculus to Partial Differential Equations", Lecture Notes in Mathematics 976 (Springer 1983) 267-382.

James Norris,
Mathematical Institute,
24-29 St. Giles,
OXFORD

Present address:
Statistical Laboratory
16 Mill Lane
CAMBRIDGE