Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Functional Analysis 218 (2005) 130–149

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Divergence theorems in path space

Denis Bell

Department of Mathematics, University of North Florida, 4567 St. Johns Bluff Road South, Jacksonville, FL 32224, USA

Received 31 July 2003; received in revised form 9 July 2004; accepted 19 July 2004
Communicated by Paul Malliavin
Available online 13 October 2004

This paper is dedicated to the memory of Phyllis Bell

Abstract

We obtain divergence theorems on the solution space of an elliptic stochastic differential equation defined on a smooth compact finite-dimensional manifold M. The resulting divergences are expressed in terms of the Ricci curvature of M with respect to a natural metric on M induced by the stochastic differential equation. The proofs of the main theorems are based on the lifting method of Malliavin together with a fundamental idea of Driver.

Keywords: Elliptic stochastic differential equation; Compact manifold; Integration by parts formula; Divergence theorem; Path space

1. Introduction

Let M denote a compact d-dimensional C^∞ manifold, o a point in M, and $X_1, \ldots, X_n$ and V smooth vector fields defined on M. Consider the following stochastic differential equation (SDE):

$$dx_t = \sum_{i=1}^{n} X_i(x_t) \circ dw_i + V(x_t)\, dt, \quad t \in [0, T],$$
$$x_0 = o. \tag{1.1}$$

E-mail address: dbell@unf.edu.

0022-1236/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jfa.2004.07.010

A *divergence theorem* (for the law of x) is a statement asserting that for each Z in a class of admissible vector fields on the path space $C_o(M)$, there exists a random variable $Div(Z)$ such that the identity

$$E[(Z\Phi)(x)] = E[\Phi(x)Div(Z)] \tag{1.2}$$

holds for a dense set of smooth functions Φ on $C_o(M)$. There are two approaches to divergence theorems on solution spaces to stochastic differential equations. Both work by reducing the problem to one of integration by parts over the (flat) Wiener space.

The first, which we shall refer to as the *lifting* approach, was introduced by Malliavin [M] in order to study the hypoellipticity of the differential operator $L = 1/2\sum_{i=1}^{n} X_i^2 + V$. In this context it suffices to obtain a special case of (1.2) where Φ is a function of the path x at an (arbitrary) fixed positive time. Let g denote the Itô map: $w \mapsto x$ defined on the space of Wiener paths w, and let $g_T = g \circ e_T$ denote its composition with the evaluation map e_T at time $T > 0$. The underlying idea is to lift smooth vector fields on M to the path space $C_0(R^n)$ by the map g_T. The Cameron–Martin space H of R^n plays a central role here. Defining the tangent bundle to $C_0(R^n)$ to be the trivial bundle $C_0(R^n) \times H$ and given a smooth vector field Z on M, one constructs a vector field r on $C_0(R^n)$ such that the following diagram commutes:

$$\begin{array}{ccc} C_0(R^n) \times H & \xrightarrow{dg_T} & TM \\ r \uparrow & & \uparrow Z \\ C_0(R^n) & \xrightarrow[g_T]{} & M \end{array}$$

(Note that since g is *non-differentiable* in the classical sense the map dg above must be interpreted in the extended sense of the Malliavin calculus). By construction we have, for smooth functions ϕ on M

$$E[(Z\phi)(x_T)] = E[r(\phi \circ g_T)(w)].$$

This can then be transformed by applying a classical divergence theorem on Wiener space (cf. e.g. [Be, Chapter 4]), resulting in the formula

$$E[(Z\phi)(x_T)] = E[\phi(x_T)Div(r)].$$

In order to construct the lift r of Z, a certain level of non-degeneracy must be assumed in the SDE (1.1). It sufficies to assume that the vector fields $X_1, \ldots, X_n$ satisfy the Hörmander condition at o (cf. e.g. [Bi,B-M] for a more general condition).

The second approach is due to Driver [D]. Assuming now that M is Riemannian, his work yields a divergence theorem on the space of paths Σ defined as the image of a Brownian motion β on T_oM under the stochastic development $s : \beta \mapsto \Sigma$. Here the admissible vector fields are those of the form $Z_t \equiv U_t h_t$, where h_t is an arbitrary path in the Cameron–Martin space of T_oM and U_t denotes (stochastic) parallel translation along the path $\{x_s,\ 0 \leqslant s \leqslant t\}$. The underlying idea is as follows: consider the push-forward $\tilde{Z}$ to $C_0(R^d)$ of the vector field Z under the antidevelopment map $a : \Sigma \mapsto \beta$ (i.e. define $\tilde{Z}(\beta) \equiv Z(a)(\Sigma)$). For smooth maps Φ on $C_o(M)$ one has

$$E[(Z\Phi)(\Sigma)] = E[\tilde{Z}(\Phi \circ s)(\beta)].$$

The essential point is that the path $\tilde{Z}$ has the form

$$\tilde{Z}_t = \int_0^t A(s)\, d\beta_s + \int_0^t B(s)\, ds, \tag{1.3}$$

where $A(s)$ is an adapted *skew-symmetric* $d \times d$ matrix-valued process defined in terms of the curvature tensor of M. It is a consequence of the infinitessimal rotational-invariance of the law of Brownian motion that the divergence of such a term exists and is zero. Thus the divergence of $\tilde{Z}$ arises purely from the drift term in the RHS of (1.3). With the aid of the Girsanov theorem, $Div(\tilde{Z})$ can be computed as an Itô integral. This leads to the formula

$$E[(Z\Phi)(\Sigma)] = E\left[\Phi(\Sigma) \int_0^T \langle B(s), d\beta_s \rangle\right],$$

where $\langle .,. \rangle$ denotes the Euclidean inner product.

Suppose now that the SDE (1.1) is *elliptic*, i.e. the vector fields $X_1, \ldots, X_n$ span T_xM, for all $x \in M$. Then these vector fields induce a metric on M with respect to which the antidevelopment of x onto T_0M has the form Brownian motion plus drift. Driver's argument thus yields a divergence theorem with respect to the paths $x \in C_o(M)$. The metric induced by $X_1, \ldots, X_n$ is defined in Section 2 and plays an important role in the present work.

The purpose of this article is to derive path-space divergence theorems by the *lifting* method. The layout is as follows. In Section 2 we establish notation, review some basic geometric constructs, and derive some (largely known) results on which the proofs in Section 3 depend. In particular, Theorem 2.3 shows that the divergence of an adapted process with values in the Cameron–Martin space exists as an Itô integral. Theorem 2.4 gives the result that the diffusion term in (1.3) is divergence free, provided the integrand A is adapted and takes values in the space of skew-symmetric matrices. These two results are easily derived from well-known theorems in stochastic analysis (the former from the Girsanov theorem and the latter from the infinitessimal rotational-invariance of the multi-dimensional Wiener process).

The main results of the paper are contained in Section 3. Theorem 3.1 gives a stochastic differential equation relating a class of paths Z in $T(C_oM)$ and their liftings to $T(C_0(R^n))$. This provides an integration by parts formula on the path space $C_o(M)$; however in general Z will not be a vector field on $C_o(M)$ for the reason given in Remark 3.3. The main results are Theorems 3.4 and 3.5, in which we obtain divergences for classes of vector fields on $C_o(M)$. These involve the Ricci curvature of M, and in this sense our results are similar to existing path-space divergence theorems (cf. [D] or [E-L]). The key point that emerges from Driver's work is that the Cameron–Martin space is too small to allow for the lifting of vector fields on $C_0(M)$. In order to obtain interesting divergence theorems in path space it is necessary to work with an augmented tangent bundle on $C_0(R^n)$ that includes stochastic integrals of adapted skew-symmetric matrix processes. The proofs of Theorems 3.4 and 3.5 exploit this fact in an essential way.

In Section 4, we consider the special case where M is an embedded submanifold of Euclidean space. In this case we are able to express our formulae in terms of stochastic parallel translation along the path x and thereby put them into a more familiar form.

We note that, although stated for the case of a compact manifold M, the methods of this paper will work in the non-compact case (e.g. if M is a Euclidean space) under additional boundedness hypotheses. The compactness condition has been assumed in order to simplify the statements of the theorems. Another point is that, although the present paper addresses the elliptic case, our work provides some insight into the non-elliptic case. (cf. Remark 3.2). This point will be developed in a later paper.

2. Notation and background material

The following notation will be assumed throughout the paper. Let each X_i, $1 \leqslant i \leqslant n$ have the expression $X_i = a_{ij}\partial/\partial x_j$ in a local coordinate system. We define a Riemannian structure $[g_{jk}]$ on M by

$$g^{jk} \equiv \sum_{i=1}^{n} a_{ij}a_{ik} \tag{2.1}$$

(note that the ellipticity condition implies that the matrix $[g^{jk}] \in Gl(d)$). From this point on, it will be assumed that whenever an index in a product is repeated, that index is summed on.

Let $\Gamma(M)$ denote the set of vector fields on M. For all $Z \in \Gamma(M)$, we have

$$Z = \langle Z, X_i \rangle X_i. \tag{2.2}$$

To see this, write $Z = Z_j\partial/\partial x_j$. Then the RHS of (2.2) is

$$Z_j a_{ik} g_{jk} a_{il} \partial/\partial x_l = Z_j g^{kl} g_{jk} \partial/\partial x_l = Z_l \partial/\partial x_l = Z.$$

In a similar fashion, we see that for $Z, W \in \Gamma(M)$

$$\langle Z, W\rangle = \langle Z, X_i\rangle\langle W, X_i\rangle. \tag{2.3}$$

Let ∇ denote the Levi–Civita covariant derivative (with respect to the above metric) and R the corresponding Riemann curvature tensor, defined for $A, B \in \Gamma(M)$ by

$$R(A, B) = \nabla_A\nabla_B - \nabla_B\nabla_A - \nabla_{[A,B]},$$

where $[A, B]$ is the Lie bracket of A and B.

The Ricci curvature tensor *Ric* is the section of $L(TM, TM)$ defined by

$$Ric(Z) = R(Z, e_i)e_i$$

for any (locally defined) orthonormal frame [1] $\{e_i, i = 1, \dots, d\}$ in *TM* and $Z \in \Gamma(M)$.

We also define the *horizontal Laplacian* Δ_H acting on $\Gamma(M)$ by

$$\Delta_H = \nabla_{e_i}\nabla_{e_i} - \nabla_{\nabla_{e_i}e_i},$$

where $\{e_i, i = 1, \dots, d\}$ are as above.

The next result expresses these objects in terms of the vector fields $X_1, \dots, X_n$.

Proposition 2.1. *For all* $Z \in G$
(i) $Ric(Z) = R(Z, X_i)X_i$.
(ii) $\Delta_H Z = \nabla_{X_i}\nabla_{X_i}Z - \nabla_{\nabla_{X_i}X_i}Z$.

Proof. To prove (i), let $\{e_i, i = 1, \dots, d\}$ in *TM* and write $X_i = c_{ij}e_j$. Then $c_{ij} = \langle X_i, e_j\rangle$. Using (2.3) we have

$$\begin{aligned} R(Z, X_i)X_i &= c_{ij}c_{ik}R(Z, e_j)e_k \\ &= \langle X_i, e_j\rangle\langle X_i, e_k\rangle R(Z, e_j)e_k \\ &= \langle e_j, e_k\rangle R(Z, e_j)e_k = Ric(Z). \end{aligned}$$

[1] In passing, we note that it is easy to construct such a frame from the local representations of the vector fields $X_1, \dots, X_n$. Let g denote the matrix $[g_{jk}]$ and let diag $\{\lambda_1, \dots, \lambda_d\} = ugu^*$ be a diagonalization of g by an orthogonal matrix u. Then the vectors $e_j \equiv u_{jk}\partial/\partial x_k, 1 \leqslant j \leqslant d$, have the property $\langle e_j, e_k\rangle = \delta_{jk}\lambda_k$, so $\{e_j/\sqrt{\lambda_j}\}$ is an orthonormal frame.

To prove (ii), we again use (2.3) to obtain

$$\begin{aligned}
&\nabla_{X_i}\nabla_{X_i} - \nabla_{\nabla_{X_i}X_i} \\
&\quad = c_{ij}\nabla_{e_j}(c_{ik}\nabla_{e_k}) - c_{ij}\nabla_{\nabla_{e_j(c_{ik}e_k)}} \\
&\quad = c_{ij}c_{ik}\nabla_{e_j}\nabla_{e_k} + c_{ij}e_j(c_{ik})\nabla_{e_k} \\
&\qquad -c_{ij}c_{ik}\nabla_{\nabla_{e_j}e_k} - c_{ij}e_j(c_{ik})\nabla_{e_k} \\
&\quad = \nabla_{e_i}\nabla_{e_i} - \nabla_{\nabla_{e_i}e_i} = \Delta_H. \qquad \square
\end{aligned}$$

The following result will be used in Section 3.

Theorem 2.2. *The following three conditions are equivalent:*

(i) *the* 1-*form* X_i^* *is closed.*

(ii) X_i *satisfies the anti-Killing condition* $\langle \nabla_Y X_i, Z\rangle = \langle \nabla_Z X_i, Y\rangle$ *for all vector fields* Y *and* Z.

(iii) *For all* $1 \leqslant s, m \leqslant d$

$$\frac{\partial}{\partial x_s}(a_{ip}g_{pm}) = \frac{\partial}{\partial x_m}(a_{ip}g_{ps}).$$

Proof. For each $1\leqslant i, j\leqslant n$

$$\langle X_i, X_j\rangle = a_{im}a_{jp}g_{mp}.$$

Applying X_k to each side of this equation and using the metric compatability of ∇ gives

$$\langle \nabla_{X_k}X_i, X_j\rangle + \langle X_i, \nabla_{X_k}X_j\rangle = X_k(a_{im}a_{jp}g_{mp}).$$

Interchanging j and k, subtracting and using the symmetry of ∇, we get

$$\begin{aligned}
&\langle \nabla_{X_k}X_i, X_j\rangle - \langle \nabla_{X_j}X_i, X_k\rangle + \langle X_i, [X_k, X_j]\rangle \\
&\quad = X_k(a_{im}a_{jp}g_{mp}) - X_j(a_{im}a_{kp}g_{mp}).
\end{aligned} \tag{2.4}$$

However, a direct computation shows that

$$(2.4) = \langle X_i, [X_k, X_j]\rangle + \frac{\partial}{\partial x_s}(a_{ip}g_{pm})(a_{ks}a_{jm} - a_{js}a_{km}).$$

Thus

$$\langle \nabla_{X_k} X_i, X_j \rangle - \langle \nabla_{X_j} X_i, X_k \rangle = \frac{\partial}{\partial x_s}(a_{ip} g_{pm})(a_{ks} a_{jm} - a_{js} a_{km}). \tag{2.5}$$

It follows that (iii) implies (ii). In order to prove the converse, let A^* and N denote the matrices $[a_{pq}]$ and $[\partial/\partial_{x_s}(a_{ip} g_{pm})]$, respectively. Then (ii) and (2.5) yield the matrix equation

$$A^*(N - N^*)A = 0. \tag{2.6}$$

Since the ellipticity condition implies the matrix AA^* is invertible, we deduce from (2.6) that $N = N^*$, i.e. (iii) holds.

Finally, the equivalence of (i) and (iii) follows easily from the fact that

$$X_i^* = a_{ij} g_{jm} dx_m,$$

which implies

$$dX_i^* = \frac{\partial}{\partial x_s}(a_{ij} g_{jm})\, dx_s \wedge dx_m.$$

Thus X_i^* is closed if and only if (iii) holds. □

We now state and prove two results concerning the existence of divergences for two classes of vector fields on the classical Wiener space.

Theorem 2.3. *Let* $h : \Omega \times [0, T] \mapsto R^n$ *denote a bounded continuous adapted path. Then*

$$Div\Big[\int_0^\cdot h_s\, ds\Big] = \int_0^T \langle h_s, dw_s \rangle. \tag{2.7}$$

Proof. This follows immediately from the Girsanov theorem, which implies that for $\Phi \in C_b^\infty(C_0(R^n))$ and $\varepsilon \in R$

$$E\Big[\Phi(w + \varepsilon \int_0^\cdot h_s\, ds)\Big] = E\Big[\Phi(w) \exp\Big\{\varepsilon \int_0^T \langle h_s, dw_s \rangle - \frac{\varepsilon^2}{2} \int_0^T ||h_s||^2\, ds.$$

Differentiating with respect to ε and setting $\varepsilon = 0$ gives (2.7). □

Theorem 2.4.[2] *Let* $A : \Omega \times [0, T] \mapsto so(n)$ (*the set of* $n \times n$ *skew symmetric matrices*) *be a bounded continuous adapted process. Then*

$$Div\Big[\int_0^{\cdot} A_s\, dw_s\Big] = 0. \tag{2.8}$$

Proof. Define a process $\theta_t^\varepsilon = \exp \varepsilon(A_t), \varepsilon \in R$, where exp denotes the matrix exponential. Then θ_t^ε is an adapted $O(n)$-valued matrix process for all ε. It follows from the (infinitessimal) rotation invariance of Brownian motion that the process

$$\int_0^{\cdot} \theta_t^\varepsilon\, dw_t$$

has the same law as w. Thus for $\Phi \in C_b^\infty(C_0(R^n))$, we have

$$E[\Phi(\theta^\varepsilon)] = E[\Phi(w)].$$

Differentiating with respect to ε and setting $\varepsilon = 0$ gives (2.8). □

3. Divergence theorems in path space

We define a tangent vector to $C_o(M)$ at a point x to be a path $Z : [0, T] \mapsto TM$ such that $Z_t \in T_{x_t}M$ for each $t \in [0, T]$. The tangent space to a path $w \in C_0(R^n)$ will be defined as the set of paths r_t of the form

$$r_t = \int_0^t A(s)\, dw_s + \int_0^t B(s)\, ds,$$

where A is a continuous adapted $so(n)$-valued process and B is a continuous adapted R^n-valued process. Following the scheme outlined in the Introduction, the objective is to construct a family of vector fields Z on $C_o(M)$ which lift to vector fields r on $C_0(R^n)$. We then compute $Div(r)$ by means of Theorems 2.3 and 2.4. This yields divergence theorems on the space of paths x.

The following result describes the lifts to $C_0(R^n)$ of a class of paths in *TM*.

Theorem 3.1. *Let* $r = (r^1, \ldots, r^n) : \Omega \times [0, T] \mapsto R^n$ *denote a semimartingale of the form*

$$dr_t^k = b_t^{kj}\, dw_j + c_t^k\, dt,$$

[2] I learned of this result and its proof from Bruce Driver.

where b_{kj} and $c^k, 1 \leqslant j, k \leqslant n$ are adapted continuous process such that

$$\sup_{t\in[0,T]} \Big[\sum_{j,k} |b_t^{kj}| + \sum_k |c_t^k| \Big] \in L^p, \quad \forall p \geqslant 1.$$

As before, let g denote the map $w \mapsto x$, where x is the solution to the SDE (1.1). *Then* (*in the extended sense of the Malliavin calculus*)

$$dg(w)r = X_i(x_t)h_t^i, \tag{3.1}$$

where $h = (h_1, \ldots, h^n)$ satisfies the SDE

$$h_t^k = r_t + \int_0^t \langle [X_j, X_i], X_k\rangle(x_s)h_s^j \circ dw_i + \int_0^t \langle \nabla_{X_j} V, X_k\rangle(x_s)h_s^j \, ds. \tag{3.2}$$

Proof. The idea is to assume that (3.1) holds and show that (3.2) gives the relationship between r and h and Set

$$\eta_t = dg(w)r_t = X_i(x_t)h_t^i.$$

We will find an equation for η_t differentiating with respect to w in the SDE (1.1) in the sense of the Malliavin calculus. To this end, define a sequence of perturbations w^ε of w by paths in the Cameron–Martin space such that $w^0 = w$ and $(dw^\varepsilon/d\varepsilon)(\varepsilon = 0) = r$. Then $x^\varepsilon \equiv g(w^\varepsilon)$ is defined and satisfies

$$dx^\varepsilon = X_i(x_t^\varepsilon)\, dw_i^\varepsilon + V(x_t^\varepsilon)\, dt. \tag{3.3}$$

Let $\nabla_t Z$ denote the covariant time differential of a vector field Z_t along x_t. Differentiating with respect to ε and setting $\varepsilon = 0$ in (3.3), we have

$$\nabla_\varepsilon/_{\varepsilon=0}\, dx^\varepsilon = \nabla_{\eta_t} X_i \circ dw_i + X_i(x_t)\, dr^i + \nabla_{\eta_t} V\, dt.$$

In view of the torsion-free property of ∇ we can write this equation as

$$\begin{aligned}\nabla_t \eta &= \nabla_{\eta_t} X_i \circ dw_i + X_i(x_t)\, dr^i + \nabla_{\eta_t} V\, dt \\ &= \nabla_{X_j} X_i h^j \circ dw_i + \nabla_{X_j} V h^j\, dt + X_i(x_t)\, dr^i. \end{aligned} \tag{3.4}$$

On the other hand, taking the covariant time differential in the equation $\eta_t = X_i(x_t)h_i$,

$$\begin{aligned}\nabla_t \eta &= \nabla_{dx_t} X_j h_j \circ dw_i + X_i dh^i \\ &= \nabla_{X_i} X_j h_j \circ dw_i + X_i dh^i.\end{aligned} \tag{3.5}$$

Equality of (3.4) and (3.5) implies

$$\begin{aligned}X_i dh^i &= (\nabla_{X_j} X_i - \nabla_{X_i} X_j) h^j \circ dw_i + \nabla_{X_j} V h^j \, dt + X_i \, dr^i \\ &= [X_j, X_i] h^j \circ dw_i + \nabla_{X_j} V h^j \, dt + X_i \, dr^i.\end{aligned} \tag{3.6}$$

Using (2.2) we can solve this equation for dh^i. The result is (3.2). □

Remark 3.2. Eq. (3.6) provides some insight into the situation in the *non-elliptic* case (when the vector fields $X_1, \ldots, X_n$ fail to span *TM* at every point of *M*). Provided the span of the Lie brackets $\{[X_j, X_i], 1 \leqslant i \leqslant n\}$ is contained within the span of $\{X_1, \ldots, X_n\}$, it is still possible to solve Eq. (3.6) and obtain an SDE for h (this point will be taken up in a later paper). On the other hand, if span $\{[X_j, X_i], 1 \leqslant i \leqslant n\}$ *exceeds* span $\{X_1, \ldots, X_n\}$, then Eq. (3.6) is *non-solvable*, except under very specific conditions on the drift *V*. Thus, in the non-elliptic case a Hörmander-type condition seems to be *detrimental*. We are inclined to believe that a divergence theorem does not exist in this situation. This is in marked contrast to the case where the central object of study is the law of x_T *at a fixed time T* (cf. Section 1), when the Hörmander condition guarantees a solution to the lifting problem in the non-elliptic setting.

Remark 3.3. In order for the ellipticity condition to be satisfied, it will generally be necessary to assume $n > d$ (e.g. if *M* has non-zero Euler characteristic, then each X_i will necessarily vanish at least one point in *M*). In particular, Eq. (1.1) cannot be solved to give *w* as a function of *x*. For this reason the path $X_i(x_t)h_t^i$ in (3.1) cannot reasonably be regarded as a vector field at *x*, since it depends on *w* via Eq. (3.2) (a similar point is made in [E-L]). The problem is that the Lie bracket term $[X_j, X_i]$ in (3.2) is *non-tensorial* in X_i and as such it cannot be combined with dw_i to create a term in *dx*. We now address this problem.

Define on *M* a set of 1-forms $T_{kj}, 1 \leqslant j, k \leqslant n$ by

$$T^{kj}(Z) = \langle \nabla_Z X_k, X_j \rangle - \langle \nabla_Z X_j, X_k \rangle \tag{3.7}$$

and a set of real-valued functions B^{kj} by

$$B^{kj} = -1/2 \big\langle \Delta_H X_k + \langle [X_i, X_k], X_p \rangle [X_i, X_p], X_j \big\rangle + \langle \nabla_{X_j} V, X_k \rangle. \tag{3.8}$$

Theorem 3.4. *Suppose that the* 1*-forms* X_i^*, $1 \leqslant i \leqslant n$, *are closed. Let* $r = (r^1, \ldots, r^n)$ *denote a Cameron–Martin path in* R^n *and define* $\eta = (\eta^1, \ldots, \eta^n)$ *to be the solution to the SDE*

$$d\eta_t^k = \{\dot{r}_t^k + B^{kj}(x_t)\eta^j\}\,dt + T^{kj}(\circ dx_t)\eta_t^j, \quad \eta_0^k = 0. \tag{3.9}$$

Let Z *denote the path* $Z_t = X_i(x_t)\eta_t^i$, *considered as a vector field on the path space* $C_o(M)$. *Then for every smooth cylindrical function* Φ *on* $C_o(M)$

$$E[(Z\Phi)(x)] = E\Big[\Phi(x)\int_0^T \big(\dot{r}_t^k + \frac{1}{2}\langle Ric(X_k)(x_t), Z_t\rangle\big)\,dw_k\Big]. \tag{3.10}$$

Proof. By Theorem 2.2 we have for all $1 \leqslant i, j, k \leqslant n$

$$\langle \nabla_{X_j} X_i, X_k\rangle = \langle \nabla_{X_k} X_i, X_j\rangle. \tag{3.11}$$

This implies

$$\langle [X_j, X_i], X_k\rangle - \langle [X_k, X_i], X_j\rangle = T^{kj}(X_i). \tag{3.12}$$

Thus we can write (3.9) in the form

$$\begin{aligned} d\eta^k =& (\dot{r}^k + C^{kj})\eta^j\,dt - \langle [X_k, X_i], X_j\rangle\eta^j \circ dw_i \\ &+ \langle [X_j, X_i], X_k\rangle\eta^j \circ dw_i + \langle \nabla_{X_j} V, X_k\rangle\eta^j\,dt, \end{aligned}$$

where

$$C^{kj} \equiv 1/2\Big\langle -\Delta_H X_k - \langle [X_i, X_k], X_p\rangle[X_i, X_p], X_j\Big\rangle.$$

Defining

$$d\tilde{r}^k = (\dot{r}^k + C^{kj})\eta^j\,dt - \langle [X_k, X_i], X_j\rangle\eta^j \circ dw_i, \tag{3.13}$$

it follows from Theorem 3.1 that

$$dg(w)\tilde{r} = X_i(x_\cdot)\eta_\cdot^i = Z.$$

Note that the term $\langle [X_k, X_i], X_j\rangle$ is *skew-symmetric* in the k and i indices. Thus the divergence $Div(\tilde{r})$ of $\tilde{r}$ exists by Theorems 2.3 and 2.4 and we obtain

$$\begin{aligned} E[Z\Phi(x)] &= E[d(\Phi\circ g)(w)\tilde{r}] \\ &= E[\Phi(x)Div(\tilde{r})]. \end{aligned}$$

In order to compute $Div(\tilde{r})$ we must convert the Stratonovich integral in (3.13) into Itô form. The correction term is

$$\begin{aligned} &-\tfrac{1}{2}d\big(\langle [X_k, X_i], X_j\rangle(x_t)\eta^j\big)\, dw_i \\ &\quad = -\tfrac{1}{2}\big(\langle \nabla_{X_i}\nabla_{X_k}X_i - \nabla_{X_i}\nabla_{X_i}X_k, X_j\rangle + (\langle [X_k, X_i], \nabla_{X_i}X_j\rangle\big)(x_t)\eta^j\, dt \\ &\qquad -\tfrac{1}{2}(\langle [X_k, X_i], X_j\rangle(x_t)\, d\eta^j\, dw_i. \end{aligned} \tag{3.14}$$

Eq. (3.9) gives

$$d\eta^j\, dw_i = \{\langle \nabla_{X_i}X_j, X_r\rangle - \langle \nabla_{X_i}X_r, X_j\rangle)\}\eta^r\, dt.$$

Thus

$$\begin{aligned} &(\langle [X_k, X_i], X_j\rangle\, d\eta^j dw_i \\ &\quad = (\langle [X_k, X_i], X_p\rangle\{\langle \nabla_{X_i}X_p, X_j\rangle - \langle \nabla_{X_i}X_j, X_p\rangle\}\eta^j\, dt \\ &\quad = \langle [X_k, X_i], X_p\rangle\{\langle \nabla_{X_i}X_p, X_j\rangle - \langle [X_k, X_i], \nabla_{X_i}X_j\rangle\}\eta^j\, dt. \end{aligned}$$

Substituting this into (3.14), we obtain for the correction term

$$-\tfrac{1}{2}\big(\langle \nabla_{X_i}\nabla_{X_k}X_i - \nabla_{X_i}\nabla_{X_i}X_k, X_j\rangle + \langle [X_k, X_i], X_p\rangle\langle \nabla_{X_i}X_p, X_j\rangle\big)\eta^j dt. \tag{3.15}$$

From the definition of $\langle ., .\rangle$ we have $\langle X_i, X_i\rangle = d$ which implies for all $1 \leqslant j \leqslant n$

$$\langle \nabla_{X_i}X_i, X_j\rangle = \langle \nabla_{X_j}X_i, X_i\rangle = 0.$$

Thus

$$\nabla_{X_i}X_i = 0.$$

This implies (cf. Proposition 2.1)

$$\nabla_{X_i}\nabla_{X_k}X_i = -Ric(X_k) + \nabla_{[X_i,X_k]}X_i$$

and

$$\nabla_{X_i}\nabla_{X_i} = \Delta_H X_k$$

so

$$\begin{aligned}(3.15) &= \tfrac{1}{2}\Big(\langle Ric(X_k) - \nabla_{[X_i,X_k]}X_i + \Delta_H X_k, X_j\rangle - \langle [X_k, X_i], X_p\rangle\langle \nabla_{X_i}X_p, X_j\rangle\Big)\eta^j dt \\ &= \tfrac{1}{2}\Big\langle Ric(X_k) + \Delta_H X_k + \langle [X_k, X_i], X_p\rangle [X_p, X_i], X_j\Big\rangle \eta^j dt.\end{aligned}$$

Substituting this into (3.13) gives the following Itô differential for $\tilde{r}$

$$\begin{aligned}d\tilde{r}^k &= \Big(\dot{r}^k + C^{kj} + \tfrac{1}{2}(\langle Ric(X_k) + \Delta_H X_k \\ &\quad - \langle [X_k, X_i], X_p\rangle [X_p, X_i], X_j\rangle\Big)\eta^j dt - \langle [X_k, X_i], X_j\rangle \eta^j\, dw_i \\ &= \dot{r}^k + \tfrac{1}{2}\langle Ric(X_k), X_j\eta^j\rangle - \langle [X_k, X_i], X_j\rangle \eta^j\, dw_i.\end{aligned}$$

It follows from Theorems 2.3 and 2.4 that

$$Div(\tilde{r}) = \int_0^T \left(\dot{r}^k + \frac{1}{2}\langle Ric(X_k), X_j\eta^j\rangle\right) dw_k$$

which gives (3.10) and completes the proof. □

In order to obtain a more general theorem, we introduce the following functions: for vectors U and $W \in T_0 M$, define

$$F_t(U, W) = \langle Y_t U, Y_t W\rangle, \tag{3.16}$$

where $Y_t : T_o M \mapsto T_{x_t} M$ is the derivative $dg_t(o)$, where $g_t : M \mapsto M$ is the stochastic flow $x_0 \mapsto x_t(w)$. The stochastic flow is a.s. a C^∞ map on M (cf. [E]). Formal differentiation in (1.1) yields the following covariant SDE for Y_t:

$$\nabla_t Y = \nabla_{Y_t} X_i \circ dw_i + \nabla_{Y_t} V dt, \tag{3.17}$$

$$Y_0 = I_{T_o M}.$$

For each $1 \leqslant j, k \leqslant n$, define

$$dG_t^{kj} = dF_t(U, W)/_{U = Y_t{}^{-1} X_k(x_t),\, W = Y_t{}^{-1} X_j(x_t)}.$$

The functions G^{kj} measure the extent to which the stochastic flow fails to be an isometry on M.

Differentiating in (3.16) and using (3.17) gives

$$dG_t^{kj} = \langle \nabla_{X_j} X_i(x_t) \circ dw_i + \nabla_{X_j} V(x_t)\, dt, X_k(x_t)\rangle \\ + \langle X_j(x_t), \nabla_{X_k} X_i(x_t) \circ dw_i + \nabla_{X_k} V(x_t)\, dt\rangle. \tag{3.18}$$

We now set

$$\tilde{T}^{kj}(Z) = \langle \nabla_Z X_k, X_j\rangle + \langle \nabla_Z X_j, X_k\rangle, \quad Z \in \Gamma(M).$$

For all $1 \leqslant j, k \leqslant n$ and $Z \in \Gamma(M)$, we now have

$$\langle [X_j, X_i], X_k\rangle = \langle [X_i, X_k], X_j\rangle + \langle \nabla_{X_j} X_i, X_k\rangle + \langle \nabla_{X_k} X_i, X_j\rangle - \tilde{T}^{kj}(X_i).$$

Combining this with (3.18) gives

$$\langle [X_j, X_i], X_k\rangle(x_t) \circ dw_i = \langle [X_i, X_k], X_j\rangle(x_t) \circ dw_i + dG_t^{jk} - \tilde{T}_{jk}(dx_t) \\ - \{\langle \nabla_{X_j} V, X_k\rangle(x_t) + \langle \nabla_{X_k} V, X_j\rangle(x_t)\}\, dt. \tag{3.19}$$

Define functions $\tilde{B}^{kj}, 1 \leqslant j, k \leqslant n$ on M by

$$\tilde{B}^{kj} = -\langle \nabla_{X_k} V, X_j\rangle + \tfrac{1}{2}\Big(\big\langle \Delta_H X_k - [X_k, \nabla_{X_i} X_i] + \langle [X_k, X_i], X_p\rangle \nabla_{X_i} X_p, X_j\big\rangle \\ - \langle \nabla_{X_j} X_i, [X_k, X_i]\rangle\Big). \tag{3.20}$$

We are now in a position to prove the following

Theorem 3.5. *Let $r = (r^1, \ldots, r^n)$ denote any Cameron–Martin path and define $\eta = (\eta^1, \ldots, \eta^n)$ where $\eta = (\eta^1, \ldots, \eta^n)$ is the solution to the SDE*

$$d\eta_k^k = \{\dot{r}_k + \tilde{B}^{kj}(x_t)\eta^j\}\, dt - \tilde{T}_{jk}(\circ\, dx_t)\eta^j + dG_t^{jk}\eta^j. \tag{3.21}$$

Let Z denote the path $Z_t = X_i(x_t)\eta_t^i$. Then for every smooth cylindrical function Φ on $C_o(M)$

$$E[(Z\Phi)(x)] = E\Big[\Phi(x) \int_0^T \big(\dot{r}_t^k - \frac{1}{2}\langle Ric(X_k)(x_t), Z_t\rangle\, dw_k\big)\Big]. \tag{3.22}$$

Proof. The proof follows similar lines to that of Theorem 3.4. Define $\tilde{r} = (\tilde{r}^1, \ldots), \tilde{r}^n$ to be the path

$$d\tilde{r}^k = (\dot{r}^k + \tilde{C}^{kj})\eta^j dt - \langle [X_i, X_k], X_j\rangle \eta^j \circ dw_i \tag{3.23}$$

where

$$\tilde{C}^{kj} = \frac{1}{2}\Big(\big\langle \Delta_H X_k - [X_k, \nabla_{X_i} X_i] + \langle [X_k, X_i], X_p\rangle \nabla_{X_i} X_p, X_j\big\rangle - \langle \nabla_{X_j} X_i, [X_k, X_i]\rangle\Big).$$

It follows from Theorem 3.1 and Eq. (3.19) that $dg(w)\tilde{r} = Z$. As before, Theorems 2.3 and 2.4 imply

$$\begin{aligned} E[Z\Phi(x)] &= E[d(\Phi \circ g)(w)\tilde{r}] \\ &= E[\Phi(x) Div(\tilde{r})]. \end{aligned}$$

The Stratonovich–Itô correction term in Eq. (3.23) can be computed by a similar calculation as before and can be shown to be

$$\frac{1}{2}\Big(\big\langle - Ric(X_k) - \Delta_H X_k + [X_k, \nabla_{X_i} X_i] - \langle [X_k, X_i], X_p\rangle \nabla_{X_i} X_p, X_j\big\rangle + \langle \nabla_{X_j} X_i, [X_k, X_i]\rangle\Big)\eta^j dt.$$

Hence the Itô form of (3.23) is

$$d\tilde{r}^k = (\dot{r}^k - \tfrac{1}{2} Ric(X_k), X_j\rangle)\eta^j dt - \langle [X_i, X_k], X_j\rangle \eta^j \, dw_i.$$

It follows from Theorems 2.3 and 2.4 that

$$Div(\tilde{r}) = \int_0^T \Big(\dot{r}_t^k - \frac{1}{2}\langle Ric(X_k)(x_t), Z_t\rangle\Big)\, dw_k.$$

The theorem follows. □

Remark 3.6. Strictly speaking, the process Z in Theorem 3.5 cannot be considered to be a vector field at x since Eq. (3.21) for η involves the functions G^{kj} which depend on the derivative of the *flow* of the original stochastic differential equation (1.1). However, Z can be considered to be a vector field defined on the flow of (1.1).

There is a special case of Theorem 3.5 when Z defines a genuine vector field at x. Suppose the vector fields $X_i, 1 \leqslant i \leqslant n$ are *Killing fields*, i.e. they have the property

$$\langle \nabla_Y X_i, Z \rangle = -\langle \nabla_Z X_i, Y \rangle, \quad \forall Y, Z \in \Gamma(M).$$

Then Eq. (3.18) shows that $dG_t^{kj} \equiv 0$ for all $1 \leqslant j, k \leqslant n$ (as is well-known, the stochastic flow is an isometry on M in this case). Hence Z is expressible in terms of x alone.

We also note that in this special case

$$\nabla_{X_i} X_i = 0$$

so formulae (3.20) for the coefficients $\tilde{B}^{kj}$ in Eq. (3.21) simplify.

4. The case of an embedded submanifold

In this section we consider the following special case of Theorem 3.4. Suppose M is an embedded submanifold of Euclidean space $\mathbf{R}^n$. Define $X_i(x) = P(x)e_i, 1 \leqslant i \leqslant n$, where $e_1, \ldots, e_n$ is the standard orthonormal basis of $\mathbf{R}^n$ and $P(x)$ is orthogonal projection onto the tangent plane $T_x M$, at each point $x \in M$. For $V_1, V_2 \in T_x M$, we have

$$\begin{aligned} dX_i^*(V_1, V_2) &= V_1[P(.)e_i.V_2] - V_2[P(.)e_i.V_1] - P(.)e_i.[V_1, V_2] \\ &= V_1[e_i.V_2] - V_2[e_i.V_1] - e_i.[V_1, V_2] = 0, \end{aligned}$$

i.e. the closure hypothesis in Theorem 3.4 is satisfied.

Integration by parts formulae in path space are usually expressed in terms of parallel translations. The purpose of this section is to reformulate Theorem 3.4 in this way, in the embedded setting. Consider the SDE

$$dx_t = \sum_{i=1}^{n} X_i(x_t) \circ dw_i, \tag{4.1}$$

where we take the drift to be 0 for simplicity. As is well-known, and can be readily checked, the infinitessimal generator of x is $1/2\Delta$, where Δ is the Laplace–Beltrami operator on M. That is, x is an (extrinsically constructed) Brownian motion in M.

The next result describes the tensors introduced in (3.7) in terms of projections.

Lemma 4.1. *Define* $Q = I - P$ *and, as in Section* 3, *set*

$$T^{kj}(V) \equiv \langle \nabla_V X_k, X_j \rangle - \langle \nabla_V X_j, X_k \rangle, \quad V \in TM, 1 \leqslant j, k \leqslant n.$$

Then

$$T^{kj}(V) = \langle e_k, (Q - P)\, dP(V) e_j \rangle. \tag{4.2}$$

Proof. Since $\nabla_V X_i = P dP(V) e_i$, we have

$$\begin{aligned} T^{kj}(V) &= \langle P dP(V) e_k, P e_j \rangle - \langle e_k, P dP(V) e_j \rangle \\ &= \langle e_k, [dP(V) P - P dP(V)] e_j \rangle. \end{aligned} \tag{4.3}$$

Differentiating in the relation $P^2 = P$ yields

$$dP(V) P + P dP(V) = dP(V).$$

Thus

$$dP(V) P - P dP(V) = dP(V) - 2 P dP(V) = (Q - P)\, dP(V).$$

Substituting this into (4.3) gives (4.2).

Let $r = (r_1, \ldots, r_n)$ denote a Cameron–Martin path, suppose $\eta^k, 1 \leqslant k \leqslant n$ solve the SDE

$$d\eta_t^k = dr_t^k + B^{kj}(x_t)\eta^j dt + T^{kj}(\circ\, dx_t)\eta_t^j,$$

where B^{kj} are as in (3.8), and define $Z_t = X_i(x_t)\eta_t^i$. Let ∇_t denote the Stratonovich covariant differential of a vector field along a curve. Then we have

$$\begin{aligned} \nabla_t Z &= (\nabla_{\circ\, dx_t} X_i)\eta_t^i + X_i(x_t) \circ d\eta_t^i \\ &= \big(\nabla_{\circ\, dx_t} X_j + X_i(x_t) T^{ij}(\circ dx_t)\big)\eta_t^j + X_i(x_t)\, dr_t^i + X_i(x_t) B^{ij}(x_t)\eta_t^j\, dt. \end{aligned} \tag{4.4}$$

Now

$$\nabla_{\circ\, dx_t} X_j = P(x_t) dP(\circ\, dx_t) e_j \tag{4.5}$$

while, using Lemma 4.1, we have

$$\begin{aligned} X_i(x_t) T^{ij}(\circ\, dx_t) &= P(x_t) e_i \langle e_i, (Q - P) dP(\circ\, dx_t) e_j \rangle \\ &= P(Q - P)(x_t)\, dP(\circ\, dx_t) e_j \\ &= -P(x_t)\, dP(\circ\, dx_t) e_j. \end{aligned}$$

Combining this with Eqs. (4.5), and (4.4) we arrive at the following expression for $\nabla_t Z$

$$\nabla_t Z = X_i(x_t)\big(dr_t^i + B^{ij}(x_t)\eta_t^j\big)\,dt. \tag{4.6}$$

Let U_t denote stochastic parallel translation along the path $x_s, 0 \leqslant s \leqslant t$, and define the path h in T_oM

$$h_t = \int_0^t U_s^{-1} X_i(x_s)(\dot{r}_s^i + B^{ij}(x_s)\eta_s^j)\,ds. \tag{4.7}$$

Then (4.6) yields

$$Z_t = U_t h_t.$$

The main result of this section is the following

Theorem 4.2. *For every smooth cylindrical function Φ on $C_o(M)$*

$$E[(Z\Phi)(x)] = E\Big[\Phi(x)\int_0^T \big(\dot{r}_t^k + \frac{1}{2}\langle Ric(X_k), U_t h_t\rangle\big)\,dw_k\Big]. \tag{4.8}$$

This follows immediately from Theorem 3.4.

We note that the right-hand side of Eq. (4.8) is expressed in terms of an Itô integral with respect to an n-dimensional Brownian motion, whereas we are working in a d-dimensional manifold, with $n \geqslant d$. As such, (4.8) contains "redundant noise" in the sense of Elworthy et al. [E-L-L]. Since the integrand in (4.8) *is a function of x*, the redundant noise can be filtered out by conditioning. To this end, we write the expectation in (4.8) as $E[\Phi(x)\rho(x)]$ where

$$\rho(x) = \int_0^T \big(\dot{r}_t^k + \frac{1}{2}\langle Ric(X_k), U_t h_t\rangle\big)\,dE[w_k/x].$$

Let $db_t \equiv U_t^{-1} \circ dx_t$ denote the antidevelopment of x onto T_0M. Then Elworthy et al. [E-L-L] have shown that w can be decomposed as

$$dw_t = U_t\,db_t + V_t\,d\beta_t,$$

where β is a Brownian motion in $T_0M^\perp$ which is independent of b (and hence of x) and V_t is parallel translation in the normal bundle to TM. An elementary calculation now shows that

$$\rho(x) = \int_0^T \langle \dot{h}_t + \frac{1}{2} U_t^{-1} Ric(U_t h_t) - g(t), db_t \rangle,$$

where

$$g(t) = U_t^{-1} X_i(x_t) B^{ij}(x_t) \eta_t^j. \tag{4.9}$$

We close with the following

Example 4.3. Let $M = S^{n-1} \subset \mathbf{R}^n$. In this case we have, for $x = (x_1, \ldots, x_n)$

$$X_i(x) = e_i - x_i x.$$

Thus, for $V \in T_x M$

$$\nabla_V X_i = P(x)(-V_i x - x_i V) = -x_i V.$$

This implies

$$\nabla_{X_i} X_j = -x_j X_i. \tag{4.10}$$

Since $\nabla_{X_i} X_i = -x_i X_i = -\langle x, e_i \rangle P e_i = -P(x) = 0$, we get

$$\begin{aligned} \Delta_H X_k &= \nabla_{X_i} \nabla_{X_i} X_k - \nabla_{\nabla_{X_i} X_i} X_k = -\nabla_{X_i}(x_k X_i) \\ &= -\langle e_k, X_i \rangle X_i - x_k \nabla_{X_i} X_i \\ &= -\langle e_k, P e_i \rangle e_i = -P e_k \\ &= -X_k. \end{aligned} \tag{4.11}$$

From (4.10) we obtain

$$[X_i, X_k] = x_i X_k - x_k X_i.$$

Since $x_i X_i = 0$ and $x_i x_i = 1$, we have

$$\begin{aligned} \langle [X_i, X_k], X_p \rangle [X_i, X_p] &= \langle x_i X_k - x_k X_i, X_p \rangle (x_i X_p - x_p X_i) \\ &= \langle x_i X_k - x_k X_i, X_p \rangle x_i X_p \\ &= \langle X_k . X_p \rangle X_p = \langle P e_k, e_p \rangle P e_p \\ &= X_k. \end{aligned} \tag{4.12}$$

Expressions (4.11) and (4.12) show that the functions B^{kj} in (3.8) and hence g in (4.9) are zero. Thus we obtain the integration by parts formula

$$E[(Z\Phi)(x)] = E\left[\Phi(x)\int_0^T \left\langle \dot{h}_t + \frac{1}{2}U_t^{-1}Ric(U_t h_t), db_t\right\rangle\right],$$

where

$$h_t = \int_0^t U_s^{-1} X_i(x_s)\dot{r}_s^i ds.$$

Acknowledgments

I would like to thank Bruce Driver for his invitation to visit San Diego and for stimulating discussions there which led to the present work. I am also indebted to Bruce for insightful suggestions which helped clarify the ideas in this paper.

References

[Be] D. Bell, Degenerate Stochastic Differential Equations and Hypoellipticity, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 79, Pitman, Longman, 1995.

[B-M] D. Bell, S. Mohammed, An extension of Hörmander's theorem for infinitely degenerate second-order operators, Duke Math. J. 78 (3) (1995) 453–475.

[Bi] J.M. Bismut, Martingales the Malliavin calculus and hypoellipticity under general Hörmander's conditions, Z. Wahrsch. Verw. Gebiete 56 (1981) 529–548.

[D] B. Driver, A Cameron–Martin type quasi-invariance theorem for Brownian motion on a compact manifold, J. Funct. Anal. 109 (1992) 272–376.

[E] K.D. Elworthy, Stochastic Differential Equations on Manifolds, Lecture Notes Series, vol. 70, Cambridge University Press, Cambridge, 1982.

[E-L] K.D. Elworthy, X.-M. Li, A class of integration by parts formulae in stochastic analysis I, Itô's Stochastic Calculus and Probability Theory, Springer, Tokyo, 1996, pp. 15–30.

[E-L-L] K.D. Elworthy, Y. Le Jan, X.-M. Li, On the Geometry of Diffusion Operators and Stochastic Flows, Lecture Notes in Mathematics, vol. 1720, Springer, Berlin, 1999.

[M] P. Malliavin, Stochastic calculus of variations and hypoelliptic operators, Proceedings of the International Conference on Stochastic Differential Equations, Kyoto, Kinokuniya and Wiley, 1976, pp. 195–263.