CURVED WIENER SPACE ANALYSIS

BRUCE K. DRIVER

Contents

Date: *File:curved.tex* Last revised: March 28, 2005.
This research was partially supported by NSF Grants DMS 9223177, DMS 96-12651, DMS 99-71036 and DMS 0202939.

1. Introduction

These notes represent a much expanded and updated version of the "mini course" that the author gave at the ETH (Zürich) and the University of Zürich in February of 1995. The purpose of these notes is to first provide some basic background to Riemannian geometry and stochastic calculus on manifolds and then to cover some of the more recent developments pertaining to analysis on "curved Wiener spaces." Essentially no differential geometry is assumed. However, it is assumed that the reader is comfortable with stochastic calculus and differential equations on Euclidean spaces. Here is a brief description of what will be covered in the text below.

Section 2 is a basic introduction to differential geometry through imbedded submanifolds. Section 3 is an introduction to the Riemannian geometry that will be needed in the sequel. Section 4 records a number of results pertaining to flows of vector fields and "Cartan's rolling map." The stochastic version of these results will be important tools in the sequel. Section 5 is a rapid introduction to stochastic calculus on manifolds and related geometric constructions. Section 6 briefly gives applications of stochastic calculus on manifolds to representation formulas for derivatives of heat kernels. Section 7 is devoted to the study of the calculus and integral geometry associated with the path space of a Riemannian manifold equipped with "Wiener measure." In particular, quasi-invariance, Poincaré and logarithmic Sobolev inequalities are developed for the Wiener measure on path spaces in this section. Section 8 is a short introduction to Malliavin's probabilistic methods for dealing with hypoelliptic diffusions. The appendix in section 9 records some basic martingale and stochastic differential equation estimates which are mostly used in section 8.

Although the majority of these notes form a survey of known results, many proofs have been cleaned up and some proofs are new. Moreover, Section 8 is written using the geometric language introduced in these notes which is not completely standard in the literature. I have also tried (without complete success) to give an overview of many of the major techniques which have been used to date in this subject. Although numerous references are given to the literature, the list is far from complete. I apologize in advance to anyone who feels cheated by not being included in the references. However, I do hope the list of references is sufficiently rich that the interested reader will be able to find additional information by looking at the related articles and the references that they contain.

Acknowledgement: It is pleasure to thank Professor A. Sznitman and the ETH for their hospitality and support and the opportunity to give the talks which started these notes. I also would like to thank Professor E. Bolthausen for his hospitality and his role in arranging the first lecture to be held at University of Zürich.

2. Manifold Primer

Conventions:

1. If A, B are linear operators on some vector space, then $[A,B] := AB - BA$ is the **commutator** of A and B.
2. If X is a topological space we will write $A \subset_o X$, $A \sqsubset X$ and $A \sqsubset\sqsubset X$ to mean A is an open, closed, and respectively a compact subset of X.
3. Given two sets A and B, the notation $f : A \to B$ will mean that f is a function from a subset $\mathcal{D}(f) \subset A$ to B. (We will allow $\mathcal{D}(f)$ to be the empty set.) The set $\mathcal{D}(f) \subset A$ is called the domain of f and the subset $\mathcal{R}(f) := f(\mathcal{D}(f)) \subset B$ is called the range of f. If f is injective, let $f^{-1} : B \to A$ denote the inverse function with domain $\mathcal{D}(f^{-1}) = \mathcal{R}(f)$ and range $\mathcal{R}(f^{-1}) = \mathcal{D}(f)$. If $f : A \to B$ and $g : B \to C$, then $g \circ f$ denotes the composite function from A to C with domain $\mathcal{D}(g\circ f) := f^{-1}(\mathcal{D}(g))$ and range $\mathcal{R}(g \circ f) := g \circ f(\mathcal{D}(g \circ f)) = g(\mathcal{R}(f) \cap \mathcal{D}(g))$.

Notation 2.1. Throughout these notes, let E and V denote finite dimensional vector spaces. A function $F : E \to V$ is said to be smooth if $\mathcal{D}(F)$ is open in E ($\mathcal{D}(F) = \emptyset$ is allowed) and $F : \mathcal{D}(F) \to V$ is infinitely differentiable. Given a **smooth** function $F : E \to V$, let $F'(x)$ denote the differential of F at $x \in \mathcal{D}(F)$. Explicitly, $F'(x) = DF(x)$ denotes the linear map from E to V determined by

$$DF(x)\,a = F'(x)a := \frac{d}{dt}|_0 F(x + ta)\ \forall\ a \in E. \tag{2.1}$$

We also let

$$F''(x)(v,w) = F''(x)(v,w) := (\partial_v\partial_w F)(x) = \frac{d}{dt}|_0 \frac{d}{ds}|_0 F(x + tv + sw)\,. \tag{2.2}$$

2.1. **Imbedded Submanifolds.** Rather than describe the most abstract setting for Riemannian geometry, for simplicity we choose to restrict our attention to imbedded submanifolds of a Euclidean space $E = \mathbb{R}^N$. [1] We will equip $\mathbb{R}^N$ with the standard inner product,

$$\langle a, b\rangle = \langle a, b\rangle_{\mathbb{R}^N} := \sum_{i=1}^N a_i b_i.$$

In general, we will denote inner products in these notes by $\langle\cdot,\cdot\rangle$.

Definition 2.2. A subset M of E (see Figure 1) is a $d-$ **dimensional imbedded submanifold** (without boundary) of E iff for all $m \in M$, there is a function $z : E \to \mathbb{R}^N$ such that:

1. $\mathcal{D}(z)$ is an open neighborhood of E containing m,
2. $\mathcal{R}(z)$ is an open subset of $\mathbb{R}^N$,
3. $z : \mathcal{D}(z) \to \mathcal{R}(z)$ is a diffeomorphism (a smooth invertible map with smooth inverse), and
4. $z(M \cap \mathcal{D}(z)) = \mathcal{R}(z) \cap (\mathbb{R}^d \times \{0\}) \subset \mathbb{R}^N$.

(We write M^d if we wish to emphasize that M is a $d-$dimensional manifold.)

Notation 2.3. Given an imbedded submanifold and diffeomorphism z as in the above definition, we will write $z = (z_<, z_>)$ where $z_<$ is the first d components of z and $z_>$ consists of the last $N - d$ components of z. Also let $x : M \to \mathbb{R}^d$ denote the function defined by $\mathcal{D}(x) := M \cap \mathcal{D}(z)$ and $x := z_<|_{\mathcal{D}(x)}$. Notice that $\mathcal{R}(x) := x(\mathcal{D}(x))$ is an open subset of $\mathbb{R}^d$ and that $x^{-1} : \mathcal{R}(x) \to \mathcal{D}(x)$, thought of as a function taking values in E, is smooth. The bijection $x : \mathcal{D}(x) \to \mathcal{R}(x)$ is called a **chart** on M. Let $\mathcal{A} = \mathcal{A}(M)$ denote the collection of charts on M. The collection of charts $\mathcal{A} = \mathcal{A}(M)$ is often referred to as an **atlas** for M.

Remark 2.4. The imbedded submanifold M is made into a topological space using the induced topology from E. With this topology, each chart $x \in \mathcal{A}(M)$ is a homeomorphism from $\mathcal{D}(x) \subset_o M$ to $\mathcal{R}(x) \subset_o \mathbb{R}^d$.

[1] Because of the Whitney imbedding theorem (see for example Theorem 6-3 in Auslander and MacKenzie [9]), this is actually not a restriction.

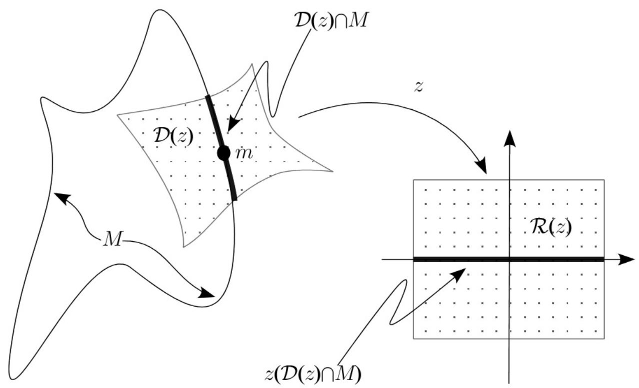

FIGURE 1. An imbedded one dimensional submanifold in $\mathbb{R}^2$.

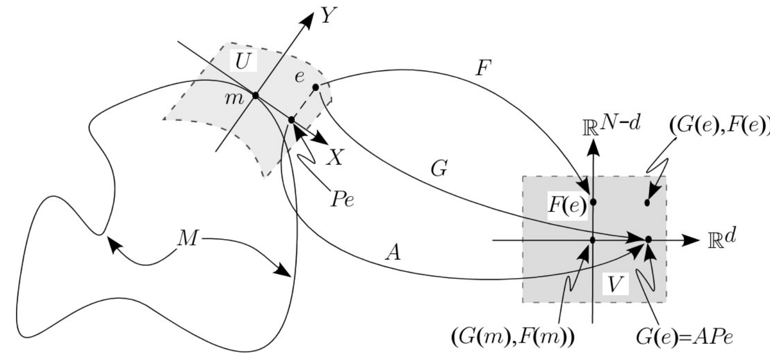

FIGURE 2. Constructing charts for M using the inverse function theorem. For simplicity of the drawing, $m \in M$ is assumed to be the origin of $E = X \oplus Y$.

Theorem 2.5 (A Basic Construction of Manifolds)**.** *Let $F : E \to \mathbb{R}^{N-d}$ be a smooth function and $M := F^{-1}(\{0\}) \subset E$ which we assume to be non-empty. Suppose that $F'(m) : E \to \mathbb{R}^{N-d}$ is surjective for all $m \in M$. Then M is a d – dimensional imbedded submanifold of E.*

Proof. Let $m \in M$, we will begin by constructing a smooth function $G : E \to \mathbb{R}^d$ such that $(G, F)'(m) : E \to \mathbb{R}^N = \mathbb{R}^d \times \mathbb{R}^{N-d}$ is invertible. To do this, let $X = \mathrm{Nul}(F'(m))$ and Y be a complementary subspace so that $E = X \oplus Y$ and let $P : E \to X$ be the associated projection map, see Figure 2. Notice that $F'(m) : Y \to \mathbb{R}^{N-d}$ is a linear isomorphism of vector spaces and hence

$$\dim(X) = \dim(E) - \dim(Y) = N - (N - d) = d.$$

In particular, X and $\mathbb{R}^d$ are isomorphic as vector spaces. Set $G(m) = APm$ where $A : X \to \mathbb{R}^d$ is an arbitrary but fixed linear isomorphism of vector spaces. Then for $x \in X$ and $y \in Y$,

$$\begin{aligned}(G, F)'(m)(x + y) &= (G'(m)(x + y), F'(m)(x + y)) \\ &= (AP(x + y), F'(m)y) = (Ax, F'(m)y) \in \mathbb{R}^d \times \mathbb{R}^{N-d}\end{aligned}$$

from which it follows that $(G, F)'(m)$ is an isomorphism.

By the inverse function theorem, there exists a neighborhood $U \subset_o E$ of m such that $V := (G, F)(U) \subset_o \mathbb{R}^N$ and $(G, F) : U \to V$ is a diffeomorphism. Let $z = (G, F)$ with $\mathcal{D}(z) = U$ and $\mathcal{R}(z) = V$. Then z is a chart of E about m satisfying the conditions of Definition 2.2. Indeed, items 1) – 3) are clear by construction. If $p \in M \cap \mathcal{D}(z)$ then $z(p) = (G(p), F(p)) = (G(p), 0) \in \mathcal{R}(z) \cap (\mathbb{R}^d \times \{0\})$. Conversely, if $p \in \mathcal{D}(z)$ is a point such that $z(p) = (G(p), F(p)) \in \mathcal{R}(z) \cap (\mathbb{R}^d \times \{0\})$, then $F(p) = 0$ and hence $p \in M \cap \mathcal{D}(z)$; so item 4) of Definition 2.2 is verified. ∎

Example 2.6. Let $gl(n, \mathbb{R})$ denote the set of all $n \times n$ real matrices. The following are examples of imbedded submanifolds.

(1) Any open subset M of E.
(2) The graph,
$$\Gamma(f) := \left\{(x, f(x)) \in \mathbb{R}^d \times \mathbb{R}^{N-d} : x \in \mathcal{D}(f)\right\} \subset \mathcal{D}(f) \times \mathbb{R}^{N-d} \subset \mathbb{R}^N,$$
of any smooth function $f : \mathbb{R}^d \to \mathbb{R}^{N-d}$ as can be seen by applying Theorem 2.5 with $F(x, y) := y - f(x)$. In this case it would be a good idea for the reader to produce an explicit chart z as in Definition 2.2 such that $\mathcal{D}(z) = \mathcal{R}(z) = \mathcal{D}(f) \times \mathbb{R}^{N-d}$.
(3) The unit sphere, $S^{N-1} := \{x \in \mathbb{R}^N : \langle x, x\rangle_{\mathbb{R}^N} = 1\}$, as is seen by applying Theorem 2.5 with $E = \mathbb{R}^{\mathbb{N}}$ and $F(x) := \langle x, x\rangle_{\mathbb{R}^N} - 1$. Alternatively, express S^{N-1} locally as the graph of smooth functions and then use item 2.
(4) $GL(n, \mathbb{R}) := \{g \in gl(n, \mathbb{R}) | \det(g) \neq 0\}$, see item 1.
(5) $SL(n, \mathbb{R}) := \{g \in gl(n, \mathbb{R}) | \det(g) = 1\}$ as is seen by taking $E = gl(n, \mathbb{R})$ and $F(g) := \det(g)$ and then applying Theorem 2.5 with the aid of Lemma 2.7 below.
(6) $O(n) := \{g \in gl(n, \mathbb{R}) | g^{\mathrm{tr}} g = I\}$ where g^{tr} denotes the transpose of g. In this case take $F(g) := g^{\mathrm{tr}} g - I$ thought of as a function from $E = gl(n, \mathbb{R})$ to $\mathcal{S}(n)$, where
$$\mathcal{S}(n) := \left\{A \in gl(n, \mathbb{R}) : A^{\mathrm{tr}} = A\right\}$$
is the subspace of symmetric matrices. To show $F'(g)$ is surjective, show
$$F'(g)(gB) = B + B^{\mathrm{tr}} \text{ for all } g \in O(n) \text{ and } B \in gl(n, \mathbb{R}).$$

(7) $SO(n) := \{g \in O(n) | \det(g) = 1\}$, an open subset of $O(n)$.
(8) $M \times N \subset E \times V$, where M and N are imbedded submanifolds of E and V respectively. The reader should verify this by constructing appropriate charts for $E \times V$ by taking "tensor" products of the charts for E and V associated to M and N respectively.
(9) The n – dimensional torus,

$$T^n := \{z \in \mathbb{C}^n : |z_i| = 1 \text{ for } i = 1, 2, \ldots, n\} = (S^1)^n,$$

where $z = (z_1, \ldots, z_n)$ and $|z_i| = \sqrt{z_i \bar{z}_i}$. This follows by induction using items 3. and 8. Alternatively apply Theorem 2.5 with $F(z) := \left(|z_1|^2 - 1, \ldots, |z_n|^2 - 1\right)$.

Lemma 2.7. *Suppose $g \in GL(n, \mathbb{R})$ and $A \in gl(n, \mathbb{R})$, then*

$$\det{}'(g)A = \det(g)\operatorname{tr}(g^{-1}A). \tag{2.3}$$

Proof. By definition we have

$$\det{}'(g)A = \frac{d}{dt}|_0 \det(g + tA) = \det(g)\frac{d}{dt}|_0 \det(I + tg^{-1}A).$$

So it suffices to prove $\frac{d}{dt}|_0 \det(I + tB) = \operatorname{tr}(B)$ for all matrices B. If B is upper triangular, then $\det(I + tB) = \prod_{i=1}^n (1 + tB_{ii})$ and hence by the product rule,

$$\frac{d}{dt}|_0 \det(I + tB) = \sum_{i=1}^n B_{ii} = \operatorname{tr}(B).$$

This completes the proof because; 1) every matrix can be put into upper triangular form by a similarity transformation, and 2) "det" and "tr" are invariant under similarity transformations. ■

Definition 2.8. Let E and V be two finite dimensional vector spaces and $M^d \subset E$ and $N^k \subset V$ be two imbedded submanifolds. A function $f : M \to N$ is said to be **smooth** if for all charts $x \in \mathcal{A}(M)$ and $y \in \mathcal{A}(N)$ the function $y \circ f \circ x^{-1} : \mathbb{R}^d \to \mathbb{R}^k$ is smooth.

Exercise 2.9. Let $M^d \subset E$ and $N^k \subset V$ be two imbedded submanifolds as in Definition 2.8.

(1) Show that a function $f : \mathbb{R}^k \to M$ is smooth iff f is smooth when thought of as a function from $\mathbb{R}^k$ to E.
(2) If $F : E \to V$ is a smooth function such that $F(M \cap \mathcal{D}(F)) \subset N$, show that $f := F|_M : M \to N$ is smooth.
(3) Show the composition of smooth maps between imbedded submanifolds is smooth.

Proposition 2.10. *Assuming the notation in Definition 2.8, a function $f : M \to N$ is smooth iff there is a smooth function $F : E \to V$ such that $f = F|_M$.*

Proof. (Sketch.) Suppose that $f : M \to N$ is smooth, $m \in M$ and $n = f(m)$. Let z be as in Definition 2.2 and w be a chart on N such that $n \in \mathcal{D}(w)$. By shrinking the domain of z if necessary, we may assume that $\mathcal{R}(z) = U \times W$ where $U \subset_o \mathbb{R}^d$ and $W \subset_o \mathbb{R}^{N-d}$ in which case $z(M \cap \mathcal{D}(z)) = U \times \{0\}$. For $\xi \in \mathcal{D}(z)$, let $F(\xi) := f(z^{-1}(z_<(\xi), 0))$ with $z = (z_<, z_>)$ as in Notation 2.3. Then $F : \mathcal{D}(z) \to N$ is a smooth function such that $F|_{M \cap \mathcal{D}(z)} = f|_{M \cap \mathcal{D}(z)}$. The function F is smooth. Indeed, letting $x = z_<|_{\mathcal{D}(z) \cap M}$,

$$w_< \circ F = w_< \circ f(z^{-1}(z_<(\xi), 0)) = w_< \circ f \circ x^{-1} \circ (z_<(\cdot), 0)$$

which, being the composition of the smooth maps $w_< \circ f \circ x^{-1}$ (smooth by assumption) and $\xi \to (z_<(\xi), 0)$, is smooth as well. Hence by definition, F is smooth as claimed. Using a standard partition of unity argument (which we omit), it is possible to piece this local argument together to construct a globally defined smooth function $F : E \to V$ such that $f = F|_M$. ■

Definition 2.11. A function $f : M \to N$ is a **diffeomorphism** if f is smooth and has a smooth inverse. The set of diffeomorphisms $f : M \to M$ is a group under composition which will be denoted by $\operatorname{Diff}(M)$.

2.2. Tangent Planes and Spaces.

Definition 2.12. Given an imbedded submanifold $M \subset E$ and $m \in M$, let $\tau_m M \subset E$ denote the collection of all vectors $v \in E$ such there exists a smooth path $\sigma : (-\varepsilon, \varepsilon) \to M$ with $\sigma(0) = m$ and $v = \frac{d}{ds}|_0 \sigma(s)$. The subset $\tau_m M$ is called the **tangent plane** to M at m and $v \in \tau_m M$ is called a **tangent vector**, see Figure 3.

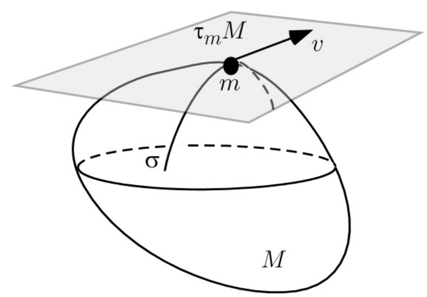

FIGURE 3. Tangent plane, $\tau_m M$, to M at m and a vector, v, in $\tau_m M$.

Theorem 2.13. *For each $m \in M$, $\tau_m M$ is a d – dimensional subspace of E. If $z : E \to \mathbb{R}^N$ is as in Definition 2.2, then $\tau_m M = \operatorname{Nul}(z_>'(m))$. If x is a chart on M such that $m \in \mathcal{D}(x)$, then*

$$\{\frac{d}{ds}|_0 x^{-1}(x(m) + se_i)\}_{i=1}^d$$

is a basis for $\tau_m M$*, where* $\{e_i\}_{i=1}^d$ *is the standard basis for* $\mathbb{R}^d$.

Proof. Let $\sigma : (-\varepsilon,\varepsilon) \to M$ be a smooth path with $\sigma(0) = m$ and $v = \frac{d}{ds}|_0\sigma(s)$ and z be a chart (for E) around m as in Definition 2.2 such that $x = z_<$. Then $z_>(\sigma(s)) = 0$ for all s and therefore,

$$0 = \frac{d}{ds}|_0 z_>(\sigma(s)) = z_>'(m)v$$

which shows that $v \in \mathrm{Nul}(z_>'(m))$, i.e. $\tau_m M \subset \mathrm{Nul}(z_>'(m))$.

Conversely, suppose that $v \in \mathrm{Nul}(z_>'(m))$. Let $w = z_<'(m)v \in \mathbb{R}^d$ and $\sigma(s) := x^{-1}(z_<(m) + sw) \in M$ – defined for s near 0. Differentiating the identity $z^{-1}\circ z = id$ at m shows

$$\left(z^{-1}\right)'(z(m))z'(m) = I.$$

Therefore,

$$\begin{aligned}
\sigma'(0) &= \frac{d}{ds}|_0 x^{-1}(z_<(m)+sw) = \frac{d}{ds}|_0 z^{-1}(z_<(m)+sw,0)\\
&= \left(z^{-1}\right)'((z_<(m),0))(z_<'(m)v,0)\\
&= \left(z^{-1}\right)'((z_<(m),0))(z_<'(m)v,z_>'(m)v)\\
&= \left(z^{-1}\right)'(z(m))z'(m)v = v,
\end{aligned}$$

and so by definition $v = \sigma'(0) \in \tau_m M$. We have now shown $\mathrm{Nul}(z_>'(m)) \subset \tau_m M$ which completes the proof that $\tau_m M = \mathrm{Nul}(z_>'(m))$.

Since $z_<'(m) : \tau_m M \to \mathbb{R}^d$ is a linear isomorphism, the above argument also shows

$$\frac{d}{ds}|_0 x^{-1}(x(m)+sw) = \left(z_<'(m)|_{\tau_m M}\right)^{-1} w \in \tau_m M \;\forall\; w\in\mathbb{R}^d.$$

In particular it follows that

$$\{\frac{d}{ds}|_0 x^{-1}(x(m)+se_i)\}_{i=1}^d = \{\left(z_<'(m)|_{\tau_m M}\right)^{-1} e_i\}_{i=1}^d$$

is a basis for $\tau_m M$, see Figure 4 below. ■

The following proposition is an easy consequence of Theorem 2.13 and the proof of Theorem 2.5.

Proposition 2.14. *Suppose that* M *is an imbedded submanifold constructed as in Theorem 2.5. Then* $\tau_m M = \mathrm{Nul}(F'(m))$.

Exercise 2.15. Show:

(1) $\tau_m M = E$, if M is an open subset of E.
(2) $\tau_g GL(n,\mathbb{R}) = gl(n,\mathbb{R})$, for all $g \in GL(n,\mathbb{R})$.
(3) $\tau_m S^{N-1} = \{m\}^\perp$ for all $m \in S^{N-1}$.
(4) Let $sl(n,\mathbb{R})$ be the traceless matrices,

$$sl(n,\mathbb{R}) := \{A \in gl(n,\mathbb{R}) |\, \mathrm{tr}(A) = 0\}. \tag{2.4}$$

Then

$$\tau_g SL(n,\mathbb{R}) = \{A \in gl(n,\mathbb{R}) | g^{-1}A \in sl(n,\mathbb{R})\}$$

and in particular $\tau_I SL(n,\mathbb{R}) = sl(n,\mathbb{R})$.
(5) Let $so\,(n,\mathbb{R})$ be the skew symmetric matrices,

$$so\,(n,\mathbb{R}) := \{A \in gl(n,\mathbb{R}) | A = -A^{\mathrm{tr}}\}.$$

Then

$$\tau_g O(n) = \{A \in gl(n,\mathbb{R}) | g^{-1}A \in so\,(n,\mathbb{R})\}$$

and in particular $\tau_I O\,(n) = so\,(n,\mathbb{R})$. **Hint:** $g^{-1} = g^{\mathrm{tr}}$ for all $g \in O(n)$.
(6) If $M \subset E$ and $N \subset V$ are imbedded submanifolds then

$$\tau_{(m,n)}(M\times N) = \tau_m M \times \tau_n N \subset E\times V.$$

It is quite possible that $\tau_m M = \tau_{m'} M$ for some $m \neq m'$, with m and m' in M (think of the sphere). Because of this, it is helpful to label each of the tangent planes with their base point.

Definition 2.16. The **tangent space** $(T_m M)$ to M at m is given by

$$T_m M := \{m\}\times\tau_m M \subset M\times E.$$

Let

$$TM := \cup_{m\in M} T_m M,$$

and call TM the **tangent space (or tangent bundle)** of M. A **tangent vector** is a point $v_m := (m,v) \in TM$ and we let $\pi : TM \to M$ denote the **canonical projection** defined by $\pi(v_m) = m$. Each tangent space is made into a vector space with the vector space operations being defined by: $c(v_m) := (cv)_m$ and $v_m + w_m := (v+w)_m$.

Exercise 2.17. Prove that TM is an imbedded submanifold of $E\times E$. **Hint:** suppose that $z : E \to \mathbb{R}^N$ is a function as in the Definition 2.2. Define $\mathcal{D}(Z) := \mathcal{D}(z)\times E$ and $Z : \mathcal{D}(Z) \to \mathbb{R}^N\times\mathbb{R}^N$ by $Z(x,a) := (z(x), z'(x)a)$. Use Z's of this type to check TM satisfies Definition 2.2.

Notation 2.18. In the sequel, given a smooth path $\sigma : (-\varepsilon,\varepsilon) \to M$, we will abuse notation and write $\sigma'(0)$ for either

$$\frac{d}{ds}|_0\sigma(s) \in \tau_{\sigma(0)}M$$

or for

$$(\sigma(0), \frac{d}{ds}|_0\sigma(s)) \in T_{\sigma(0)}M = \{\sigma(0)\}\times\tau_{\sigma(0)}M.$$

Also given a chart $x = (x^1, x^2, \ldots, x^d)$ on M and $m \in \mathcal{D}(x)$, let $\partial/\partial x^i|_m$ denote the element $T_m M$ determined by $\partial/\partial x^i|_m = \sigma'(0)$, where $\sigma(s) := x^{-1}(x(m) + se_i)$, i.e.

$$\frac{\partial}{\partial x^i}|_m = (m, \frac{d}{ds}|_0 x^{-1}(x(m) + se_i)), \tag{2.5}$$

see Figure 4.

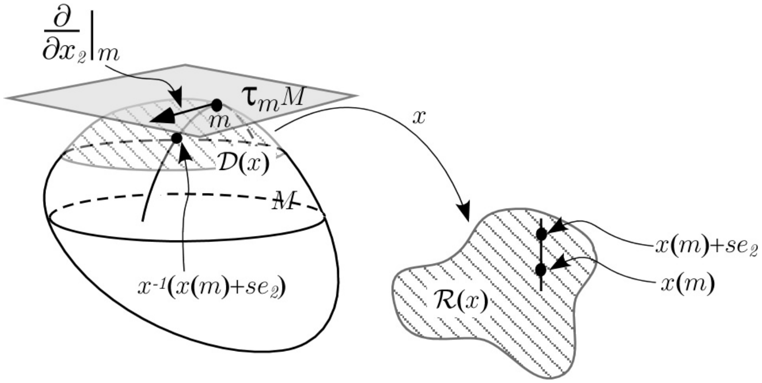

FIGURE 4. Forming a basis of tangent vectors.

The reason for the strange notation in Eq. (2.5) will be explained after Notation 2.20. By definition, every element of $T_m M$ is of the form $\sigma'(0)$ where σ is a smooth path into M such that $\sigma(0) = m$. Moreover by Theorem 2.13, $\{\partial/\partial x^i|_m\}_{i=1}^d$ is a basis for $T_m M$.

Definition 2.19. Suppose that $f : M \to V$ is a smooth function, $m \in \mathcal{D}(f)$ and $v_m \in T_m M$. Write

$$v_m f = df(v_m) := \frac{d}{ds}|_0 f(\sigma(s)),$$

where σ is any smooth path in M such that $\sigma'(0) = v_m$. The function $df : TM \to V$ will be called the **differential of** f.

Notation 2.20. If M and N are two manifolds $f : M \times N \to V$ is a smooth function, we will write $d_M f(\cdot, n)$ to indicate that we are computing the differential of the function $m \in M \to f(m, n) \in V$ for fixed $n \in N$.

To understand the notation in (2.5), suppose that $f = F \circ x = F(x^1, x^2, \ldots, x^d)$ where $F : \mathbb{R}^d \to \mathbb{R}$ is a smooth function and x is a chart on M. Then

$$\frac{\partial f(m)}{\partial x^i} := \frac{\partial}{\partial x^i}|_m f = (D_i F)(x(m)),$$

where D_i denotes the i^{th} – partial derivative of F. Also notice that $dx^j\left(\frac{\partial}{\partial x^i}|_m\right) = \delta_{ij}$ so that $\left\{dx^i|_{T_m M}\right\}_{i=1}^d$ is the dual basis of $\{\partial/\partial x^i|_m\}_{i=1}^d$ and therefore if $v_m \in T_m M$ then

$$v_m = \sum_{i=1}^d dx^i(v_m) \frac{\partial}{\partial x^i}|_m. \tag{2.6}$$

This explicitly exhibits v_m as a first order differential operator acting on "germs" of smooth functions defined near $m \in M$.

Remark 2.21 (Product Rule). Suppose that $f : M \to V$ and $g : M \to \operatorname{End}(V)$ are smooth functions, then

$$v_m(gf) = \frac{d}{ds}|_0 [g(\sigma(s))f(\sigma(s))] = v_m g \cdot f(m) + g(m) v_m f$$

or equivalently

$$d(gf)(v_m) = dg(v_m) f(m) + g(m) df(v_m).$$

This last equation will be abbreviated as $d(gf) = dg \cdot f + g df$.

Definition 2.22. Let $f : M \to N$ be a smooth map of imbedded submanifolds. Define the **differential,** f_*, of f by

$$f_* v_m = (f \circ \sigma)'(0) \in T_{f(m)} N,$$

where $v_m = \sigma'(0) \in T_m M$, and $m \in \mathcal{D}(f)$.

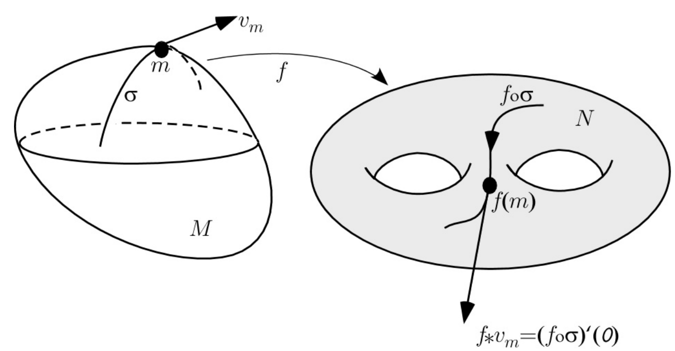

FIGURE 5. The differential of f.

Lemma 2.23. *The differentials defined in Definitions 2.19 and 2.22 are well defined linear maps on $T_m M$ for each $m \in \mathcal{D}(f)$.*

Proof. I will only prove that f_* is well defined, since the case of df is similar. By Proposition 2.10, there is a smooth function $F : E \to V$, such that $f = F|_M$. Therefore by the chain rule

$$f_* v_m = (f \circ \sigma)'(0) := \left[\frac{d}{ds}|_0 f(\sigma(s)) \right]_{f(\sigma(0))} = [F'(m)v]_{f(m)}, \tag{2.7}$$

where σ is a smooth path in M such that $\sigma'(0) = v_m$. It follows from (2.7) that $f_* v_m$ does not depend on the choice of the path σ. It is also clear from (2.7), that f_* is linear on $T_m M$. ■

Remark 2.24. Suppose that $F : E \to V$ is a smooth function and that $f := F|_M$. Then as in the proof of Lemma 2.23,

$$df(v_m) = F'(m)v \tag{2.8}$$

for all $v_m \in T_m M$, and $m \in \mathcal{D}(f)$. Incidentally, since the left hand sides of (2.7) and (2.8) are defined "intrinsically," the right members of (2.7) and (2.8) are independent of the possible choices of functions F which extend f.

Lemma 2.25 (Chain Rules)**.** *Suppose that M, N, and P are imbedded submanifolds and V is a finite dimensional vector space. Let $f : M \to N$, $g : N \to P$, and $h : N \to V$ be smooth functions. Then:*

$$(g \circ f)_* v_m = g_*(f_* v_m), \quad \forall\ v_m \in TM \tag{2.9}$$

and

$$d(h \circ f)(v_m) = dh(f_* v_m), \quad \forall\ v_m \in TM. \tag{2.10}$$

These equations will be written more concisely as $(g \circ f)_* = g_* f_*$ and $d(h \circ f) = dh f_*$ respectively.

Proof. Let σ be a smooth path in M such that $v_m = \sigma'(0)$. Then, see Figure 6,

$$\begin{aligned}(g \circ f)_* v_m &:= (g \circ f \circ \sigma)'(0) = g_*(f \circ \sigma)'(0) \\ &= g_* f_* \sigma'(0) = g_* f_* v_m.\end{aligned}$$

Similarly,

$$\begin{aligned}d(h \circ f)(v_m) &:= \frac{d}{ds}|_0 (h \circ f \circ \sigma)(s) = dh((f \circ \sigma)'(0)) \\ &= dh(f_* \sigma'(0)) = dh(f_* v_m).\end{aligned}$$

■

If $f : M \to V$ is a smooth function, x is a chart on M, and $m \in \mathcal{D}(f) \cap \mathcal{D}(x)$, we will write $\partial f(m)/\partial x^i$ for $df\left(\partial/\partial x^i|_m\right)$. Combining this notation with Eq.

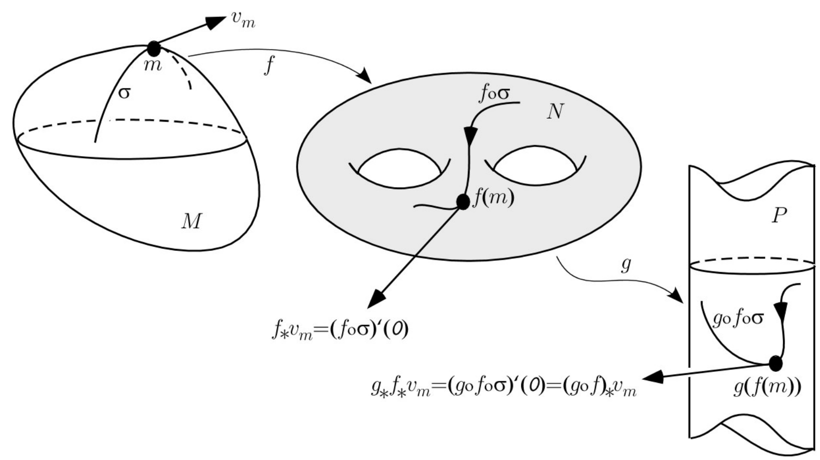

FIGURE 6. The chain rule.

(2.6) leads to the pleasing formula,

$$df = \sum_{i=1}^{d} \frac{\partial f}{\partial x^i} dx^i, \tag{2.11}$$

by which we mean

$$df(v_m) = \sum_{i=1}^{d} \frac{\partial f(m)}{\partial x^i} dx^i(v_m).$$

Suppose that $f : M^d \to N^k$ is a smooth map of imbedded submanifolds, $m \in M$, x is a chart on M such that $m \in \mathcal{D}(x)$, and y is a chart on N such that $f(m) \in \mathcal{D}(y)$. Then the matrix of

$$f_{*m} := f_*|_{T_m M} : T_m M \to T_{f(m)} N$$

relative to the bases $\{\partial/\partial x^i|_m\}_{i=1}^d$ of $T_m M$ and $\{\partial/\partial y^j|_{f(m)}\}_{j=1}^k$ of $T_{f(m)} N$ is $(\partial(y^j \circ f)(m)/\partial x^i)$. Indeed, if $v_m = \sum_{i=1}^d v^i \partial/\partial x^i|_m$, then

$$\begin{aligned}
f_*v_m &= \sum_{j=1}^{k} dy^j(f_*v_m)\partial/\partial y^j|_{f(m)} \\
&= \sum_{j=1}^{k} d(y^j\circ f)(v_m)\partial/\partial y^j|_{f(m)} \qquad \text{(by Eq. (2.10))} \\
&= \sum_{j=1}^{k}\sum_{i=1}^{d} \frac{\partial(y^j\circ f)(m)}{\partial x^i}\cdot dx^i(v_m)\partial/\partial y^j|_{f(m)} \qquad \text{(by Eq. (2.11))} \\
&= \sum_{j=1}^{k}\sum_{i=1}^{d} \frac{\partial(y^j\circ f)(m)}{\partial x^i} v^i\partial/\partial y^j|_{f(m)}.
\end{aligned}$$

Example 2.26. Let $M = O(n)$, $k \in O(n)$, and $f : O(n) \to O(n)$ be defined by $f(g) := kg$. Then f is a smooth function on $O(n)$ because it is the restriction of a smooth function on $gl(n,\mathbb{R})$. Given $A_g \in T_gO(n)$, by Eq. (2.7),

$$f_*A_g = (kg, kA) = (kA)_{kg}$$

(In the future we denote f by L_k; L_k is **left translation** by $k \in O(n)$.)

Definition 2.27. A **Lie group** is a manifold, G, which is also a group such that the group operations are smooth functions. The tangent space, $\mathfrak{g} := \mathrm{Lie}(G) := T_eG$, to G at the identity $e \in G$ is called the **Lie algebra** of G.

Exercise 2.28. Verify that $GL(n,\mathbb{R})$, $SL(n,\mathbb{R})$, $O(n)$, $SO(n)$ and T^n (see Example 2.6) are all Lie groups and

$$\begin{aligned}
\mathrm{Lie}(GL(n,\mathbb{R})) &\cong gl(n,\mathbb{R}), \\
\mathrm{Lie}(SL(n,\mathbb{R}))) &\cong sl(n,\mathbb{R}) \\
\mathrm{Lie}(O(n))) = \mathrm{Lie}(SO(n))) &\cong so(n,\mathbb{R}) \text{ and} \\
\mathrm{Lie}(T^n)) &\cong (i\mathbb{R})^n \subset \mathbb{C}^n.
\end{aligned}$$

See Exercise 2.15 for the notation being used here.

Exercise 2.29 (Continuation of Exercise 2.17)**.** Show for each chart x on M that the function

$$\phi(v_m) := (x(m), dx(v_m)) = x_*v_m$$

is a chart on TM. Note that $\mathcal{D}(\phi) := \cup_{m\in\mathcal{D}(x)}T_mM$.

The following lemma gives an important example of a smooth function on M which will be needed when we consider M as a "Riemannian manifold."

Lemma 2.30. *Suppose that $(E, \langle\cdot,\cdot\rangle)$ is an inner product space and the $M \subset E$ is an imbedded submanifold. For each $m \in M$, let $P(m)$ denote the orthogonal projection of E onto τ_mM and $Q(m) := I - P(m)$ denote the orthogonal projection onto $\tau_mM^\perp$. Then P and Q are smooth functions from M to $gl(E)$, where $gl(E)$ denotes the vector space of linear maps from E to E.*

Proof. Let $z : E \to \mathbb{R}^N$ be as in Definition 2.2. To simplify notation, let $F(p) := z_>(p)$ for all $p \in \mathcal{D}(z)$, so that $\tau_mM = \mathrm{Nul}\,(F'(m))$ for $m \in \mathcal{D}(x) = \mathcal{D}(z)\cap M$. Since $F'(m) : E \to \mathbb{R}^{N-d}$ is surjective, an elementary exercise in linear algebra shows

$$(F'(m)F'(m)^*) : \mathbb{R}^{N-d} \to \mathbb{R}^{N-d}$$

is invertible for all $m \in \mathcal{D}(x)$. The orthogonal projection $Q(m)$ may be expressed as;

$$Q(m) = F'(m)^*(F'(m)F'(m)^*)^{-1}F'(m). \tag{2.12}$$

Since being invertible is an open condition, $(F'(\cdot)F'(\cdot)^*)$ is invertible in an open neighborhood $\mathcal{N} \subset E$ of $\mathcal{D}(x)$. Hence Q has a smooth extension $\tilde{Q}$ to $\mathcal{N}$ given by

$$\tilde{Q}(x) := F'(x)^*(F'(x)F'(x)^*)^{-1}F'(x).$$

Since $Q|_{\mathcal{D}(x)} = \tilde{Q}|_{\mathcal{D}(x)}$ and $\tilde{Q}$ is smooth on $\mathcal{N}$, $Q|_{\mathcal{D}(x)}$ is also smooth. Since z as in Definition 2.2 was arbitrary and smoothness is a local property, it follows that Q is smooth on M. Clearly, $P := I - Q$ is also a smooth function on M. ■

Definition 2.31. A **local vector field** Y on M is a smooth function $Y : M \to TM$ such that $Y(m) \in T_mM$ for all $m \in \mathcal{D}(Y)$, where $\mathcal{D}(Y)$ is assumed to be an open subset of M. Let $\Gamma(TM)$ denote the collection of globally defined (i.e. $\mathcal{D}(Y) = M$) smooth vector-fields Y on M.

Note that $\partial/\partial x^i$ are local vector-fields on M for each chart $x \in \mathcal{A}(M)$ and $i = 1, 2, \ldots, d$. The next exercise asserts that these vector fields are smooth.

Exercise 2.32. Let Y be a vector field on M, $x \in \mathcal{A}(M)$ be a chart on M and $Y^i := dx^i(Y)$. Then

$$Y(m) := \sum_{i=1}^{d} Y^i(m)\,\partial/\partial x^i|_m \;\forall\; m \in \mathcal{D}(x),$$

which we abbreviate as $Y = \sum_{i=1}^d Y^i\partial/\partial x^i$. Show the condition that Y is smooth translates into the statement that each of the functions Y^i is smooth.

Exercise 2.33. Let $Y : M \to TM$, be a vector field. Then

$$Y(m) = (m, y(m)) = y(m)_m$$

for some function $y : M \to E$ such that $y(m) \in \tau_mM$ for all $m \in \mathcal{D}(Y) = \mathcal{D}(y)$. Show that Y is smooth iff $y : M \to E$ is smooth.

Example 2.34. Let $M = SL(n,\mathbb{R})$ and $A \in sl(n,\mathbb{R}) = \tau_ISL(n,\mathbb{R})$, i.e. A is a $n\times n$ real matrix such that $\mathrm{tr}\,(A) = 0$. Then $\tilde{A}(g) := L_{g*}A_e = (g, gA)$ for $g \in M$ is a smooth vector field on M.

Example 2.35. Keep the notation of Lemma 2.30. Let $y : M \to E$ be any smooth function. Then $Y(m) := (m, P(m)y(m))$ for all $m \in M$ is a smooth vector-field on M.

Definition 2.36. Given $Y \in \Gamma(TM)$ and $f \in C^\infty(M)$, let $Yf \in C^\infty(M)$ be defined by $(Yf)(m) := df(Y(m))$, for all $m \in \mathcal{D}(f) \cap \mathcal{D}(Y)$. In this way the vector-field Y may be viewed as a first order differential operator on $C^\infty(M)$.

Notation 2.37. The **Lie bracket** of two smooth vector fields, Y and W, on M is the vector field $[Y,W]$ which acts on $C^\infty(M)$ by the formula

$$[Y,W]f := Y(Wf) - W(Yf), \quad \forall f \in C^\infty(M). \tag{2.13}$$

(In general one might suspect that $[Y,W]$ is a second order differential operator, however this is not the case, see Exercise 2.38.) Sometimes it will be convenient to write $L_Y W$ for $[Y,W]$.

Exercise 2.38. Show that $[Y,W]$ is again a first order differential operator on $C^\infty(M)$ coming from a vector-field. In particular, if x is a chart on M, $Y = \sum_{i=1}^d Y^i \partial/\partial x^i$ and $W = \sum_{i=1}^d W^i \partial/\partial x^i$, then on $\mathcal{D}(x)$,

$$[Y,W] = \sum_{i=1}^d (YW^i - WY^i)\partial/\partial x^i. \tag{2.14}$$

Proposition 2.39. *If $Y(m) = (m, y(m))$ and $W(m) = (m, w(m))$ and $y, w : M \to E$ are smooth functions such that $y(m), w(m) \in \tau_m M$, then we may express the Lie bracket, $[Y,W](m)$, as*

$$[Y,W](m) = (m, (Yw - Wy)(m)) = (m, dw(Y(m)) - dy(W(m))). \tag{2.15}$$

Proof. Let f be a smooth function M which we may take, by Proposition 2.10, to be the restriction of a smooth function on E. Similarly we we may assume that y and w are smooth functions on E such that $y(m), w(m) \in \tau_m M$ for all $m \in M$. Then

$$\begin{aligned}(YW - WY)f &= Y[f'w] - W[f'y] \\ &= f''(y,w) - f''(w,y) + f'(Yw) - f'(Wy) \\ &= f'(Yw - Wy)\end{aligned} \tag{2.16}$$

wherein the last equality we have use the fact that mixed partial derivatives commute to conclude

$$f''(u,v) - f''(v,u) := (\partial_u\partial_v - \partial_v\partial_u) f = 0 \ \forall\, u, v \in E.$$

Taking $f = z_>$ in Eq. (2.16) with $z = (z_<, z_>)$ being a chart on E as in Definition 2.2, shows

$$0 = (YW - WY)z_>(m) = z'_>(dw(Y(m)) - dy(W(m)))$$

and thus $(m, dw(Y(m)) - dy(W(m))) \in T_m M$. With this observation, we then have

$$f'(Yw - Wy) = df((m, dw(Y(m)) - dy(W(m))))$$

which combined with Eq. (2.16) verifies Eq. (2.15). ∎

Exercise 2.40. Let $M = SL(n,\mathbb{R})$ and $A, B \in sl(n,\mathbb{R})$ and $\tilde{A}$ and $\tilde{B}$ be the associated left invariant vector fields on M as introduced in Example 2.34. Show $\left[\tilde{A}, \tilde{B}\right] = \widetilde{[A,B]}$ where $[A,B] := AB - BA$ is the matrix commutator of A and B.

2.3. **More References.** The reader wishing to learn about manifolds is referred to [1, 9, 19, 41, 42, 95, 111, 112, 113, 114, 115, 164]. The texts by Kobayashi and Nomizu are very thorough while the books by Klingenberg give an idea of why differential geometers are interested in loop spaces. There is a vast literature on Lie groups and there representations. Here are just two books which I have found very useful, [24, 178].

3. Riemannian Geometry Primer

This section introduces the following objects: 1) Riemannian metrics, 2) Riemannian volume forms, 3) gradients, 4) divergences, 5) Laplacians, 6) covariant derivatives, 7) parallel translations, and 8) curvatures.

3.1. Riemannian Metrics.

Definition 3.1. A **Riemannian metric**, $\langle\cdot,\cdot\rangle$ (also denoted by g), on M is a smoothly varying choice of inner product, $g_m = \langle\cdot,\cdot\rangle_m$, on each of the tangent spaces $T_m M$, $m \in M$. The smoothness condition is the requirement that the function $m \in M \to \langle X(m), Y(m)\rangle_m \in \mathbb{R}$ is smooth for all smooth vector fields X and Y on M.

It is customary to write ds^2 for the function on TM defined by

$$ds^2(v_m) := \langle v_m, v_m\rangle_m = g_m(v_m, v_m). \tag{3.1}$$

By polarization, the Riemannian metric $\langle\cdot,\cdot\rangle$ is uniquely determined by the function ds^2. Given a chart x on M and $v \in T_m M$, by Eqs. (3.1) and (2.6) we have

$$ds^2(v_m) = \sum_{i,j=1}^d \langle \partial/\partial x^i|_m, \partial/\partial x^j|_m\rangle_m dx^i(v_m) dx^j(v_m). \tag{3.2}$$

We will abbreviate this equation in the future by writing

$$ds^2 = \sum_{i,j=1}^d g^x_{ij} dx^i dx^j \tag{3.3}$$

where

$$g^x_{i,j}(m) := \langle \partial/\partial x^i|_m, \partial/\partial x^j|_m \rangle_m = g\left(\partial/\partial x^i|_m, \partial/\partial x^j|_m\right).$$

Typically $g^x_{i,j}$ will be abbreviated by g_{ij} if no confusion is likely to arise.

Example 3.2. Let $M = \mathbb{R}^N$ and let $x = (x^1, x^2, \ldots, x^N)$ denote the standard chart on M, i.e. $x(m) = m$ for all $m \in M$. The standard Riemannian metric on $\mathbb{R}^N$ is determined by

$$ds^2 = \sum_{i=1}^N (dx^i)^2 = \sum_{i=1}^N dx^i \cdot dx^i,$$

and so g^x is the identity matrix here. The general Riemannian metric on $\mathbb{R}^N$ is determined by $ds^2 = \sum_{i,j=1}^N g_{ij} dx^i dx^j$, where $g = (g_{ij})$ is a smooth $gl(N, \mathbb{R})$ – valued function on $\mathbb{R}^N$ such that $g(m)$ is positive definite matrix for all $m \in \mathbb{R}^N$.

Let M be an imbedded submanifold of a finite dimensional inner product space $(E, \langle \cdot, \cdot \rangle)$. The manifold M **inherits** a metric from E determined by

$$ds^2(v_m) = \langle v, v \rangle \ \forall \ v_m \in TM.$$

It is a well known deep fact that **all** finite dimensional Riemannian manifolds may be constructed in this way, see Nash [143] and Moser [138, 139, 140]. To simplify the exposition, in the sequel we will usually assume that $(E, \langle \cdot, \cdot \rangle)$ is an inner product space, $M^d \subset E$ is an imbedded submanifold, and the Riemannian metric on M is determined in this way, i.e.

$$\langle v_m, w_m \rangle = \langle v, w \rangle_{\mathbb{R}^N}, \quad \forall \ v_m, w_m \in T_m M \text{ and } m \in M.$$

In this setting the components $g^x_{i,j}$ of the metric ds^2 relative to a chart x may be computed as $g^x_{i,j}(m) = \langle \phi_{;i}(x(m)), \phi_{;j}(x(m)) \rangle$, where $\{e_i\}_{i=1}^d$ is the standard basis for $\mathbb{R}^d$,

$$\phi := x^{-1} \text{ and } \phi_{;i}(a) := \frac{d}{dt}|_0 \phi(a + te_i).$$

Example 3.3. Let $M = G := SL(n, \mathbb{R})$ and $A_g \in T_g M$.

(1) Then

$$ds^2(A_g) := \operatorname{tr}(A^* A) \tag{3.4}$$

defines a Riemannian metric on G. This metric is the inherited metric from the inner product space $E = gl(n, \mathbb{R})$ with inner product $\langle A, B \rangle := \operatorname{tr}(A^* B)$.

(2) A more "natural" choice of a metric on G is

$$ds^2(A_g) := \operatorname{tr}((g^{-1}A)^* g^{-1}A). \tag{3.5}$$

This metric is invariant under left translations, i.e. $ds^2(L_{k*}A_g) = ds^2(A_g)$, for all $k \in G$ and $A_g \in TG$. According to the imbedding theorem of Nash and Moser, it would be possible to find another imbedding of G into a Euclidean space, E, so that the metric in Eq. (3.5) is inherited from an inner product on E.

Example 3.4. Let $M = \mathbb{R}^3$ be equipped with the standard Riemannian metric and (r, φ, θ) be spherical coordinates on M, see Figure 7. Here r, φ, and θ

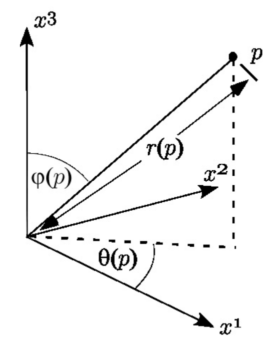

FIGURE 7. Defining the spherical coordinates, (r, θ, ϕ) on $\mathbb{R}^3$.

are taken to be functions on $\mathbb{R}^3 \setminus \{p \in \mathbb{R}^3 : p_2 = 0 \text{ and } p_1 > 0\}$ defined by $r(p) = |p|$, $\varphi(p) = \cos^{-1}(p_3/|p|) \in (0, \pi)$, and $\theta(p) \in (0, 2\pi)$ is given by $\theta(p) = \tan^{-1}(p_2/p_1)$ if $p_1 > 0$ and $p_2 > 0$ with similar formulas for (p_1, p_2) in the other three quadrants of $\mathbb{R}^2$. Since $x^1 = r \sin\varphi \cos\theta$, $x^2 = r \sin\varphi \sin\theta$, and $x^3 = r\cos\varphi$, it follows using Eq. (2.11) that,

$$\begin{aligned} dx^1 &= \frac{\partial x^1}{\partial r} dr + \frac{\partial x^1}{\partial \varphi} d\varphi + \frac{\partial x^1}{\partial \theta} d\theta \\ &= \sin\varphi\cos\theta dr + r\cos\varphi\cos\theta d\varphi - r\sin\varphi\sin\theta d\theta, \end{aligned}$$

$$dx^2 = \sin\varphi\sin\theta dr + r\cos\varphi\sin\theta d\varphi + r\sin\varphi\cos\theta d\theta,$$

and

$$dx^3 = \cos\varphi dr - r\sin\varphi d\varphi.$$

An elementary calculation now shows that

$$ds^2 = \sum_{i=1}^3 (dx^i)^2 = dr^2 + r^2 d\varphi^2 + r^2 \sin^2\varphi d\theta^2. \tag{3.6}$$

From this last equation, we see that

$$g^{(r,\varphi,\theta)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\varphi \end{bmatrix}. \tag{3.7}$$

Exercise 3.5. Let $M := \{m \in \mathbb{R}^3 : |m|^2 = \rho^2\}$, so that M is a sphere of radius ρ in $\mathbb{R}^3$. Since $r = \rho$ on M and $dr(v) = 0$ for all $v \in T_mM$, it follows from Eq. (3.6) that the induced metric ds^2 on M is given by

$$ds^2 = \rho^2 d\varphi^2 + \rho^2 \sin^2\varphi d\theta^2, \tag{3.8}$$

and hence

$$g^{(\varphi,\theta)} = \begin{bmatrix} \rho^2 & 0 \\ 0 & \rho^2 \sin^2\varphi \end{bmatrix}. \tag{3.9}$$

3.2. Integration and the Volume Measure.

Definition 3.6. Let $f \in C_c^\infty(M)$ (the smooth functions on M^d with compact support) and assume the support of f is contained in $\mathcal{D}(x)$, where x is some chart on M. Set

$$\int_M f dx = \int_{\mathcal{R}(x)} f \circ x^{-1}(a) da,$$

where da denotes Lebesgue measure on $\mathbb{R}^d$.

The problem with this notion of integration is that (as the notation indicates) $\int_M f dx$ depends on the choice of chart x. To remedy this, consider a small cube $C(\delta)$ of side δ contained in $\mathcal{R}(x)$, see Figure 8. We wish to estimate "the volume" of $\phi(C(\delta))$ where $\phi := x^{-1} : \mathcal{R}(x) \to \mathcal{D}(x)$. Heuristically, we expect the volume of $\phi(C(\delta))$ to be approximately equal to the volume of the parallelepiped, $\tilde{C}(\delta)$, in the tangent space T_mM determined by

$$\tilde{C}(\delta) := \left\{ \sum_{i=1}^d s_i\delta \cdot \phi_{;i}(x(m)) | 0 \le s_i \le 1, \text{ for } i = 1, 2, \dots, d \right\}, \tag{3.10}$$

where we are using the notation proceeding Example 3.3, see Figure 8. Since T_mM is an inner product space, the volume of $\tilde{C}(\delta)$ is well defined. For example choose an isometry $\theta : T_mM \to \mathbb{R}^d$ and define the volume of $\tilde{C}(\delta)$ to be $m\left(\theta(\tilde{C}(\delta))\right)$ where m is Lebesgue measure on $\mathbb{R}^d$. The next elementary lemma will be used to give a formula for the volume of $\tilde{C}(\delta)$.

Lemma 3.7. *If V is a finite dimensional inner product space, $\{v_i\}_{i=1}^{\dim V}$ is any basis for V and $A : V \to V$ is a linear transformation, then*

$$\det(A) = \frac{\det[\langle Av_i, v_j\rangle]}{\det[\langle v_i, v_j\rangle]}, \tag{3.11}$$

where $\det[\langle Av_i, v_j\rangle]$ is the determinant of the matrix with i-j^{th} – entry being $\langle Av_i, v_j\rangle$. Moreover if

$$\tilde{C}(\delta) := \left\{ \sum_{i=1}^d \delta s_i \cdot v_i : 0 \le s_i \le 1, \text{ for } i = 1, 2, \dots, d \right\}$$

then the volume of $\tilde{C}(\delta)$ is $\delta^d \sqrt{\det[\langle v_i, v_j\rangle]}$.

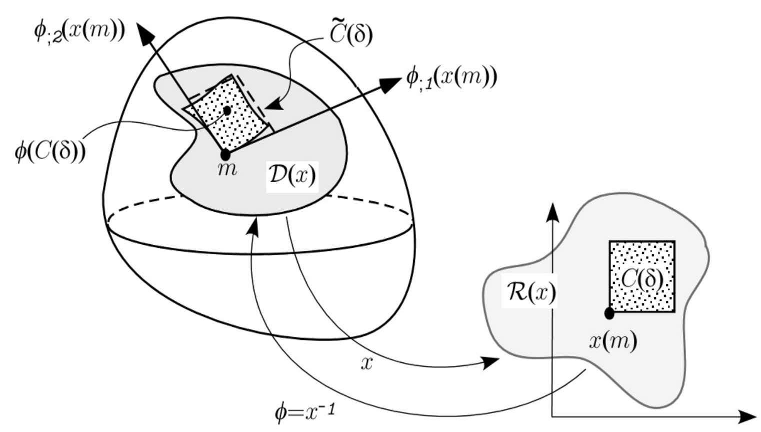

FIGURE 8. Defining the Riemannian "volume element."

Proof. Let $\{e_i\}_{i=1}^{\dim V}$ be an orthonormal basis for V, then

$$\langle Av_i, v_j\rangle = \sum_{l,k} \langle v_i, e_l\rangle \langle Ae_l, e_k\rangle \langle e_k, v_j\rangle$$

and therefore by the multiplicative property of the determinant,

$$\begin{aligned} \det[\langle Av_i, v_j\rangle] &= \det[\langle v_i, e_l\rangle] \det[\langle Ae_l, e_k\rangle] \det[\langle e_k, v_j\rangle] \\ &= \det(A) \det[\langle v_i, e_l\rangle] \cdot \det[\langle e_k, v_j\rangle]. \end{aligned} \tag{3.12}$$

Taking $A = I$ in this equation then shows

$$\det[\langle v_i, v_j\rangle] = \det[\langle v_i, e_l\rangle] \cdot \det[\langle e_k, v_j\rangle]. \tag{3.13}$$

Dividing Eq. (3.13) into Eq. (3.12) proves Eq. (3.11).

For the second assertion, it suffices to assume $V = \mathbb{R}^d$ with the usual inner-product. Define $T : \mathbb{R}^d \to \mathbb{R}^d$ so that $Te_i = v_i$ where $\{e_i\}_{i=1}^d$ is the standard basis for $\mathbb{R}^d$, then $\tilde{C}(\delta) = T\left([0,\delta]^d\right)$ and hence

$$\begin{aligned} m\left(\tilde{C}(\delta)\right) &= |\det T| \, m\left([0,\delta]^d\right) = \delta^d |\det T| = \delta^d \sqrt{\det(T^{\mathrm{tr}}T)} \\ &= \delta^d \sqrt{\det[\langle T^{\mathrm{tr}}Te_i, e_j\rangle]} = \delta^d \sqrt{\det[\langle Te_i, Te_j\rangle]} = \delta^d \sqrt{\det[\langle v_i, v_j\rangle]}. \end{aligned}$$

■

Using the second assertion in Lemma 3.7, the volume of $\tilde{C}(\delta)$ in Eq. (3.10) is $\delta^d \sqrt{\det g^x(m)}$, where $g_{ij}^x(m) = \langle \phi_{;i}(x(m)), \phi_{;j}(x(m))\rangle_m$. Because of the above computations, it is reasonable to try to define a new integral on $\mathcal{D}(x) \subset M$ by

$$\int_{\mathcal{D}(x)} f \, d\lambda_{\mathcal{D}(x)} := \int_{\mathcal{D}(x)} f \sqrt{g^x} dx,$$

i.e. let $\lambda_{\mathcal{D}(x)}$ be the measure satisfying

$$d\lambda_{\mathcal{D}(x)} = \sqrt{g^x}dx, \tag{3.14}$$

where $\sqrt{g^x}$ is shorthand for $\sqrt{\det g^x}$.

Lemma 3.8. *Suppose that y and x are two charts on M, then*

$$g^y_{l,k} = \sum_{i,j=1}^d g^x_{i,j}\frac{\partial x^i}{\partial y^k}\frac{\partial x^j}{\partial y^l}. \tag{3.15}$$

Proof. Inserting the identities

$$dx^i = \sum_{k=1}^d \frac{\partial x^i}{\partial y^k}dy^k \text{ and } dx^j = \sum_{l=1}^d \frac{\partial x^j}{\partial y^l}dy^l$$

and into the formula $ds^2 = \sum_{i,j=1}^d g^x_{i,j}dx^i dx^j$ gives

$$ds^2 = \sum_{i,j,k,l=1}^d g^x_{i,j}\frac{\partial x^i}{\partial y^k}\frac{\partial x^j}{\partial y^l}dy^l dy^k$$

from which (3.15) follows. ∎

Exercise 3.9. Suppose that x and y are two charts on M and $f \in C_c^\infty(M)$ such that the support of f is contained in $\mathcal{D}(x)\cap\mathcal{D}(y)$. Using Lemma 3.8 and the change of variable formula show,

$$\int_{\mathcal{D}(x)\cap\mathcal{D}(y)} f\sqrt{g^x}dx = \int_{\mathcal{D}(x)\cap\mathcal{D}(y)} f\sqrt{g^y}dy.$$

Theorem 3.10 (Riemann Volume Measure)**.** *There exists a unique measure, λ_M on the Borel σ – algebra of M such that for any chart x on M,*

$$d\lambda_M(x) = d\lambda_{\mathcal{D}(x)} = \sqrt{g^x}dx \text{ on } \mathcal{D}(x). \tag{3.16}$$

Proof. Choose a countable collection of charts, $\{x_i\}_{i=1}^\infty$ such that $M = \cup_{i=1}^\infty \mathcal{D}(x_i)$ and let $U_1 := \mathcal{D}(x_1)$ and $U_i := \mathcal{D}(x_i)\setminus(\cup_{j=1}^{i-1}\mathcal{D}(x_j))$ for $i \geq 1$. Then if $B \subset X$ is a Borel set, define the measure $\lambda_M(B)$ by

$$\lambda_M(B) := \sum_{i=1}^\infty \lambda_{\mathcal{D}(x_i)}(B\cap U_i). \tag{3.17}$$

If x is any chart on M and $B \subset \mathcal{D}(x)$, then $B\cap U_i \subset \mathcal{D}(x_i)\cap\mathcal{D}(x)$ and so by Exercise 3.9, $\lambda_{\mathcal{D}(x_i)}(B\cap U_i) = \lambda_{\mathcal{D}(x)}(B)$. Using this identity in Eq. (3.17) implies

$$\lambda_M(B) := \sum_{i=1}^\infty \lambda_{\mathcal{D}(x)}(B\cap U_i) = \lambda_{\mathcal{D}(x)}(B)$$

and hence we have proved the existence of λ_M. The uniqueness assertion is easy and will be left to the reader. ∎

Example 3.11. Let $M = \mathbb{R}^3$ with the standard Riemannian metric, and let x denote the standard coordinates on M determined by $x(m) = m$ for all $m \in M$. Then $\lambda_{\mathbb{R}^3}$ is Lebesgue measure which in spherical coordinates may be written as

$$d\lambda_{\mathbb{R}^3} = r^2\sin\varphi dr d\varphi d\theta$$

because $\sqrt{g^{(r,\varphi,\theta)}} = r^2\sin\varphi$ by Eq. (3.7). Similarly using Eq. (3.9),

$$d\lambda_M = \rho^2\sin\varphi d\varphi d\theta$$

when $M \subset \mathbb{R}^3$ is the sphere of radius ρ centered at $0 \in \mathbb{R}^3$.

Exercise 3.12. Compute the "volume element," $d\lambda_{\mathbb{R}^3}$, for $\mathbb{R}^3$ in cylindrical coordinates.

Theorem 3.13 (Change of Variables Formula)**.** *Let $(M, \langle\cdot,\cdot\rangle_M)$ and $(N, \langle\cdot,\cdot\rangle_N)$ be two Riemannian manifolds, $\psi : M \to N$ be a diffeomorphism and $\rho \in C^\infty(M,(0,\infty))$ be determined by the equation*

$$\rho(m) = \sqrt{\det[\psi_{*m}^{\mathrm{tr}}\psi_{*m}]} \text{ for all } m \in M,$$

where ψ_{*m}^{tr} denotes the adjoint of ψ_{*m} relative to Riemannian inner products on T_mM and $T_{\psi(m)}N$. If $f : N \to \mathbb{R}_+$ is a positive Borel measurable function, then

$$\int_N f d\lambda_N = \int_M \rho\cdot(f\circ\psi)\,d\lambda_M.$$

In particular if ψ is an isometry, i.e. $\psi_{*m} : T_mM \to T_{\psi(m)}N$ is orthogonal for all m, then

$$\int_N f d\lambda_N = \int_M f\circ\psi \; d\lambda_M.$$

Proof. By a partition of unity argument (see the proof of Theorem 3.10), it suffices to consider the case where f has "small" support, i.e. we may assume that the support of $f\circ\psi$ is contained in $\mathcal{D}(x)$ for some chart x on M. Letting $\phi := x^{-1}$, by Eq. (3.11) of Lemma 3.7,

$$\begin{aligned}
&\frac{\det[\langle\partial_i(\psi\circ\phi)(t), \partial_j(\psi\circ\phi)(t)\rangle_N]}{\det[\langle\partial_i\phi(t), \partial_j\phi(t)\rangle_M]} \\
&\quad = \frac{\det[\langle\psi_*\partial_i\phi(t), \psi_*\partial_j\phi(t)\rangle_N]}{\det[\langle\partial_i\phi(t), \partial_j\phi(t)\rangle_M]} = \frac{\det[\langle\psi_*^{\mathrm{tr}}\psi_*\partial_i\phi(t), \partial_j\phi(t)\rangle_M]}{\det[\langle\partial_i\phi(t), \partial_j\phi(t)\rangle_M]} \\
&\quad = \det\left[\psi_{*\phi(t)}^{\mathrm{tr}}\psi_{*\phi(t)}\right] = \rho^2(\phi(t)).
\end{aligned}$$

This implies

$$\int_N f d\lambda_N = \int_{\mathcal{R}(x)} f \circ (\psi \circ \phi)(t) \sqrt{\det [\langle \partial_i (\psi \circ \phi)(t), \partial_j (\psi \circ \phi)(t) \rangle_N]} dt$$
$$= \int_{\mathcal{R}(x)} (f \circ \psi) \circ \phi(t) \cdot \rho(\phi(t)) \sqrt{\det [\langle \partial_i \phi(t), \partial_j \phi(t) \rangle_M]} dt$$
$$= \int_{\mathcal{D}(x)} (f \circ \psi) \cdot \rho \cdot \sqrt{g^x} dx = \int_M \rho \cdot f \circ \psi \, d\lambda_M.$$

■

Example 3.14. Let $M = SL(n, \mathbb{R})$ as in Example 3.3 and let $\langle \cdot, \cdot \rangle_M$ be the metric given by Eq. (3.5). Because $L_g : M \to M$ is an isometry, Theorem 3.13 implies

$$\int_{SL(n,\mathbb{R})} f(gx) \, d\lambda_G(x) = \int_{SL(n,\mathbb{R})} f(x) \, d\lambda_G(x) \text{ for all } g \in G.$$

That is λ_G is invariant under left translations by elements of G and such an invariant left invariant measure is called a "**left Haar**" measure on G.

Similarly if $G = O(n)$ with Riemannian metric determined by Eq. (3.5), then, since $g \in G$ is orthogonal, we have

$$ds^2(A_g) := \mathrm{tr}((g^{-1}A)^* g^{-1}A) = \mathrm{tr}((g^*A)^* g^{-1}A) = \mathrm{tr}(A^* g g^{-1} A) = \mathrm{tr}(A^* A)$$

and

$$\mathrm{tr}((Ag^{-1})^* A g^{-1}) = \mathrm{tr}(g A^* A g^{-1}) = \mathrm{tr}(A^* A g^{-1} g) = \mathrm{tr}(A^* A).$$

Therefore, both left and right translations by element $g \in G$ are isometries for this Riemannian metric on $O(m)$ and so by Theorem 3.13,

$$\int_{O(n)} f(gx) \, d\lambda_G(x) = \int_{O(n)} f(x) \, d\lambda_G(x) = \int_{O(n)} f(xg) \, d\lambda_G(x)$$

for all $g \in G$.

3.3. **Gradients, Divergence, and Laplacians.** In the sequel, let M be a Riemannian manifold, x be a chart on M, $g_{ij} := \langle \partial/\partial x^i, \partial/\partial x^j \rangle$, and $ds^2 = \sum_{i,j=1}^d g_{ij} dx^i dx^j$.

Definition 3.15. Let g^{ij} denote the i-j^{th} – matrix element for the inverse matrix to the matrix, (g_{ij}).

Given $f \in C^\infty(M)$ and $m \in M$, $df_m := df|_{T_m M}$ is a linear functional on $T_m M$. Hence there is a unique vector $v_m \in T_m M$ such that $df_m = \langle v_m, \cdot \rangle_m$.

Definition 3.16. The vector v_m above is called the **gradient** of f at m and will be denoted by either $\mathrm{grad}\, f(m)$ or $\vec{\nabla} f(m)$.

Exercise 3.17. If x is a chart on M and $m \in \mathcal{D}(x)$ then

$$\vec{\nabla} f(m) = \mathrm{grad} f(m) = \sum_{i,j=1}^d g^{ij}(m) \frac{\partial f(m)}{\partial x^i} \frac{\partial}{\partial x^j}|_m, \tag{3.18}$$

where as usual, $g_{ij} = g^x_{ij}$ and $g^{ij} = (g_{ij})^{-1}$. Notice from Eq. (3.18) that $\vec{\nabla} f$ is a smooth vector field on M.

Exercise 3.18. Suppose $M \subset \mathbb{R}^N$ is an imbedded submanifold with the induced Riemannian structure. Let $F : \mathbb{R}^N \to \mathbb{R}$ be a smooth function and set $f := F|_M$. Then $\mathrm{grad}\, f(m) = (P(m) \vec{\nabla} F(m))_m$, where $\vec{\nabla} F(m)$ denotes the usual gradient on $\mathbb{R}^N$, and $P(m)$ denotes orthogonal projection of $\mathbb{R}^N$ onto $\tau_m M$.

We now introduce the divergence of a vector field Y on M.

Lemma 3.19 (Divergence)**.** *To every smooth vector field Y on M there is a unique smooth function, $\vec{\nabla} \cdot Y = \mathrm{div}\, Y$, on M such that*

$$\int_M Y f \, d\lambda_M = -\int_M \mathrm{div} Y \cdot f \, d\lambda_M, \quad \forall f \in C_c^\infty(M). \tag{3.19}$$

*(The function, $\vec{\nabla} \cdot Y = \mathrm{div}\, Y$, is called the **divergence** of Y.) Moreover if x is a chart on M, then on its domain, $\mathcal{D}(x)$,*

$$\vec{\nabla} \cdot Y = \mathrm{div}\, Y = \sum_{i=1}^d \frac{1}{\sqrt{g}} \frac{\partial(\sqrt{g} Y^i)}{\partial x^i} = \sum_{i=1}^d \{ \frac{\partial Y^i}{\partial x^i} + \frac{\partial \log \sqrt{g}}{\partial x^i} Y^i \} \tag{3.20}$$

where $Y^i := dx^i(Y)$ and $\sqrt{g} = \sqrt{g^x} = \sqrt{\det(g^x_{ij})}$.

Proof. (Sketch) Suppose that $f \in C_c^\infty(M)$ such that the support of f is contained in $\mathcal{D}(x)$. Because $Yf = \sum_{i=1}^d Y^i \partial f / \partial x^i$,

$$\int_M Y f \, d\lambda_M = \int_M \sum_{i=1}^d Y^i \partial f / \partial x^i \cdot \sqrt{g} dx = -\int_M \sum_{i=1}^d f \frac{\partial(\sqrt{g}\, Y^i)}{\partial x^i} dx$$
$$= -\int_M f \sum_{i=1}^d \frac{1}{\sqrt{g}} \frac{\partial(\sqrt{g} Y^i)}{\partial x^i} \, d\lambda_M,$$

where the second equality follows by an integration by parts. This shows that if $\mathrm{div}\, Y$ exists it must be given on $\mathcal{D}(x)$ by Eq. (3.20). This proves the uniqueness assertion. Using what we have already proved, it is easy to conclude that the formula for $\mathrm{div}\, Y$ is chart independent. Hence we may define smooth function $\mathrm{div}\, Y$ on M using Eq. (3.20) in each coordinate chart x on M. It is then possible to show (again using a smooth partition of unity argument) that this function satisfies Eq. (3.19). ■

Remark 3.20. We may write Eq. (3.19) as

$$\int_M \langle Y, \operatorname{grad} f\rangle \, d\lambda_M = -\int_M \operatorname{div} Y \cdot f \, d\lambda_M, \quad \forall\, f \in C_c^\infty(M), \tag{3.21}$$

so that "div" is the negative of the formal adjoint of "grad."

Exercise 3.21 (Product Rule)**.** If $f \in C^\infty(M)$ and $Y \in \Gamma(TM)$ then

$$\vec{\nabla} \cdot (fY) = \langle \vec{\nabla} f, Y\rangle + f\, \vec{\nabla} \cdot Y.$$

Lemma 3.22 (Integration by Parts)**.** *Suppose that* $Y \in \Gamma(TM)$, $f \in C_c^\infty(M)$, *and* $h \in C^\infty(M)$, *then*

$$\int_M Yf \cdot h \, d\lambda_M = \int_M f\{-Yh - h \cdot \operatorname{div} Y\} \, d\lambda_M.$$

Proof. By the definition of $\operatorname{div} Y$ and the product rule,

$$\int_M fh \operatorname{div} Y \, d\lambda_M = -\int_M Y(fh) \, d\lambda_M = -\int_M \{hYf + fYh\} \, d\lambda_M.$$

■

Definition 3.23. The **Laplacian** on M is the second order differential operator, $\Delta : C^\infty(M) \to C^\infty(M)$, defined by

$$\Delta f := \operatorname{div}(\operatorname{grad} f) = \vec{\nabla} \cdot \vec{\nabla} f. \tag{3.22}$$

In local coordinates,

$$\Delta f = \frac{1}{\sqrt{g}} \sum_{i,j=1}^d \partial_i \{\sqrt{g} g^{ij} \partial_j f\}, \tag{3.23}$$

where $\partial_i = \partial/\partial x^i$, $g = g^x$, $\sqrt{g} = \sqrt{\det g}$, and $(g^{ij}) = (g^x_{ij})^{-1}$.

Remark 3.24. The Laplacian, Δf, may be characterized by the equation:

$$\int_M \Delta f \cdot h \, d\lambda_M = -\int_M \langle \vec{\nabla} f, \vec{\nabla} h\rangle \, d\lambda_M,$$

which is to hold for all $f \in C^\infty(M)$ and $h \in C_c^\infty(M)$.

Example 3.25. Suppose that $M = \mathbb{R}^N$ with the standard Riemannian metric $ds^2 = \sum_{i=1}^N (dx^i)^2$, then the standard formulas:

$$\operatorname{grad} f = \sum_{i=1}^N \partial f/\partial x^i \cdot \partial/\partial x^i, \ \operatorname{div} Y = \sum_{i=1}^N \partial Y^i/\partial x^i \text{ and } \Delta f = \sum_{i=1}^N \frac{\partial^2 f}{(\partial x^i)^2}$$

are easily verified, where f is a smooth function on $\mathbb{R}^N$ and $Y = \sum_{i=1}^N Y^i \partial/\partial x^i$ is a smooth vector-field.

Exercise 3.26. Let $M = \mathbb{R}^3$, (r, φ, θ) be spherical coordinates on $\mathbb{R}^3$, $\partial_r = \partial/\partial r$, $\partial_\varphi = \partial/\partial\varphi$, and $\partial_\theta = \partial/\partial\theta$. Given a smooth function f and a vector-field $Y = Y_r \partial_r + Y_\varphi \partial_\varphi + Y_\theta \partial_\theta$ on $\mathbb{R}^3$ verify:

$$\operatorname{grad} f = (\partial_r f)\partial_r + \frac{1}{r^2}(\partial_\varphi f)\partial_\varphi + \frac{1}{r^2 \sin^2\varphi}(\partial_\theta f)\partial_\theta,$$

$$\begin{aligned} \operatorname{div} Y &= \frac{1}{r^2 \sin\varphi}\{\partial_r(r^2 \sin\varphi Y_r) + \partial_\varphi(r^2 \sin\varphi Y_\varphi) + r^2 \sin\varphi \partial_\theta Y_\theta\} \\ &= \frac{1}{r^2}\partial_r(r^2 Y_r) + \frac{1}{\sin\varphi}\partial_\varphi(\sin\varphi Y_\varphi) + \partial_\theta Y_\theta, \end{aligned}$$

and

$$\Delta f = \frac{1}{r^2}\partial_r(r^2 \partial_r f) + \frac{1}{r^2 \sin\varphi}\partial_\varphi(\sin\varphi \partial_\varphi f) + \frac{1}{r^2 \sin^2\varphi}\partial_\theta^2 f.$$

Example 3.27. Let $M = G = O(n)$ with Riemannian metric determined by Eq. (3.5) and for $A \in \mathfrak{g} := T_e G$ let $\tilde{A} \in \Gamma(TG)$ be the left invariant vector field,

$$\tilde{A}(x) := L_{x*}A = \frac{d}{dt}|_0 x e^{tA}$$

as was done for $SL(n, \mathbb{R})$ in Example 2.34. Using the invariance of $d\lambda_G$ under right translations established in Example 3.14, we find for $f, h \in C^1(G)$ that

$$\begin{aligned} \int_G \tilde{A} f(x) \cdot h(x) \, d\lambda_G(x) &= \int_G \frac{d}{dt}|_0 f\left(x e^{tA}\right) \cdot h(x) \, d\lambda_G(x) \\ &= \frac{d}{dt}|_0 \int_G f\left(x e^{tA}\right) \cdot h(x) \, d\lambda_G(x) \\ &= \frac{d}{dt}|_0 \int_G f(x) \cdot h\left(x e^{-tA}\right) d\lambda_G(x) \\ &= \int_G f(x) \cdot \frac{d}{dt}|_0 h\left(x e^{-tA}\right) d\lambda_G(x) \\ &= -\int_G f(x) \cdot \tilde{A} h(x) \, d\lambda_G(x). \end{aligned}$$

Taking $h \equiv 1$ implies

$$\begin{aligned} 0 = \int_G \tilde{A} f(x) \, d\lambda_G(x) &= \int_G \left\langle \tilde{A}(x), \vec{\nabla} f(x)\right\rangle d\lambda_G(x) \\ &= -\int_G \vec{\nabla} \cdot \tilde{A}(x) \cdot f(x) \, d\lambda_G(x) \end{aligned}$$

from which we learn $\vec{\nabla} \cdot \tilde{A} = 0$.

Now letting $S_0 \subset \mathfrak{g}$ be an orthonormal basis for $\mathfrak{g}$, because L_{g*} is an isometry, $\{\tilde{A}(g) : A \in S_0\}$ is an orthonormal basis for $T_g G$ for all $g \in G$. Hence

$$\vec{\nabla} f(g) = \sum_{A \in S_0} \left\langle \vec{\nabla} f(g), \tilde{A}(g) \right\rangle \tilde{A}(g) = \sum_{A \in S_0} \left(\tilde{A} f \right)(g) \tilde{A}(g).$$

and, by the product rule and $\vec{\nabla} \cdot \tilde{A} = 0$,

$$\Delta f = \vec{\nabla} \cdot \vec{\nabla} f = \sum_{A \in S_0} \vec{\nabla} \cdot \left[\left(\tilde{A} f \right) \tilde{A} \right] = \sum_{A \in S_0} \left\langle \vec{\nabla} \tilde{A} f, \tilde{A} \right\rangle = \sum_{A \in S_0} \tilde{A}^2 f.$$

3.4. **Covariant Derivatives and Curvature.**

Definition 3.28. We say a smooth path $s \to V(s)$ in TM is a **vector-field along a smooth path** $s \to \sigma(s)$ in M if $\pi \circ V(s) = \sigma(s)$, i.e. $V(s) \in T_{\sigma(s)} M$ for all s. (Recall that π is the canonical projection defined in Definition 2.16.)

Note: if V is a smooth path in TM then V is a vector-field along $\sigma := \pi \circ V$. This section is motivated by the desire to have the notion of the derivative of a smooth path $V(s) \in TM$. On one hand, since TM is a manifold, we may write $V'(s)$ as an element of TTM. However, this is not what we will want for later purposes. We would like the derivative of V to again be a path back in TM, not in TTM. In order to define such a derivative, we will need to use more than just the manifold structure of M, see Definition 3.31 below.

Notation 3.29. In the sequel, we assume that M^d is an imbedded submanifold of an inner product space $(E = \mathbb{R}^N, \langle \cdot, \cdot \rangle)$, and that M is equipped with the inherited Riemannian metric. Also let $P(m)$ denote orthogonal projection of E onto $\tau_m M$ for all $m \in M$ and $Q(m) := I - P(m)$ be orthogonal projection onto $(\tau_m M)^\perp$.

The following elementary lemma will be used throughout the sequel.

Lemma 3.30. *The differentials of the orthogonal projection operators, P and Q, satisfy*

$$\begin{aligned} 0 &= dP + dQ, \\ PdQ &= -dPQ = dQQ \text{ and} \\ QdP &= -dQP = dPP. \end{aligned}$$

In particular,

$$QdPQ = QdQQ = PdPP = PdQP = 0.$$

Proof. The first equality comes from differentiating the identity, $I = P + Q$, the second from differentiating $0 = PQ$ and the third from differentiating $0 = QP$. ∎

Definition 3.31 (Levi-Civita Covariant Derivative)**.** Let $V(s) = (\sigma(s), v(s)) = v(s)_{\sigma(s)}$ be a smooth path in TM (see Figure 9), then the **covariant derivative**, $\nabla V(s)/ds$, is the vector field along σ defined by

$$\frac{\nabla V(s)}{ds} := (\sigma(s), P(\sigma(s)) \frac{d}{ds} v(s)). \tag{3.24}$$

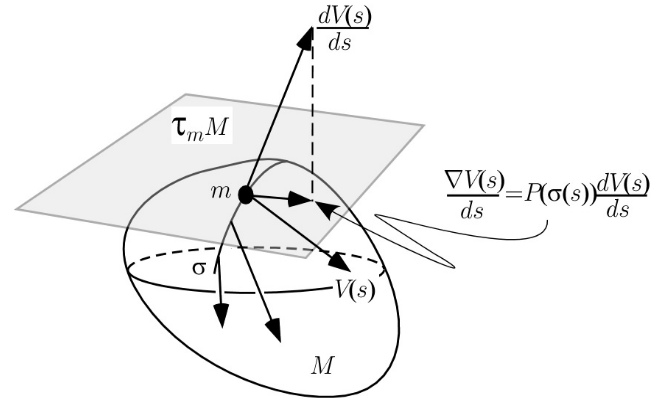

FIGURE 9. The Levi-Civita covariant derivative.

Proposition 3.32 (Properties of ∇/ds)**.** *Let $W(s) = (\sigma(s), w(s))$ and $V(s) = (\sigma(s), v(s))$ be two smooth vector fields along a path σ in M. Then:*

(1) *$\nabla W(s)/ds$ may be computed as:*

$$\frac{\nabla W(s)}{ds} := (\sigma(s), \frac{d}{ds} w(s) + (dQ(\sigma'(s))) w(s)). \tag{3.25}$$

(2) *∇ is* ***metric compatible****, i.e.*

$$\frac{d}{ds} \langle W(s), V(s) \rangle = \langle \frac{\nabla W(s)}{ds}, V(s) \rangle + \langle W(s), \frac{\nabla V(s)}{ds} \rangle. \tag{3.26}$$

Now suppose that $(s,t) \to \sigma(s,t)$ is a smooth function into M, $W(s,t) = (\sigma(s,t), w(s,t))$ is a smooth function into TM, $\sigma'(s,t) := (\sigma(s,t), \frac{d}{ds}\sigma(s,t))$ and $\dot{\sigma}(s,t) = (\sigma(s,t), \frac{d}{dt}\sigma(s,t))$. (Notice by assumption that $w(s,t) \in T_{\sigma(s,t)} M$ for all (s,t).)

(3) *∇ has* ***zero torsion****, i.e.*

$$\frac{\nabla \sigma'}{dt} = \frac{\nabla \dot{\sigma}}{ds}. \tag{3.27}$$

(4) *If R is the* ***curvature tensor*** *of ∇ defined by*

$$R(u_m, v_m) w_m = (m, [dQ(u_m), dQ(v_m)] w), \tag{3.28}$$

then

$$\left[\frac{\nabla}{dt},\frac{\nabla}{ds}\right]W:=(\frac{\nabla}{dt}\frac{\nabla}{ds}-\frac{\nabla}{ds}\frac{\nabla}{dt})W=R(\dot\sigma,\sigma')W. \tag{3.29}$$

Proof. Differentiate the identity, $P(\sigma(s))w(s)=w(s)$, relative to s implies

$$(dP(\sigma'(s)))w(s)+P(\sigma(s))\frac{d}{ds}w(s)=\frac{d}{ds}w(s)$$

from which Eq. (3.25) follows.

For Eq. (3.26) just compute:

$$\begin{aligned}\frac{d}{ds}\langle W(s),V(s)\rangle &= \frac{d}{ds}\langle w(s),v(s)\rangle\\ &=\left\langle \frac{d}{ds}w(s),v(s)\right\rangle+\left\langle w(s),\frac{d}{ds}v(s)\right\rangle\\ &=\left\langle \frac{d}{ds}w(s),P(\sigma(s))v(s)\right\rangle+\left\langle P(\sigma(s))w(s),\frac{d}{ds}v(s)\right\rangle\\ &=\left\langle P(\sigma(s))\frac{d}{ds}w(s),v(s)\right\rangle+\left\langle w(s),P(\sigma(s))\frac{d}{ds}v(s)\right\rangle\\ &=\left\langle \frac{\nabla W(s)}{ds},V(s)\right\rangle+\left\langle W(s),\frac{\nabla V(s)}{ds}\right\rangle,\end{aligned}$$

where the third equality relies on $v(s)$ and $w(s)$ being in $\tau_{\sigma(s)}M$ and the fourth equality relies on $P(\sigma(s))$ being an orthogonal projection.

From the definitions of σ', $\dot\sigma$, ∇/dt, ∇/ds and the fact that mixed partial derivatives commute,

$$\begin{aligned}\frac{\nabla\sigma'(s,t)}{dt}&=\frac{\nabla}{dt}(\sigma(t,s),\sigma'(s,t))=(\sigma(t,s),P(\sigma(s,t))\frac{d}{dt}\frac{d}{ds}\sigma(t,s))\\ &=(\sigma(t,s),P(\sigma(s,t))\frac{d}{ds}\frac{d}{dt}\sigma(t,s))=\nabla\dot\sigma(s,t)/ds,\end{aligned}$$

which proves Eq. (3.27).

For Eq. (3.29) we observe,

$$\begin{aligned}\frac{\nabla}{dt}\frac{\nabla}{ds}W(s,t)&=\frac{\nabla}{dt}(\sigma(s,t),\frac{d}{ds}w(s,t)+dQ(\sigma'(s,t))w(s,t))\\ &=(\sigma(s,t),\eta_+(s,t))\end{aligned}$$

where (with the arguments (s,t) suppressed from the notation)

$$\begin{aligned}\eta_+&=\frac{d}{dt}\left[\frac{d}{ds}w+dQ(\sigma')w\right]+dQ(\dot\sigma)\left[\frac{d}{ds}w+dQ(\sigma')w\right]\\ &=\frac{d}{dt}\frac{d}{ds}w+\left(\frac{d}{dt}\left[dQ(\sigma')\right]\right)w+dQ(\sigma')\frac{d}{dt}w+dQ(\dot\sigma)\frac{d}{ds}w+dQ(\dot\sigma)dQ(\sigma')w.\end{aligned}$$

Therefore

$$\left[\frac{\nabla}{dt},\frac{\nabla}{ds}\right]W=(\sigma,\eta_+-\eta_-),$$

where η_- is defined the same as η_+ with all s and t derivatives interchanged. Hence, it follows (using again $\frac{d}{dt}\frac{d}{ds}w=\frac{d}{ds}\frac{d}{dt}w$) that

$$\left[\frac{\nabla}{dt},\frac{\nabla}{ds}\right]W=(\sigma,[\frac{d}{dt}(dQ(\sigma'))]w-[\frac{d}{ds}(dQ(\dot\sigma))]w+[dQ(\dot\sigma),dQ(\sigma')]w).$$

The proof of Eq. (3.28) is finished because

$$\frac{d}{dt}(dQ(\sigma'))-\frac{d}{ds}(dQ(\dot\sigma))=\frac{d}{dt}\frac{d}{ds}(Q\circ\sigma)-\frac{d}{ds}\frac{d}{dt}(Q\circ\sigma)=0.$$

■

Example 3.33. Let $M=\{m\in\mathbb{R}^N:|m|=\rho\}$ be the sphere of radius ρ. In this case $Q(m)=\frac{1}{\rho^2}mm^{\mathrm{tr}}$ for all $m\in M$. Therefore

$$dQ(v_m)=\frac{1}{\rho^2}\{vm^{\mathrm{tr}}+mv^{\mathrm{tr}}\}\ \forall\ v_m\in T_mM$$

and hence

$$\begin{aligned}dQ(u_m)dQ(v_m)&=\frac{1}{\rho^4}\{um^{\mathrm{tr}}+mu^{\mathrm{tr}}\}\{vm^{\mathrm{tr}}+mv^{\mathrm{tr}}\}\\ &=\frac{1}{\rho^4}\{\rho^2uv^{\mathrm{tr}}+\langle u,v\rangle Q(m)\}.\end{aligned}$$

So the curvature tensor is given by

$$R(u_m,v_m)w_m=(m,\frac{1}{\rho^2}\{uv^{\mathrm{tr}}-vu^{\mathrm{tr}}\}w)=(m,\frac{1}{\rho^2}\{\langle v,w\rangle u-\langle u,w\rangle v\}).$$

Exercise 3.34. Show the curvature tensor of the cylinder

$$M=\{(x,y,z)\in\mathbb{R}^3:x^2+y^2=1\}$$

is zero.

Definition 3.35 (Covariant Derivative on $\Gamma(TM)$)**.** Suppose that Y is a vector field on M and $v_m\in T_mM$. Define $\nabla_{v_m}Y\in T_mM$ by

$$\nabla_{v_m}Y:=\frac{\nabla Y(\sigma(s))}{ds}|_{s=0},$$

where σ is any smooth path in M such that $\sigma'(0)=v_m$.

If $Y(m)=(m,y(m))$, then

$$\nabla_{v_m}Y=(m,P(m)dy(v_m))=(m,dy(v_m)+dQ(v_m)y(m)),$$

from which it follows $\nabla_{v_m}Y$ is well defined, i.e. $\nabla_{v_m}Y$ is independent of the choice of σ such that $\sigma'(0)=v_m$. The following proposition relates curvature and torsion to the covariant derivative ∇ on vector fields.

Proposition 3.36. *Let $m \in M$, $v \in T_mM$, $X, Y, Z \in \Gamma(TM)$, and $f \in C^\infty(M)$, then the following relations hold.*

1. **Product Rule:** $\nabla_v(f \cdot X) = df(v) \cdot X(m) + f(m) \cdot \nabla_v X$.
2. **Zero Torsion:** $\nabla_X Y - \nabla_Y X - [X, Y] = 0$.
3. **Zero Torsion:** *For all $v_m, w_m \in T_mM$, $dQ(v_m)w_m = dQ(w_m)v_m$.*
4. **Curvature Tensor:** $R(X, Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z$, *where*

$$[\nabla_X, \nabla_Y]Z := \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z).$$

Moreover if $u, v, w, z \in T_mM$, then R has the following symmetries

a: $R(u_m, v_m) = -R(v_m, u_m)$

b: $[R(u_m, v_m)]^{\mathrm{tr}} = -R(u_m, v_m)$ *and*

c: *if $z_m \in \tau_mM$, then*

$$\langle R(u_m, v_m)w_m, z_m\rangle = \langle R(w_m, z_m)u_m, v_m\rangle. \tag{3.30}$$

5. **Ricci Curvature Tensor:** *For each $m \in M$, let $\mathrm{Ric}_m : T_mM \to T_mM$ be defined by*

$$\mathrm{Ric}_m v_m := \sum_{a \in S} R(v_m, a)a, \tag{3.31}$$

where $S \subset T_mM$ is an orthonormal basis. Then $\mathrm{Ric}_m^{\mathrm{tr}} = \mathrm{Ric}_m$ and Ric_m may be computed as

$$\langle \mathrm{Ric}_m u, v\rangle = \mathrm{tr}(dQ(dQ(u)v) - dQ(v)dQ(u)) \text{ for all } u, v \in T_mM. \tag{3.32}$$

Proof. The product rule is easily checked and may be left to the reader. For the second and third items, write $X(m) = (m, x(m))$, $Y(m) = (m, y(m))$, and $Z(m) = (m, z(m))$ where $x, y, z : M \to \mathbb{R}^N$ are smooth functions such that $x(m)$, $y(m)$, and $z(m)$ are in τ_mM for all $m \in M$. Then using Eq. (2.15), we have

$$\begin{aligned}(\nabla_X Y - \nabla_Y X)(m) &= (m, P(m)(dy(X(m)) - dx(Y(m))))\\ &= (m, (dy(X(m)) - dx(Y(m)))) = [X, Y](m),\end{aligned} \tag{3.33}$$

which proves the second item. Since $(\nabla_X Y)(m)$ may also be written as

$$(\nabla_X Y)(m) = (m, dy(X(m)) + dQ(X(m))y(m)),$$

Eq. (3.33) may be expressed as $dQ(X(m))y(m) = dQ(Y(m))x(m)$ which implies the third item.

Similarly for fourth item:

$$\begin{aligned}\nabla_X \nabla_Y Z &= \nabla_X(\cdot, Yz + (YQ)z)\\ &= (\cdot, XYz + (XYQ)z + (YQ)Xz + (XQ)(Yz + (YQ)z)),\end{aligned}$$

where $YQ := dQ(Y)$ and $Yz := dz(Y)$. Interchanging X and Y in this last expression and then subtracting gives:

$$\begin{aligned}[\nabla_X, \nabla_Y]Z &= (\cdot, [X, Y]z + ([X, Y]Q)z + [XQ, YQ]z)\\ &= \nabla_{[X,Y]}Z + R(X, Y)Z.\end{aligned}$$

The anti-symmetry properties in items 4a) and 4b) follow easily from Eq. (3.28). For example for 4b), $dQ(u_m)$ and $dQ(v_m)$ are symmetric operators and hence

$$\begin{aligned}[R(u_m, v_m)]^{\mathrm{tr}} &= [dQ(u_m), dQ(v_m)]^{\mathrm{tr}} = [dQ(v_m)^{\mathrm{tr}}, dQ(u_m)^{\mathrm{tr}}]\\ &= [dQ(v_m), dQ(u_m)] = -[dQ(u_m), dQ(v_m)] = -R(u_m, v_m).\end{aligned}$$

To prove Eq. (3.30) we make use of the zero - torsion condition $dQ(v_m)w_m = dQ(w_m)v_m$ and the fact that $dQ(u_m)$ is symmetric to learn

$$\begin{aligned}\langle R(u_m, v_m)w, z\rangle &= \langle [dQ(u_m), dQ(v_m)]w, z\rangle\\ &= \langle [dQ(u_m)dQ(v_m) - dQ(v_m)dQ(u_m)]w, z\rangle\\ &= \langle dQ(v_m)w, dQ(u_m)z\rangle - \langle dQ(u_m)w, dQ(v_m)z\rangle\\ &= \langle dQ(w)v, dQ(z)u\rangle - \langle dQ(w)u, dQ(z)v\rangle \qquad (3.34)\\ &= \langle [dQ(z), dQ(w)]\, v, u\rangle = \langle R(z, w)\, v, u\rangle = \langle R(w, z)\, u, v\rangle\end{aligned}$$

where we have used the anti-symmetry properties in 4a. and 4b. By Eq. (3.34) with $v = w = a$,

$$\begin{aligned}\langle \mathrm{Ric}\, u, z\rangle &= \sum_{a \in S} \langle R(u, a)a, z\rangle\\ &= \sum_{a \in S} [\langle dQ(a)a, dQ(u)z\rangle - \langle dQ(u)a, dQ(a)z\rangle]\\ &= \sum_{a \in S} [\langle a, dQ(a)dQ(u)z\rangle - \langle dQ(u)a, dQ(z)a\rangle]\\ &= \sum_{a \in S} [\langle a, dQ(dQ(u)z)a\rangle - \langle dQ(z)dQ(u)a, a\rangle]\\ &= \mathrm{tr}(dQ(dQ(u)z) - dQ(z)dQ(u))\end{aligned}$$

which proves Eq. (3.32). The assertion that $\mathrm{Ric}_m : T_mM \to T_mM$ is a symmetric operator follows easily from this formula and item 3. ■

Notation 3.37. To each $v \in \mathbb{R}^N$, let ∂_v denote the vector field on $\mathbb{R}^N$ defined by

$$\partial_v(\text{at } x) = v_x = \frac{d}{dt}|_0(x + tv).$$

So if $F \in C^\infty(\mathbb{R}^N)$, then

$$(\partial_v F)(x) := \frac{d}{dt}|_0 F(x + tv) = F'(x)\, v$$

and

$$(\partial_v \partial_w F)(x) = F''(x)(v, w),$$

see Notation 2.1.

Notice that if $w : \mathbb{R}^N \to \mathbb{R}^N$ is a function and $v \in \mathbb{R}^N$, then

$$(\partial_v \partial_w F)(x) = \partial_v [F'(\cdot) w(\cdot)](x) = F'(x) \partial_v w(x) + F''(x)(v, w(x)).$$

The following variant of item 4. of Proposition 3.36 will be useful in proving the key Bochner-Weitenböck identity in Theorem 3.49 below.

Proposition 3.38. *Suppose that $Z \in \Gamma(TM)$, $v, w \in T_m M$ and let $X, Y \in \Gamma(TM)$ such that $X(m) = v$ and $Y(m) = w$. Then*

(1) *$\nabla^2_{v \otimes w} Z$ defined by*

$$\nabla^2_{v\otimes w} Z := (\nabla_X \nabla_Y Z - \nabla_{\nabla_X Y} Z)(m) \tag{3.35}$$

is well defined, independent of the possible choices for X and Y.

(2) *If $Z(m) = (m, z(m))$ with $z : \mathbb{R}^N \to \mathbb{R}^N$ a smooth function such $z(m) \in \tau_m M$ for all $m \in M$, then*

$$\nabla^2_{v\otimes w} Z = dQ(v)\, dQ(w)\, z(m) + P(m)\, z''(m)(v, w) - P(m)\, z'(m)[dQ(v)\, w]. \tag{3.36}$$

(3) *The curvature tensor $R(v, w)$ may be computed as*

$$\nabla^2_{v\otimes w} Z - \nabla^2_{w\otimes v} Z = R(v, w)\, Z(m). \tag{3.37}$$

(4) *If V is a smooth vector field along a path $\sigma(s)$ in M, then the following product rule holds,*

$$\frac{\nabla}{ds}(\nabla_{V(s)} Z) = \left(\nabla_{\frac{\nabla}{ds} V(s)} Z\right) + \nabla^2_{\sigma'(s)\otimes V(s)} Z. \tag{3.38}$$

Proof. We will prove items 1. and 2. by showing the right sides of Eq. (3.35) and Eq. (3.36) are equal. To do this write $X(m) = (m, x(m))$, $Y(m) = (m, y(m))$, and $Z(m) = (m, z(m))$ where $x, y, z : \mathbb{R}^N \to \mathbb{R}^N$ are smooth functions such that $x(m)$, $y(m)$, and $z(m)$ are in $\tau_m M$ for all $m \in M$. Then, suppressing m from the notation,

$$\begin{aligned}
\nabla_X \nabla_Y Z - \nabla_{\nabla_X Y} Z &= P \partial_x [P \partial_y z] - P \partial_{P \partial_x y} z \\
&= P(\partial_x P) \partial_y z + P \partial_x \partial_y z - P \partial_{P \partial_x y} z \\
&= P(\partial_x P) \partial_y z + P z''(x, y) + P z'[\partial_x y - P \partial_x y] \\
&= (\partial_x P) Q \partial_y z + P z''(x, y) + P z'[Q \partial_x y].
\end{aligned}$$

Differentiating the identity, $Qy = 0$ on M shows $Q \partial_x y = -(\partial_x Q) y$ which combined with the previous equation gives

$$\begin{aligned}
\nabla_X \nabla_Y Z - \nabla_{\nabla_X Y} Z &= (\partial_x P) Q \partial_y z + P z''(x, y) - P z'[(\partial_x Q) Y] \\
&= -(\partial_x P)(\partial_y Q) z + P z''(X, Y) - P z'[(\partial_x Q) Y].
\end{aligned} \tag{3.39}$$

Evaluating this expression at m proves the right side of Eq. (3.36).

Equation (3.37) now follows from Eqs. (3.36) and (3.28), item 3. of Proposition 3.36 and the fact the $z''(v, w) = z''(w, v)$ because mixed partial derivatives commute.

We give two proofs of Eq. (3.38). For the first proof, choose local vector fields $\{E_i\}_{i=1}^d$ defined in a neighborhood of $\sigma(s)$ such that $\{E_i(\sigma(s))\}_{i=1}^d$ is a basis for $T_{\sigma(s)} M$ for each s. We may then write $V(s) = \sum_{i=1}^d V_i(s) E_i(\sigma(s))$ and therefore,

$$\frac{\nabla}{ds} V(s) = \sum_{i=1}^d \left\{ V_i'(s) E_i(\sigma(s)) + V_i(s) \nabla_{\sigma'(s)} E_i \right\} \tag{3.40}$$

and

$$\begin{aligned}
\frac{\nabla}{ds}(\nabla_{V(s)} Z) &= \frac{\nabla}{ds}\left(\sum_{i=1}^d V_i(s) (\nabla_{E_i} Z)(\sigma(s)) \right) \\
&= \sum_{i=1}^d V_i'(s) (\nabla_{E_i} Z)(\sigma(s)) + \sum_{i=1}^d V_i(s) \nabla_{\sigma'(s)} (\nabla_{E_i} Z).
\end{aligned}$$

Using Eq. (3.35),

$$\nabla_{\sigma'(s)} (\nabla_{E_i} Z) = \nabla^2_{\sigma'(s)\otimes E_i(\sigma(s))} Z + \left(\nabla_{\nabla_{\sigma'(s)} E_i} Z \right)$$

and using this in the previous equation along with Eq. (3.40) shows

$$\begin{aligned}
\frac{\nabla}{ds}(\nabla_{V(s)} Z) &= \nabla_{\sum_{i=1}^d \left\{ V_i'(s) E_i(\sigma(s)) + V_i(s) \nabla_{\sigma'(s)} E_i \right\}} Z + \sum_{i=1}^d V_i(s) \nabla^2_{\sigma'(s)\otimes E_i(\sigma(s))} Z \\
&= \left(\nabla_{\frac{\nabla}{ds} V(s)} Z \right) + \nabla^2_{\sigma'(s)\otimes V(s)} Z.
\end{aligned}$$

For the second proof, write $V(s) = (\sigma(s), v(s)) = v(s)_{\sigma(s)}$ and $p(s) := P(\sigma(s))$, then

$$\begin{aligned}
\frac{\nabla}{ds}(\nabla_V Z) - \left(\nabla_{\frac{\nabla}{ds} V} Z \right) &= p \frac{d}{ds}(p z'(v)) - p z'(p v') \\
&= p[p' z'(v) + p z''(\sigma', v) + p z'(v')] - p z'(p v') \\
&= p p' z'(v) + p z''(\sigma', v) + p z'(q v') \\
&= p' q z'(v) + p z''(\sigma', v) - p z'(q' v) \\
&= \nabla^2_{\sigma'(s)\otimes V(s)} Z
\end{aligned}$$

wherein the last equation we have made use of Eq. (3.39). ■

3.5. **Formulas for the Divergence and the Laplacian.**

Theorem 3.39. *Let Y be a vector field on M, then*

$$\operatorname{div} Y = \operatorname{tr}(\nabla Y). \tag{3.41}$$

(Note: $(v_m \to \nabla_{v_m} Y) \in \operatorname{End}(T_m M)$ for each $m \in M$, so it makes sense to take the trace.) Consequently, if f is a smooth function on M, then

$$\Delta f = \operatorname{tr}(\nabla \operatorname{grad} f). \tag{3.42}$$

Proof. Let x be a chart on M, $\partial_i := \partial/\partial x^i$, $\nabla_i := \nabla_{\partial_i}$, and $Y^i := dx^i(Y)$. Then by the product rule and the fact that ∇ is Torsion free (item 2. of the Proposition 3.36),

$$\nabla_i Y = \sum_{j=1}^d \nabla_i (Y^j \partial_j) = \sum_{j=1}^d (\partial_i Y^j \partial_j + Y^j \nabla_i \partial_j),$$

and $\nabla_i \partial_j = \nabla_j \partial_i$. Hence,

$$\begin{aligned}\operatorname{tr}(\nabla Y) &= \sum_{i=1}^d dx^i(\nabla_i Y) = \sum_{i=1}^d \partial_i Y^i + \sum_{i,j=1}^d dx^i(Y^j \nabla_i \partial_j) \\ &= \sum_{i=1}^d \partial_i Y^i + \sum_{i,j=1}^d dx^i(Y^j \nabla_j \partial_i).\end{aligned}$$

Therefore, according to Eq. (3.20), to finish the proof it suffices to show that

$$\sum_{i=1}^d dx^i(\nabla_j \partial_i) = \partial_j \log \sqrt{g}.$$

From Lemma 2.7,

$$\partial_j \log \sqrt{g} = \frac{1}{2} \partial_j \log(\det g) = \frac{1}{2} \operatorname{tr}(g^{-1} \partial_j g) = \frac{1}{2} \sum_{k,l=1}^d g^{kl} \partial_j g_{kl},$$

and using Eq. (3.26) we have

$$\partial_j g_{kl} = \partial_j \langle \partial_k, \partial_l \rangle = \langle \nabla_j \partial_k, \partial_l \rangle + \langle \partial_k, \nabla_j \partial_l \rangle.$$

Combining the last two equations along with the symmetry of g^{kl} implies

$$\partial_j \log \sqrt{g} = \sum_{k,l=1}^d g^{kl} \langle \nabla_j \partial_k, \partial_l \rangle = \sum_{k=1}^d dx^k(\nabla_j \partial_k),$$

where we have used

$$\sum_{k=1}^d g^{kl} \langle \cdot, \partial_l \rangle = dx^k.$$

This last equality is easily verified by applying both sides of this equation to ∂_i for $i = 1, 2, \ldots, n$. ■

Definition 3.40 (One forms)**.** A **one form** ω on M is a smooth function $\omega : TM \to \mathbb{R}$ such that $\omega_m := \omega|_{T_m M}$ is linear for all $m \in M$. Note: if x is a chart of M with $m \in \mathcal{D}(x)$, then

$$\omega_m = \sum_{i=1}^d \omega_i(m) dx^i|_{T_m M},$$

where $\omega_i := \omega(\partial/\partial x^i)$. The condition that ω is smooth is equivalent to the condition that each of the functions ω_i is smooth on M. Let $\Omega^1(M)$ denote the smooth one-forms on M.

Given a one form, $\omega \in \Omega^1(M)$, there is a unique vector field X on M such that $\omega_m = \langle X(m), \cdot \rangle_m$ for all $m \in M$. Using this observation, we may extend the definition of ∇ to one forms by requiring

$$\nabla_{v_m} \omega := \langle \nabla_{v_m} X, \cdot \rangle \in T_m^* M := (T_m M)^*. \tag{3.43}$$

Lemma 3.41 (Product Rule)**.** *Keep the notation of the above paragraph. Let $Y \in \Gamma(TM)$, then*

$$v_m [\omega(Y)] = (\nabla_{v_m} \omega)(Y(m)) + \omega(\nabla_{v_m} Y). \tag{3.44}$$

Moreover, if $\theta : M \to (\mathbb{R}^N)^*$ is a smooth function and

$$\omega(v_m) := \theta(m) v$$

for all $v_m \in TM$, then

$$(\nabla_{v_m} \omega)(w_m) = d\theta(v_m) w - \theta(m) dQ(v_m) w = (d(\theta P)(v_m)) w, \tag{3.45}$$

where $(\theta P)(m) := \theta(m) P(m) \in (\mathbb{R}^N)^*$.

Proof. Using the metric compatibility of ∇,

$$\begin{aligned} v_m(\omega(Y)) = v_m(\langle X, Y \rangle) &= \langle \nabla_{v_m} X, Y(m) \rangle + \langle X(m), \nabla_{v_m} Y \rangle \\ &= (\nabla_{v_m} \omega)(Y(m)) + \omega(\nabla_{v_m} Y). \end{aligned}$$

Writing $Y(m) = (m, y(m)) = y(m)_m$ and using Eq. (3.44), it follows that

$$\begin{aligned} (\nabla_{v_m} \omega)(Y(m)) &= v_m(\omega(Y)) - \omega(\nabla_{v_m} Y) \\ &= v_m(\theta(\cdot) y(\cdot)) - \theta(m)(dy(v_m) + dQ(v_m) y(m)) \\ &= (d\theta(v_m)) y(m) - \theta(m)(dQ(v_m)) y(m). \end{aligned}$$

Choosing Y such that $Y(m) = w_m$ proves the first equality in Eq. (3.45). The second equality in Eq. (3.45) is a simple consequence of the formula

$$d(\theta P) = d\theta(\cdot) P + \theta dP = d\theta(\cdot) P - \theta dQ.$$

■

Before continuing, let us record the following useful corollary of the previous proof.

Corollary 3.42. *To every one – form ω on M, there exists $f_i, g_i \in C^\infty(M)$ for $i = 1, 2, \ldots, N$ such that*

$$\omega = \sum_{i=1}^{N} f_i dg_i. \tag{3.46}$$

Proof. Let $f_i(m) := \theta(m)P(m)e_i$ and $g_i(m) = x^i(m) = \langle m, e_i\rangle_{\mathbb{R}^N}$ where $\{e_i\}_{i=1}^N$ is the standard basis for $\mathbb{R}^N$ and $P(m)$ is orthogonal projection of $\mathbb{R}^N$ onto $\tau_m M$ for each $m \in M$. ■

Definition 3.43. For $f \in C^\infty(M)$ and v_m, w_m in $T_m M$, let

$$\nabla df(v_m, w_m) := (\nabla_{v_m} df)(w_m),$$

so that

$$\nabla df : \cup_{m\in M}(T_m M \times T_m M) \to \mathbb{R}.$$

We call ∇df the **Hessian** of f.

Lemma 3.44. *Let $f \in C^\infty(M)$, $F \in C^\infty(\mathbb{R}^N)$ such that $f = F|_M$, $X, Y \in \Gamma(TM)$ and $v_m, w_m \in T_m M$. Then:*

1. $\nabla df(X, Y) = XYf - df(\nabla_X Y)$.
2. $\nabla df(v_m, w_m) = F''(m)(v, w) - F'(m)dQ(v_m)w$.
3. $\nabla df(v_m, w_m) = \nabla df(w_m, v_m)$ *– another manifestation of zero torsion.*

Proof. Using the product rule (see Eq. (3.44)):

$$XYf = X(df(Y)) = (\nabla_X df)(Y) + df(\nabla_X Y),$$

and hence

$$\nabla df(X, Y) = (\nabla_X df)(Y) = XYf - df(\nabla_X Y).$$

This proves item 1. From this last equation and Proposition 3.36 (∇ has zero torsion), it follows that

$$\nabla df(X, Y) - \nabla df(Y, X) = [X, Y]f - df(\nabla_X Y - \nabla_Y X) = 0.$$

This proves the third item upon choosing X and Y such that $X(m) = v_m$ and $Y(m) = w_m$. Item 2 follows easily from Lemma 3.41 applied with $\theta := F'$. ■

Definition 3.45. Given a point $m \in M$, a **local orthonormal frame** $\{E_i\}_{i=1}^d$ at m is a collection of local vector fields defined near m such that $\{E_i(p)\}_{i=1}^d$ is an orthonormal basis for $T_p M$ for all p near m.

Corollary 3.46. *Suppose that $F \in C^\infty(\mathbb{R}^N)$, $f := F|_M$, and $m \in M$. Let $\{e_i\}_{i=1}^d$ be an orthonormal basis for $\tau_m M$ and let $\{E_i\}_{i=1}^d$ be an orthonormal frame near $m \in M$. Then*

$$\Delta f(m) = \sum_{i=1}^{d} \nabla df(E_i(m), E_i(m)), \tag{3.47}$$

$$\Delta f(m) = \sum_{i=1}^{d} \{E_i E_i f)(m) - df(\nabla_{E_i(m)} E_i)\}, \tag{3.48}$$

and

$$\Delta f(m) = \sum_{i=1}^{d} F''(m)(e_i, e_i) - F'(m)(dQ(E_i(m))e_i) \tag{3.49}$$

where $E_i(m) := (m, e_i)$.

Proof. By Theorem 3.39, $\Delta f = \sum_{i=1}^d \langle \nabla_{E_i} \operatorname{grad} f, E_i\rangle$ and by Eq. (3.43), $\nabla_{E_i} df = \langle \nabla_{E_i} \operatorname{grad} f, \cdot\rangle$. Therefore

$$\Delta f = \sum_{i=1}^{d} (\nabla_{E_i} df)(E_i) = \sum_{i=1}^{d} \nabla df(E_i, E_i),$$

which proves Eq. (3.47). Equations (3.48) and (3.49) follows from Eq. (3.47) and Lemma 3.44. ■

Notation 3.47. Let $\{e_i\}_{i=1}^N$ be the standard basis on $\mathbb{R}^N$ and define $X_i(m) := P(m)e_i$ for all $m \in M$ and $i = 1, 2, \ldots, N$.

In the next proposition we will express the gradient, divergence and the Laplacian in terms of the vector fields, $\{X_i\}_{i=1}^N$. These formula will prove very useful when we start discussing Brownian motion on M.

Proposition 3.48. *Let $f \in C^\infty(M)$ and $Y \in \Gamma(TM)$ then*

1. $v_m = \sum_{i=1}^N \langle v_m, X_i(m)\rangle X_i(m)$ *for all* $v_m \in T_m M$.
2. $\vec{\nabla} f = \operatorname{grad} f = \sum_{i=1}^N X_i f \cdot X_i$
3. $\vec{\nabla} \cdot Y = \operatorname{div}(Y) = \sum_{i=1}^N \langle \nabla_{X_i} Y, X_i\rangle$
4. $\sum_{i=1}^N \nabla_{X_i} X_i = 0$
5. $\Delta f = \sum_{i=1}^N X_i^2 f$.

Proof. 1. The main point is to show

$$\sum_{i=1}^{N} X_i(m) \otimes X_i(m) = \sum_{i=1}^{d} u_i \otimes u_i \tag{3.50}$$

where $\{u_i\}_{i=1}^d$ is an orthonormal basis for $T_m M$. But this is easily proved since

$$\sum_{i=1}^{N} X_i(m) \otimes X_i(m) = \sum_{i=1}^{N} P(m)e_i \otimes P(m)e_i$$

and the latter expression is independent of the choice of orthonormal basis $\{e_i\}_{i=1}^N$ for $\mathbb{R}^N$. Hence if we choose $\{e_i\}_{i=1}^N$ so that $e_i = u_i$ for $i = 1, \ldots, d$, then

$$\sum_{i=1}^{N} P(m)e_i \otimes P(m)e_i = \sum_{i=1}^{d} u_i \otimes u_i$$

as desired. Since $\sum_{i=1}^{N}\langle v_m, X_i(m)\rangle X_i(m)$ is quadratic in X_i, it now follows that

$$\sum_{i=1}^{N}\langle v_m, X_i(m)\rangle X_i(m) = \sum_{i=1}^{d}\langle v_m, u_i\rangle u_i = v_m.$$

2. This is an immediate consequence of item 1:

$$\operatorname{grad} f(m) = \sum_{i=1}^{N}\langle \operatorname{grad} f(m), X_i(m)\rangle X_i(m) = \sum_{i=1}^{N} X_i f(m)\cdot X_i(m).$$

3. Again $\sum_{i=1}^{N}\langle \nabla_{X_i} Y, X_i\rangle(m)$ is quadratic in X_i and so by Eq. (3.50) and Theorem 3.39,

$$\sum_{i=1}^{N}\langle \nabla_{X_i} Y, X_i\rangle(m) = \sum_{i=1}^{d}\langle \nabla_{u_i} Y, u_i\rangle(m) = \operatorname{div}(Y).$$

4. By definition of X_i and ∇ and using Lemma 3.30,

$$\sum_{i=1}^{N}(\nabla_{X_i} X_i)(m) = \sum_{i=1}^{N} P(m)\,dP(X_i(m))\,e_i = \sum_{i=1}^{N} dP(P(m)e_i)\,Q(m)\,e_i. \tag{3.51}$$

The latter expression is independent of the choice of orthonormal basis $\{e_i\}_{i=1}^{N}$ for $\mathbb{R}^N$. So again we may choose $\{e_i\}_{i=1}^{N}$ so that $e_i = u_i$ for $i = 1,\ldots,d$, in which case $P(m)e_j = 0$ for $j > d$ and so each summand in the right member of Eq. (3.51) is zero.

5. To compute Δf, use items 2.– 4., the definition of $\vec{\nabla} f$ and the product rule to find

$$\begin{aligned}\Delta f = \vec{\nabla}\cdot(\vec{\nabla} f) &= \sum_{i=1}^{N}\langle \nabla_{X_i}\vec{\nabla} f, X_i\rangle \\ &= \sum_{i=1}^{N} X_i\langle \vec{\nabla} f, X_i\rangle - \sum_{i=1}^{N}\langle \vec{\nabla} f, \nabla_{X_i} X_i\rangle = \sum_{i=1}^{N} X_i X_i f.\end{aligned}$$

■

The following commutation formulas are at the heart of many of the results to appear in the latter sections of these note.

Theorem 3.49 (The Bochner-Weitenböck Identity)**.** *Let $f \in C^\infty(M)$ and $a,b,c \in T_mM$, then*

$$\langle \nabla^2_{a\otimes b}\vec{\nabla} f, c\rangle = \langle \nabla^2_{a\otimes c}\vec{\nabla} f, b\rangle \tag{3.52}$$

and if $S \subset T_mM$ is an orthonormal basis, then

$$\sum_{a\in S}\nabla^2_{a\otimes a}\vec{\nabla} f = (\operatorname{grad}\ \Delta f)(m) + \operatorname{Ric}\vec{\nabla} f(m). \tag{3.53}$$

This result is the first indication that the Ricci tensor is going to play an important role in later developments. The proof will be given after the next technical lemma which will be helpful in simplifying the proof of the theorem.

Lemma 3.50. *Given $m \in M$ and $v \in T_mM$ there exists $V \in \Gamma(TM)$ such that $V(m) = v$ and $\nabla_w V = 0$ for all $w \in T_mM$. Moreover if $\{e_i\}_{i=1}^{d}$ is an orthonormal basis for T_mM, there exists a local orthonormal frame $\{E_i\}_{i=1}^{d}$ near m such that $\nabla_w E_i = 0$ for all $w \in T_mM$.*

Proof. In the proof to follow it is assume that V, Q and P have all been extended off M to smooth function on the ambient space. If V is to exist, we must have

$$0 = \nabla_w V = V'(m)\,w + \partial_w Q(m)\,v,$$

i.e.

$$V'(m)\,w = -\partial_w Q(m)\,v \text{ for all } w \in T_mM.$$

This helps to motivate defining V by

$$V(x) := P(x)\,(v - (\partial_{x-m}Q)(m)\,v) \in T_xM \text{ for all } x \in M.$$

By construction, $V(m) = v$ and making use of the identities in Lemma 3.30,

$$\begin{aligned}\nabla_w V &= \partial_w\left[P(x)\,(v - (\partial_{x-m}Q)(m)\,v)\right]|_{x=m} + (\partial_w Q)(m)\,v \\ &= (\partial_w P)(m)\,v - P(m)\,(\partial_w Q)(m)\,v + (\partial_w Q)(m)\,v \\ &= (\partial_w P)(m)\,v + Q(m)\,(\partial_w Q)(m)\,v = (\partial_w P)(m)\,v + (\partial_w Q)(m)\,v = 0\end{aligned}$$

as desired.

For the second assertion, choose a local frame $\{V_i\}_{i=1}^{d}$ such that $V_i(m) = e_i$ and $\nabla_w V_i = 0$ for all i and $w \in T_mM$. The desired frame $\{E_i\}_{i=1}^{d}$ is now constructed by performing Gram-Schmidt orthogonalization on $\{V_i\}_{i=1}^{d}$. The resulting orthonormal frame, $\{E_i\}_{i=1}^{d}$, still satisfies $\nabla_w E_i = 0$ for all $w \in T_mM$. For example, $E_1 = \langle V_1, V_1\rangle^{-1/2} V_1$ and since

$$w\langle V_1, V_1\rangle = 2\langle \nabla_w V_1, V_1(m)\rangle = 0$$

it follows that

$$\nabla_w E_1 = w\left(\langle V_1, V_1\rangle^{-1/2}\right)\cdot V_1(m) + \langle V_1, V_1\rangle^{-1/2}(m)\,\nabla_w V_1(m) = 0.$$

The similar verifications that $\nabla_w E_j = 0$ for $j = 2,\ldots,d$ will be left to the reader. ■

Proof. (*Proof of Theorem 3.49.*) Let $a,b,c \in T_mM$ and suppose $A,B,C \in \Gamma(TM)$ have been chosen as in Lemma 3.50, so that $A(m) = a$, $B(m) = b$ and $C(m) = c$ with $\nabla_w A = \nabla_w B = \nabla_w C = 0$ for all $w \in T_mM$. Then

$$\begin{aligned}ABCf &= AB\langle \vec{\nabla} f, C\rangle = A\langle \nabla_B \vec{\nabla} f, C\rangle + A\langle \vec{\nabla} f, \nabla_B C\rangle \\ &= \langle \nabla_A \nabla_B \vec{\nabla} f, C\rangle + \langle \nabla_B \vec{\nabla} f, \nabla_A C\rangle + A\langle \vec{\nabla} f, \nabla_B C\rangle\end{aligned}$$

which evaluated at m gives

$$\begin{aligned}(ABCf)(m) &= \left(\langle \nabla_A \nabla_B \vec{\nabla} f, C\rangle + A\langle \vec{\nabla} f, \nabla_B C\rangle\right)(m) \\ &= \langle \nabla^2_{a\otimes b} \vec{\nabla} f, c\rangle + \left(A\langle \vec{\nabla} f, \nabla_B C\rangle\right)(m)\end{aligned}$$

wherein the last equality we have used $(\nabla_A B)(m) = 0$. Interchanging B and C in this equation and subtracting then implies

$$\begin{aligned}(A[B,C]f)(m) &= \langle \nabla^2_{a\otimes b} \vec{\nabla} f, c\rangle - \langle \nabla^2_{a\otimes c} \vec{\nabla} f, b\rangle + \left(A\langle \vec{\nabla} f, \nabla_B C - \nabla_C B\rangle\right)(m) \\ &= \langle \nabla^2_{a\otimes b} \vec{\nabla} f, c\rangle - \langle \nabla^2_{a\otimes c} \vec{\nabla} f, b\rangle + \left(A\langle \vec{\nabla} f, [B,C]\rangle\right)(m) \\ &= \langle \nabla^2_{a\otimes b} \vec{\nabla} f, c\rangle - \langle \nabla^2_{a\otimes c} \vec{\nabla} f, b\rangle + (A[B,C]f)(m)\end{aligned}$$

and this equation implies Eq. (3.52).

Now suppose that $\{E_i\}_{i=1}^d \subset T_m M$ is an orthonormal frame as in Lemma 3.50 and $e_i = E_i(m)$. Then, using Proposition 3.38,

(3.54)

$$\sum_{i=1}^d \langle \nabla^2_{e_i\otimes e_i} \vec{\nabla} f, c\rangle = \sum_{i=1}^d \langle \nabla^2_{e_i\otimes c} \vec{\nabla} f, e_i\rangle = \sum_{i=1}^d \langle \nabla^2_{c\otimes e_i} \vec{\nabla} f + R(e_i, c)\vec{\nabla} f(m), e_i\rangle.$$

Since

$$\begin{aligned}\sum_{i=1}^d \langle \nabla^2_{c\otimes e_i} \vec{\nabla} f, e_i\rangle &= \sum_{i=1}^d \left(\langle \nabla_C \nabla_{E_i} \vec{\nabla} f, E_i\rangle\right)(m) = \sum_{i=1}^d \left(C\langle \nabla_{E_i} \vec{\nabla} f, E_i\rangle\right)(m) \\ &= (C\Delta f)(m) = \langle \left(\vec{\nabla}\Delta f\right)(m), c\rangle\end{aligned}$$

and (using $R(e_i, c)^{\mathrm{tr}} = R(c, e_i)$)

$$\begin{aligned}\sum_{i=1}^d \langle R(e_i, c)\vec{\nabla} f(m), e_i\rangle &= \sum_{i=1}^d \langle \vec{\nabla} f(m), R(c, e_i) e_i\rangle \\ &= \langle \vec{\nabla} f(m), \mathrm{Ric}\, c\rangle = \langle \mathrm{Ric}\, \vec{\nabla} f(m), c\rangle,\end{aligned}$$

Eq. (3.54) is implies

$$\sum_{i=1}^d \langle \nabla^2_{e_i\otimes e_i} \vec{\nabla} f, c\rangle = \langle \left(\vec{\nabla}\Delta f\right)(m) + \mathrm{Ric}\, \vec{\nabla} f(m), c\rangle$$

which proves Eq. (3.53) since $c \in T_m M$ was arbitrary. ∎

3.6. Parallel Translation.

Definition 3.51. Let V be a smooth path in TM. V is said to **parallel** or **covariantly constant** if $\nabla V(s)/ds \equiv 0$.

Theorem 3.52. *Let σ be a smooth path in M and $(v_0)_{\sigma(0)} \in T_{\sigma(0)}M$. Then there exists a unique smooth vector field V along σ such that V is parallel and $V(0) = (v_0)_{\sigma(0)}$. Moreover if $V(s)$ and $W(s)$ are parallel along σ, then $\langle V(s), W(s)\rangle = \langle V(0), W(0)\rangle$ for all s.*

Proof. If V and W are parallel, then

$$\frac{d}{ds}\langle V(s), W(s)\rangle = \left\langle \frac{\nabla}{ds}V(s), W(s)\right\rangle + \left\langle V(s), \frac{\nabla}{ds}W(s)\right\rangle = 0$$

which proves the last assertion of the theorem. If a parallel vector field $V(s) = (\sigma(s), v(s))$ along $\sigma(s)$ is to exist, then

$$dv(s)/ds + dQ(\sigma'(s))v(s) = 0 \quad \text{and} \quad v(0) = v_0. \tag{3.55}$$

By existence and uniqueness of solutions to ordinary differential equations, there is exactly one solution to Eq. (3.55). Hence, if V exists it is unique.

Now let v be the unique solution to Eq. (3.55) and set $V(s) := (\sigma(s), v(s))$. To finish the proof it suffices to show that $v(s) \in \tau_{\sigma(s)}M$. Equivalently, we must show that $w(s) := q(s)v(s)$ is identically zero, where $q(s) := Q(\sigma(s))$. Letting $v'(s) = dv(s)/ds$ and $p(s) = P(\sigma(s))$, then Eq. (3.55) states $v' = -q'v$ and from Lemma 3.30 we have $pq' = q'q$. Thus the function w satisfies

$$w' = q'v + qv' = q'v - qq'v = pq'v = q'qv = q'w$$

with $w(0) = 0$. But this linear ordinary differential equation has $w \equiv 0$ as its unique solution. ∎

Definition 3.53 (Parallel Translation)**.** Given a smooth path σ, let $//_s(\sigma) : T_{\sigma(0)}M \to T_{\sigma(s)}M$ be defined by $//_s(\sigma)(v_0)_{\sigma(0)} = V(s)$, where V is the unique parallel vector field along σ such that $V(0) = (v_0)_{\sigma(0)}$. We call $//_s(\sigma)$ **parallel translation** along σ up to time s.

Remark 3.54. Notice that $//_s(\sigma)v_{\sigma(0)} = (u(s)v)_{\sigma(0)}$, where $s \to u(s) \in \mathrm{Hom}(\tau_{\sigma(0)}M, \mathbb{R}^N)$ is the unique solution to the differential equation

$$u'(s) + dQ(\sigma'(s))u(s) = 0 \quad \text{with} \quad u(0) = P(\sigma(0)). \tag{3.56}$$

Because of Theorem 3.52, $u(s) : \tau_{\sigma(0)}M \to \mathbb{R}^N$ is an isometry for all s and the range of $u(s)$ is $\tau_{\sigma(s)}M$. Moreover, if we let $\bar{u}(s)$ denote the solution to

$$\bar{u}'(s) - \bar{u}(s)dQ(\sigma'(s)) = 0 \text{ with } \bar{u}(0) = P(\sigma(0)), \tag{3.57}$$

then

$$\begin{aligned}\frac{d}{ds}[\bar{u}(s)u(s)] &= \bar{u}'(s)u(s) + \bar{u}(s)u'(s) \\ &= \bar{u}(s)dQ(\sigma'(s))u(s) - \bar{u}(s)dQ(\sigma'(s))u(s) = 0.\end{aligned}$$

Hence $\bar{u}(s)u(s) = P(\sigma(0))$ for all s and therefore $\bar{u}(s)$ is the inverse to $u(s)$ thought of as an linear operator from $\tau_{\sigma(0)}M$ to $\tau_{\sigma(s)}M$. See also Lemma 3.57 below.

The following techniques for computing covariant derivatives will be useful in the sequel.

Lemma 3.55. *Suppose $Y \in \Gamma(TM)$, $\sigma(s)$ is a path in M, $W(s) = (\sigma(s), w(s))$ is a vector field along σ and let $//_s = //_s(\sigma)$ be parallel translation along σ. Then*

(1) $\frac{\nabla}{ds}W(s) = //_s \frac{d}{ds}\left[//_s^{-1}W(s)\right].$
(2) *For any $v \in T_{\sigma(0)}M$,*

$$\frac{\nabla}{ds}\nabla_{//_s v}Y = \nabla^2_{\sigma'(s)\otimes //_s v}Y. \tag{3.58}$$

where $\nabla^2_{\sigma'(s)\otimes//_s v}Y$ was defined in Proposition 3.38.

Proof. Let $\bar{u}$ be as in Eq. (3.57). From Eq. (3.25),

$$\frac{\nabla W(s)}{ds} = \left(\frac{d}{ds}w(s) + dQ(\sigma'(s))w(s)\right)_{\sigma(s)}$$

while, using Remark 3.54,

$$\begin{aligned}\frac{d}{ds}\left[//_s^{-1}W(s)\right] &= \left(\frac{d}{ds}\left[\bar{u}(s)w(s)\right]\right)_{\sigma(s)}\\ &= \left(\bar{u}'(s)W(s) + \bar{u}(s)w'(s)\right)_{\sigma(s)}\\ &= \left(\bar{u}(s)dQ(\sigma'(s))w(s) + \bar{u}(s)w'(s)\right)_{\sigma(s)}\\ &= //_s^{-1}\frac{\nabla W(s)}{ds}.\end{aligned}$$

This proves the first item. We will give two proves of the second item, the first proof being extrinsic while the second will be intrinsic. In each of these proofs there will be an implied sum on repeated indices.

First proof. Let $\{X_i\}_{i=1}^N \subset \Gamma(TM)$ be as in Notation 3.47, then by Proposition 3.48,

$$//_s v = \langle //_s v, X_i(\sigma(s))\rangle X_i(\sigma(s)) = \langle v, //_s^{-1}X_i(\sigma(s))\rangle X_i(\sigma(s)) \tag{3.59}$$

and therefore,

$$\begin{aligned}\frac{\nabla}{ds}\nabla_{//_s v}Y &= \frac{\nabla}{ds}\left[\langle //_s v, X_i(\sigma(s))\rangle \cdot (\nabla_{X_i}Y)(\sigma(s))\right]\\ &= \langle //_s v, X_i(\sigma(s))\rangle \cdot \nabla_{\sigma'(s)}(\nabla_{X_i}Y) + \langle //_s v, \nabla_{\sigma'(s)}X_i\rangle \cdot (\nabla_{X_i}Y)(\sigma(s)).\end{aligned} \tag{3.60}$$

Now

$$\nabla_{\sigma'(s)}(\nabla_{X_i}Y) = \nabla^2_{\sigma'(s)\otimes X_i}Y + \nabla_{\sigma'(s)X_i}Y$$

and so again using Proposition 3.48,

$$\langle //_s v, X_i(\sigma(s))\rangle \cdot \nabla_{\sigma'(s)}(\nabla_{X_i}Y) = \nabla^2_{\sigma'(s)\otimes //_s v}Y + \langle //_s v, X_i(\sigma(s))\rangle \cdot \nabla_{\sigma'(s)X_i}Y. \tag{3.61}$$

Taking ∇/ds of Eq. (3.59) shows

$$0 = \langle //_s v, \nabla_{\sigma'(s)}X_i\rangle X_i(\sigma(s)) + \langle //_s v, X_i(\sigma(s))\rangle \nabla_{\sigma'(s)}X_i.$$

and so

$$\langle //_s v, X_i(\sigma(s))\rangle \cdot \nabla_{\sigma'(s)X_i}Y = -\langle //_s v, \nabla_{\sigma'(s)}X_i\rangle \cdot (\nabla_{X_i}Y)(\sigma)(s). \tag{3.62}$$

Assembling Eqs. (3.59), (3.61) and (3.62) proves Eq. (3.58).

Second proof. Let $\{E_i\}_{i=1}^d$ be an orthonormal frame near $\sigma(s)$, then

$$\begin{aligned}\frac{\nabla}{ds}\nabla_{//_s v}Y &= \frac{\nabla}{ds}\left[\langle //_s v, E_i(\sigma(s))\rangle \cdot (\nabla_{E_i}Y)(\sigma(s))\right]\\ &= \langle //_s v, \nabla_{\sigma'(s)}E_i\rangle \cdot (\nabla_{E_i}Y)(\sigma(s)) + \langle //_s v, E_i(\sigma(s))\rangle \cdot \nabla_{\sigma'(s)}\nabla_{E_i}Y.\end{aligned} \tag{3.63}$$

Working as in the first proof,

$$\begin{aligned}\langle //_s v, E_i(\sigma(s))\rangle \cdot \nabla_{\sigma'(s)}\nabla_{E_i}Y &= \langle //_s v, E_i(\sigma(s))\rangle \cdot \left(\nabla^2_{\sigma'(s)\otimes E_i}Y + \nabla_{\nabla_{\sigma'(s)}E_i}Y\right)\\ &= \nabla^2_{\sigma'(s)\otimes //_s v}Y + \nabla_{\langle //_s v, E_i(\sigma(s))\rangle \nabla_{\sigma'(s)}E_i}Y\end{aligned}$$

and using

$$0 = \frac{\nabla}{ds}//_s v = \langle //_s v, \nabla_{\sigma'(s)}E_i\rangle \cdot E_i(\sigma(s)) + \langle //_s v, E_i(\sigma(s))\rangle \cdot \nabla_{\sigma'(s)}E_i$$

we learn

$$\langle //_s v, E_i(\sigma(s))\rangle \cdot \nabla_{\sigma'(s)}\nabla_{E_i}Y = \nabla^2_{\sigma'(s)\otimes //_s v}Y - \langle //_s v, \nabla_{\sigma'(s)}E_i\rangle \cdot (\nabla_{E_i}Y)(\sigma(s)).$$

This equation combined with Eq. (3.63) again proves Eq. (3.58). ■

The remainder of this section discusses a covariant derivative on $M \times \mathbb{R}^N$ which "extends" ∇ defined above. This will be needed in Section 5, where it will be convenient to have a covariant derivative on the normal bundle:

$$N(M) := \cup_{m\in M}(\{m\} \times \tau_m M^{\perp}) \subset M \times \mathbb{R}^N.$$

Analogous to the definition of ∇ on TM, it is reasonable to extend ∇ to the normal bundle $N(M)$ by setting

$$\frac{\nabla V(s)}{ds} = (\sigma(s), Q(\sigma(s))v'(s)) = (\sigma(s), v'(s) + dP(\sigma'(s))v(s)),$$

for all smooth paths $s \to V(s) = (\sigma(s), v(s))$ in $N(M)$. Then this covariant derivative on the normal bundle satisfies analogous properties to ∇ on the tangent bundle TM. The covariant derivatives on TM and $N(M)$ can be put together to make a covariant derivative on $M \times \mathbb{R}^N$. Explicitly, if $V(s) = (\sigma(s), v(s))$ is a smooth path in $M \times \mathbb{R}^N$, let $p(s) := P(\sigma(s))$, $q(s) := Q(\sigma(s))$ and then define

$$\frac{\nabla V(s)}{ds} := (\sigma(s), p(s)\frac{d}{ds}\{p(s)v(s)\} + q(s)\frac{d}{ds}\{q(s)v(s)\}).$$

Since

$$\frac{\nabla V(s)}{ds} = (\sigma(s), \frac{d}{ds}\{p(s)v(s)\} + q'(s)p(s)v(s) \\ + \frac{d}{ds}\{q(s)v(s)\} + p'(s)q(s)v(s)) \\ = (\sigma(s), v'(s) + q'(s)p(s)v(s) + p'(s)q(s)v(s)) \\ = (\sigma(s), v'(s) + dQ(\sigma'(s))P(\sigma(s))v(s) + dP(\sigma'(s))Q(\sigma(s))v(s))$$

we may write $\nabla V(s)/ds$ as

$$\frac{\nabla V(s)}{ds} = (\sigma(s), v'(s) + \Gamma(\sigma'(s))v(s)) \tag{3.64}$$

where

$$\Gamma(w_m)v := dQ(w_m)P(m)v + dP(w_m)Q(m)v \tag{3.65}$$

for all $w_m \in TM$ and $v \in \mathbb{R}^N$.

It should be clear from the above computation that the covariant derivative defined in (3.64) agrees with those already defined on TM and $N(M)$. Many of the properties of the covariant derivative on TM follow quite naturally from this fact and Eq. (3.64).

Lemma 3.56. *For each $w_m \in TM$, $\Gamma(w_m)$ is a skew symmetric $N \times N$ – matrix. Hence, if $u(s)$ is the solution to the differential equation*

$$u'(s) + \Gamma(\sigma'(s))u(s) = 0 \quad \text{with} \quad u(0) = I, \tag{3.66}$$

then u is an orthogonal matrix for all s.

Proof. Since $\Gamma = dQP + dPQ$ and P and Q are orthogonal projections and hence symmetric, the adjoint Γ^{tr} of Γ is given by

$$\Gamma^{\mathrm{tr}} = PdQ + QdP = -dPQ - dQP = -\Gamma.$$

where Lemma 3.30 was used in the second equality. Hence Γ is a skew-symmetric valued one form. Now let u denote the solution to (3.66) and $A(s) := \Gamma(\sigma'(s))$. Then

$$\frac{d}{ds}u^{\mathrm{tr}}u = (-Au)^{\mathrm{tr}}u + u^{\mathrm{tr}}(-Au) = u^{\mathrm{tr}}(A - A)u = 0,$$

which shows that $u^{\mathrm{tr}}(s)u(s) = u^{\mathrm{tr}}(0)u(0) = I$. ■

Lemma 3.57. *Let u be the solution to (3.66). Then*

$$u(s)(\tau_{\sigma(0)}M) = \tau_{\sigma(s)}M \tag{3.67}$$

and

$$u(s)(\tau_{\sigma(0)}M)^{\perp} = \tau_{\sigma(s)}M^{\perp}. \tag{3.68}$$

In particular, if $v \in \tau_{\sigma(0)}M$ ($v \in \tau_{\sigma(0)}M^{\perp}$) then $V(s) := (\sigma(s), u(s)v)$ is the parallel vector field along σ in TM ($N(M)$) such that $V(0) = v_{\sigma(0)}$.

Proof. By the product rule,

$$\frac{d}{ds}\{u^{\mathrm{tr}}P(\sigma)\,u\} = u^{\mathrm{tr}}\{\Gamma(\sigma')\,P(\sigma) + dP(\sigma') - P(\sigma)\,\Gamma(\sigma')\}u. \tag{3.69}$$

Moreover, making use of Lemma 3.30,

$$\Gamma(\sigma')\,P(\sigma) - P(\sigma)\,\Gamma(\sigma') + dP(\sigma') \\ = dP(\sigma') + [dQ(\sigma')P(\sigma) + dP(\sigma')Q(\sigma)]\,P(\sigma) \\ - P(\sigma)\,[dQ(\sigma')P(\sigma) + dP(\sigma')Q(\sigma)] \\ = dP(\sigma') + dQ(\sigma')P(\sigma) - dP(\sigma')Q(\sigma) \\ = dP(\sigma') + dQ(\sigma') = 0,$$

which combined with Eq. (3.69) shows $\frac{d}{ds}\{u^{\mathrm{tr}}P(\sigma)\,u\} = 0$. Therefore,

$$u^{\mathrm{tr}}(s)P(\sigma(s))\,u(s) = P(\sigma(0))$$

for all s. Combining this with Lemma 3.56, shows

$$P(\sigma(s))\,u(s) = u(s)P(\sigma(0)).$$

This last equation is equivalent to Eq. (3.67). Eq. (3.68) has completely analogous proof or can be seen easily from the fact that $P + Q = I$. ■

3.7. More References.

I recommend [86] and [42] for more details on Riemannian geometry. The references, [1, 19, 41, 42, 86, 95, 111, 112, 113, 114, 115, 149] and the complete five volume set of Spivak's books on differential geometry starting with [164] are also very useful.

4. Flows and Cartan's Development Map

The results of this section will serve as a warm-up for their stochastic counter parts. These types of theorems will be crucial for the path space analysis results to be developed in Sections 7 and 8 below.

4.1. Time - Dependent Smooth Flows.

Notation 4.1. Given a smooth **time dependent vector** field, $(t, m) \to X_t(m) \in T_mM$ on a manifold M, let $T_t^X(m)$ denote the solution to the ordinary differential equation,

$$\frac{d}{dt}T_t^X(m) = X_t \circ T_t^X(m) \text{ with } T_0^X(m) = m.$$

If X is **time independent** we will write $e^{tX}(m)$ for $T_t^X(m)$. We call T^X the **flow** of X. See Figure 10.

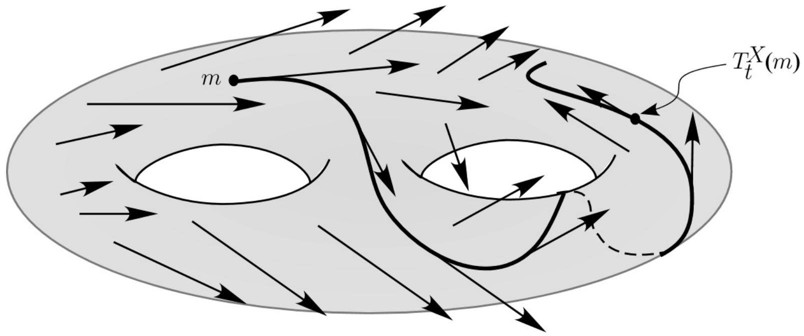

FIGURE 10. Going with the flow. Here we suppose that X is a time independent vector field which is indicated by the arrows in the picture and the curve is the corresponding flow line starting at $m \in M$.

Theorem 4.2 (Flow Theorem)**.** *Suppose that X_t is a smooth time dependent vector field on M. Then for each $m \in M$, there exists a maximal open interval $J_m \subset \mathbb{R}$ such that $0 \in J_m$ and $t \to T_t^X(m)$ exists for $t \in J_m$. Moreover the set $\mathcal{D}(X) := \cup_m (J_m \times \{m\}) \subset \mathbb{R} \times M$ is open and the map $(t,m) \in \mathcal{D}(X) \to T_t^X(m) \in M$ is a smooth map.*

Proof. Let Y_t be a smooth extension of X_t to a vector field on E where E is the Euclidean space in which M is imbedded. The stated results with X replaced by Y follows from the standard theory of ordinary differential equations on Euclidean spaces. Let T_t^Y denote the flow of Y on E. We will construct T^X by setting $T_t^X(m) := T_t^Y(m)$ for all $m \in M$ and $t \in J_m$. In order for this to work we must show that $T_t^Y(m) \in M$ whenever $m \in M$.

To verify this last assertion, let x be a chart on M such that $m \in \mathcal{D}(x)$, then $\sigma(t)$ solves $\dot{\sigma}(t) = X_t(\sigma(t))$ with $\sigma(0) = m$ iff

$$\frac{d}{dt}[x \circ \sigma(t)] = dx(\dot{\sigma}(t)) = dx(X_t(\sigma(t))) = dx\left(X_t \circ x^{-1}(x \circ \sigma(t))\right)$$

with $x \circ \sigma(0) = m$. Since this is a differential equation for $x \circ \sigma(t) \in \mathcal{R}(z)$ and $\mathcal{R}(z)$ is an open subset $\mathbb{R}^d$, the standard local existence theorem for ordinary differential equations implies $x \circ \sigma(t)$ exists for small time. This then implies $\sigma(t) \in M$ exists for small t and satisfies

$$\dot{\sigma}(t) = X_t(\sigma(t)) = Y_t(\sigma(t)) \text{ with } \sigma(0) = m.$$

By uniqueness of solutions to ordinary differential equations, we must have $T_t^Y(m) = \sigma(t)$ for small t and in particular $T_t^Y(m) \in M$ for small t. Let

$$\tau := \sup\left\{t \in J_m : T_s^Y(m) \in M \text{ for } 0 \le s \le t\right\}$$

and for sake of contradiction suppose that $[0,\tau] \subset J_m$. Then by continuity, $T_\tau^Y(m) \in M$ and by repeating the above argument using a chart x on M centered at $T_\tau^Y(m)$, we would find that $T_t^Y(m) \in M$ for t in a neighborhood of τ. This contradicts the definition of τ and hence we may conclude that τ is the right end point of J_m. A similar argument works for $t \in J_m$ with $t < 0$ and hence $T_t^Y(m) \in M$ for all $t \in J_m$. ∎

Assumption 1 (Completeness)**.** For simplicity in these notes it will always be assumed that X is **complete**, i.e. $J_m = \mathbb{R}$ for all $m \in M$ and hence $\mathcal{D}(X) = \mathbb{R} \times M$. This will be the case if, for example, M is compact or M is imbedded in $\mathbb{R}^N$ and the vector field X satisfies a Lipschitz condition. (Later we will restrict to the compact case.)

Notation 4.3. For $g, h \in \mathrm{Diff}(M)$ let $Ad_g h := g \circ h \circ g^{-1}$. We will also write Ad_g for the linear transformation on $\Gamma(TM)$ defined by

$$Ad_g Y = \frac{d}{ds}|_0 Ad_g e^{sY} = \frac{d}{ds}|_0 g \circ e^{sY} \circ g^{-1} = g_*\left(Y \circ g^{-1}\right)$$

for all $Y \in \Gamma(TM)$. (The vector space $\Gamma(TM)$ should be interpreted as the Lie algebra of the diffeomorphism group, $\mathrm{Diff}(M)$.)

In order to verify T_t^X is invertible, let $T_{t,s}^X$ denote the solution to

$$\frac{d}{dt}T_{t,s}^X = X_t \circ T_{t,s}^X \text{ with } T_{s,s}^X = id.$$

Lemma 4.4. *Suppose that X_t is a complete time dependent vector field on M, then $T_t^X \in \mathrm{Diff}(M)$ for all t and*

$$\left(T_t^X\right)^{-1} = T_{0,t}^X = T_t^{-Ad_{(T^X)^{-1}}X}, \tag{4.1}$$

where

$$\left(Ad_{(T^X)^{-1}}X\right)_t := Ad_{\left(T_t^X\right)^{-1}}X_t.$$

Proof. If $s,t,u \in \mathbb{R}$, then $S_t := T_{t,s}^X \circ T_{s,u}^X$ solves

$$\dot{S}_t = X_t \circ S_t \text{ with } S_s = T_{s,u}^X$$

which is the same equation that $t \to T_{t,u}^X$ solves and therefore $T_{t,u}^X = T_{t,s}^X \circ T_{s,u}^X$. In particular, $T_{0,t}^X$ is the inverse to T_t^X. Moreover if we let $T_t := T_t^X$ and $S_t := T_t^{-1}$ then

$$0 = \frac{d}{dt}id = \frac{d}{dt}[T_t \circ S_t] = X_t \circ T_t \circ S_t + T_{t*}\dot{S}_t.$$

So it follows that S_t solves

$$\dot{S}_t = -T_{t*}^{-1}X_t \circ T_t \circ S_t = -\left(Ad_{T_t^{-1}}X_t\right) \circ S_t$$

which proves the second equality in Eq. (4.1). ∎

4.2. **Differentials of T_t^X.** In the later sections of this article, we will make heavy use of the stochastic analogues of the following two differentiation theorems.

Theorem 4.5 (Differentiating $m \to T_t^X(m)$)**.** *Suppose ∇ is the Levi-Civita*[2] *covariant derivative on TM and $T_t = T_t^X$ as above, then*

$$\frac{\nabla}{dt} T_{t*} v = \nabla_{T_{t*}v} X_t \text{ for all } v \in TM. \tag{4.2}$$

If we further let $m \in M$, $//_t = //_t(\tau \to T_\tau(m))$ be parallel translation relative to ∇ along the flow line $\tau \to T_\tau(m)$ and $z_t := //_t^{-1} T_{t*m}$, then

$$\frac{d}{dt} z_t v = //_t^{-1} \nabla_{//_t z_t v} X_t \text{ for all } v \in T_m M. \tag{4.3}$$

(This is a linear differential equation for $z_t \in \operatorname{End}(T_m M)$.)

Proof. Let $\sigma(s)$ be smooth path in M such that $\sigma'(0) = v$, then

$$\begin{aligned} \frac{\nabla}{dt} T_{t*} v &= \frac{\nabla}{dt} \frac{d}{ds}|_0 T_t(\sigma(s)) = \frac{\nabla}{ds}|_0 \frac{d}{dt} T_t(\sigma(s)) \\ &= \frac{\nabla}{ds}|_0 X_t(T_t(\sigma(s))) = \nabla_{T_{t*}v} X_t \end{aligned}$$

wherein the second equality we have used ∇ has zero torsion. Eq. (4.3) follows directly from Eq. (4.2) using $\frac{\nabla}{dt} = //_t \frac{d}{dt} //_t^{-1}$, see Lemma 3.55. ■

Remark 4.6. As a warm up for writing the stochastic version of Eq. (4.3) in Itô form let us pause to compute $\frac{\nabla}{dt}(\nabla_{T_{t*}v} Y)$ for $Y \in \Gamma(TM)$. Using Eqs. (3.38), (3.37) and (3.35) of Proposition 3.38,

$$\begin{aligned} \frac{\nabla}{dt} \nabla_{T_{t*}v} Y &= \nabla^2_{\dot{T}_t(m) \otimes T_{t*}v} Y + \nabla_{\frac{\nabla}{dt} T_{t*}v} Y = \nabla^2_{X_t(T_t(m)) \otimes T_{t*}v} Y + \nabla_{\nabla_{T_{t*}v} X_t} Y \\ &= \nabla^2_{T_{t*}v \otimes X_t(T_t(m))} Y + R^\nabla(X_t(T_t(m)), T_{t*}v) Y(T_t(m)) + \nabla_{\nabla_{T_{t*}v} X_t} Y \\ &= R^\nabla(X_t(T_t(m)), T_{t*}v) Y(T_t(m)) + \nabla_{T_{t*}v}(\nabla_{X_t} Y). \end{aligned} \tag{4.4}$$

Theorem 4.7 (Differentiating T_t^X in X)**.** *Suppose $(t,m) \to X_t(m)$ and $(t,m) \to Y_t(m)$ are smooth time dependent vector fields on M and let*

$$\partial_Y T_t^X := \frac{d}{ds}|_0 T_t^{X+sY}. \tag{4.5}$$

Then

$$\partial_Y T_t^X = T_{t*}^X \int_0^t \left(T_{\tau*}^X\right)^{-1} Y_\tau \circ T_\tau^X d\tau = T_{t*}^X \int_0^t Ad_{T_\tau^X}^{-1} Y_\tau d\tau. \tag{4.6}$$

This formula may also be written as

$$\partial_Y T_t^X = \left(\int_0^t Ad_{T_{t,\tau}^X} Y_\tau d\tau\right) \circ T_t^X = \left(\int_0^t Ad_{T_t^X \circ (T_\tau^X)^{-1}} Y_\tau d\tau\right) \circ T_t^X. \tag{4.7}$$

[2] Actually, for those in the know, any torsion zero covariant derivative could be used here.

Proof. To simplify notation, let $T_t := T_t^X$ and define $V_t := \left(T_{t*}^X\right)^{-1} \partial_Y T_t^X$. Then $V_0 = 0$ and $\partial_Y T_t^X = T_{t*}^X V_t$ or equivalently, for all $f \in C^\infty(M)$,

$$\frac{d}{ds}|_0 f \circ T_t^{X+sY} = \left(T_{t*}^X V_t\right) f = V_t\left(f \circ T_t^X\right).$$

Given $f \in C^\infty(M)$, on one hand we have

$$\begin{aligned} \frac{d}{dt}\frac{d}{ds}|_0 f \circ T_t^{X+sY} &= \frac{d}{dt}\left[V_t(f \circ T_t^X)\right] = \dot{V}_t(f \circ T_t^X) + V_t(X_t f \circ T_t^X) \\ &= \left(T_{t*}^X \dot{V}_t\right) f + V_t(X_t f \circ T_t^X) \end{aligned}$$

while on the other hand

$$\begin{aligned} \frac{d}{ds}|_0 \frac{d}{dt} f \circ T_t^{X+sY} &= \frac{d}{ds}|_0 \left[((X_t + sY_t) f) \circ T_t^{X+sY}\right] = (Y_t f) \circ T_t^X + V_t\left(X_t f \circ T_t^X\right) \\ &= \left(Y_t \circ T_t^X\right) f + V_t\left(X_t f \circ T_t^X\right). \end{aligned}$$

Since $\left[\frac{d}{dt}, \frac{d}{ds}|_0\right] = 0$, the previous two displayed equations imply $\left(T_{t*}^X \dot{V}_t\right) f = \left(Y_t \circ T_t^X\right) f$ and because this holds for all $f \in C^\infty(M)$,

$$T_{t*}^X \dot{V}_t = Y_t \circ T_t^X. \tag{4.8}$$

Solving Eq. (4.8) for $\dot{V}_t$ and then integrating on t shows

$$V_t = \int_0^t \left(T_{\tau*}^X\right)^{-1} Y_\tau \circ T_\tau^X d\tau.$$

which along with the relation, $\partial_Y T_t^X = T_{t*}^X V_t$, implies Eq. (4.6).

We may now rewrite the formula in Eq. (4.6) as

$$\begin{aligned} \partial_Y T_t^X &= T_{t*}^X \left(\int_0^t Ad_{T_\tau^X}^{-1} Y_\tau d\tau\right) \circ \left(T_t^X\right)^{-1} \circ T_t^X = Ad_{T_t^X}\left(\int_0^t Ad_{T_\tau^X}^{-1} Y_\tau d\tau\right) \circ T_t^X \\ &= \left(\int_0^t Ad_{T_t^X} Ad_{T_\tau^X}^{-1} Y_\tau d\tau\right) \circ T_t^X = \left(\int_0^t Ad_{T_t^X \circ (T_\tau^X)^{-1}} Y_\tau d\tau\right) \circ T_t^X \\ &= \left(\int_0^t Ad_{T_{t,\tau}^X} Y_\tau d\tau\right) \circ T_t^X \end{aligned}$$

which gives Eq. (4.7). ■

Example 4.8. Suppose that G is a Lie group, $\mathfrak{g} := \operatorname{Lie}(G)$, A_t and B_t are two smooth $\mathfrak{g}$ – valued functions and $g_t^A \in G$ solves the equation

$$\frac{d}{dt} g_t^A = \tilde{A}_t\left(g_t^A\right) \text{ with } g_0^A = e \in G$$

where $\tilde{A}_t(x) := L_{x*} A_t$ is the **left invariant** vector field on G associated to $A_t \in \mathfrak{g}$, see Examples 2.34 and 3.27. Then

$$\partial_B g_t^A = R_{g_t^A *} \int_0^t Ad_{g_\tau^A} B_\tau d\tau$$

where

$$Ad_gA = R_{g^{-1}*}L_{g*}A \text{ for all } g \in G \text{ and } A \in \mathfrak{g}.$$

Proof. Let T_t^A denote the flow of A_t. Because A_t is left invariant,

$$T_t^A(x) = xg_t^A = R_{g_t^A}x$$

as the reader should verify. Thus

$$\begin{aligned}\partial_B g_t^A = \partial_B T_t^A(e) &= R_{g_t^A *}\int_0^t \left(R_{g_\tau^A *}\right)^{-1}\tilde{B}_\tau \circ R_{g_\tau^A}(e)\,d\tau \\ &= R_{g_t^A *}\int_0^t \left(R_{g_\tau^A *}\right)^{-1}\tilde{B}_\tau\left(g_\tau^A\right)d\tau = R_{g_t^A *}\int_0^t \left(R_{g_\tau^A *}\right)^{-1}L_{g_\tau^A *}B_\tau d\tau \\ &= R_{g_t^A *}\int_0^t Ad_{g_\tau^A}B_\tau d\tau.\end{aligned}$$

■

The next theorem expresses $[X_t, Y]$ using the flow T^X. The stochastic analog of this theorem is a key ingredient in the "Malliavin calculus," see Proposition 8.14 below.

Theorem 4.9. *If X_t and T_t^X are as above and $Y \in \Gamma(TM)$, then*

$$\frac{d}{dt}\left[\left(T_{t*}^X\right)^{-1}Y \circ T_t^X\right] = \left(T_{t*}^X\right)^{-1}[X_t, Y]\circ T_t^X \tag{4.9}$$

or equivalently put

$$\frac{d}{dt}Ad_{T_t^X}^{-1} = Ad_{T_t^X}^{-1}L_{X_t} \tag{4.10}$$

where $L_XY := [X, Y]$.

Proof. Let $V_t := \left(T_{t*}^X\right)^{-1}Y \circ T_t^X$ which is equivalent to $T_{t*}^X V_t = Y \circ T_t^X$, or more explicitly to

$$Yf \circ T_t^X = \left(Y \circ T_t^X\right)f = \left(T_{t*}^X V_t\right)f = V_t\left(f \circ T_t^X\right) \text{ for all } f \in C^\infty(M).$$

Differentiating this equation in t then shows

$$\begin{aligned}(X_tYf)\circ T_t^X &= \dot{V}_t\left(f \circ T_t^X\right) + V_t\left(X_tf \circ T_t^X\right) \\ &= \left(T_{t*}^X\dot{V}_t\right)f + \left(T_{t*}^X V_t\right)X_tf \\ &= \left(T_{t*}^X\dot{V}_t\right)f + \left(Y \circ T_t^X\right)X_tf \\ &= \left(T_{t*}^X\dot{V}_t\right)f + (YX_tf)\circ T_t^X.\end{aligned}$$

Therefore

$$\left(T_{t*}^X\dot{V}_t\right)f = ([X_t, Y]f)\circ T_t^X$$

from which we conclude $T_{t*}^X\dot{V}_t = [X_t, Y]\circ T_t^X$ and therefore

$$\dot{V}_t = \left(T_{t*}^X\right)^{-1}[X_t, Y]\circ T_t^X.$$

■

4.3. **Cartan's Development Map.** For this section assume that M is compact[3] Riemannian manifold and let $W^\infty(T_0M)$ be the collection of piecewise smooth paths, $b : [0,1] \to T_oM$ such that $b(0) = 0_o \in T_oM$ and let $W_o^\infty(M)$ be the collection of piecewise smooth paths, $\sigma : [0,1] \to M$ such that $\sigma(0) = o \in M$.

Theorem 4.10 (Development Map)**.** *To each $b \in W^\infty(T_0M)$ there is a unique $\sigma \in W_o^\infty(M)$ such that*

$$\sigma'(s) := (\sigma(s), d\sigma(s)/ds) = //_s(\sigma)b'(s) \quad \text{and} \quad \sigma(0) = o, \tag{4.11}$$

where $//_s(\sigma)$ denotes parallel translation along σ.

Proof. Suppose that σ is a solution to Eq. (4.11) and $//_s(\sigma)v_o = (o, u(s)v)$, where $u(s) : \tau_oM \to \mathbb{R}^N$. Then u satisfies the differential equation

$$u'(s) + dQ(\sigma'(s))u(s) = 0 \quad \text{with} \quad u(0) = u_0, \tag{4.12}$$

where $u_0v := v$ for all $v \in \tau_oM$, see Remark 3.54. Hence Eq. (4.11) is equivalent to the following pair of coupled ordinary differential equations:

$$\sigma'(s) = u(s)b'(s) \quad \text{with} \quad \sigma(0) = o, \tag{4.13}$$

and

$$u'(s) + dQ((\sigma(s), u(s)b'(s))u(s) = 0 \quad \text{with} \quad u(0) = u_0. \tag{4.14}$$

Therefore the uniqueness assertion follows from standard uniqueness theorems for ordinary differential equations. The slickest prove of existence to Eq. (4.11) is to first introduce the orthogonal frame bundle, $O(M)$, on M defined by $O(M) := \cup_{m\in M}O_m(M)$ where $O_m(M)$ is the set of all isometries, $u : T_oM \to T_mM$. It is then possible to show that $O(M)$ is an imbedded submanifold in $\mathbb{R}^N \times \text{Hom}\left(\tau_oM, \mathbb{R}^N\right)$ and that coupled pair of ordinary differential equations (4.13) and (4.14) may be viewed as a flow equation on $O(M)$. Hence the existence of solutions may be deduced from the Theorem 4.2, see, for example, [47] for details of this method. Here I will sketch a proof which does not require us to develop the frame bundle formalism in detail.

Looking at the proof of Lemma 2.30, Q has an extension to a neighborhood in $\mathbb{R}^N$ of $m \in M$ in such a way that $Q(x)$ is still an orthogonal projection onto $\text{Nul}(F'(x))$, where $F(x) = z_>(x)$ is as in Lemma 2.30. Hence for small s, we may define σ and u to be the unique solutions to Eq. (4.13) and Eq. (4.14) with values in $\mathbb{R}^N$ and $\text{Hom}(\tau_oM, \mathbb{R}^N)$ respectively. The key point now is to show that $\sigma(s) \in M$ and that the range of $u(s)$ is $\tau_{\sigma(s)}M$.

[3]It would actually be sufficient to assume that M is a "complete" Riemannian manifold for this section.

Using the same proof as in Theorem 3.52, $w(s) := Q(\sigma(s))u(s)$ satisfies,

$$w' = dQ(\sigma')u + Q(\sigma)u' = dQ(\sigma')u - Q(\sigma)dQ(\sigma')u$$
$$= P(\sigma)dQ(\sigma')u = dQ(\sigma')Q(\sigma)u = dQ(\sigma')w,$$

where Lemma 3.30 was used in the last equality. Since $w(0) = 0$, it follows by uniqueness of solutions to linear ordinary differential equations that $w \equiv 0$ and hence

$$\mathrm{Ran}\,[u(s)] \subset \mathrm{Nul}\,[Q(\sigma(s))] = \mathrm{Nul}\,[F'(\sigma(s))].$$

Consequently

$$dF(\sigma(s))/ds = F'(\sigma(s))d\sigma(s)/ds = F'(\sigma(s))u(s)b'(s) = 0$$

for small s and since $F(\sigma(0)) = F(o) = 0$, it follows that $F(\sigma(s)) = 0$, i.e. $\sigma(s) \in M$. So we have shown that there is a solution (σ, u) to (4.13) and (4.14) for small s such that σ stays in M and $u(s)$ is parallel translation along s. By standard ordinary differential equation methods, there is a maximal solution (σ, u) with these properties. Notice that (σ, u) is a path in $M \times \mathrm{Iso}(T_oM, \mathbb{R}^N)$, where $\mathrm{Iso}(T_oM, \mathbb{R}^N)$ is the set of isometries from T_oM to $\mathbb{R}^N$. Since $M \times \mathrm{Iso}(T_oM, \mathbb{R}^N)$ is a compact space, (σ, u) can not explode. Therefore (σ, u) is defined on the same interval where b is defined. ■

The geometric interpretation of Cartan's map is to roll the manifold M along a freshly painted curve b in T_oM to produce a curve σ on M, see Figure 11.

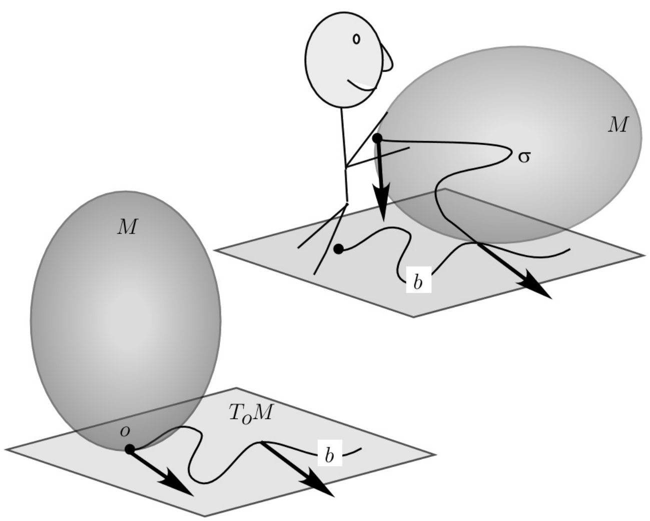

FIGURE 11. Monsieur Cartan is shown here rolling, without "slipping," a manifold M along a curve, b, in T_oM to produce a curve, σ, on M.

Notation 4.11. Let $\phi : W^\infty(T_0M) \to W_o^\infty(M)$ be the map $b \to \sigma$, where σ is the solution to (4.11). It is easy to construct the inverse map $\Psi := \phi^{-1}$. Namely, $\Psi(\sigma) = b$, where

$$\Psi_s(\sigma) = b(s) := \int_0^s //_r(\sigma)^{-1}\sigma'(r)dr.$$

We now conclude this section by computing the differentials of Ψ and ϕ. For more details on computations of this nature the reader is referred to [46, 47] and the references therein.

Theorem 4.12 (Differential of Ψ)**.** *Let $(t, s) \to \Sigma(t, s)$ be a smooth map into M such that $\Sigma(t, \cdot) \in W_o^\infty(M)$ for all t. Let*

$$H(s) := \dot\Sigma(0, s) := (\Sigma(0, s), d\Sigma(t, s)/dt|_{t=0}),$$

so that H is a vector-field along $\sigma := \Sigma(0, \cdot)$. One should view H as an element of the "tangent space" to $W_o^\infty(M)$ at σ, see Figure 12. Let $u(s) := //_s(\sigma)$, $h(s) := //_s(\sigma)^{-1}H(s)$ $b := \Psi_s(\sigma)$ and, for all $a, c \in T_oM$, let

$$(R_u(a, c))(s) := u(s)^{-1}R(u(s)a, u(s)c)u(s). \tag{4.15}$$

Then

$$d\Psi(H) = d\Psi(\Sigma(t, \cdot))/dt|_{t=0} = h + \int_0 \left(\int_0 R_u(h, \delta b)\right)\delta b, \tag{4.16}$$

where $\delta b(s)$ is short hand notation for $b'(s)ds$, and $\int_0 f\delta b$ denotes the function $s \to \int_0^s f(r)b'(r)dr$ when f is a path of matrices.

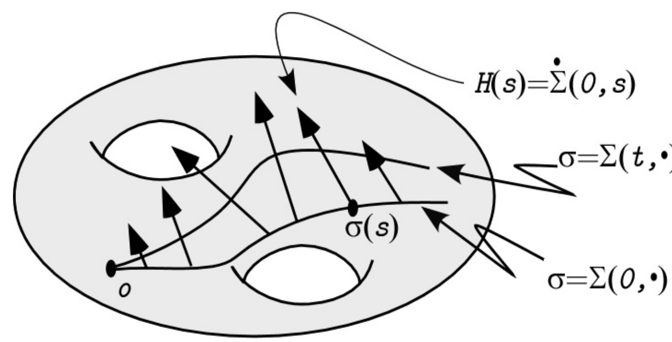

FIGURE 12. A variation of σ giving rise to a vector field along σ.

Proof. To simplify notation let " $\cdot$ "$= \frac{d}{dt}|_0$, " $'$ "$= \frac{d}{ds}$, $B(t, s) := \Psi(\Sigma(t, \cdot))(s)$, $U(t, s) := //_s(\Sigma(t, \cdot))$, $u(s) := //_s(\sigma) = U(0, s)$ and

$$\dot b(s) := (d\Psi(H))(s) := dB(t, s)/dt|_{t=0}.$$

I will also suppress (t,s) from the notation when possible. With this notation

$$\Sigma' = UB', \quad \dot{\Sigma} = H = uh, \tag{4.17}$$

and

$$\frac{\nabla U}{ds} = 0. \tag{4.18}$$

In Eq. (4.18), $\frac{\nabla U}{ds} : T_oM \to T_\Sigma M$ is defined by $\frac{\nabla U}{ds} = P(\Sigma)\, U'$ or equivalently by

$$\frac{\nabla U}{ds} a := \frac{\nabla (Ua)}{ds} \text{ for all } a \in T_oM.$$

Taking ∇/dt of (4.17) at $t = 0$ gives, with the aid of Proposition 3.32,

$$\frac{\nabla U}{dt}|_{t=0} b' + u\dot{b}' = \nabla\Sigma'/dt|_{t=0} = \nabla\dot{\Sigma}/ds = uh'.$$

Therefore,

$$\dot{b}' = h' + Ab', \tag{4.19}$$

where $A := -U^{-1}\frac{\nabla U}{dt}|_{t=0}$, i.e.

$$\frac{\nabla U}{dt}(0,\cdot) = -uA.$$

Taking ∇/ds of this last equation and using $\nabla u/ds = 0$ along with Proposition 3.32 gives

$$-uA' = \frac{\nabla}{ds}\frac{\nabla}{dt}U\Big|_{t=0} = \left[\frac{\nabla}{ds}, \frac{\nabla}{dt}\right] U\Big|_{t=0} = R(\sigma', H)u$$

and hence $A' = R_u(h, b')$. By integrating this identity using $A(0) = 0$ ($\nabla U(t,0)/dt = 0$ since $U(t,0) := //_0(\Sigma(t,\cdot)) = I$ is independent of t) shows

$$A = \int_0 R_u(h, \delta b) \tag{4.20}$$

The theorem now follows by integrating (4.19) relative to s making use of Eq. (4.20) and the fact that $\dot{b}(0) = 0$. ■

Theorem 4.13 (Differential of ϕ)**.** *Let $b, k \in W^\infty(T_0M)$ and $(t,s) \to B(t,s)$ be a smooth map into T_oM such that $B(t,\cdot) \in W^\infty(T_0M)$, $B(0,s) = b(s)$, and $\dot{B}(0,s) = k(s)$. (For example take $B(t,s) = b(s) + tk(s)$.) Then*

$$\phi_*(k_b) := \frac{d}{dt}|_0 \phi(B(t,\cdot)) = //.(\sigma)h,$$

where $\sigma := \phi(b)$ and h is the first component in the solution (h, A) to the pair of coupled differential equations:

$$k' = h' + Ab', \quad \text{with} \quad h(0) = 0 \tag{4.21}$$

and

$$A' = R_u(h, b') \quad \text{with} \quad A(0) = 0. \tag{4.22}$$

Proof. This theorem has an analogous proof to that of Theorem 4.12. We can also deduce the result from Theorem 4.12 by defining Σ by $\Sigma(t,s) := \phi_s(B(t,\cdot))$. We now assume the same notation used in Theorem 4.12 and its proof. Then $B(t,\cdot) = \Psi(\Sigma(t,\cdot))$ and hence by Theorem 4.13

$$k = \frac{d}{dt}|_0 \Psi(\Sigma(t,\cdot)) = d\Psi(H) = h + \int_0 (\int_0 R_u(h, \delta b))\delta b.$$

Therefore, defining $A := \int_0 R_u(h, \delta b)$ and differentiating this last equation relative to s, it follows that A solves (4.22) and that h solves (4.21). ■

The following theorem is a mild extension of Theorem 4.12 to include the possibility that $\Sigma(t,\cdot) \notin W_o^\infty(M)$ when $t \neq 0$, i.e. the base point may change.

Theorem 4.14. *Let $(t,s) \to \Sigma(t,s)$ be a smooth map into M such that $\sigma := \Sigma(0,\cdot) \in W_o^\infty(M)$. Define $H(s) := d\Sigma(t,s)/dt|_{t=0}$, $\sigma := \Sigma(0,\cdot)$, and $h(s) := //_s(\sigma)^{-1}H(s)$. (**Note:** $H(0)$ and $h(0)$ are no longer necessarily equal to zero.) Let*

$$U(t,s) := //_s(\Sigma(t,\cdot))//_t(\Sigma(\cdot,0)) : T_oM \to T_{\Sigma(t,s)}M,$$

so that $\nabla U(t,0)/dt = 0$ and $\nabla U(t,s)/ds \equiv 0$. Set $B(t,s) := \int_0^s U(t,r)^{-1}\Sigma'(t,r)dr$, then

$$\dot{b}(s) := \frac{d}{dt}|_0 B(t,s) = h_s + \int_0^s \left(\int_0 R_u(h, \delta b)\right)\delta b, \tag{4.23}$$

where as before $b := \Psi(\sigma)$.

Proof. The proof is almost identical to the proof of Theorem 4.12 and hence will be omitted. ■

5. Stochastic Calculus on Manifolds

In this section and the rest of the text the reader is assumed to be well versed in stochastic calculus in the Euclidean context.

Notation 5.1. In the sequel we will always assume there is any underlying filtered probability space $(\Omega, \{\mathcal{F}_s\}_{s\geq 0}, \mathcal{F}, \mu)$ satisfying the "usual hypothesis." Namely, $\mathcal{F}$ is μ – complete, $\mathcal{F}_s$ contains all of the null sets in $\mathcal{F}$, and $\mathcal{F}_s$ is right continuous. As usual $\mathbb{E}$ will be used to denote the expectation relative to the probability measure μ.

Definition 5.2. For simplicity, we will call a function $\Sigma : \mathbb{R}_+ \times \Omega \to V$ (V a vector space) a **process** if $\Sigma_s = \Sigma(s) := \Sigma(s,\cdot)$ is $\mathcal{F}_s$ – measurable for all $s \in \mathbb{R}_+ := [0,\infty)$, i.e. a process will mean an adapted process unless otherwise stated. As above, we will always assume that M is an imbedded submanifold of $\mathbb{R}^N$ with the induced Riemannian structure. An M – **valued semi-martingale** is a **continuous** $\mathbb{R}^N$-valued semi-martingale (Σ) such that $\Sigma(s,\omega) \in M$ for all $(s,\omega) \in \mathbb{R}_+ \times \Omega$. It will be convenient to let λ be the distinguished process: $\lambda(s) = \lambda_s := s$.

Since $f \in C^\infty(M)$ is the restriction of a smooth function F on $\mathbb{R}^N$, it follows by Itô's lemma that $f \circ \Sigma = F \circ \Sigma$ is a real-valued semi-martingale if Σ is an M – valued semi-martingale. Conversely, if Σ is an M – valued process and $f \circ \Sigma$ is a real-valued semi-martingale for all $f \in C^\infty(M)$ then Σ is an M – valued semi-martingale. Indeed, let $x = (x^1, \ldots, x^N)$ be the standard coordinates on $\mathbb{R}^N$, then $\Sigma^i := x^i \circ \Sigma$ is a real semi-martingale for each i, which implies that Σ is a $\mathbb{R}^N$- valued semi-martingale.

Notation 5.3 (Fisk-Stratonovich Integral)**.** Suppose V is a finite dimensional vector space and
$$\pi = \{0 = s_0 < s_1 < s_2 < \cdots\}$$
is a partition of $\mathbb{R}_+$ with $\lim_{n\to\infty} s_n = \infty$. To such a partition π, let $|\pi| := \sup_i |s_{i+1} - s_i|$ be the **mesh size** of π and $s \wedge s_i := \min\{s, s_i\}$. To each $\mathrm{Hom}\left(\mathbb{R}^N, V\right)$ – valued semi-martingale Z_t and each M – valued semi-martingale Σ_t, the **Fisk-Stratonovich integral** of Z relative to Σ is defined by
$$\begin{aligned}\int_0^s Z\delta\Sigma &= \lim_{|\pi|\to 0} \sum_{i=0}^{\infty} \frac{1}{2}\left(Z_{s\wedge s_i} + Z_{s\wedge s_{i+1}}\right)\left(\Sigma_{s\wedge s_{i+1}} - \Sigma_{s\wedge s_i}\right) \\ &= \int_0^s Z d\Sigma + \frac{1}{2}\int_0^s dZ d\Sigma \in V\end{aligned}$$
where
$$\int_0^s Z d\Sigma = \lim_{|\pi|\to 0} \sum_{i=0}^{\infty} Z_{s\wedge s_i}\left(\Sigma_{s\wedge s_{i+1}} - \Sigma_{s\wedge s_i}\right) \in V$$
is the **Itô integral** and
$$[Z, \Sigma]_s = \int_0^s dZ d\Sigma := \lim_{|\pi|\to 0} \sum_{i=0}^{\infty} \left(Z_{s\wedge s_i} - Z_{s\wedge s_{i+1}}\right)\left(\Sigma_{s\wedge s_{i+1}} - \Sigma_{s\wedge s_i}\right) \in V$$
is the **mutual variation** of Z and Σ. (All limits may be taken in the sense of uniform convergence on compact subsets of $\mathbb{R}_+$ in probability.)

5.1. Stochastic Differential Equations on Manifolds.

Notation 5.4. Suppose that $\{X_i\}_{i=0}^n \subset \Gamma(TM)$ are vector fields on M. For $a \in \mathbb{R}^n$ let
$$X_a(m) := \mathbf{X}(m)\, a := \sum_{i=1}^n a_i X_i(m)$$
With this notation, $X(m) : \mathbb{R}^n \to T_m M$ is a linear map for each $m \in M$.

Definition 5.5. Given an $\mathbb{R}^n$ – valued semi-martingale, β_s, we say an M – valued semi-martingale Σ_s solves the Fisk-Stratonovich stochastic differential equation
$$\delta\Sigma_s = \mathbf{X}(\Sigma_s)\,\delta\beta_s + X_0(\Sigma_s)\, ds := \sum_{i=1}^n X_i(\Sigma_s)\,\delta\beta_s^i + X_0(\Sigma_s)\, ds \tag{5.1}$$
if for all $f \in C^\infty(M)$,
$$\delta f(\Sigma_s) = \sum_{i=1}^n (X_i f)(\Sigma_s)\,\delta\beta_s^i + X_0 f(\Sigma_s)\, ds,$$
i.e. if
$$f(\Sigma_s) = f(\Sigma_0) + \sum_{i=1}^n \int_0^s (X_i f)(\Sigma_r)\,\delta\beta_r^i + \int_0^s X_0 f(\Sigma_r)\, dr.$$

Lemma 5.6 (Itô Form of Eq. (5.1))**.** *Suppose that $\beta = B$ is an $\mathbb{R}^n$ – valued Brownian motion and let $L := \frac{1}{2}\sum_{i=1}^n X_i^2 + X_0$. Then an M – valued semi-martingale Σ_s solves Eq. (5.1) iff*
$$f(\Sigma_s) = f(\Sigma_0) + \sum_{i=1}^n \int_0^s (X_i f)(\Sigma_r)\, dB_r^i + \int_0^s Lf(\Sigma_r)\, dr \tag{5.2}$$
for all $f \in C^\infty(M)$.

Proof. Suppose that Σ_s solves Eq. (5.1), then
$$\begin{aligned}d\left[(X_i f)(\Sigma_r)\right] &= \sum_{j=1}^n (X_j X_i f)(\Sigma_r)\,\delta B_s^j + X_0 X_i f(\Sigma_s)\, ds \\ &= \sum_{j=1}^n (X_j X_i f)(\Sigma_r)\, dB_s^j + d(BV)\end{aligned}$$
where BV denotes a process of bounded variation. Hence
$$\begin{aligned}\int_0^s (X_i f)(\Sigma_r)\,\delta B_r^i &= \sum_{i=1}^n \int_0^s (X_i f)(\Sigma_r)\, dB_r^i + \frac{1}{2}\int_0^s d\left[(X_i f)(\Sigma_r)\right] dB_r^i \\ &= \sum_{i=1}^n \int_0^s (X_i f)(\Sigma_r)\, dB_r^i + \frac{1}{2}\sum_{i,j=1}^n \int_0^s (X_j X_i f)(\Sigma_r)\, dB_s^j dB_r^i \\ &= \sum_{i=1}^n \int_0^s (X_i f)(\Sigma_r)\, dB_r^i + \frac{1}{2}\int_0^s \sum_{i=1}^n X_i^2 f(\Sigma_r)\, dr.\end{aligned}$$

Similarly if Eq. (5.2) holds for all $f \in C^\infty(M)$ we have
$$d\left[(X_i f)(\Sigma_r)\right] = (X_j X_i f)(\Sigma_r)\, dB_s^j + L X_i f(\Sigma_s)\, ds$$
and so as above
$$\int_0^s (X_i f)(\Sigma_r)\,\delta B_r^i = \sum_{i=1}^n \int_0^s (X_i f)(\Sigma_r)\, dB_r^i + \frac{1}{2}\int_0^s \sum_{i=1}^n X_i^2 f(\Sigma_r)\, dr.$$

Solving for $\int_0^s (X_i f)(\Sigma_r)\,dB_r^i$ and putting the result into Eq. (5.2) shows

$$\begin{aligned} f(\Sigma_s) &= f(\Sigma_0) + \sum_{i=1}^n \int_0^s (X_i f)(\Sigma_r)\,\delta B_r^i - \frac{1}{2}\int_0^s \sum_{i=1}^n X_i^2 f(\Sigma_r)\,dr + \int_0^s Lf(\Sigma_r)\,dr \\ &= f(\Sigma_0) + \sum_{i=1}^n \int_0^s (X_i f)(\Sigma_r)\,\delta B_r^i + \int_0^s X_0 f(\Sigma_r)\,dr. \end{aligned}$$

■

To avoid technical problems with possible explosions of stochastic differential equations in the sequel, we make the following assumption.

Assumption 2. Unless otherwise stated, in the remainder of these notes, M will be a compact manifold imbedded in $E := \mathbb{R}^N$.

To shortcut the development of a number of issues here it is useful to recall the following Wong and Zakai type approximation theorem for solutions to Fisk-Stratonovich stochastic differential equations.

Notation 5.7. Let $\{B_s\}_{s\in[0,T]}$ be a standard $\mathbb{R}^n$—valued Brownian motion. Given a partition

$$\pi = \{0 = s_0 < s_1 < s_2 < ... < s_k = T\}$$

of $[0,T]$, let

$$|\pi| = \max\{s_i - s_{i-1} : i = 1,2,\dots,k\}$$

and

$$B_\pi(s) = B(s_{i-1}) + (s - s_{i-1})\frac{\Delta_i B}{\Delta_i s} \text{ if } s \in (s_{i-1}, s_i],$$

where $\Delta_i B := B(s_i) - B(s_{i-1})$ and $\Delta_i s := s_i - s_{i-1}$. Notice that $B_\pi(s)$ is a continuous piecewise linear path in $\mathbb{R}^n$.

Theorem 5.8 (Wong-Zakai type approximation theorem)**.** *Let $a \in \mathbb{R}^N$,*

$$f : \mathbb{R}^n \times \mathbb{R}^N \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N) \text{ and } f_0 : \mathbb{R}^n \times \mathbb{R}^N \to \mathbb{R}^N$$

be twice differentiable functions with bounded continuous derivatives. Let π and B_π be as in Notation 5.7 and $\xi_\pi(s)$ denote the solution to the ordinary differential equation:

$$\xi_\pi'(s) = f(B_\pi(s), \xi_\pi(s))B_\pi'(s) + f_0(B_\pi(s), \xi_\pi(s)), \qquad \xi_\pi(0) = a \tag{5.3}$$

and ξ denote the solution to the Fisk-Stratonovich stochastic differential equation,

$$d\xi_s = f(B_s, \xi_s)\delta B_s + f_0(B_s, \xi_s)ds, \qquad \xi_0 = a. \tag{5.4}$$

Then, for any $\gamma \in (0, \frac{1}{2})$ and $p \in [1,\infty)$, there is a constant $C(p,\gamma) < \infty$ such that

$$\lim_{|\pi|\to 0} \mathbb{E}\left[\sup_{s\le T} |\xi_\pi(s) - \xi_s|^p\right] \le C(p,\gamma)|\pi|^{\gamma p}. \tag{5.5}$$

This theorem is a special case of Theorem 5.7.3 and Example 5.7.4 in Kunita [116]. Theorems of this type have a long history starting with Wong and Zakai [180, 181]. The reader may also find this and related results in the following *partial* list of references: [7, 10, 11, 20, 22, 44, 68, 94, 103, 107, 108, 118, 117, 126, 129, 132, 134, 135, 141, 142, 151, 166, 174, 167, 175, 177]. Also see [8, 53] and the references therein for more of the geometry associated to the Wong and Zakai approximation scheme.

Remark 5.9 (Transfer Principle). Theorem 5.8 is a manifestation of the **transfer principle** (coined by Malliavin) which loosely states: to get a correct stochastic formula one should take the corresponding deterministic smooth formula and replace all derivatives by Fisk-Stratonovich differentials. We will see examples of this principle over and over again in the sequel.

Theorem 5.10. *Given a point $m \in M$ there exits a unique M – valued semi martingale Σ which solves Eq. (5.1) with the initial condition, $\Sigma_0 = m$. We will write $T_s(m)$ for Σ_s if we wish to emphasize the dependence of the solution on the initial starting point $m \in M$.*

Proof. Existence. If for the moment we assumed that the Brownian motion B_s were differentiable in s, Eq. (5.1) could be written as

$$\Sigma_s' = X_s(\Sigma_s) \text{ with } \Sigma_0 = m$$

where

$$X_s(m) := \sum_{i=1}^n X_i(m)\left(B^i\right)'(s) + X_0(m)$$

and the existence of Σ_s could be deduced from Theorem 4.2. We will make this rigorous with an application of Theorem 5.8.

Let $\{Y_i\}_{i=0}^n$ be smooth vector fields on E with compact support such that $Y_i = X_i$ on M for each i and let $B_\pi(s)$ be as in Notation 5.7 and define

$$X_s^\pi(m) := \sum_{i=1}^n X_i(m)\left(B_\pi^i\right)'(s) + X_0(m) \text{ and}$$

$$Y_s^\pi(m) := \sum_{i=1}^n Y_i(m)\left(B_\pi^i\right)'(s) + Y_0(m).$$

Then by Theorem 4.2 we may use X^π and Y^π to generate (random) flows $T^\pi := T^{X^\pi}$ on M and $\tilde{T}^\pi := T^{Y^\pi}$ on E respectively. Moreover, as in the proof of Theorem 4.2 we know $T_s^\pi(m) = \tilde{T}_s^\pi(m)$ for all $m \in M$. An application of Theorem 5.8 now shows that $\Sigma_s := \tilde{T}_s(m) := \lim_{|\pi|\to 0} \tilde{T}_s^\pi(m) = \lim_{|\pi|\to 0} T_s^\pi(m) \in M$

exists[4] and satisfies the Fisk-Stratonovich differential equation **on** E,

$$d\Sigma_s = \sum_{i=1}^{n} Y_i(\Sigma_s)\,\delta B_s^i + Y_0(\Sigma_s)\,ds \text{ with } \Sigma_0 = m. \tag{5.6}$$

Given $f \in C^\infty(M)$, let $F \in C^\infty(E)$ be chosen so that $f = F|_M$. Then Eq. (5.6) implies

$$d[F(\Sigma_s)] = \sum_{i=1}^{n} Y_i F(\Sigma_s)\,\delta B_s^i + Y_0 F(\Sigma_s)\,ds. \tag{5.7}$$

Since we have already seen $\Sigma_s \in M$ and by construction $Y_i = X_i$ on M, we have $F(\Sigma_s) = f(\Sigma_s)$ and $Y_i F(\Sigma_s) = X_i f(\Sigma_s)$. Therefore Eq. (5.7) implies

$$d[f(\Sigma_s)] = \sum_{i=1}^{n} X_i f(\Sigma_s)\,\delta B_s^i + Y_0 F(\Sigma_s)\,ds,$$

i.e. Σ_s solves Eq. (5.1) as desired.

Uniqueness. If Σ is a solution to Eq. (5.1), then for $F \in C^\infty(E)$, we have

$$\begin{aligned} dF(\Sigma_s) &= \sum_{i=1}^{n} X_i F(\Sigma_s)\,\delta B_s^i + X_0 F(\Sigma_s)\,ds \\ &= \sum_{i=1}^{n} Y_i F(\Sigma_s)\,\delta B_s^i + Y_0 F(\Sigma_s)\,ds \end{aligned}$$

which shows, by taking F to be the standard linear coordinates on E, Σ_s also solves Eq. (5.6). But this is a stochastic differential equation on a Euclidean space E with smooth compactly supported coefficients and therefore has a **unique** solution. ■

5.2. **Line Integrals.** For $a, b \in \mathbb{R}^N$, let $\langle a, b\rangle_{\mathbb{R}^N} := \sum_{i=1}^N a_i b_i$ denote the standard inner product on $\mathbb{R}^N$. Also let $\mathfrak{gl}(N) = \mathfrak{gl}(N, \mathbb{R})$ be the set of $N \times N$ real matrices. (It is not necessary to assume M is compact for most of the results in this section.)

Theorem 5.11. *As above, for $m \in M$, let $P(m)$ and $Q(m)$ denote orthogonal projection or $\mathbb{R}^N$ onto $\tau_m M$ and $\tau_m M^\perp$ respectively. Then for any M – valued semi-martingale Σ,*

$$0 = Q(\Sigma)\delta\Sigma \text{ and } d\Sigma = P(\Sigma)\,\delta\Sigma,$$

i.e.

$$\Sigma_s - \Sigma_0 = \int_0^s P(\Sigma_r)\delta\Sigma_r.$$

[4]Here we have used the fact that M is a closed subset of $\mathbb{R}^N$.

Proof. We will first assume that M is the level set of a function F as in Theorem 2.5. Then we may assume that

$$Q(x) = \phi(x)F'(x)^*(F'(x)F'(x)^*)^{-1}F'(x),$$

where ϕ is smooth function on $\mathbb{R}^N$ such that $\phi := 1$ in a neighborhood of M and the support of ϕ is contained in the set: $\{x \in \mathbb{R}^N | F'(x) \text{ is surjective}\}$. By Itô's lemma

$$0 = d0 = d(F(\Sigma)) = F'(\Sigma)\delta\Sigma.$$

The lemma follows in this special case by multiplying the above equation through by $\phi(\Sigma)F'(\Sigma)^*(F'(\Sigma)F'(\Sigma)^*)^{-1}$, see the proof of Lemma 2.30.

For the general case, choose two open covers $\{V_i\}$ and $\{U_i\}$ of M such that each $\bar{V}_i$ is compactly contained in U_i, there is a smooth function $F_i \in C_c^\infty(U_i \to \mathbb{R}^{N-d})$ such that $V_i \cap M = V_i \cap \{F_i^{-1}(\{0\})\}$ and F_i has a surjective differential on $V_i \cap M$. Choose $\phi_i \in C_c^\infty(\mathbb{R}^N)$ such that the support of ϕ_i is contained in V_i and $\sum \phi_i = 1$ on M, with the sum being locally finite. (For the existence of such covers and functions, see the discussion of partitions of unity in any reasonable book about manifolds.) Notice that $\phi_i \cdot F_i \equiv 0$ and that $F_i \cdot \phi_i' \equiv 0$ on M so that

$$\begin{aligned} 0 &= d\{\phi_i(\Sigma)F_i(\Sigma)\} = (\phi_i'(\Sigma)\delta\Sigma)F_i(\Sigma) + \phi_i(\Sigma)F_i'(\Sigma)\delta\Sigma \\ &= \phi_i(\Sigma)F_i'(\Sigma)\delta\Sigma. \end{aligned}$$

Multiplying this equation by $\Psi_i(\Sigma)F_i'(\Sigma)^*(F_i'(\Sigma)F_i'(\Sigma)^*)^{-1}$, where each Ψ_i is a smooth function on $\mathbb{R}^N$ such that $\Psi_i \equiv 1$ on the support of ϕ_i and the support of Ψ_i is contained in the set where F_i' is surjective, we learn that

$$0 = \phi_i(\Sigma)F_i'(\Sigma)^*(F_i'(\Sigma)F_i'(\Sigma)^*)^{-1}F_i'(\Sigma)\delta\Sigma = \phi_i(\Sigma)Q(\Sigma)\delta\Sigma \tag{5.8}$$

for all i. By a stopping time argument we may assume that Σ never leaves a compact set, and therefore we may choose a finite subset I of the indices $\{i\}$ such that $\sum_{i\in I}\phi_i(\Sigma)Q(\Sigma) = Q(\Sigma)$. Hence summing over $i \in I$ in equation (5.8) shows that $0 = Q(\Sigma)\delta\Sigma$. Since $Q + P = I$, it follows that

$$d\Sigma = I\delta\Sigma = [Q(\Sigma) + P(\Sigma)]\,\delta\Sigma = P(\Sigma)\,\delta\Sigma.$$

■

The following notation will be needed to define line integrals along a semi-martingale Σ.

Notation 5.12. Let $P(m)$ be orthogonal projection of $\mathbb{R}^N$ onto $\tau_m M$ as above.

(1) Given a one-form α on M let $\tilde{\alpha} : M \to (\mathbb{R}^N)^*$ be defined by

$$\tilde{\alpha}(m)v := \alpha((P(m)v)_m) \tag{5.9}$$

for all $m \in M$ and $v \in \mathbb{R}^N$.

(2) Let $\Gamma(T^*M \otimes T^*M)$ denote the set of functions $\rho : \cup_{m\in M} T_mM \otimes T_mM \to \mathbb{R}$ such that $\rho_m := \rho|_{T_mM\otimes T_mM}$ is linear, and $m \to \rho(X(m) \otimes Y(m))$ is a smooth function on M for all smooth vector-fields $X, Y \in \Gamma(TM)$. (Riemannian metrics and Hessians of smooth functions are examples of elements of $\Gamma(T^*M \otimes T^*M)$.)

(3) For $\rho \in \Gamma(T^*M \otimes T^*M)$, let $\tilde{\rho} : M \to (\mathbb{R}^N \otimes \mathbb{R}^N)^*$ be defined by

$$\tilde{\rho}(m)(v \otimes w) := \rho((P(m)v)_m \otimes (P(m)w)_m). \tag{5.10}$$

Definition 5.13. Let α be a one form on M, $\rho \in \Gamma(T^*M \otimes T^*M)$, and Σ be an M − valued semi-martingale. Then the **Fisk-Stratonovich** integral of α along Σ is:

$$\int_0^\cdot \alpha(\delta\Sigma) := \int_0^\cdot \tilde{\alpha}(\Sigma)\delta\Sigma, \tag{5.11}$$

and the **Itô** integral is given by:

$$\int_0^\cdot \alpha(\bar{d}\Sigma) := \int_0^\cdot \tilde{\alpha}(\Sigma)d\Sigma, \tag{5.12}$$

where the stochastic integrals on the right hand sides of Eqs. (5.11) and (5.12) are Fisk-Stratonovich and Itô integrals respectively. Formally, $\bar{d}\Sigma := P(\Sigma)d\Sigma$. We also define **quadratic integral**:

$$\int_0^\cdot \rho(d\Sigma \otimes d\Sigma) := \int_0^\cdot \tilde{\rho}(\Sigma)(d\Sigma \otimes d\Sigma) := \sum_{i,j=1}^N \int_0^\cdot \tilde{\rho}(\Sigma)(e_i \otimes e_j)d[\Sigma^i, \Sigma^j], \tag{5.13}$$

where $\{e_i\}_{i=1}^N$ is an orthonormal basis for $\mathbb{R}^N$, $\Sigma^i := \langle e_i, \Sigma\rangle$, and $d[\Sigma^i, \Sigma^j]$ is the differential of the mutual quadratic variation of Σ^i and Σ^j.

So as not to confuse $[\Sigma^i, \Sigma^j]$ with a commutator or a Lie bracket, in the sequel we will write $d\Sigma^i d\Sigma^j$ for $d[\Sigma^i, \Sigma^j]$.

Remark 5.14. The above definitions may be generalized as follows. Suppose that α is now a T^*M − valued semi-martingale and Σ is the M valued semi-martingale such that $\alpha_s \in T^*_{\Sigma_s}M$ for all s. Then we may define

$$\tilde{\alpha}_s v := \alpha_s((P(\Sigma_s)v)_{\Sigma_s}),$$

$$\int_0^\cdot \alpha(\delta\Sigma) := \int_0^\cdot \tilde{\alpha}\delta\Sigma, \tag{5.14}$$

and

$$\int_0^\cdot \alpha(\bar{d}\Sigma) := \int_0^\cdot \tilde{\alpha}d\Sigma. \tag{5.15}$$

Similarly, if ρ is a process in $T^*M \otimes T^*M$ such that $\rho_s \in T^*_{\Sigma_s}M \otimes T^*_{\Sigma_s}M$, let

$$\int_0^\cdot \rho(d\Sigma \otimes d\Sigma) = \int_0^\cdot \tilde{\rho}(d\Sigma \otimes d\Sigma), \tag{5.16}$$

where

$$\tilde{\rho}_s(v \otimes w) := \rho_s((P(\Sigma_s)v)_{\Sigma_s} \otimes (P(\Sigma_s)v)_{\Sigma_s})$$

and

$$d\Sigma \otimes d\Sigma = \sum_{i,j=1}^N e_i \otimes e_j d\Sigma^i d\Sigma^j \tag{5.17}$$

as in Eq. (5.13).

Lemma 5.15. *Suppose that $\alpha = fdg$ for some functions $f, g \in C^\infty(M)$, then*

$$\int_0^\cdot \alpha(\delta\Sigma) = \int_0^\cdot f(\Sigma)\delta[g(\Sigma)].$$

Since, by Corollary 3.42, any one form α on M may be written as $\alpha = \sum_{i=1}^N f_i dg_i$ with $f_i, g_i \in C^\infty(M)$, it follows that the Fisk-Stratonovich integral is intrinsically defined independent of how M is imbedded into a Euclidean space.

Proof. Let G be a smooth function on $\mathbb{R}^N$ such that $g = G|_M$. Then $\tilde{\alpha}(m) = f(m)G'(m)P(m)$, so that

$$\begin{aligned}\int_0^\cdot \alpha(\delta\Sigma) &= \int_0^\cdot f(\Sigma)G'(\Sigma)P(\Sigma)\delta\Sigma \\ &= \int_0^\cdot f(\Sigma)G'(\Sigma)\delta\Sigma \qquad \text{(by Theorem 5.11)} \\ &= \int_0^\cdot f(\Sigma)\delta[G(\Sigma)] \qquad \text{(by Itô's Lemma)} \\ &= \int_0^\cdot f(\Sigma)\delta[g(\Sigma)]. \qquad (g(\Sigma) = G(\Sigma))\end{aligned}$$

■

Lemma 5.16. *Suppose that $\rho = fdh \otimes dg$, where $f, g, h \in C^\infty(M)$, then*

$$\int_0^\cdot \rho(d\Sigma \otimes d\Sigma) = \int_0^\cdot f(\Sigma)d[h(\Sigma), g(\Sigma)] =: \int_0^\cdot f(\Sigma)d\,[h(\Sigma)]\,d\,[g(\Sigma)]\,.$$

Since, by an argument similar to that in Corollary 3.42, any $\rho \in \Gamma(T^*M \otimes T^*M)$ may be written as a finite linear combination $\rho = \sum_i f_i dh_i \otimes dg_i$ with $f_i, h_i, g_i \in C^\infty(M)$, it follows that the quadratic integral is intrinsically defined independent of the imbedding.

Proof. By Theorem 5.11, $\delta\Sigma = P(\Sigma)\delta\Sigma$, so that

$$\begin{aligned}\Sigma_s^i &= \Sigma_0^i + \int_0^\cdot (e_i, P(\Sigma)d\Sigma) + B.V. \\ &= \Sigma_0^i + \sum_k \int_0^\cdot (e_i, P(\Sigma)e_k)d\Sigma^k + B.V.,\end{aligned}$$

where $B.V.$ denotes a process of bounded variation. Therefore

$$d[\Sigma^i, \Sigma^j] = \sum_{k,l}(e_i, P(\Sigma)e_k)(e_i, P(\Sigma)e_l)d\Sigma^k d\Sigma^l. \tag{5.18}$$

Now let H and G be in $C^\infty(\mathbb{R}^N)$ such that $h = H|_M$ and $g = G|_M$. By Itô's lemma and Eq. (5.18),

$$\begin{aligned} d[h(\Sigma), g(\Sigma)] &= \sum_{i,j}(H'(\Sigma)e_i)(G'(\Sigma)e_j)d[\Sigma^i, \Sigma^j] \\ &= \sum_{i,j,k,l}(H'(\Sigma)e_i)(G'(\Sigma)e_j)(e_i, P(\Sigma)e_k)(e_i, P(\Sigma)e_l)d\Sigma^k d\Sigma^l \\ &= \sum_{k,l}(H'(\Sigma)P(\Sigma)e_k)(G'(\Sigma)P(\Sigma)e_l)d\Sigma^k d\Sigma^l. \end{aligned}$$

Since

$$\tilde{\rho}(m) = f(m) \cdot (H'(m)P(m)) \otimes (G'(m)P(m)),$$

it follows from Eq. (5.13) and the two above displayed equations that

$$\begin{aligned} \int_0^\cdot f(\Sigma)d[h(\Sigma), g(\Sigma)] &:= \int_0^\cdot \sum_{k,l} f(\Sigma)(H'(\Sigma)P(\Sigma)e_k)(G'(\Sigma)P(\Sigma)e_l)d\Sigma^k d\Sigma^l \\ &= \int_0^\cdot \tilde{\rho}(\Sigma)(d\Sigma \otimes d\Sigma) =: \int_0^\cdot \rho(d\Sigma \otimes d\Sigma). \end{aligned}$$

■

Theorem 5.17. *Let α be a one form on M, and Σ be a M – valued semi-martingale. Then*

$$\int_0^\cdot \alpha(\delta\Sigma) = \int_0^\cdot \alpha(\bar{d}\Sigma) + \frac{1}{2}\int_0^\cdot \nabla\alpha(d\Sigma \otimes d\Sigma), \tag{5.19}$$

where $\nabla\alpha(v_m \otimes w_m) := (\nabla_{v_m}\alpha)(w_m)$ and $\nabla\alpha$ is defined in Definition 3.40, also see Lemma 3.41. (This shows that the Itô integral depends not only on the manifold structure of M but on the geometry of M as reflected in the Levi-Civita covariant derivative ∇.)

Proof. Let $\tilde{\alpha}$ be as in Eq. (5.9). For the purposes of the proof, suppose that $\tilde{\alpha} : M \to (\mathbb{R}^N)^*$ has been extended to a smooth function from $\mathbb{R}^N \to (\mathbb{R}^N)^*$. We still denote this extension by $\tilde{\alpha}$. Then using Eq. (5.18),

$$\begin{aligned} \int_0^\cdot \alpha(\delta\Sigma) &:= \int_0^\cdot \tilde{\alpha}(\Sigma)\delta\Sigma \\ &= \int_0^\cdot \tilde{\alpha}(\Sigma)d\Sigma + \frac{1}{2}\int_0^\cdot \tilde{\alpha}'(\Sigma)(d\Sigma)d\Sigma \\ &= \int_0^\cdot \alpha(\bar{d}\Sigma) + \frac{1}{2}\sum_{i,j,k,l}\int_0^\cdot \tilde{\alpha}'(\Sigma)(e_i)e_j(e_i, P(\Sigma)e_k)(e_i, P(\Sigma)e_l)d\Sigma^k d\Sigma^l \\ &= \int_0^\cdot \alpha(\bar{d}\Sigma) + \frac{1}{2}\sum_{k,l}\int_0^\cdot \tilde{\alpha}'(\Sigma)(P(\Sigma)e_k)P(\Sigma)e_l d\Sigma^k d\Sigma^l \\ &= \int_0^\cdot \alpha(\bar{d}\Sigma) + \frac{1}{2}\sum_{k,l}\int_0^\cdot d\tilde{\alpha}((P(\Sigma)e_k)_\Sigma)P(\Sigma)e_l d\Sigma^k d\Sigma^l. \end{aligned}$$

But by Eq. (3.45), we know for all $v_m, w_m \in TM$ that

$$\nabla\alpha(v_m \otimes w_m) = d\tilde{\alpha}(v_m)w$$

which combined with the previous equation implies

$$\begin{aligned} \int_0^\cdot \alpha(\delta\Sigma) &= \int_0^\cdot \alpha(\bar{d}\Sigma) + \frac{1}{2}\sum_{k,l}\int_0^\cdot \nabla\alpha((P(\Sigma)e_k)_\Sigma \otimes (P(\Sigma)e_l)_\Sigma)d\Sigma^k d\Sigma^l \\ &= \int_0^\cdot \alpha(\bar{d}\Sigma) + \frac{1}{2}\sum_{k,l}\int_0^\cdot \nabla\alpha(d\Sigma \otimes d\Sigma). \end{aligned}$$

■

Corollary 5.18 (Itô's Lemma for Manifolds)**.** *If $u \in C^\infty((0,T) \times M)$ and Σ is an $M-$valued semi-martingale, then*

$$\begin{aligned} d\left[u\left(s, \Sigma_s\right)\right] &= \left(\partial_s u\right)\left(s, \Sigma_s\right)ds \\ &\quad + d_M\left[u\left(s, \cdot\right)\right]\left(\bar{d}\Sigma_s\right) + \frac{1}{2}\left(\nabla d_M u\left(s, \cdot\right)\right)\left(d\Sigma_s \otimes d\Sigma_s\right), \end{aligned} \tag{5.20}$$

where, as in Notation 2.20, $d_M u(s, \cdot)$ is being used to denote the differential of the map: $m \in M \to u(s, m)$.

Proof. Let $U \in C^\infty((0,T) \times \mathbb{R}^N)$ such that $u(s, \cdot) = U(s, \cdot)|_M$. Then by Itô's lemma and Theorem 5.11,

$$\begin{aligned} d\left[u\left(s, \Sigma_s\right)\right] = d\left[U\left(s, \Sigma_s\right)\right] &= \left(\partial_s U\right)\left(s, \Sigma_s\right)ds + D_\Sigma U(s, \Sigma_s)\delta\Sigma_s \\ &= \left(\partial_s U\right)\left(s, \Sigma_s\right)ds + D_\Sigma U(s, \Sigma_s)P(\Sigma_s)\delta\Sigma_s \\ &= \left(\partial_s u\right)\left(s, \Sigma_s\right)ds + d_M\left[u\left(s, \cdot\right)\right]\left(\delta\Sigma_s\right) \\ &= \left(\partial_s u\right)\left(s, \Sigma_s\right)ds + d_M\left[u\left(s, \cdot\right)\right]\left(\bar{d}\Sigma_s\right) \\ &\quad + \frac{1}{2}\left(\nabla d_M u\left(s, \cdot\right)\right)\left(d\Sigma_s \otimes d\Sigma_s\right), \end{aligned}$$

wherein the last equality is a consequence of Theorem 5.17. ■

5.3. M – valued Martingales and Brownian Motions.

Definition 5.19. An M – valued semi-martingale Σ is said to be a (local) **martingale** (more precisely a ∇-martingale) if

$$\int_0^{\cdot} df(\bar{d}\Sigma) = f(\Sigma) - f(\Sigma_0) - \frac{1}{2}\int_0^{\cdot} \nabla df(d\Sigma \otimes d\Sigma) \tag{5.21}$$

is a (local) martingale for all $f \in C^\infty(M)$. (See Theorem 5.17 for the truth of the equality in Eq. (5.21).) The process Σ is said to be a **Brownian motion** if

$$f(\Sigma) - f(\Sigma_0) - \frac{1}{2}\int_0^{\cdot} \Delta f(\Sigma) d\lambda \tag{5.22}$$

is a local martingale for all $f \in C^\infty(M)$, where $\lambda(s) := s$ and $\int_0^{\cdot} \Delta f(\Sigma) d\lambda$ denotes the process $s \to \int_0^s \Delta f(\Sigma) d\lambda$.

Theorem 5.20 (Projection Construction of Brownian Motion)**.** *Suppose that $B = \left(B^1, B^2, \ldots, B^N\right)$ is an N – dimensional Brownian motion. The there is a unique M – valued semi-martingale Σ which solves the Fisk-Stratonovich stochastic differential equation,*

$$\delta\Sigma = P(\Sigma)\delta B \quad \text{with} \quad \Sigma_0 = o \in M, \tag{5.23}$$

see Figure 13. Moreover, Σ is an M – valued Brownian motion.

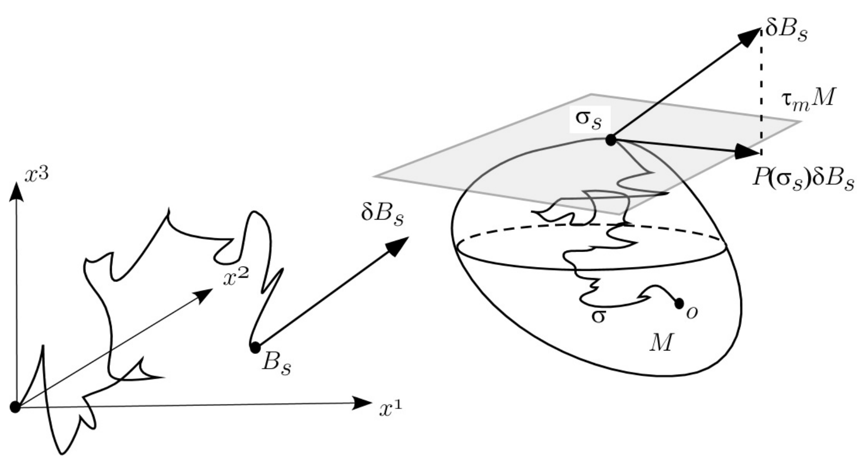

FIGURE 13. Projection construction of Brownian motion on M.

Proof. Let $\{e_i\}_{i=1}^N$ be the standard basis for $\mathbb{R}^N$ and $X_i(m) := P(m)e_i \in T_mM$ for each $i = 1, 2, \ldots, N$ and $m \in M$. Then Eq. (5.23) is equivalent to the Stochastic differential equation.,

$$\delta\Sigma = \sum_{i=1}^N X_i(\Sigma)\delta B^i \quad \text{with} \quad \Sigma_0 = o \in M$$

which has a unique solution by Theorem 5.10. Using Lemma 5.6, this equation may be rewritten in Itô form as

$$d\left[f(\Sigma)\right] = \sum_{i=1}^N X_i f(\Sigma) dB^i + \frac{1}{2}\sum_{i=1}^N X_i^2 f(\Sigma)\, ds \text{ for all } f \in C^\infty(M).$$

This completes the proof since $\sum_{i=1}^N X_i^2 = \Delta$ by Proposition 3.48. ■

Lemma 5.21 (Lévy's Criteria)**.** *For each $m \in M$, let $\mathcal{I}(m) := \sum_{i=1}^d E_i \otimes E_i$, where $\{E_i\}_{i=1}^d$ is an orthonormal basis for T_mM. An M – valued semi-martingale, Σ, is a Brownian motion iff Σ is a martingale and*

$$d\Sigma \otimes d\Sigma = \mathcal{I}(\Sigma) d\lambda. \tag{5.24}$$

More precisely, this last condition is to be interpreted as:

$$\int_0^{\cdot} \rho(d\Sigma \otimes d\Sigma) = \int_0^{\cdot} \rho(\mathcal{I}(\Sigma)) d\lambda \;\forall\; \rho \in \Gamma(T^*M \otimes T^*M). \tag{5.25}$$

Proof. ($\Rightarrow$) Suppose that Σ is a Brownian motion on M (so Eq. (5.22) holds) and $f, g \in C^\infty(M)$. Then on one hand

$$\begin{aligned} d(f(\Sigma)g(\Sigma)) &= d\left[f(\Sigma)\right] \cdot g(\Sigma) + f(\Sigma) d\left[g(\Sigma)\right] + d[f(\Sigma), g(\Sigma)] \\ &\cong \frac{1}{2}\{\Delta f(\Sigma) g(\Sigma) + f(\Sigma)\Delta g(\Sigma)\} d\lambda + d[f(\Sigma), g(\Sigma)], \end{aligned}$$

where "$\cong$" denotes equality up to the differential of a martingale. On the other hand,

$$\begin{aligned} d(f(\Sigma)g(\Sigma)) &\cong \frac{1}{2}\Delta(fg)(\Sigma) d\lambda \\ &= \frac{1}{2}\{\Delta f(\Sigma) g(\Sigma) + f(\Sigma)\Delta g(\Sigma) + 2\langle \operatorname{grad} f, \operatorname{grad} g\rangle(\Sigma)\} d\lambda. \end{aligned}$$

Comparing the above two equations implies that

$$d[f(\Sigma), g(\Sigma)] = \langle \operatorname{grad} f, \operatorname{grad} g\rangle(\Sigma) d\lambda = df \otimes dg(\mathcal{I}(\Sigma)\rangle d\lambda.$$

Therefore by Lemma 5.16, if $\rho = h \cdot df \otimes dg$ then

$$\begin{aligned} \int_0^{\cdot} \rho(d\Sigma \otimes d\Sigma) &= \int_0^{\cdot} h(\Sigma) d[f(\Sigma), g(\Sigma)] \\ &= \int_0^{\cdot} h(\Sigma)(df \otimes dg)(\mathcal{I}(\Sigma)) d\lambda = \int_0^{\cdot} \rho(\mathcal{I}(\Sigma)) d\lambda. \end{aligned}$$

Since the general element ρ of $\Gamma(T^*M\otimes T^*M)$ is a finite linear combination of expressions of the form $hdf\otimes dg$, it follows that Eq. (5.24) holds. Moreover, Eq. (5.24) implies

$$(\nabla df)(d\Sigma\otimes d\Sigma)=(\nabla df)(\mathcal{I}(\Sigma))d\lambda=\Delta f(\Sigma)d\lambda \tag{5.26}$$

and therefore,

$$\begin{aligned}f(\Sigma)-f(\Sigma_0)-\frac{1}{2}\int_0^\cdot \nabla df(d\Sigma\otimes d\Sigma)\\ =f(\Sigma)-f(\Sigma_0)-\frac{1}{2}\int_0^\cdot \Delta f(\Sigma)d\lambda\end{aligned} \tag{5.27}$$

is a martingale and so by definition Σ is a martingale.

Conversely assume Σ is a martingale and Eq. (5.24) holds. Then Eq. (5.26) and Eq. (5.27) hold and they imply Σ is a Brownian motion, see Definition 5.19. ∎

Definition 5.22 ($\delta^\nabla V:=P\delta V$)**.** Suppose α is a one form on M and V is a $TM-$valued semi-martingale, i.e. $V_s=(\Sigma_s,v_s)$, where Σ is an M – valued semi-martingale and v is a $\mathbb{R}^N$-valued semi-martingale such that $v_s\in\tau_{\Sigma_s}M$ for all s. Then we define:

$$\int_0^\cdot \alpha(\delta^\nabla V):=\int_0^\cdot \tilde\alpha(\Sigma)\delta v=\int_0^\cdot \alpha(\Sigma)\left(P\left(\Sigma\right)\delta v\right). \tag{5.28}$$

Remark 5.23. Suppose that $\alpha(v_m)=\theta(m)v$, where $\theta:M\to(\mathbb{R}^N)^*$ is a smooth function. Then

$$\int_0^\cdot \alpha(\delta^\nabla V):=\int_0^\cdot \theta(\Sigma)P(\Sigma)\delta v=\int_0^\cdot \theta(\Sigma)\{\delta v+dQ(\delta\Sigma)v\},$$

where we have used the identity:

$$\delta^\nabla V=P(\Sigma)\delta v=\delta v+dQ(\delta\Sigma)v. \tag{5.29}$$

This last identity follows by taking the differential of the identity, $v=P(\Sigma)v$, as in the proof of Proposition 3.32.

Proposition 5.24 (Product Rule)**.** *Keeping the notation of above, we have*

$$\delta(\alpha(V))=\nabla\alpha(\delta\Sigma\otimes V)+\alpha(\delta^\nabla V), \tag{5.30}$$

where $\nabla\alpha(\delta\Sigma\otimes V):=\gamma(\delta\Sigma)$ and γ is the T^*M – valued semi-martingale defined by

$$\gamma_s(w):=\nabla\alpha(w\otimes V_s)=(\nabla_w\alpha)(V_s)\ \textit{for any}\ w\in T_{\Sigma_s}M.$$

Proof. Let $\theta:\mathbb{R}^N\to(\mathbb{R}^N)^*$ be a smooth map such that $\tilde\alpha(m)=\theta(m)|_{\tau_m M}$ for all $m\in M$. By Lemma 5.15, $\delta(\theta(\Sigma)P(\Sigma))=d(\theta P)(\delta\Sigma)$ and hence by Lemma 3.41, $\delta(\theta(\Sigma)P(\Sigma))v=\nabla\alpha(\delta\Sigma\otimes V)$, where $\nabla\alpha(v_m\otimes w_m):=(\nabla_{v_m}\alpha)(w_m)$ for all $v_m,w_m\in TM$. Therefore:

$$\begin{aligned}\delta(\alpha(V))&=\delta(\theta(\Sigma)v)=\delta(\theta(\Sigma)P(\Sigma)v)=(d(\theta P)(\delta\Sigma))v+\theta(\Sigma)P(\Sigma)\delta v\\ &=(d(\theta P)(\delta\Sigma))v+\tilde\alpha(\Sigma)\delta v=\nabla\alpha(\delta\Sigma\otimes V)+\alpha(\delta^\nabla V).\end{aligned}$$

∎

5.4. Stochastic Parallel Translation and Development Maps.

Definition 5.25. A TM – valued semi-martingale V is said to be **parallel** if $\delta^\nabla V\equiv 0$, i.e. $\int_0^\cdot \alpha(\delta^\nabla V)\equiv 0$ for all one forms α on M.

Proposition 5.26. *A TM – valued semi-martingale $V=(\Sigma,v)$ is parallel iff*

$$\int_0^\cdot P(\Sigma)\delta v=\int_0^\cdot \{\delta v+dQ(\delta\Sigma)v\}\equiv 0. \tag{5.31}$$

Proof. Let $x=(x^1,\dots,x^N)$ denote the standard coordinates on $\mathbb{R}^N$. If V is parallel then,

$$0\equiv\int_0^\cdot dx^i(\delta^\nabla V)=\int_0^\cdot \langle e_i,P(\Sigma)\delta v\rangle$$

for each i which implies Eq. (5.31). The converse follows from Remark 5.23. ∎

In the following theorem, V_0 is said to be a measurable vector-field on M if $V_0(m)=(m,v(m))$ with $v:M\to\mathbb{R}^N$ being a measurable function such that $v(m)\in\tau_m M$ for all $m\in M$.

Theorem 5.27 (Stochastic Parallel Translation on $M\times\mathbb{R}^N$)**.** *Let Σ be an M – valued semi-martingale, and $V_0(m)=(m,v(m))$ be a measurable vector-field on M, then there is a unique parallel TM-valued semi-martingale V such that $V_0=V_0(\Sigma_0)$ and $V_s\in T_{\Sigma_s}M$ for all s. Moreover, if u denotes the solution to the stochastic differential equation:*

$$\delta u+\Gamma(\delta\Sigma)u=0\quad \textit{with}\quad u_0=I\in O(N), \tag{5.32}$$

(where $O(N)$ is as in Example 2.6 and Γ is as in Eq. (3.65)) then $V_s=(\Sigma_s,u_s v(\Sigma_0))$. The process u defined in (5.32) is orthogonal for all s and satisfies $P(\Sigma_s)u_s=u_sP(\Sigma_0)$. Moreover if $\Sigma_0=o\in M$ a.e. and $v\in\tau_oM$ and $w\perp\tau_oM$, then u_sv and u_sw satisfy

$$\delta\left[u_sv\right]+dQ\left(\delta\Sigma\right)u_sv=P\left(\Sigma\right)\delta\left[u_sv\right]=0 \tag{5.33}$$

and

$$\delta\left[u_sw\right]+dP\left(\delta\Sigma\right)u_sv=Q\left(\Sigma\right)\delta\left[u_sv\right]=0. \tag{5.34}$$

Proof. The assertions prior to Eq. (5.33) are the stochastic analogs of Lemmas 3.56 and 3.57. The proof may be given by replacing $\frac{d}{ds}$ everywhere in the proofs of Lemmas 3.56 and 3.57 by δ_s to get a proof in this stochastic setting. Eqs. (5.33) and (5.34) are now easily verified, for example using and $P(\Sigma)uv = uv$, we have

$$\delta[uv] = \delta[P(\Sigma)uv] = P(\delta\Sigma)uv + P(\Sigma)\delta[uv]$$

which proves the first equality in Eq. (5.33). For the second equality in Eq. (5.33),

$$\begin{aligned}P(\Sigma)\delta[uv] &= -P(\Sigma)\Gamma(\delta\Sigma)[uv] \\ &= -P(\Sigma)[dQ(\delta\Sigma)P(\Sigma) + dP(\delta\Sigma)Q(\Sigma)][uv] \\ &= -dQ(\delta\Sigma)Q(\Sigma)P(\Sigma)\delta[uv] = 0\end{aligned}$$

where Lemma 3.30 was used in the third equality. The proof of Eq. (5.34) is completely analogous. The skeptical reader is referred to Section 3 of Driver [47] for more details. ∎

Definition 5.28 (Stochastic Parallel Translation)**.** Given $v \in \mathbb{R}^N$ and an M – valued semi-martingale Σ, let $//_s(\Sigma)v_{\Sigma_0} = (\Sigma_s, u_s v)$, where u solves (5.32). (Note: $V_s = //_s(\Sigma)V_0$.)

In the remainder of these notes, I will often abuse notation and write u_s instead of $//_s := //_s(\Sigma)$ and v_s rather than $V_s = (\Sigma_s, v_s)$. For example, the reader should sometimes interpret $u_s v$ as $//_s(\Sigma)v_{\Sigma_0}$ depending on the context. Essentially, we will be identifying $\tau_m M$ with $T_m M$ when no particular confusion will arise.

Convention. Let us now fix a **base point** $o \in M$ and unless otherwise noted, we will assume that all M – valued semi-martingales, Σ, start of $o \in M$, i.e. $\Sigma_0 = o$ a.e.

To each M – valued semi-martingale, Σ, let $\Psi(\Sigma) := b$ where

$$b := \int_0^\cdot //^{-1}\delta\Sigma = \int_0^\cdot u^{-1}\delta\Sigma = \int_0^\cdot u^{\mathrm{tr}}\delta\Sigma.$$

Then $b = \Psi(\Sigma)$ is a T_oM – valued semi-martingale such that $b_0 = 0_o \in T_oM$. The converse holds as well.

Theorem 5.29 (Stochastic Development Map)**.** *Suppose that $o \in M$ is given and b is a T_oM – valued semi-martingale. Then there exists a unique M – valued semi-martingale Σ such that*

$$\delta\Sigma_s = //_s\delta b_s = u_s\delta b_s \quad \textit{with} \quad \Sigma_0 = o \tag{5.35}$$

where u solves (5.32). .

Proof. This theorem is a stochastic analog of Theorem 4.10 and the reader is again referred to Figure 11. To prove the existence and uniqueness, we may follow the method in the proof of Theorem 4.10. Namely, the pair $(\Sigma, u) \in M \times O(N)$ solves an Stochastic differential equation. of the form

$$\begin{aligned}&\delta\Sigma = u\delta b \quad \text{with} \quad \Sigma_0 = o \\ &\delta u = -\Gamma(\delta\Sigma)u = -\Gamma(u\delta b)u \quad \text{with} \quad u_0 = I \in O(N)\end{aligned}$$

which after a little effort can be expressed in a form for which Theorem 5.10 may be applied. The details will be left to the reader, or see (for example) Section 3 of Driver [47]. ∎

Notation 5.30. As in the smooth case, define $\Sigma = \phi(b)$, so that

$$\Psi(\Sigma) := \phi^{-1}(b) = \int_0^\cdot //_r(\Sigma)^{-1}\delta\Sigma_r.$$

In what follows, we will assume that b_s, u_s (or equivalently $//_s(\Sigma)$), and Σ_s are related by Equations (5.35) and (5.32), i.e. $\Sigma = \phi(b)$ and $u = // = //(\Sigma)$. Recall that $\bar{d}\Sigma = P(\Sigma)d\Sigma$ is the Itô differential of Σ, see Definition 5.13.

Proposition 5.31. *Let $\Sigma = \phi(b)$, then*

$$\bar{d}\Sigma = P(\Sigma)d\Sigma = udb. \tag{5.36}$$

Also

$$d\Sigma \otimes d\Sigma = udb \otimes udb := \sum_{i,j=1}^d ue_i \otimes ue_j db^i db^j, \tag{5.37}$$

where $\{e_i\}_{i=1}^d$ is an orthonormal basis for T_oM and $b = \sum_{i=1}^d b^i e_i$. More precisely

$$\int_0^\cdot \rho(d\Sigma \otimes d\Sigma) = \int_0^\cdot \sum_{i,j=1}^d \rho(ue_i \otimes ue_j)db^i db^j,$$

for all $\rho \in \Gamma(T^*M \otimes T^*M)$.

Proof. Consider the identity:

$$\begin{aligned}d\Sigma = u\delta b &= udb + \frac{1}{2}dudb \\ &= udb - \frac{1}{2}\Gamma(\delta\Sigma)udb = udb - \frac{1}{2}\Gamma(udb)udb\end{aligned}$$

where Γ is as defined in Eq. (3.65). Hence

$$\bar{d}\Sigma = P(\Sigma)d\Sigma = udb - \frac{1}{2}\sum_{i,j=1}^d P(\Sigma)\Gamma((ue_i)_\Sigma)ue_j db^i db^j.$$

The proof of Eq. (5.36) is finished upon observing,

$$P\Gamma P = P\{dQP + dPQ\}P = PdQP = PQdQ = 0.$$

The proof of Eq. (5.37) is easy and will be left for the reader. ∎

Fact 5.32. If (M,g) is a complete Riemannian manifold and the Ricci curvature tensor is bounded from below[5], then $\Delta = \Delta_g$ acting on $C_c^\infty(M)$ is essentially self-adjoint, i.e. the closure $\bar{\Delta}$ of Δ is an unbounded self-adjoint operator on $L^2(M,d\lambda)$. (Here $d\lambda = \sqrt{g}dx^1 \dots dx^n$ is being used to denote the Riemann volume measure on M.) Moreover, the semi-group $e^{t\bar{\Delta}/2}$ has a smooth integral kernel, $p_t(x,y)$, such that

$$p_t(x,y) \geq 0 \text{ for all } x,y \in M$$

$$\int_M p_t(x,y)d\lambda(y) = 1 \text{ for all } x \in M \text{ and}$$

$$\left(e^{t\bar{\Delta}/2}f\right)(x) = \int_M p_t(x,y)f(y)d\lambda(y) \text{ for all } f \in L^2(M).$$

If $f \in C_c^\infty(M)$, the function $u(t,x) := e^{t\bar{\Delta}/2}f(x)$ is smooth for $t > 0$ and $x \in M$ and $Le^{t\bar{\Delta}/2}f(x)$ is continuous for $t \geq 0$ and $x \in M$ for any smooth linear differential operator L on $C^\infty(M)$. For these results, see for example Strichartz [165], Dodziuk [43] and Davies [41].

Theorem 5.33 (Stochastic Rolling Constructions)**.** *Assume M is compact and let Σ, $u_s = //_s$, and b be as in Theorem 5.29, then:*

1. *Σ is a martingale iff b is a T_oM – valued martingale.*
2. *Σ is a Brownian motion iff b is a T_oM – valued Brownian motion.*

Furthermore if Σ is a Brownian motion, $T \in (0,\infty)$ and $f \in C^\infty(M)$, then

$$M_s := \left(e^{(T-s)\bar{\Delta}/2}f\right)(\Sigma_s)$$

is a martingale for $s \in [0,T]$ and

$$dM_s = \left(de^{(T-s)\bar{\Delta}/2}f\right)(u_s db_s)_{\Sigma_s} = \left(de^{(T-s)\bar{\Delta}/2}f\right)(//_s db_s). \tag{5.38}$$

Proof. Keep the same notation as in Proposition 5.31 and let $f \in C^\infty(M)$. By Proposition 5.31, if b is a martingale, then $\int_0^\cdot df(\bar{d}\Sigma) = \int_0^\cdot df(udb)$ is also a martingale and hence Σ is a martingale, see Definition 5.19. Combining this with Corollary 5.18 and Proposition 5.31,

$$\begin{aligned} d[f(\Sigma)] &= df(\bar{d}\Sigma) + \frac{1}{2}\nabla df(d\Sigma \otimes d\Sigma) \\ &= df(udb) + \frac{1}{2}\nabla df(udb \otimes udb). \end{aligned}$$

Since u is an isometry, if and b is a Brownian motion then $udb \otimes udb = \mathcal{I}(\Sigma)d\lambda$. Hence

$$d[f(\Sigma)] = df(udb) + \frac{1}{2}\Delta f(\Sigma)d\lambda$$

from which it follows that Σ is a Brownian motion.

Conversely, if Σ is a M – valued martingale, then

$$N := \sum_{i=1}^N \int_0^\cdot dx^i(\bar{d}\Sigma)e_i = \sum_{i=1}^N \int_0^\cdot \langle e_i, udb\rangle e_i = \int_0^\cdot udb \tag{5.39}$$

is a martingale, where $x = (x^1,\dots,x^N)$ are standard coordinates on $\mathbb{R}^N$ and $\{e_i\}_{i=1}^N$ is the standard basis for $\mathbb{R}^N$. From Eq. (5.39), it follows that $b = \int_0^\cdot u^{-1}dN$ is also a martingale.

Now suppose that Σ is an M – valued Brownian motion, then we have already proved that b is a martingale. To finish the proof it suffices by Lévy's criteria (Lemma 5.21) to show that $db \otimes db = \mathcal{I}(o)d\lambda$. But $\Sigma = N +$(bounded variation) and hence

$$\begin{aligned} db \otimes db &= u^{-1}d\Sigma \otimes u^{-1}d\Sigma = u^{-1}dN \otimes u^{-1}dN \\ &= (u^{-1} \otimes u^{-1})(d\Sigma \otimes d\Sigma) \\ &= (u^{-1} \otimes u^{-1})\mathcal{I}(\Sigma)d\lambda = \mathcal{I}(o)d\lambda, \end{aligned}$$

wherein Eq. (5.24) was used in the fourth equality and the orthogonality of u was used in the last equality.

To prove Eq. (5.38), let $M_s = u(s,\Sigma_s)$ where $u(s,x) := \left(e^{(T-s)\bar{\Delta}/2}f\right)(x)$ which satisfies

$$\partial_s u(s,x) + \frac{1}{2}\Delta u(s,x) = 0 \text{ with } u(T,x) = f(x)$$

By Itô's Lemma (see Corollary 5.18) along with Lemma 5.21 and Proposition 5.31,

$$\begin{aligned} dM_s &= \partial_s u(s,\Sigma_s)ds + d_M[u(s,\cdot)](\bar{d}\Sigma_s) + \frac{1}{2}\nabla d_M[u(s,\cdot)](d\Sigma_s \otimes d\Sigma_s) \\ &= \partial_s u(s,\Sigma_s)ds + \frac{1}{2}\Delta u(s,\Sigma_s)ds + \left(d_M e^{(T-s)\bar{\Delta}/2}f\right)((u_s db_s)_{\Sigma_s}) \\ &= \left(d_M e^{(T-s)\bar{\Delta}/2}f\right)((u_s db_s)_{\Sigma_s}). \end{aligned}$$

■

The rolling construction of Brownian motion seems to have first been discovered by Eells and Elworthy [63] who used ideas of Gangolli [87]. The relationship of the stochastic development map to stochastic differential equations on the orthogonal frame bundle $O(M)$ of M is pointed out in Elworthy [66, 67, 68]. The frame bundle point of view has also been extensively developed by Malliavin, see for example [130, 129, 131]. For a more detailed history of the stochastic development map, see pp. 156–157 in Elworthy [68]. The reader may also wish to consult [74, 103, 116, 132, 171, 101].

Corollary 5.34. *If Σ is a Brownian motion on M,*

$$\pi = \{0 = s_0 < s_1 < \cdots < s_n = T\}$$

[5]These assumptions are always satisfied when M is compact.

is a partition of $[0,T]$ *and* $f \in C^\infty(M^n)$ *, then*

$$\mathbb{E}f(\Sigma_{s_1},\dots,\Sigma_{s_n}) = \int_{M^n} f(x_1,x_2,\dots,x_n)\prod_{i=1}^n p_{\Delta_i s}(x_{i-1},x_i)\,d\lambda(x_i) \tag{5.40}$$

where $\Delta_i s := s_i - s_{i-1}$, $x_0 := o$ *and* $\lambda := \lambda_M$. *In particular* Σ *is a Markov process relative to the filtration,* $\{\mathcal{F}_s\}$ *where* $\mathcal{F}_s$ *is the* σ *– algebra generated by* $\{\Sigma_\tau : \tau \le s\}$.

Proof. By standard measure theoretic arguments, it suffices to prove Eq. (5.40) when f is a product function of the form $f(x_1,x_2,\dots,x_n) = \prod_{i=1}^n f_i(x_i)$ with $f_i \in C^\infty(M)$. By Theorem 5.33, $M_s := e^{(T-s)\bar{\Delta}/2} f_n(\Sigma_s)$ is a martingale for $s \le T$ and therefore

$$\begin{aligned}\mathbb{E}[f(\Sigma_{s_1},\dots,\Sigma_{s_n})] &= \mathbb{E}\left[\prod_{i=1}^{n-1} f_i(\Sigma_{s_i})\cdot M_T\right] = \mathbb{E}\left[\prod_{i=1}^{n-1} f_i(\Sigma_{s_i})\cdot M_{s_{n-1}}\right]\\ &= \mathbb{E}\left[\prod_{i=1}^{n-1} f_i(\Sigma_{s_i})\cdot (P_{\Delta_n s} f_n)(\Sigma_{s_{n-1}})\right]. \end{aligned}\tag{5.41}$$

In particular if $n=1$, it follows that

$$\mathbb{E}[f_1(\Sigma_T)] = \mathbb{E}\left[\left(e^{T\bar{\Delta}/2} f_1\right)(\Sigma_0)\right] = \int_M p_T(o,x_1) f_1(x_1)\,d\lambda(x_1).$$

Now assume we have proved Eq. (5.40) with n replaced by $n-1$ and to simplify notation let $g(x_1,x_2,\dots,x_{n-1}) := \prod_{i=1}^{n-1} f_i(x_i)$. It would then follow from Eq. (5.41) that

$$\begin{aligned}&\mathbb{E}[f(\Sigma_{s_1},\dots,\Sigma_{s_n})]\\ &= \int_{M^{n-1}} g(x_1,x_2,\dots,x_{n-1})\left(e^{\frac{s_n-s_{n-1}}{2}\bar{\Delta}} f_n\right)(x_{n-1})\prod_{i=1}^{n-1} p_{\Delta_i s}(x_{i-1},x_i)\,d\lambda(x_i)\\ &= \int_{M^{n-1}} g(x_1,x_2,\dots,x_{n-1})\left[\int_M f_n(x_n)\,p_{\Delta_n s}(x_{n-1},x_n)\,d\lambda(x_n)\right]\times\\ &\qquad\qquad\times\prod_{i=1}^{n-1} p_{\Delta_i s}(x_{i-1},x_i)\,d\lambda(x_i)\\ &= \int_{M^n} f(x_1,x_2,\dots,x_n)\prod_{i=1}^n p_{\Delta_i s}(x_{i-1},x_i)\,d\lambda(x_i).\end{aligned}$$

This completes the induction step and hence also the proof of the theorem. ■

5.5. More Constructions of Semi-Martingales and Brownian Motions.

Let Γ be the one form on M with values in the skew symmetric $N \times N$ matrices defined by $\Gamma = dQP + dPQ$ as in Eq. (3.65). Given an M−valued semi-martingale Σ, let u denote parallel translation along Σ as defined in Eq. (5.32) of Theorem 5.27.

Lemma 5.35 (Orthogonality Lemma)**.** *Suppose that* B *is an* $\mathbb{R}^N$ *– valued semi-martingale and* Σ *is the solution to*

$$\delta\Sigma = P(\Sigma)\delta B \quad \textit{with} \quad \Sigma_0 = o \in M. \tag{5.42}$$

Let $\{e_i\}_{i=1}^N$ *be any orthonormal basis for* $\mathbb{R}^N$ *and define* $B^i := \langle e_i, B\rangle$ *then*

$$P(\Sigma)dB \otimes Q(\Sigma)dB := \sum_{i,j=1}^N P(\Sigma)e_i \otimes Q(\Sigma)e_j\left(dB^i dB^j\right) = 0.$$

Proof. Suppose $\{v_i\}_{i=1}^N$ is another orthonormal basis for $\mathbb{R}^N$. Using the bilinearity of the joint quadratic variation,

$$\begin{aligned}[\langle e_i,B\rangle,\langle e_j,B\rangle] &= \sum_{k,l}[\langle e_i,v_k\rangle\langle v_k,B\rangle,\langle e_j,v_l\rangle\langle v_l,B\rangle]\\ &= \sum_{k,l}\langle e_i,v_k\rangle\langle e_j,v_l\rangle[\langle v_k,B\rangle,\langle v_l,B\rangle].\end{aligned}$$

Therefore,

$$\begin{aligned}&\sum_{i,j=1}^N P(\Sigma)e_i\otimes Q(\Sigma)e_j\cdot d\left[B^i,B^j\right]\\ &\quad= \sum_{i,j,k,l=1}^N [P(\Sigma)e_i\otimes Q(\Sigma)e_j]\,\langle e_i,v_k\rangle\langle e_j,v_l\rangle d[\langle v_k,B\rangle,\langle v_l,B\rangle]\\ &\quad= \sum_{k,l=1}^N [P(\Sigma)v_k\otimes Q(\Sigma)v_l]\,d[\langle v_k,B\rangle,\langle v_l,B\rangle]\end{aligned}$$

which shows $P(\Sigma)dB \otimes Q(\Sigma)dB$ is well defined.

Now define

$$\tilde{B} := \int_0^\cdot u^{-1}dB \text{ and } \tilde{B}^i := \langle e_i,\tilde{B}\rangle = \int_0^\cdot \langle ue_i, dB\rangle$$

where u is parallel translation along Σ in $M\times\mathbb{R}^N$ as defined in Eq. (5.32). Then

$$\begin{aligned}P(\Sigma)dB\otimes Q(\Sigma)dB &= \sum_{i,j,k,l=1}^{N} P(\Sigma)ue_k\otimes Q(\Sigma)ue_l\langle e_i,ue_k\rangle\langle e_j,ue_l\rangle\left(dB^idB^j\right)\\ &=\sum_{k,l=1}^{N} P(\Sigma)ue_k\otimes Q(\Sigma)ue_l\left(d\tilde{B}^kd\tilde{B}^l\right)\\ &=\sum_{k,l=1}^{N} uP(o)e_k\otimes uQ(o)e_l\left(d\tilde{B}^kd\tilde{B}^l\right)\end{aligned}$$

wherein we have used $P(\Sigma)u=uP(o)$ and $Q(\Sigma)u=uQ(o)$, see Theorem 5.27. This last expression is easily seen to be zero by choosing $\{e_i\}$ such that $P(o)e_i=e_i$ for $i=1,2,\dots,d$ and $Q(o)e_j=e_j$ for $j=d+1,\dots,N$. ■

The next proposition is a stochastic analogue of Lemma 3.55 and the proof is very similar to that of Lemma 3.55.

Proposition 5.36. *Suppose that V is a TM – valued semi-martingale, $\Sigma=\pi(V)$ so that Σ is an M – valued semi-martingale and $V_s\in T_{\Sigma_s}M$ for all $s\geq 0$. Then*

$$//_s\delta_s\left[//_s^{-1}V_s\right]=\delta_s^{\nabla}V_s=:P(\Sigma_s)\,\delta V_s \tag{5.43}$$

where $//_s$ is stochastic parallel translation along Σ. If $Y_s\in\Gamma(TM)$ is a time dependent vector field, then

$$\delta_s\left[//_s^{-1}Y_s(\Sigma_s)\right]=//_s^{-1}\left(\frac{d}{ds}Y_s\right)(\Sigma_s)\,ds+//_s^{-1}\nabla_{\delta\Sigma_s}Y_s \tag{5.44}$$

and for $w\in T_oM$,

$$\begin{aligned}//_s^{-1}\delta_s^{\nabla}\left[\nabla_{//_sw}Y_s\right]&=\delta_s\left[//_s^{-1}\nabla_{//_sw}Y_s\right]\\ &=//_s^{-1}\nabla^2_{\delta\Sigma_s\otimes//_sw}Y_s+//_s^{-1}\left[\nabla_{//_sw}\left(\frac{d}{ds}Y_s\right)\right]ds.\end{aligned} \tag{5.45}$$

Furthermore if Σ_s is a Brownian motion, then

$$\begin{aligned}d\left[//_s^{-1}Y_s(\Sigma_s)\right]=&//_s^{-1}\nabla_{//_sdb_s}Y_s+//_s^{-1}\left(\frac{d}{ds}Y_s\right)(\Sigma_s)\,ds\\ &+\frac{1}{2}\sum_{i=1}^{d}//_s^{-1}\nabla^2_{//_se_i\otimes//_se_i}Y_sds\end{aligned} \tag{5.46}$$

where $\{e_i\}_{i=1}^d$ is an orthonormal basis for T_oM.

Proof. We will use the convention of summing on repeated indices and write u_s for stochastic parallel translation, $//_s$, in TM along Σ. Recall that u_s solves

$$\delta u_s+dQ(\delta\Sigma_s)\,u_s=0\text{ with }u_0=I_{T_oM}.$$

Define $\bar{u}_s$ as the solution to:

$$\delta\bar{u}_s=\bar{u}_sdQ(\delta\Sigma_s)\text{ with }\bar{u}_0=I_{T_oM}.$$

Then

$$\delta(\bar{u}_su_s)=-\bar{u}_sdQ(\delta\Sigma_s)\,u_s+\bar{u}_sdQ(\delta\Sigma_s)\,u_s=0$$

from which it follows that $\bar{u}_su_s=I$ for all s and hence $\bar{u}_s=u_s^{-1}$. This proves Eq. (5.43) since

$$\begin{aligned}u_s\delta_s\left[u_s^{-1}V_s\right]&=u_s\left[u_s^{-1}dQ(\delta\Sigma_s)\,V_s+u_s^{-1}\delta V_s\right]\\ &=dQ(\delta\Sigma_s)\,V_s+\delta V_s=\delta^{\nabla}V_s,\end{aligned}$$

where the last equality comes from Eq. (5.29).

Applying Eq. (5.43) to $V_s:=Y_s(\Sigma_s)$ gives

$$\begin{aligned}\delta_s\left[//_s^{-1}Y_s(\Sigma_s)\right]&=//_s^{-1}P(\Sigma_s)\,\delta_s[Y_s(\Sigma_s)]\\ &=//_s^{-1}P(\Sigma_s)\left(\frac{d}{ds}Y_s\right)(\Sigma_s)\,ds+//_s^{-1}P(\Sigma_s)\,Y_s'(\Sigma_s)\,\delta_s\Sigma_s\\ &=//_s^{-1}\left(\frac{d}{ds}Y_s\right)(\Sigma_s)\,ds+//_s^{-1}\nabla_{\delta_s\Sigma_s}Y_s,\end{aligned}$$

which proves Eq. (5.44).

To prove Eq. (5.45), let $X_i(m)=P(m)\,e_i$ for $i=1,2,\dots,N$. By Proposition 3.48,

$$\begin{aligned}\nabla_{//_sw}Y_s&=\langle//_sw,X_i(\Sigma_s)\rangle\,(\nabla_{X_i}Y_s)(\Sigma_s)\\ &=\langle w,//_s^{-1}X_i(\Sigma_s)\rangle\,(\nabla_{X_i}Y_s)(\Sigma_s)\end{aligned} \tag{5.47}$$

and

$$//_sw=\langle//_sw,X_i(\Sigma_s)\rangle X_i(\Sigma_s)=\langle w,//_s^{-1}X_i(\Sigma_s)\rangle X_i(\Sigma_s)$$

or equivalently,

$$w=\langle w,//_s^{-1}X_i(\Sigma_s)\rangle//_s^{-1}X_i(\Sigma_s)\,. \tag{5.48}$$

Taking the covariant differential of Eq. (5.47), making use of Eq. (5.44), gives

$$\begin{aligned}
&\delta_s^\nabla\left[\nabla_{//_s w} Y_s\right]\\
&= \langle //_s w, \nabla_{\delta_s\Sigma_s} X_i\rangle \left(\nabla_{X_i} Y_s\right)(\Sigma_s) + \langle //_s w, X_i(\Sigma_s)\rangle \nabla_{\delta_s\Sigma_s}\nabla_{X_i} Y_s\\
&\quad + \langle //_s w, X_i(\Sigma_s)\rangle \left(\nabla_{X_i}\left(\frac{d}{ds}Y_s\right)\right)(\Sigma_s)\, ds\\
&= \langle //_s w, \nabla_{\delta_s\Sigma_s} X_i\rangle \left(\nabla_{X_i} Y_s\right)(\Sigma_s) + \langle //_s w, X_i(\Sigma_s)\rangle \nabla^2_{\delta_s\Sigma_s\otimes X_i} Y_s\\
&\quad + \langle //_s w, X_i(\Sigma_s)\rangle \nabla_{\nabla_{\delta_s\Sigma_s} X_i} Y_s + \left(\nabla_{//_s w}\left(\frac{d}{ds}Y_s\right)\right)(\Sigma_s)\, ds\\
&= \left(\nabla_{\langle //_s w, \nabla_{\delta_s\Sigma_s} X_i\rangle X_i(\Sigma_s) + \langle //_s w, X_i(\Sigma_s)\rangle \nabla_{\delta_s\Sigma_s} X_i} Y_s\right)(\Sigma_s)\\
&\quad + \nabla^2_{\delta_s\Sigma_s\otimes //_s w} Y_s + \left(\nabla_{//_s w}\left(\frac{d}{ds}Y_s\right)\right)(\Sigma_s)\, ds, \qquad (5.49)
\end{aligned}$$

Taking the differential of Eq. (5.48) implies

$$0 = \delta w = \langle w, //_s^{-1}\nabla_{\delta_s\Sigma_s} X_i\rangle //_s^{-1} X_i(\Sigma_s) + \langle w, //_s^{-1} X_i(\Sigma_s)\rangle //_s^{-1}\nabla_{\delta_s\Sigma_s} X_i$$

which upon multiplying by $//_s$ shows

$$\langle //_s w, \nabla_{\delta_s\Sigma_s} X_i\rangle X_i(\Sigma_s) + \langle //_s w, X_i(\Sigma_s)\rangle \nabla_{\delta_s\Sigma_s} X_i = 0.$$

Using this identity in Eq. (5.49) completes the proof of Eq. (5.45).

Now suppose that Σ_s is a Brownian motion and $b_s = \Psi_s(\Sigma)$ is the anti-developed T_oM – valued Brownian motion associated to Σ. Then by Eq. (5.44),

$$\begin{aligned}
d\left[//_s^{-1} Y_s(\Sigma_s)\right] &= //_s^{-1}\left(\frac{d}{ds}Y_s\right)(\Sigma_s)\, ds + //_s^{-1}\nabla_{//_s \delta b_s} Y_s\\
&= //_s^{-1}\left(\frac{d}{ds}Y_s\right)(\Sigma_s)\, ds + \left(//_s^{-1}\nabla_{//_s e_i} Y_s\right)\delta b_s^i.
\end{aligned}$$

Using Eq. (5.45),

$$\begin{aligned}
\left(//_s^{-1}\nabla_{//_s e_i} Y_s\right)\delta b_s^i &= \left(//_s^{-1}\nabla_{//_s e_i} Y_s\right) db_s^i + \frac{1}{2} d\left(//_s^{-1}\nabla_{//_s e_i} Y_s\right) db_s^i\\
&= //_s^{-1}\nabla_{//_s db_s} Y_s + \frac{1}{2}//_s^{-1}\nabla^2_{\delta\Sigma_s\otimes //_s e_i} Y_s db_s^i\\
&= //_s^{-1}\nabla_{//_s db_s} Y_s + \frac{1}{2}//_s^{-1}\nabla^2_{//_s e_j\otimes //_s e_i} Y_s db_s^i db_s^j\\
&= //_s^{-1}\nabla_{//_s db_s} Y_s + \frac{1}{2}//_s^{-1}\nabla^2_{//_s e_i\otimes //_s e_i} Y_s ds.
\end{aligned}$$

Combining the last two equations proves Eq. (5.46). ∎

Theorem 5.37. *Let Σ_s denote the solution to Eq. (5.1) with $\Sigma_0 = o \in M$, $\beta = B$ and $b_s = \Psi_s(\Sigma) \in T_oM$. Then*

$$\begin{aligned}
b_s &= \int_0^s //_r^{-1}(\Sigma)\left[\mathbf{X}(\Sigma_r)\,\delta B_r + X_0(\Sigma_r)\, dr\right]\\
&= \int_0^s //_r^{-1}(\Sigma)\,\mathbf{X}(\Sigma_r)\, dB_r\\
&\quad + \int_0^s //_r^{-1}\left[\frac{1}{2}\sum_{i,j=1}^n \left(\nabla_{X_i} X_j\right)(\Sigma_r)\, dB_r^i dB_r^j + X_0(\Sigma_r)\, dr\right]. \qquad (5.50)
\end{aligned}$$

Hence if B is a Brownian motion, then

$$\begin{aligned}
b_s &= \int_0^s //_r^{-1}(\Sigma)\,\mathbf{X}(\Sigma_r)\, dB_r\\
&\quad + \int_0^s //_r^{-1}\left[\frac{1}{2}\sum_{i=1}^n \left(\nabla_{X_i} X_i\right)(\Sigma_r) + X_0(\Sigma_r)\right] dr. \qquad (5.51)
\end{aligned}$$

Proof. By the definition of b,

$$\begin{aligned}
db_s &= //_s^{-1}(\Sigma)\left[\mathbf{X}(\Sigma_s)\,\delta B_s + X_0(\Sigma_s)\, ds\right]\\
&= //_s^{-1}(\Sigma)\left[\mathbf{X}(\Sigma_s)\, dB_s + X_0(\Sigma_s)\, ds\right] + \frac{1}{2} d\left[//_s^{-1}(\Sigma)\,\mathbf{X}(\Sigma_s)\right] dB_s\\
&= //_s^{-1}(\Sigma)\left[\mathbf{X}(\Sigma_s)\, dB_s + X_0(\Sigma_s)\, ds\right] + \frac{1}{2}\left[//_s^{-1}(\Sigma)\,\nabla_{\mathbf{X}(\Sigma_s) dB_s}\mathbf{X}\right] dB_s\\
&= //_s^{-1}(\Sigma)\left[\mathbf{X}(\Sigma_s)\, dB_s + ds\right] + \frac{1}{2}//_s^{-1}(\Sigma)\sum_{i,j=1}^n \left(\nabla_{X_i} X_j\right)(\Sigma_s)\, dB_s^i dB_s^j
\end{aligned}$$

which combined with the identity,

$$\begin{aligned}
d\left[//_s^{-1}(\Sigma)\,\mathbf{X}(\Sigma_s)\right] dB_s &= \left[//_s^{-1}(\Sigma)\,\nabla_{d\Sigma_s}\mathbf{X}\right] dB_s = \left[//_s^{-1}(\Sigma)\,\nabla_{\mathbf{X}(\Sigma_s) dB_s}\mathbf{X}\right] dB_s\\
&= \sum_{i,j=1}^n \left(\nabla_{X_i} X_j\right)(\Sigma_s)\, dB_s^i dB_s^j
\end{aligned}$$

proves Eq. (5.50). ∎

Corollary 5.38. *Suppose B_s is an $\mathbb{R}^n$ – valued Brownian motion, Σ_s is the solution to Eq. (5.1) with $\beta = B$ and $\frac{1}{2}\sum_{k=1}^n \left(\nabla_{X_k} X_k\right) + X_0 = 0$, then Σ is an M – valued martingale with quadratic variation,*

$$d\Sigma_s \otimes d\Sigma_s = \sum_{k=1}^n X_k(\Sigma_s)\otimes X_k(\Sigma_s)\, ds. \qquad (5.52)$$

Proof. By Eq. (5.51) and Theorem 5.33, Σ is a martingale and from Eq. (5.1),

$$d\Sigma^i d\Sigma^j = \sum_{k,l=1}^n X_k^i(\Sigma) X_l^j(\Sigma) dB^k dB^l = \sum_{k=1}^n X_k^i(\Sigma) X_k^j(\Sigma) ds$$

where $\{e_i\}_{i=1}^N$ is the standard basis for $\mathbb{R}^N$, $\Sigma^i := \langle \Sigma, e_i\rangle$ and $X_k^i(\Sigma) = \langle X_k(\Sigma), e_i\rangle$. Using this identity in Eq. (5.17), shows

$$d\Sigma_s \otimes d\Sigma_s = \sum_{i,j=1}^N \sum_{k=1}^n e_i \otimes e_j X_k^i(\Sigma) X_k^j(\Sigma) ds = \sum_{k=1}^n X_k(\Sigma_s) \otimes X_k(\Sigma_s) ds.$$

■

Corollary 5.39. *Suppose now that B_s is an $\mathbb{R}^N$ – valued semi-martingale and Σ_s is the solution to Eq. (5.42) in Lemma 5.35. If B is a martingale, then Σ is a martingale and if B is a Brownian motion, then Σ is a Brownian motion.*

Proof. Solving Eq. (5.42) is the same as solving Eq. (5.1) with $n = N$, $\beta = B$, $X_0 \equiv 0$ and $X_i(m) = P(m) e_i$ for all $i = 1, 2, \ldots, N$. Since

$$\nabla_{X_i} X_j = P dP(X_i) e_j = dP(X_i) Q e_j = dP(P e_i) Q e_j,$$

it follows from orthogonality Lemma 5.35 that

$$\sum_{i,j=1}^n (\nabla_{X_i} X_j)(\Sigma_r) dB_r^i dB_r^j = 0.$$

Therefore from Eq. (5.50), $b_s := \int_0^s //_r^{-1} \delta\Sigma_r$ is a T_oM – martingale which is equivalent to Σ_s being a M – valued martingale. Finally if B is a Brownian motion, then from Eq. (5.52), Σ has quadratic variation given by

$$d\Sigma_s \otimes d\Sigma_s = \sum_{i=1}^N P(\Sigma_s) e_i \otimes P(\Sigma_s) e_i ds \tag{5.53}$$

Since $\sum_{i=1}^N P(m)e_i \otimes P(m)e_i$ is independent of the choice of orthonormal basis for $\mathbb{R}^N$, we may choose $\{e_i\}$ such that $\{e_i\}_{i=1}^d$ is an orthonormal basis for $\tau_m M$ to learn

$$\sum_{i=1}^N P(m)e_i \otimes P(m)e_i = \mathcal{I}(m).$$

Using this in Eq. (5.53) we learn that $d\Sigma_s \otimes d\Sigma_s = \mathcal{I}(\Sigma_s) ds$ and hence Σ is a Brownian motion on M by the Lévy criteria, see Lemma 5.21. ■

Theorem 5.40. *Let B be any $\mathbb{R}^N$-valued semi-martingale, Σ be the solution to Eq. (5.42),*

$$b := \int_0^{\cdot} u^{-1}\delta\Sigma = \int_0^{\cdot} u^{-1} P(\Sigma)\delta B \tag{5.54}$$

be the anti-development of Σ and

$$\beta := \int_0^{\cdot} u^{-1} Q(\Sigma) dB = Q(o)\int_0^{\cdot} u^{-1} dB \tag{5.55}$$

be the "normal" process. Then

$$b = \int_0^{\cdot} u^{-1} P(\Sigma) dB = P(o)\int_0^{\cdot} u^{-1} dB, \tag{5.56}$$

i.e. the Fisk-Stratonovich integral may be replaced by the Itô integral. Moreover if B is a standard $\mathbb{R}^N$ – valued Brownian motion then (b,β) is also a standard $\mathbb{R}^N$ – valued Brownian and the processes, b_s, Σ_s and $//_s$ are all independent of β.

Proof. Let $p = P(\Sigma)$ and u be parallel translation on $M \times \mathbb{R}^N$ (see Eq. (5.32)), then

$$\begin{aligned} d(u^{-1}P(\Sigma)) \cdot dB &= u^{-1}\left[\Gamma(\delta\Sigma)P(\Sigma)dB + dP(\delta\Sigma)dB\right] \\ &= u^{-1}\left[(dQ(\delta\Sigma)P(\Sigma) + dP(\delta\Sigma)Q(\Sigma))P(\Sigma)dB + dP(\delta\Sigma)dB\right] \\ &= u^{-1}\left[dQ(\delta\Sigma)P(\Sigma)dB - dQ(\delta\Sigma)dB\right] \\ &= -u^{-1}dQ(\delta\Sigma)Q(\Sigma)dB = -u^{-1}dQ(P(\Sigma)dB)Q(\Sigma)dB = 0 \end{aligned}$$

where we have again used $P(\Sigma) dB \otimes Q(\Sigma) dB = 0$. This proves (5.56).

Now suppose that B is a Brownian motion. Since $(b,\beta) = \int_0^{\cdot} u^{-1} dB$ and u is an orthogonal process, it easily follow's using Lévy's criteria that (b,β) is a standard Brownian motion and in particular, β is independent of b. Since (Σ, u) satisfies the coupled pair of stochastic differential equations

$$d\Sigma = u\delta b \text{ and } du + \Gamma(u\delta b)u = 0 \text{ with}$$
$$\Sigma_0 = o \text{ and } u_0 = I \in \mathrm{End}(\mathbb{R}^N),$$

it follows that (Σ, u) is a functional of b and hence the process (Σ, u) are independent of β. ■

5.6. **The Differential in the Starting Point of a Stochastic Flow.** In this section let B_s be an $\mathbb{R}^n$ – valued Brownian motion and for each $m \in M$ let $T_s(m) = \Sigma_s$ where Σ_s is the solution to Eq. (5.1) with $\Sigma_0 = m$. It is well known, see Kunita [116] that there is a version of $T_s(m)$ which is continuous in s and smooth in m, moreover the differential of $T_s(m)$ relative to m solves the stochastic differential equation found by differentiating Eq. (5.1). Let

$$Z_s := T_{s*o} \text{ and } z_s := //_s^{-1} Z_s \in \mathrm{End}(T_oM) \tag{5.57}$$

where $//_s$ is stochastic parallel translation along $\Sigma_s := T_s(o)$.

Theorem 5.41. *For all $v \in T_oM$*

$$\delta_s^{\nabla} Z_s v = (\nabla_{Z_s v}\mathbf{X})\,\delta B_s + (\nabla_{Z_s v} X_0)\, ds \text{ with } Z_0 v = v. \tag{5.58}$$

Alternatively z_s satisfies

$$dz_s v = //_s^{-1}\left(\nabla_{//_s z_s v}\mathbf{X}\right)\delta B_s + //_s^{-1}\left(\nabla_{//_s z_s v}X_0\right)ds. \tag{5.59}$$

Proof. Equations (5.58) and (5.59) are the formal analogues Eqs. (4.2) and (4.3) respectively. Because of Proposition 5.36, Eq. (5.58) is equivalent to Eq. (5.59). To prove Eq. (5.58), differentiate Eq. (5.1) in m in the direction $v \in T_oM$ to find

$$\delta_s Z_s v = DX_i(\Sigma_s) Z_s v \circ \delta B_s^i + DX_0(\Sigma_s) Z_s v ds \text{ with } Z_0 v = v.$$

Multiplying this equation through by $P(\Sigma_s)$ on the left then gives Eq. (5.58). ■

Notation 5.42. The pull back, $\mathrm{Ric}_{//_s}$, of the Ricci tensor by parallel translation is defined by

$$\mathrm{Ric}_{//_s} := //_s^{-1}\mathrm{Ric}_{\Sigma_s} //_s. \tag{5.60}$$

Theorem 5.43 (Itô form of Eq. (5.59))**.** *The Itô form of Eq. (5.59) is*

$$dz_s v = //_s^{-1}\left(\nabla_{//_s z_s v}\mathbf{X}\right)dB_s + \alpha_s ds \tag{5.61}$$

where

$$\alpha_s := //_s^{-1}\left[\nabla_{//_s z_s v}\left(\sum_{i=1}^n \nabla_{X_i}X_i + X_0\right) - \frac{1}{2}\sum_{i=1}^n R^\nabla\left(//_s z_s v, X_i(\Sigma_s)\right)X_i(\Sigma_s)\right]ds. \tag{5.62}$$

If we further assume that $n = N$ and $X_i(m) = P(m)e_i$ (so that Eq. (5.1) is equivalent to Eq. (5.42) if $X_0 \equiv 0$), then $\alpha_s = -\frac{1}{2}\mathrm{Ric}_{//_s} z_s v ds$, i.e. Eq. (5.59) is equivalent to

$$dz_s v = //_s^{-1}P(\Sigma_s)dP(//_s z_s v)dB_s + \left[//_s^{-1}\nabla_{//_s z_s v}X_0 - \frac{1}{2}\mathrm{Ric}_{//_s} z_s v\right]ds. \tag{5.63}$$

Proof. In this proof there will always be an implied sum on repeated indices. Using Proposition 5.36,

$$\begin{aligned} d\left[//_s^{-1}\left(\nabla_{//_s z_s v}\mathbf{X}\right)\right]dB_s &= //_s^{-1}\left[\nabla^2_{\mathbf{X}(\Sigma_s)dB_s \otimes //_s z_s v}\mathbf{X} + \nabla_{//_s dz_s v}\mathbf{X}\right]dB_s \\ &= //_s^{-1}\left[\nabla^2_{\mathbf{X}(\Sigma_s)dB_s \otimes //_s z_s v}\mathbf{X} + \nabla_{\left(\nabla_{//_s z_s v}\mathbf{X}\right)dB_s}\mathbf{X}\right]dB_s \\ &= //_s^{-1}\left[\nabla^2_{X_i(\Sigma_s)\otimes //_s z_s v}X_i + \nabla_{\left(\nabla_{//_s z_s v}X_i\right)}X_i\right]ds. \end{aligned} \tag{5.64}$$

Now by Proposition 3.38,

$$\begin{aligned} \nabla^2_{X_i(\Sigma_s)\otimes //_s z_s v}X_i &= \nabla^2_{//_s z_s v \otimes X_i(\Sigma_s)}X_i ds + R^\nabla\left(X_i(\Sigma_s), //_s z_s v\right)X_i(\Sigma_s) \\ &= \nabla^2_{//_s z_s v \otimes X_i(\Sigma_s)}X_i ds - R^\nabla\left(//_s z_s v, X_i(\Sigma_s)\right)X_i(\Sigma_s) \\ &= \left[\nabla_{//_s z_s v}\nabla_{X_i}X_i - \nabla_{\nabla_{//_s z_s v}X_i}X_i\right] \\ &\quad - R^\nabla\left(//_s z_s v, X_i(\Sigma_s)\right)X_i(\Sigma_s) \end{aligned}$$

which combined with Eq. (5.64) implies

$$d\left[//_s^{-1}\left(\nabla_{//_s z_s v}\mathbf{X}\right)\right]dB_s = //_s^{-1}\left[\nabla_{//_s z_s v}\nabla_{X_i}X_i - R^\nabla\left(//_s z_s v, X_i(\Sigma_s)\right)X_i(\Sigma_s)\right]ds. \tag{5.65}$$

Eq. (5.61) is now a follows directly from this equation and Eq. (5.59).

If we further assume $n = N$, $X_i(m) = P(m)e_i$ and $X_0(m) = 0$, then

$$\left(\nabla_{//_s z_s v}\mathbf{X}\right)dB_s = //_s^{-1}P(\Sigma_s)dP(//_s z_s v)dB_s. \tag{5.66}$$

Moreover, from the definition of the Ricci tensor in Eq. (3.31) and making use of Eq. (3.50) in the proof of Proposition 3.48 we have

$$R^\nabla\left(//_s z_s v, X_i(\Sigma_s)\right)X_i(\Sigma_s) = \mathrm{Ric}_{//_s} //_s z_s v. \tag{5.67}$$

Combining Eqs. (5.66) and (5.67) along with $\nabla_{X_i}X_i = 0$ (from Proposition 3.48) with Eqs. (5.61) and (5.62) implies Eq. (5.63). ■

In the next result, we will filter out the "redundant noise" in Eq. (5.63). This is useful for deducing intrinsic formula from their extrinsic cousins, see, for example, Corollary 6.4 and Theorem 7.39 below.

Theorem 5.44 (Filtering out the Redundant Noise)**.** *Keep the same setup in Theorem 5.43 with $n = N$ and $X_i(m) = P(m)e_i$. Further let $\mathcal{M}$ be the σ – algebra generated by the solution $\Sigma = \{\Sigma_s : s \geq 0\}$. Then there is a version, $\bar{z}_s$, of $\mathbb{E}[z_s|\mathcal{M}]$ such that $s \to \bar{z}_s$ is continuous and $\bar{z}$ satisfies,*

$$\bar{z}_s v = v + \int_0^s \left[//_r^{-1}\left(\nabla_{//_r \bar{z}_r v}X_0\right) - \frac{1}{2}\mathrm{Ric}_{//_r}\bar{z}_r v\right]dr. \tag{5.68}$$

In particular if $X_0 = 0$, then

$$\frac{d}{ds}\bar{z}_s = -\frac{1}{2}\mathrm{Ric}_{//_s}\bar{z}_s \text{ with } \bar{z}_0 = id, \tag{5.69}$$

Proof. In this proof, we let b_s be the martingale part of the anti-development map, $\Psi_s(\Sigma)$, i.e.

$$b_s := \int_0^s //_r^{-1}P(\Sigma_r)\delta B_r = \int_0^s //_r^{-1}P(\Sigma_r)dB_r.$$

Since (Σ_s, u_s) solves the stochastic differential equation,

$$\begin{aligned} \delta\Sigma_s &= u_s \delta b_s + X_0(\Sigma_s)ds \text{ with } \Sigma_0 = o \\ \delta u &= -\Gamma(\delta\Sigma)u = -\Gamma(u\delta b)u \text{ with } u_0 = I \in O(N) \end{aligned}$$

it follows that (Σ, u) may be expressed as a function of the Brownian motion, b. Therefore by the martingale representation property, see Corollary 7.20 below, any measurable function, $f(\Sigma)$, of Σ may be expressed as

$$f(\Sigma) = f_0 + \int_0^1 \langle a_r, db_r\rangle = f_0 + \int_0^1 \langle a_r, //_r^{-1}\left[P(\Sigma_r)dB_r\right]\rangle.$$

Hence, using $PdP = dPQ$, the previous equation and the isometry property of the Itô integral,

$$\mathbb{E}\left\{\int_0^s \left[P\left(\Sigma_r\right) dP\left(//_r z_r v\right) dB_r\right] f\left(\Sigma\right)\right\}$$
$$= \mathbb{E}\left\{\int_0^s \left[dP\left(//_r z_r v\right) Q\left(\Sigma_r\right) dB_r\right] \int_0^1 \left\langle P\left(\Sigma_r\right) //_r a_r, dB_r\right\rangle\right\}$$
$$= \mathbb{E}\left\{\int_0^s \left[dP\left(//_r z_r v\right) Q\left(\Sigma_r\right) P\left(\Sigma_r\right) //_r a_r\right] dr\right\} = 0.$$

This shows that

$$\mathbb{E}\left[\int_0^s P\left(\Sigma_r\right) dP\left(//_r z_r v\right) dB_r | \mathcal{M}\right] = 0$$

and hence taking the conditional expectation, $\mathbb{E}\left[\cdot|\mathcal{M}\right]$, of the integrated version of Eq. (5.63) implies Eq. (5.68). In performing this operation we have used the fact that $(\Sigma, //)$ is $\mathcal{M}$ – measurable and that z_s appears linearly in Eq. (5.63). I have also glossed over the technicality of passing the conditional expectation past the integrals involving a ds term. For this detail and a much more general presentation of these ideas the reader is referred to Elworthy, Li and Le Jan [71]. ■

5.7. **More References.** For more details on the sorts of results in this section, the books by Elworthy [69], Emery [74], and Ikeda and Watanabe [104], Malliavin [132], Stroock [171], and Hsu [101] are highly recommended. The following articles and books are also relevant, [14, 20, 21, 40, 64, 63, 65, 110, 129, 137, 144, 154, 155, 156, 179].

6. Heat Kernel Derivative Formula

In this short section we will illustrate how to derive Bismut type formulas for derivatives of heat kernels. For more details and much more general formula see, for example, Driver and Thalmaier [58], Elworthy, Le Jan and Li [71], Stroock and Turetsky [173, 172] and Hsu [99] and the references therein. Throughout this section Σ_s will be an M – valued semi-martingale, $//_s$ will be stochastic parallel translation along Σ and

$$b_s = \Psi_s\left(\Sigma\right) := \int_0^s //_r^{-1} \delta \Sigma_r.$$

Furthermore, let Q_s denote the unique solution to the differential equation:

$$\frac{dQ_s}{ds} = -\frac{1}{2} Q_s \mathrm{Ric}_{//_s} \text{ with } Q_0 = I. \tag{6.1}$$

See Eq. (5.60) for the definition of $\mathrm{Ric}_{//_s}$.

Lemma 6.1. *Let $f : M \to \mathbb{R}$ be a smooth function, $t > 0$ and for $s \in [0, t]$ let*

$$F(s, m) := (e^{(t-s)\bar{\Delta}/2} f)(m). \tag{6.2}$$

If Σ_s is an M – valued Brownian motion, then the process $s \in [0, t] \to Q_s //_s^{-1} \vec{\nabla} F(s, \Sigma_s)$ is a martingale and

$$d\left[Q_s //_s^{-1} \vec{\nabla} F(s, \Sigma_s)\right] = Q_s //_s^{-1} \nabla_{//_s db_s} \vec{\nabla} F(s, \cdot). \tag{6.3}$$

Proof. Let $W_s := //_s^{-1} \vec{\nabla} F(s, \Sigma_s)$. Then by Proposition 5.36 and Theorem 3.49,

$$dW_s = \left[//_s^{-1} \vec{\nabla} \partial_s F(s, \Sigma_s) + \frac{1}{2} //_s^{-1} \nabla^2_{//_s e_i \otimes //_s e_i} \vec{\nabla} F(s, \cdot)\right] ds$$
$$+ //_s^{-1} \nabla_{//_s e_i} \vec{\nabla} F(s, \cdot) db_s^i$$
$$= \frac{1}{2} //_s^{-1} \left[\nabla^2_{//_s e_i \otimes //_s e_i} \vec{\nabla} F(s, \cdot) - \left(\vec{\nabla} \Delta F(s, \cdot)\right)(\Sigma_s)\right] ds$$
$$+ //_s^{-1} \nabla_{//_s e_i} \vec{\nabla} F(s, \cdot) db_s^i$$
$$= \frac{1}{2} //_s^{-1} \mathrm{Ric}\, \vec{\nabla} F(s, \Sigma_s) ds + //_s^{-1} \nabla_{//_s e_i} \vec{\nabla} F(s, \cdot) db_s^i$$
$$= \frac{1}{2} \mathrm{Ric}_{//_s} W_s ds + //_s^{-1} \nabla_{//_s e_i} \vec{\nabla} F(s, \cdot) db_s^i$$

where $\{e_i\}_{i=1}^d$ is an orthonormal basis for $T_o M$ and there is an implied sum on repeated indices. Hence if Q solves Eq. (6.1), then

$$d\left[Q_s W_s\right] = -\frac{1}{2} Q_s \mathrm{Ric}_{//_s} W_s ds + Q_s \left[\frac{1}{2} \mathrm{Ric}_{//_s} W_s ds + //_s^{-1} \nabla_{//_s e_i} \vec{\nabla} F(s, \cdot) db_s^i\right]$$
$$= Q_s //_s^{-1} \nabla_{//_s e_i} \vec{\nabla} F(s, \cdot) db_s^i$$

which proves Eq. (6.3) and shows that $Q_s W_s$ is a martingale as desired. ■

Theorem 6.2 (Bismut)**.** *Let $f : M \to \mathbb{R}$ be a smooth function and Σ be an M – valued Brownian motion with $\Sigma_0 = o$, then for $0 < t_0 \le t < \infty$,*

$$\vec{\nabla}(e^{t\Delta/2} f)(o) = \frac{1}{t_0} E\left[\left(\int_0^{t_0} Q_r db_r\right) f(\Sigma_t)\right]. \tag{6.4}$$

Proof. The proof given here is modelled on Remark 6 on p. 84 in Bismut [21] and the proof of Theorem 2.1 in Elworthy and Li [72]. Also see Norris [145, 144, 146]. For $(s, m) \in [0, t] \times M$ let F be defined as in Eq. (6.2). We wish to compute the differential of $k_s := \left(\int_0^s Q_r db_r\right) F(s, \Sigma_s)$. By Eq. (5.38),

$d\left[F(s,\Sigma_s)\right] = \langle\vec{\nabla}(F(s,\cdot))(\Sigma_s), //_s db_s\rangle$ and therefore:

$$dk_s = F(s,\Sigma_s)Q_s db_s + \left(\int_0^s Q_r dr_r\right)\langle\vec{\nabla}(F(s,\cdot))(\Sigma_s), //_s db_s\rangle$$
$$+ \sum_{i=1}^d \langle\vec{\nabla}(F(s,\cdot))(\Sigma_s), //_s e_i\rangle Q_s e_i\, ds.$$

From this we conclude that

$$\begin{aligned}\mathbb{E}\left[k_{t_0}\right] &= \mathbb{E}\left[k_0\right] + \mathbb{E}\int_0^{t_0}\sum_{i=1}^d \langle //_s^{-1}\vec{\nabla}(F(s,\cdot))(\Sigma_s), e_i\rangle Q_s e_i\, ds \\ &= \int_0^{t_0}\mathbb{E}\left[Q_s //_s^{-1}\vec{\nabla}(F(s,\cdot))(\Sigma_s)\right] ds \\ &= \int_0^{t_0}\mathbb{E}\left[Q_0 //_0^{-1}\vec{\nabla}(F(0,\cdot))(\Sigma_0)\right] ds = t_0\vec{\nabla}(e^{t\Delta/2}f)(o)\end{aligned}$$

wherein the the third equality we have used (by Lemma 6.1) that $s \to Q_s //_s^{-1}\vec{\nabla}(F(s,\cdot))(\Sigma_s)$ is a martingale. Hence

$$\vec{\nabla}(e^{t\Delta/2}f)(o) = \frac{1}{t_0}\mathbb{E}\left[\left(\int_0^{t_0} Q_s db_s\right)(e^{(t-t_0)\Delta/2}f)(\Sigma_{t_0})\right]$$

from which Eq. (6.4) follows using either the Markov property of Σ_s or the fact that $s \to \left(e^{(t-s)\Delta/2}f\right)(\Sigma_s)$ is a martingale. ■

The following theorem is an non-intrinsic form of Theorem 6.2. In this theorem we will be using the notation introduced before Theorem 5.41. Namely, let $\{X_i\}_{i=0}^n \subset \Gamma(TM)$ be as in Notation 5.4, B_s be an $\mathbb{R}^n$ – valued Brownian motion, and $T_s(m) = \Sigma_s$ where Σ_s is the solution to Eq. (5.1) with $\Sigma_s = m \in M$ and $\beta = B$.

Theorem 6.3 (Elworthy - Li)**.** *Assume that $\mathbf{X}(m) : \mathbb{R}^n \to T_m M$ (recall $\mathbf{X}(m)a := \sum_{i=1}^n X_i(m)a_i$) is surjective for all $m \in M$ and let*

$$\mathbf{X}(m)^{\#} = \left[\mathbf{X}(m)|_{\mathrm{Nul}(\mathbf{X}(m))^{\perp}}\right]^{-1} : T_m M \to \mathbb{R}^n, \tag{6.5}$$

where the orthogonal complement is taken relative to the standard inner product on $\mathbb{R}^n$. (See Lemma 7.38 below for more on $\mathbf{X}(m)^{\#}$.) Then for all $v \in T_o M$, $0 < t_o < t < \infty$ and $f \in C(M)$ we have

$$v\left(e^{tL/2}f\right) = \frac{1}{t_0}\mathbb{E}\left[f(\Sigma_t)\int_0^{t_0}\langle\mathbf{X}(\Sigma_s)^{\#} Z_s v, dB_s\rangle\right] \tag{6.6}$$

where $Z_s = T_{s*o}$ as in Eq. (5.57).

Proof. Let $L = \sum_{i=1}^n X_i^2 + 2X_0$ be the generator of the diffusion, $\{T_s(m)\}_{s\geq 0}$. Since $\mathbf{X}(m) : \mathbb{R}^n \to T_m M$ is surjective for all $m \in M$, L is an elliptic operator on $C^\infty(M)$. So, using results similar to those in Fact 5.32, it makes sense to define $F_s(m) := \left(e^{(t-s)L/2}f\right)(m)$ and $N_s^m = F_s(T_s(m))$. Then

$$\partial_s F_s + \frac{1}{2}LF_s = 0 \text{ with } F_t = f$$

and by Itô's lemma,

$$dN_s^m = d\left[F_s(T_s(m))\right] = \sum_{i=1}^n (X_i F_s)(T_s(m))dB_s^i. \tag{6.7}$$

This shows N_s^m is a martingale for all $m \in M$ and, upon integrating Eq. (6.7) on s, that

$$f(T_t(m)) = e^{tL/2}f(m) + \sum_{i=1}^n \int_0^t (X_i F_s)(T_s(m))dB_s^i.$$

Hence if $a_s \in \mathbb{R}^n$ is a predictable process such that $\mathbb{E}\int_0^t |a_s|^2 ds < \infty$, then by the Itô isometry property,

$$\begin{aligned}\mathbb{E}\left[f(T_t(m))\int_0^t\langle a, dB\rangle\right] &= \int_0^t \mathbb{E}\left[(X_i F_s)(T_s(m))a_i(s)\right] ds \\ &= \int_0^t \mathbb{E}\left[(d_M F_s)(\mathbf{X}(T_s(m))a_s)\right] ds.\end{aligned} \tag{6.8}$$

Suppose that $\ell_s \in \mathbb{R}$ is a continuous piecewise differentiable function and let $a_s := \ell_s'\mathbf{X}(\Sigma_s)^{\#} Z_s v$. Then form Eq. (6.8) we have

$$\mathbb{E}\left[f(\Sigma_t)\int_0^t\langle\ell_s'\mathbf{X}(\Sigma_s)^{\#} Z_s v, dB_s\rangle\right] = \int_0^t \ell_s'\mathbb{E}\left[(d_M F_s)(Z_s v)\right] ds. \tag{6.9}$$

Since $N_s^m = F_s(T_s(m))$ is a martingale for all m, we may deduce that

$$v(m \to N_s^m) = d_M F_s(T_{s*o}v) = d_M F_s(Z_s v) \tag{6.10}$$

is a martingale as well for any $v \in T_o M$. In particular, $s \in [0,t] \to \mathbb{E}\left[(d_M F_s)(Z_s v)\right]$ is constant and evaluating this expression at $s = 0$ and $s = t$ implies

$$\mathbb{E}\left[(d_M F_s)(Z_s v)\right] = v\left(e^{tL/2}f\right) = \mathbb{E}\left[(d_M f)(Z_t v)\right]. \tag{6.11}$$

Using Eq. (6.11) in Eq. (6.9) then shows

$$\mathbb{E}\left[f(\Sigma_t)\int_0^t\langle\ell_s'\mathbf{X}(\Sigma_s)^{\#} Z_s v, dB_s\rangle\right] = (\ell_t - \ell_0)\,v\left(e^{tL/2}f\right)$$

which, by taking $\ell_s = s \wedge t_0$, implies Eq. (6.6). ■

Corollary 6.4. *Theorem 6.3 may be used to deduce Theorem 6.2.*

Proof. Apply Theorem 6.3 with $n=N$, $X_0\equiv 0$ and $X_i(m)=P(m)e_i$ for $i=1,\dots,N$ to learn

(6.12)
$$v\left(e^{t\Delta/2}f\right)=\frac{1}{t_0}\mathbb{E}\left[f(\Sigma_t)\int_0^{t_0}\langle Z_s v,dB_s\rangle\right]=\frac{1}{t_0}\mathbb{E}\left[f(\Sigma_t)\int_0^{t_0}\langle //_s z_s v,dB_s\rangle\right]$$

where we have used $L=\Delta$ (see Proposition 3.48) and $\mathbf{X}(m)^{\#}=P(m)$ in this setting. By Theorem 5.40,

$$\int_0^{t_0}\langle //_s z_s v,dB_s\rangle=\int_0^{t_0}\langle //_s z_s v,P(\Sigma_s)dB_s\rangle$$
$$=\int_0^{t_0}\langle z_s v,//_s^{-1}P(\Sigma_s)dB_s\rangle=\int_0^{t_0}\langle z_s v,db_s\rangle$$

and therefore Eq. (6.12) may be written as

$$v\left(e^{t\Delta/2}f\right)=\frac{1}{t_0}\mathbb{E}\left[f(\Sigma_t)\int_0^{t_0}\langle z_s v,db_s\rangle\right].$$

Using Theorem 5.44 to factor out the redundant noise, this may also be expressed as

$$v\left(e^{t\Delta/2}f\right)=\frac{1}{t_0}\mathbb{E}\left[f(\Sigma_t)\int_0^{t_0}\langle \bar z_s v,db_s\rangle\right]=\frac{1}{t_0}\mathbb{E}\left[f(\Sigma_t)\int_0^{t_0}\langle v,\bar z_s^{\mathrm{tr}}db_s\rangle\right] \tag{6.13}$$

where $\bar z_s$ solves Eq. (5.69). By taking transposes of Eq. (5.69) it follows that $\bar z_s^{\mathrm{tr}}$ satisfies Eq. (6.1) and hence $\bar z_s^{\mathrm{tr}}=Q_s$. Since $v\in T_oM$ was arbitrary, Equation (6.4) is now an easy consequence of Eq. (6.13) and the definition of $\vec\nabla(e^{t\Delta/2}f)(o)$. ■

7. Calculus on $W(M)$

In this section, (M,o) is assumed to be either a compact Riemannian manifold equipped with a fixed point $o\in M$ or $M=\mathbb{R}^d$ with $o=0$.

Notation 7.1. We will be interested in the following path spaces:

$$W(T_oM):=\{\omega\in C([0,1]\to T_oM)|\omega(0)=0_o\in T_oM\},$$
$$H(T_oM):=\{h\in W(T_oM):h(0)=0,\ \&\ \langle h,h\rangle_H:=\int_0^1|h'(s)|^2_{T_oM}ds<\infty\}$$

and

$$W(M):=\{\sigma\in C([0,1]\to M):\sigma(0)=0\in M\}.$$

(By convention $\langle h,h\rangle_H=\infty$ if $h\in W(T_oM)$ is not absolutely continuous.) We refer to $W(T_oM)$ as **Wiener space**, $W(M)$ as **curved Wiener space** and $H(T_oM)$ or $H(\mathbb{R}^d)$ as the **Cameron-Martin Hilbert space**.

Definition 7.2. Let μ and $\mu_{W(M)}$ denote the Wiener measures on $W(T_oM)$ and $W(M)$ respectively, i.e. $\mu=\mathrm{Law}(b)$ and $\mu_{W(M)}=\mathrm{Law}(\Sigma)$ where b and Σ are Brownian motions on T_oM and M starting at $0\in T_oM$ and $o\in M$ respectively.

Notation 7.3. The probability space in this section will often be $\left(W(M),\mathcal{F},\mu_{W(M)}\right)$, where $\mathcal{F}$ is the completion of the $\sigma-$algebra generated by the projection maps, $\Sigma_s:W(M)\to M$ defined by $\Sigma_s(\sigma)=\sigma_s$ for $s\in[0,1]$. We make this into a filtered probability space by taking $\mathcal{F}_s$ to be the $\sigma-$algebra generated by $\{\Sigma_r:r\le s\}$ and the null sets in $\mathcal{F}_s$. Also let $//_s$ be stochastic parallel translation along Σ.

Definition 7.4. A function $F:W(M)\to\mathbb{R}$ is called a C^k- **cylinder function** if there exists a partition

$$\pi:=\{0=s_0<s_1<s_2\cdots<s_n=1\} \tag{7.1}$$

of $[0,1]$ and $f\in C^k(M^n)$ such that

$$F(\sigma)=f(\sigma_{s_1},\dots,\sigma_{s_n})\text{ for all }\sigma\in W(M). \tag{7.2}$$

If $M=\mathbb{R}^d$, we further require that f and all of its derivatives up to order k have at most polynomial growth at infinity. The collection of C^k-cylinder functions will be denoted by $\mathcal{F}C^k(W(M))$.

Definition 7.5. The **continuous tangent space** to $W(M)$ at $\sigma\in W(M)$ is the set $CT_\sigma W(M)$ of continuous vector-fields along σ which are zero at $s=0$:

(7.3)
$$CT_\sigma W(M)=\{X\in C([0,1],TM)|X_s\in T_{\sigma_s}M\ \forall\ s\in[0,1]\text{ and }X(0)=0\}.$$

To motivate the above definition, consider a differentiable path in $\gamma\in W(M)$ going through σ at $t=0$. Writing $\gamma(t)(s)$ as $\gamma(t,s)$, the derivative $X_s:=\frac{d}{dt}|_0\gamma(t,s)\in T_{\sigma(s)}M$ of such a path should, by definition, be a tangent vector to $W(M)$ at σ.

We now wish to define a "Riemannian metric" on $W(M)$. It turns out that the continuous tangent space $CT_\sigma W(M)$ is too large for our purposes, see for example the Cameron-Martin Theorem 7.13 below. To remedy this we will introduce a Riemannian structure on a an a.e. defined "sub-bundle" of $CTW(M)$.

Definition 7.6. A **Cameron-Martin process**, h, is a T_oM-valued process on $W(M)$ such that $s\to h(s)$ is in H, $\mu_{W(M)}-$a.e. Contrary to our earlier assumptions, we do **not** assume that h is adapted unless explicitly stated.

Definition 7.7. Suppose that X is a $TM-$valued process on $\left(W(M),\mu_{W(M)}\right)$ such that the process $\pi(X_s)=\Sigma_s\in M$. We will say X is a **Cameron-Martin vector-field** if

$$h_s:=//_s^{-1}X_s \tag{7.4}$$

is a Cameron-Martin valued process and

$$\langle X,X\rangle_{\mathcal{X}}:=\mathbb{E}[\langle h,h\rangle_H]<\infty. \tag{7.5}$$

A Cameron-Martin vector field X is said to be adapted if $h:=//^{-1}X$ is adapted. The set of Cameron-Martin vector-fields will be denoted by $\mathcal{X}$ and those which are adapted will be denoted by $\mathcal{X}_a$.

Remark 7.8. Notice that $\mathcal{X}$ is a Hilbert space with the inner product determined by $\langle\cdot,\cdot\rangle_{\mathcal{X}}$ in (7.5). Furthermore, $\mathcal{X}_a$ is a Hilbert-subspace of $\mathcal{X}$.

Notation 7.9. Given a Cameron-Martin process h, let $X^h:=//h$. In this way we may identify Cameron-Martin processes with Cameron-Martin vector fields.

We define a "metric", G,[6] on $\mathcal{X}$ by

$$G(X^h,X^h)=\langle h,h\rangle_H. \tag{7.6}$$

With this notation we have $\langle X,X\rangle_{\mathcal{X}}=\mathbb{E}\left[G(X,X)\right].$

Remark 7.10. Notice, if σ is a smooth path then the expression in (7.6) could be written as

$$G(X,X)=\int_0^1 g\left(\frac{\nabla}{ds}X(s),\frac{\nabla}{ds}X(s)\right)ds,$$

where $\frac{\nabla}{ds}$ denotes the covariant derivative along the path σ which is induced from the covariant derivative ∇. This is a typical metric used by differential geometers on path and loop spaces.

Notation 7.11. Given a Cameron-Martin vector field X on $\left(W(M),\mu_{W(M)}\right)$ and a cylinder function $F\in\mathcal{F}C^1(W(M))$ as in Eq. (7.2), let XF denote the random variable

$$XF(\sigma):=\sum_{i=1}^n(\operatorname{grad}_iF(\sigma),X_{s_i}(\sigma)), \tag{7.7}$$

where

$$\operatorname{grad}_iF(\sigma):=(\operatorname{grad}_if)(\sigma_{s_1},\dots,\sigma_{s_n}) \tag{7.8}$$

and $(\operatorname{grad}_i f)$ denotes the gradient of f relative to the i^{th} variable.

Notation 7.12. The **gradient**, DF, of a smooth cylinder functions, F, on $W(M)$ is the unique Cameron-Martin process such that $G(DF,X)=XF$ for all $X\in\mathcal{X}$. The explicit formula for D, as the reader should verify, is

$$(DF)_s=//_s\left(\sum_{i=1}^n s\wedge s_i//_{s_i}{}^{-1}\operatorname{grad}_iF(\sigma)\right). \tag{7.9}$$

The formula in Eq. (7.9) defines a densely defined operator, $D:L^2(\mu)\to\mathcal{X}$ with $\mathcal{D}(D)=\mathcal{F}C^1(W(M))$ as its domain.

[6] The function G is to be loosely interpreted as a Riemannian metric on $W(M)$.

7.1. **Classical Wiener Space Calculus.** In this subsection (which is a warm up for the sequel) we will specialize to the case where $M=\mathbb{R}^d$, $o=0\in\mathbb{R}^d$. To simplify notation let $W:=W(\mathbb{R}^d)$, $H:=H\left(\mathbb{R}^d\right)$, $\mu=\mu_{W(\mathbb{R}^d)}$, $b_s(\omega)=\omega_s$ for all $s\in[0,1]$ and $\omega\in W$. Recall that $\{\mathcal{F}_s:s\in[0,1]\}$ is the filtration on W as explained in Notation 7.3 where we are now writing b for Σ. Cameron and Martin [25, 26, 28, 27] and Cameron [28] began the study of calculus on this classical Wiener space. They proved the following two results, see Theorem 2, p. 387 of [26] and Theorem II, p. 919 of [28] respectively. (There have been many extensions of these results partly initiated by Gross' work in [90, 91].)

Theorem 7.13 (Cameron & Martin 1944)**.** *Let $(W,\mathcal{F},\mu)$ be the classical Wiener space described above and for $h\in W$, define $T_h:W\to W$ by $T_h(\omega)=\omega+h$ for all $\omega\in W$. If h is C^1, then μT_h^{-1} is absolutely continuous relative to μ.*

This theorem was extended by Maruyama [133] and Girsanov [88] to allow the same conclusion for $h\in H$ and more general Cameron-Martin processes. Moreover it is now well known $\mu T_h^{-1}\ll\mu$ iff $h\in H$. From the Cameron and Martin theorem one may prove Cameron's integration by parts formula.

Theorem 7.14 (Cameron 1951)**.** *Let $h\in H$ and $F,G\in L^{\infty-}(\mu):=\cap_{1\le p<\infty}L^p(\mu)$ such that $\partial_hF:=\frac{d}{d\varepsilon}F\circ T_{\varepsilon h}|_{\varepsilon=0}$ and $\partial_hG:=\frac{d}{d\varepsilon}G\circ T_{\varepsilon h}|_{\varepsilon=0}$ where the derivatives are supposed to exist*[7] *in $L^p(\mu)$ for all $1\le p<\infty$. Then*

$$\int_W\partial_hF\cdot G\,d\mu=\int_W F\partial_h^*G\,d\mu,$$

where $\partial_h^*G=-\partial_hG+z_hG$ and $z_h:=\int_0^1\langle h'(s),db_s\rangle_{\mathbb{R}^d}$.

In this flat setting parallel translation is trivial, i.e. $//_s=id$ for all s. Hence the gradient operator D in Eq. (7.9) reduces to the equation,

$$(DF)_s(\omega)=\left(\sum_{i=1}^n s\wedge s_i\operatorname{grad}_iF(\omega_s)\right).$$

Similarly the association of a Cameron-Martin vector field X on $W(\mathbb{R}^d)$ with a Cameron-Martin valued process h in Eq. (7.4) is simply that $X=h$.

We will now recall that adapted Cameron-Martin vector fields, $X=h$, are in the domain of D^*. From this fact it will easily follow that D^* is densely defined.

Theorem 7.15. *Let h be an adapted Cameron-Martin process (vector field) on W. Then $h\in\mathcal{D}(D^*)$ and*

$$D^*h=\int_0^1\langle h',db\rangle.$$

[7] The notion of derivative stated here is weaker than the notion given in [28]. Nevertheless Cameron's proof covers this case without any essential change.

Proof. We start by proving the theorem under the additional assumption that

$$\sup_{s\in[0,1]}|h_s'|\le C, \tag{7.10}$$

where C is a non-random constant. For each $t\in\mathbb{R}$ let $b(t,s)=b_s(t)=b_s+th_s$. By Girsanov's theorem, $s\to b_s(t)$ (for fixed t) is a Brownian motion relative to $Z_t\cdot\mu$, where

$$Z_t:=\exp\left(-\int_0^1 t\langle h_s',db_s\rangle-\frac{1}{2}t^2\int_0^1\langle h_s',h_s'\rangle ds\right).$$

Hence if F is a smooth cylinder function on W,

$$\mathbb{E}\left[F\left(b(t,\cdot)\right)\cdot Z_t\right]=\mathbb{E}\left[F(b)\right].$$

Differentiating this equation in t at $t=0$, using

$$\langle DF,h\rangle_H=\frac{d}{dt}|_0F\left(b\left(t,\cdot\right)\right)\text{ and }\frac{d}{dt}|_0Z_t=-\int_0^1\langle h',db\rangle,$$

shows

$$\mathbb{E}\left[\langle DF,h\rangle_H\right]-\mathbb{E}\left[F\int_0^1\langle h',db\rangle\right]=0.$$

From this equation it follows that $h\in\mathcal{D}(D^*)$ and $D^*h=\int_0^1\langle h',db\rangle$. So it now only remains to remove the restriction placed on h in Eq. (7.10).

Let h be a general adapted Cameron-Martin vector-field and for each $n\in\mathbb{N}$, let

$$h_n(s):=\int_0^s h'(r)\cdot 1_{|h'(r)|\le n}dr. \tag{7.11}$$

(Notice that h_n is still adapted.) By the special case above we know that $h_n\in\mathcal{D}(D^*)$ and $D^*h_n=\int_0^1\langle h_n',db\rangle$. Therefore,

$$\mathbb{E}\left|D^*(h_m-h_n)\right|^2=\mathbb{E}\int_0^1|h_m'-h_n'|^2ds\to 0\text{ as }m,n\to\infty$$

from which it follows that D^*h_n is convergent. Because D^* is a closed operator, $h\in\mathcal{D}(D^*)$ and

$$D^*h=\lim_{n\to\infty}D^*h_n=\lim_{n\to\infty}\int_0^1\langle h_n',db\rangle=\int_0^1\langle h',db\rangle.$$

■

Corollary 7.16. *The operator D^* is densely defined and hence D is closable. (Let $\bar{D}$ denote the closure of D.)*

Proof. Let $h\in H$ and F and K be smooth cylinder functions. Then, by the product rule,

$$\begin{aligned}\langle DF,Kh\rangle_{\mathcal{X}}&=\mathbb{E}[\langle KDF,h\rangle_H]=\mathbb{E}[\langle D\left(KF\right)-FDK,h\rangle_H]\\&=\mathbb{E}[F\cdot KD^*h-F\langle DK,h\rangle_H].\end{aligned}$$

Therefore $Kh\in\mathcal{D}(D^*)$ ($\mathcal{D}(D^*)$ is the domain of D^*) and

$$D^*(Kh)=KD^*h-\langle DK,h\rangle_H.$$

Since the subspace,

$$\{Kh|h\in H\text{ and }K\text{ is a smooth cylinder function}\},$$

is a dense subspace of $\mathcal{X}$, D^* is densely defined. ■

7.1.1. *Martingale Representation Property and the Clark-Ocone Formula.*

Lemma 7.17. *Let $F(b)=f(b_{s_1},\dots,b_{s_n})$ be the smooth cylinder function on W as in Definition 7.4, then*

$$F=\mathbb{E}F+\int_0^1\langle a_s,db_s\rangle, \tag{7.12}$$

where a_s is a bounded, piecewise-continuous (in s) and predictable process. Furthermore, the jumps points of a_s are contained in the set $\{s_1,\dots,s_n\}$ and $a_s\equiv 0$ is $s\ge s_n$.

Proof. The proof will be by induction on n. First assume that $n=1$, so that $F(b)=f(b_t)$ for some $0<t\le 1$. Let $H(s,m):=(e^{(t-s)\Delta/2}f)(m)$ for $0\le s\le t$ and $m\in\mathbb{R}^d$. Then, by Itô's formula (or see Eq. (5.38)),

$$dH(s,b_s)=\langle\operatorname{grad}H(s,b_s),db_s\rangle$$

which upon integrating on $s\in[0,t]$ gives

$$F(b)=(e^{t\Delta/2}f)(o)+\int_0^t\langle\operatorname{grad}H(s,b_s),db_s\rangle=\mathbb{E}F+\int_0^1\langle a_s,db_s\rangle,$$

where $a_s=1_{s\le t}//_s^{-1}\operatorname{grad}H(s,b_s)$. This proves the $n=1$ case. To finish the proof it suffices to show that we may reduce the assertion of the lemma at the level n to the assertion at the level $n-1$.

Let $F(b)=f(b_{s_1},\dots,b_{s_n})$,

$$\begin{aligned}(\Delta_nf)(x_1,x_2,\dots,x_n)&=(\Delta g)(x_n)\text{ and}\\(\operatorname{grad}_n f)(x_1,x_2,\dots,x_n)&=\vec{\nabla}g\left(x_n\right)\end{aligned}$$

where $g(x):=f(x_1,x_2,\dots,x_{n-1},x)$. (So Δ_nf and $\operatorname{grad}_n f$ is the Laplacian and the gradient of f in the n^{th} – variable.) Itô's lemma applied to the process,

$$s\in[s_{n-1},s_n]\to H(s,b):=(e^{(s_n-s)\Delta_n/2}f)(b_{s_1},\dots,b_{s_{n-1}},b_s)$$

gives

$$dH(s,b) = \langle \mathrm{grad}_n e^{(s_n - s)\Delta_n/2} f)(b_{s_1}, \dots, b_{s_{n-1}}, b_s, db_s\rangle$$

and hence

$$\begin{aligned} F(b) &= (e^{(s_n - s_{n-1})\Delta_n/2} f)(b_{s_1}, \dots, b_{s_{n-1}}, b_{s_{n-1}}) \\ &\quad + \int_{s_{n-1}}^{s_n} \langle \mathrm{grad}_n e^{(s_n - s)\Delta_n/2} f)(b_{s_1}, \dots, b_{s_{n-1}}, b_s, db_s\rangle \\ &= (e^{(s_n - s_{n-1})\Delta_n/2} f)(b_{s_1}, \dots, b_{s_{n-1}}, b_{s_{n-1}}) + \int_{s_{n-1}}^{s_n} \langle \alpha_s, db_s\rangle, \end{aligned} \tag{7.13}$$

where $\alpha_s := (\mathrm{grad}_n\, e^{(s_n - s)\Delta_n/2} f)(b_{s_1}, \dots, b_{s_{n-1}}, b_s)$ for $s \in (s_{n-1}, s_n)$. By induction we know that the smooth cylinder function

$$(e^{(s_n - s_{n-1})\Delta_n/2} f)(b_{s_1}, \dots, b_{s_{n-1}}, b_{s_{n-1}})$$

may be written as a constant plus $\int_0^1 \langle a_s, db_s\rangle$, where a_s is bounded and piecewise continuous and $a_s \equiv 0$ if $s \ge s_{n-1}$. Hence it follows by replacing a_s by $a_s + 1_{(s_{n-1}, s_n)s}\alpha_s$ that

$$F(b) = C + \int_0^{s_n} \langle a_s, db_s\rangle$$

for some constant C. Taking expectations of both sides of this equation then shows $C = \mathbb{E}[F(b)]$. ■

Remark 7.18. By being more careful in the proof of the Lemma 7.17 (as is done in more generality later in Theorem 7.47) it is possible to show a_s in Eq. (7.12) may be written as

$$a_s = \mathbb{E}\left[\sum_{i=1}^n 1_{s \le s_i} \mathrm{grad}_i f(b_{s_1}, \dots, b_{s_n}) \middle| \mathcal{F}_s \right]. \tag{7.14}$$

This will also be explained, by indirect means, in Theorem 7.21 below.

Corollary 7.19. *Let F be a smooth cylinder function on W, then there is a predictable, piecewise continuously differentiable Cameron-Martin process h such that $F = \mathbb{E}F + D^*h$.*

Proof. Let $h_s := \int_0^s a_r dr$ where a is the process as in Lemma 7.17. ■

Corollary 7.20 (Martingale Representation Property)**.** *Let $F \in L^2(\mu)$, then there is a predictable process, a_s, such that $\mathbb{E}\int_0^1 |a_s|^2 ds < \infty$, and*

$$F = \mathbb{E}F + \int_0^1 \langle a, db\rangle. \tag{7.15}$$

Proof. Choose a sequence of smooth cylinder functions $\{F_n\}$ such that $F_n \to F$ as $n \to \infty$. By replacing F by $F - \mathbb{E}F$ and F_n by $F_n - \mathbb{E}F_n$, we may assume that $\mathbb{E}F = 0$ and $\mathbb{E}F_n = 0$. Let a^n be predictable processes such that $F_n = \int_0^1 \langle a^n, db\rangle$ for all n. Notice that

$$\mathbb{E}\int_0^1 |a_s^n - a_s^m|^2 ds = \mathbb{E}(F_n - F_m)^2 \to 0 \text{ as } m, n \to \infty.$$

Hence, if $a := L^2(ds \times d\mu) - \lim_{n\to\infty} a^n$, then

$$F_n = \int_0^1 a^n \cdot db \to \int_0^1 \langle a, db\rangle \text{ as } n \to \infty.$$

This show that $F = \int_0^1 \langle a, db\rangle$. ■

Theorem 7.21 (Clark – Ocone Formula)**.** *Suppose that $F \in \mathcal{D}(\bar{D})$, then*[8]

$$F = \mathbb{E}F + \int_0^1 \left\langle \mathbb{E}\left[\frac{d}{ds}\left(\bar{D}F\right)_s(b) \middle| \mathcal{F}_s \right], db_s \right\rangle. \tag{7.16}$$

In particular if $F = f(b_{s_1}, \dots, b_{s_n})$ is a smooth cylinder function on $W(M)$ then

$$F = \mathbb{E}F + \int_0^1 \left\langle \mathbb{E}\left[\sum_{i=1}^n 1_{s \le s_i} \mathrm{grad}_i f(b_{s_1}, \dots, b_{s_n}) \middle| \mathcal{F}_s \right], db_s \right\rangle. \tag{7.17}$$

Proof. Let h be a predictable Cameron-Martin valued process such that $\mathbb{E}\int_0^1 |h_s'|^2 ds < \infty$. Then using Theorem 7.15 and the Itô isometry property,

$$\begin{aligned} \mathbb{E}\langle \bar{D}F, h\rangle_H &= \mathbb{E}[FD^*h] = \mathbb{E}\left[F \int_0^1 \langle h_s', db_s\rangle \right] \\ &= \mathbb{E}\left[\left(\mathbb{E}F + \int_0^1 \langle a, db\rangle \right) \int_0^1 \langle h_s', db_s\rangle \right] = \mathbb{E}\left[\int_0^1 \langle a_s, h_s'\rangle ds \right] \end{aligned} \tag{7.18}$$

where a is the predictable process in Corollary 7.20. Since h is predictable,

$$\begin{aligned} \mathbb{E}\langle \bar{D}F, h\rangle_H &= \mathbb{E}\left[\int_0^1 \left\langle \frac{d}{ds}\left(\bar{D}F\right)_s, h_s' \right\rangle ds \right] \\ &= \mathbb{E}\left[\int_0^1 \left\langle \mathbb{E}\left[\frac{d}{ds}\left(\bar{D}F\right)_s \middle| \mathcal{F}_s \right], h_s' \right\rangle ds \right]. \end{aligned} \tag{7.19}$$

Since h is an arbitrary predictable Cameron-Martin valued process, comparing Eqs. (7.18) and (7.18) shows

$$a_s = \mathbb{E}\left[\frac{d}{ds}\left(\bar{D}F\right)_s \middle| \mathcal{F}_s \right]$$

which combined with Eq. (7.12) completes the proof. ■

[8]Here we are abusing notation and writing $\mathbb{E}\left[\frac{d}{ds}\bar{D}F_s(b) \middle| \mathcal{F}_s\right]$ for the "predictable" projection of the process $s \to \frac{d}{ds}\bar{D}F_s(b)$. Since we will only really use Eq. (7.17) in these notes, this technicality need not concern us here.

Remark 7.22. As mentioned in Remark 7.18 it is possible to prove Eq. (7.17) by an inductive procedure. On the other hand if we were to know that Eq. (7.17) was valid for all $F \in \mathcal{F}C^1(W)$, then for $h \in \mathcal{X}_a$,

$$\begin{aligned}\mathbb{E}\left[F\int_0^1 \langle h_s', db_s\rangle\right] &= \mathbb{E}\left[\left(\mathbb{E}F + \int_0^1 \left\langle \mathbb{E}\left[\frac{d}{ds}\, DF_s|\, \mathcal{F}_s\right], db_s\right\rangle\right)\int_0^1 \langle h_s', db_s\rangle\right] \\ &= \mathbb{E}\left[\int_0^1 \left\langle \mathbb{E}\left[\frac{d}{ds}\, DF_s|\, \mathcal{F}_s\right], h_s'\right\rangle ds\right] \\ &= \mathbb{E}\left[\int_0^1 \left\langle \frac{d}{ds}DF_s, h_s'\right\rangle ds\right] = \langle DF, h\rangle_{\mathcal{X}}.\end{aligned}$$

This identity shows $h \in \mathcal{D}(D^*)$ and that $D^*h = \int_0^1 \langle h_s', db_s\rangle$, i.e. we have recovered Theorem 7.15. In this way we see that the Clark-Ocone formula may be used to recover integration by parts on Wiener space.

Let $\mathcal{L}$ be the infinite dimensional Ornstein-Uhlenbeck operator defined as the self-adjoint operator on $L^2(\mu)$ given by $\mathcal{L} = D^*\bar{D}$. The following spectral gap inequality for $\mathcal{L}$ has been known since the early days of quantum mechanics. This is because $\mathcal{L}$ is unitarily equivalent to a "harmonic oscillator Hamiltonian" for which the full spectrum may be found, see for example [162]. However, these explicit computations will not in general be available when we consider analogous spectral gap inequalities when $\mathbb{R}^d$ is replaced by a general compact Riemannian manifold M.

Theorem 7.23 (Ornstein Uhlenbeck Spectral Gap Inequality)**.** *The null-space of $\mathcal{L}$ consists of the constant functions on W and $\mathcal{L}$ has a spectral gap of size* 1, *i.e.*

$$\langle \mathcal{L}F, F\rangle_{L^2(\mu)} \ge \langle F, F\rangle_{L^2(\mu)} \tag{7.20}$$

for all $F \in \mathcal{D}(\mathcal{L})$ such that $F \in \mathrm{Nul}(\mathcal{L})^\perp = \{1\}^\perp$.

Proof. Let $F \in \mathcal{D}(\bar{D})$, then by the Clark-Ocone formula in Eq. (7.16), the isometry property of the Itô integral and the contractive properties of conditional expectation,

$$\begin{aligned}\mathbb{E}(F - \mathbb{E}F)^2 &= \mathbb{E}\left[\int_0^1 \left\langle \mathbb{E}\left[\frac{d}{ds}\, \bar{D}F_s(b)\middle|\, \mathcal{F}_s\right], db_s\right\rangle\right]^2 \\ &= \mathbb{E}\left[\int_0^1 \left|\mathbb{E}\left[\frac{d}{ds}\, \bar{D}F_s(b)\middle|\, \mathcal{F}_s\right]\right|^2 ds\right] \\ &\le \mathbb{E}\left[\int_0^1 \left(\mathbb{E}\left[\left|\frac{d}{ds}\bar{D}F_s(b)\right| |\mathcal{F}_s\right]\right)^2 ds\right] \\ &\le \mathbb{E}\left[\int_0^1 \mathbb{E}\left[\left|\frac{d}{ds}\bar{D}F_s(b)\right|^2 |\mathcal{F}_s\right] ds\right] \\ &= \mathbb{E}\left[\int_0^1 \left|\frac{d}{ds}\bar{D}F_s(b)\right|^2 ds\right] = \langle \bar{D}F, \bar{D}F\rangle_{\mathcal{X}}.\end{aligned}$$

In particular if $F \in \mathcal{D}(\mathcal{L})$, then $\langle \bar{D}F, \bar{D}F\rangle_{\mathcal{X}} = \mathbb{E}[\mathcal{L}F \cdot F]$, and hence

$$\langle \mathcal{L}F, F\rangle_{L^2(\mu)} \ge \langle F - \mathbb{E}F, F - \mathbb{E}F\rangle_{L^2(\mu)}. \tag{7.21}$$

Therefore, if $F \in \mathrm{Nul}(\mathcal{L})$, it follows that $F = \mathbb{E}F$, i.e. F is a constant. Moreover if $F \perp 1$ (i.e. $\mathbb{E}F = 0$) then Eq. (7.20) becomes Eq. (7.21). ■

It turns out that using a method which is attributed to Maurey and Neveu in [29], it is possible to use the Clark-Ocone formula as the starting point for a proof of Gross' logarithmic Sobolev inequality which by general theory is known to be stronger than the spectral gap inequality in Theorem 7.23.

Theorem 7.24 (Gross' Logarithmic Sobolev Inequality for $W(\mathbb{R}^d)$)**.** *For all $F \in \mathcal{D}(\bar{D})$,*

$$\mathbb{E}\left[F^2 \log F^2\right] \le 2\mathbb{E}\left[\langle DF, DF\rangle_H\right] + \mathbb{E}F^2 \cdot \log \mathbb{E}F^2. \tag{7.22}$$

Proof. Let $F \in \mathcal{F}C^1(W)$, $\varepsilon > 0$, $H_\varepsilon := F^2 + \varepsilon \in \mathcal{D}(\bar{D})$ and $a_s = \mathbb{E}\left[\frac{d}{ds}(DH_\varepsilon)_s |\mathcal{F}_s\right]$. By Theorem 7.21,

$$H_\varepsilon = \mathbb{E}H_\varepsilon + \int_0^1 \langle a, db\rangle$$

and hence

$$M_s := \mathbb{E}[H_\varepsilon|\mathcal{F}_s] = \mathbb{E}\left[F^2 + \varepsilon|\mathcal{F}_s\right] \ge \varepsilon$$

is a positive martingale which may be written as

$$M_s := M_0 + \int_0^s \langle a, db\rangle$$

where $M_0 = \mathbb{E}H_\varepsilon$.

Let $\phi(x) = x\ln x$ so that $\phi'(x) = \ln x + 1$ and $\phi''(x) = x^{-1}$. Then by Itô's formula,

$$\begin{aligned} d[\phi(M_s)] &= \phi(M_0) + \phi'(M_s)\,dM_s + \frac{1}{2}\phi''(M_s)\,|a_s|^2\,ds \\ &= \phi(M_0) + \phi'(M_s)\,dM_s + \frac{1}{2}\frac{1}{M_s}\,|a_s|^2\,ds. \end{aligned}$$

Integrating this equation on s and then taking expectations shows

$$\mathbb{E}[\phi(M_1)] = \phi(\mathbb{E}M_1) + \frac{1}{2}\mathbb{E}\left[\int_0^1 \frac{1}{M_s}\,|a_s|^2\,ds\right]. \tag{7.23}$$

Since $\bar{D}H_\varepsilon = 2F\bar{D}F$, Eq. (7.23) is equivalent to

$$\mathbb{E}[\phi(H_\varepsilon)] = \phi(\mathbb{E}H_\varepsilon) + \frac{1}{2}\mathbb{E}\left[\int_0^1 \frac{1}{\mathbb{E}[H_\varepsilon|\mathcal{F}_s]}\left|\mathbb{E}\left[2F\left(\bar{D}F\right)_s'|\mathcal{F}_s\right]\right|^2 ds\right].$$

Using the Cauchy-Schwarz inequality and the contractive properties of conditional expectations,

$$\begin{aligned} \left|\mathbb{E}\left[2F\frac{d}{ds}\left(\bar{D}F\right)_s|\mathcal{F}_s\right]\right|^2 &\le 4\left(\mathbb{E}\left[F\left|\frac{d}{ds}\left(\bar{D}F\right)_s\right||\mathcal{F}_s\right]\right)^2 \\ &\le 4\mathbb{E}\left[F^2|\mathcal{F}_s\right]\cdot\mathbb{E}\left[\left|\frac{d}{ds}\left(\bar{D}F\right)_s\right|^2|\mathcal{F}_s\right]. \end{aligned}$$

Combining the last two equations, using

$$\frac{\mathbb{E}\left[F^2|\mathcal{F}_s\right]}{\mathbb{E}[H_\varepsilon|\mathcal{F}_s]} = \frac{\mathbb{E}\left[F^2|\mathcal{F}_s\right]}{\mathbb{E}[F^2|\mathcal{F}_s] + \varepsilon} \le 1 \tag{7.24}$$

gives,

$$\begin{aligned} \mathbb{E}[\phi(H_\varepsilon)] &\le \phi(\mathbb{E}H_\varepsilon) + 2\mathbb{E}\int_0^1 \mathbb{E}\left[\left|\frac{d}{ds}\left(\bar{D}F\right)_s\right|^2|\mathcal{F}_s\right]ds \\ &= \phi(\mathbb{E}H_\varepsilon) + 2\mathbb{E}\int_0^1 \left|\frac{d}{ds}\left(\bar{D}F\right)_s\right|^2 ds. \end{aligned}$$

We may now let $\varepsilon \downarrow 0$ in this inequality to find Eq. (7.22) is valid for $F \in \mathcal{F}C^1(W)$. Since $\mathcal{F}C^1(W)$ is a core for $\bar{D}$, standard limiting arguments show that Eq. (7.22) is valid in general.

The main objective for the rest of this section is to generalize the previous theorems to the setting of general compact Riemannian manifolds. Before doing this we need to record the stochastic analogues of the differentiation formula in Theorems 4.7, 4.12, and 4.13. ■

7.2. **Differentials of Stochastic Flows and Developments.**

Notation 7.25. Let $T_s^\beta(m) = \Sigma_s$ where Σ_s is the solution to Eq. (5.1) with $\Sigma_0 = m$ and β_s is an $\mathbb{R}^n$ – valued semi-martingale, i.e.

$$\delta\Sigma_s = \sum_{i=1}^n X_i(\Sigma_s)\,\delta\beta_s^i + X_0(\Sigma_s)\,ds \text{ with } \Sigma_0 = m.$$

Theorem 7.26 (Differentiating Σ in B)**.** *Let $\beta_s = B_s$ be an $\mathbb{R}^n$ – valued Brownian motion and h be an adapted Cameron-Martin process, $h_s \in \mathbb{R}^n$ with $|h_s'|$ bounded. Then there is a version of $T_s^{B+th}(m)$ which is continuous in s and differentiable in (t,m). Moreover if we define $\partial_h T_s^B(o) := \frac{d}{ds}|_0 T_s^{B+sh}(o)$, then*

$$\partial_h T_s^B(o) = Z_s\int_0^s Z_r^{-1}X_{h_r'}(\Sigma_r)\,dr = //_s z_s\int_0^s z_r^{-1}//_r^{-1}X_{h_r'}(\Sigma_r)\,dr \tag{7.25}$$

where $Z_s := \left(T_s^B\right)_{*o}$, $//_s$ is stochastic parallel translation along Σ, and $z_s := //_s^{-1}Z_s$. (See Theorem 5.41 for more on the processes Z and z.) Recall from Notation 5.4 that

$$X_a(m) := \sum_{i=1}^n a_iX_i(m) = \mathbf{X}(m)\,a.$$

Proof. This is a stochastic analogue of Theorem 4.7. Formally, if B_s were piecewise differentiable it would follow from Theorem 4.7 with $s = t$,

$$X_s(m) = \mathbf{X}(m)\,B_s' + X_0(m) \text{ and } Y_s(m) = \mathbf{X}(m)\,h_s'.$$

(Notice that $\frac{d}{dt}|_0\left[\mathbf{X}(m)\left(B_s' + th_s'\right) + X_0(m)\right] = Y_s$.) For a rigorous proof of this theorem in the flat case, which is essentially applicable here because of M is an imbedded submanifold, see Bell [12] or Nualart [148] for example. For this theorem in this geometric context see Bismut [20] or Driver [47] for example. ■

Notation 7.27. Let b be an $T_oM \cong \mathbb{R}^d$ – valued Brownian motion. A T_oM – valued semi-martingale Y is called an **adapted vector field** or **tangent process** to b if Y can be written as

$$Y_s = \int_0^s q_r db_r + \int_0^s \alpha_r dr \tag{7.26}$$

where q_r is an $so(d)$ – valued adapted process and α_s is a T_oM such that

$$\int_0^1 |\alpha_s|^2 ds < \infty \text{ a.e.}$$

A key point of a tangent process Y as above is that it gives rise to natural perturbations of the underlying Brownian motion b. Namely, following Bismut (also see Fang and Malliavin [78]), for $t \in \mathbb{R}$ let b_s^t be the process given by:

$$b_s^t := \int_0^s e^{tq_r}b_r + t\int_0^s \alpha_r dr. \tag{7.27}$$

Then (under some integrability restrictions on α) by Lévy's criteria and Girsanov's theorem, the law of b^t is absolutely continuous relative to the law of b. Moreover $b^0 = b$ and, with some additional integrability assumptions on q_r, $\frac{d}{dt}|_0 b^t = Y$.

Let b be an $T_oM \cong \mathbb{R}^d$ – valued Brownian motion, $\Sigma := \phi(b)$ be the stochastic development map as in Notation 5.30 and suppose that $X^h = //h$ is a Cameron-Martin vector field on $W(M)$. Using Theorem 4.12 as motivation (see Eq. (4.16)), the pull back of X under the stochastic development map should be the process Y defined by

$$Y_s = h_s + \int_0^s \left(\int_0^r R_{//_\rho}(h_\rho, \delta b_\rho)\right)\delta b_r \tag{7.28}$$

where

$$R_{//_s}(h_s, \delta b_s) = //_s^{-1} R(//_s h_s, //_s \delta b_s)//_s \tag{7.29}$$

like in Eq. (4.15). Since

$$\begin{aligned}\left(\int_0^r R_{//_\rho}(h_\rho, \delta b_\rho)\right)\delta b_r &= \left(\int_0^r R_{//_\rho}(h_\rho, \delta b_\rho)\right) db_r + \frac{1}{2} R_{//_\rho}(h_\rho, db_\rho) db_\rho \\ &= \left(\int_0^r R_{//_\rho}(h_\rho, \delta b_\rho)\right) db_r + \frac{1}{2}\sum_{i=1}^d R_{//_\rho}(h_\rho, e_i) e_i d\rho\end{aligned}$$

where $\{e_i\}_{i=1}^d$ is an orthonormal basis for T_oM, Eq. (7.28) may be written in Itô form as

$$Y_\cdot = \int_0^\cdot C_s db_s + \int_0^\cdot r_s ds, \tag{7.30}$$

where

$$C_s := \int_0^s R_{//_\sigma}(h_\sigma, \delta b_\sigma), \quad r_s = h_s' + \frac{1}{2}\operatorname{Ric}_{//_s} h_s \text{ and} \tag{7.31}$$

$$\operatorname{Ric}_{//_s} a := //_s^{-1} \operatorname{Ric} //_s a \ \forall\, a \in T_oM. \tag{7.32}$$

By the symmetry property in item 4b of Proposition 3.36, the matrix C_s is skew symmetric and therefore Y is a tangent process. Here is a theorem which relates Y in Eq. (7.30) to $X^h = //h$.

Theorem 7.28 (Differential of the development map)**.** *Assume M is compact manifold, $o \in M$ is fixed, b is $T_oM \cong \mathbb{R}^d$ – valued Brownian motion, $\Sigma := \phi(b)$, h is a Cameron-Martin process with $|h_s'| \le K < \infty$ (K is a non-random constant) and Y is as in Eq. (7.30). As in Eq. (7.27) let*

$$b_s^t := \int_0^s e^{tC_r} db_r + t\int_0^s r_r dr. \tag{7.33}$$

Then there exists a version of $\phi_s(b^t)$ which is continuous in (s,t), differentiable in t and $\frac{d}{dt}|_0 \phi(b^t) = X^h$.

Proof. For the proof of this theorem and its generalization to more general h, the reader is referred to Section 3.1 of [45] and to [47]. Let me just point out here that formally the proof is very analogous to the deterministic version in Theorems 4.12 and 4.13. ■

7.3. **Quasi – Invariance Flow Theorem for $W(M)$.** In this section, we will discuss the $W(M)$ analogues of Theorems 7.13 and 7.14.

Theorem 7.29 (Cameron-Martin Theorem for M)**.** *Let $h \in H(T_oM)$ and X^h be the $\mu_{W(M)}$ – a.e. well defined vector field on $W(M)$ given by*

$$X_s^h(\sigma) = //_s(\sigma) h_s \text{ for } s \in [0,1], \tag{7.34}$$

where $//_s(\sigma)$ is stochastic parallel translation along $\sigma \in W(M)$. Then X^h admits a flow e^{tX^h} on $W(M)$ (see Figure 14) and this flow leaves the Wiener measure, $\mu_{W(M)}$, quasi-invariant.

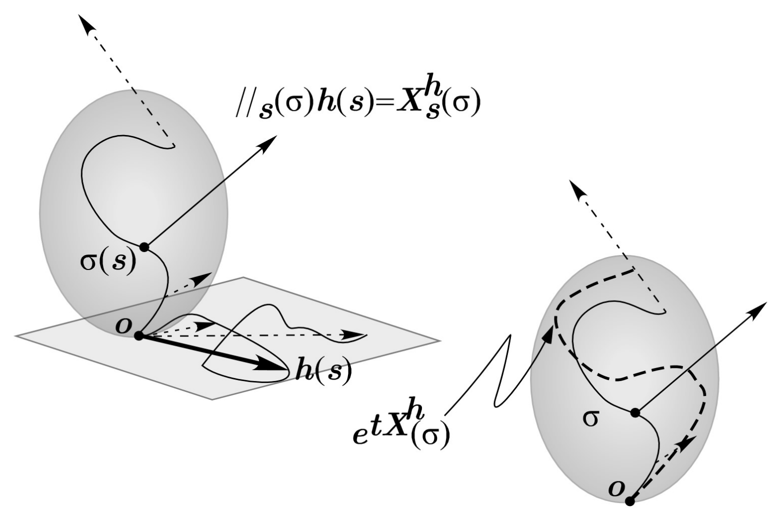

FIGURE 14. Constructing a vector field, X^h, on $W(M)$ from a vector field h on $W(T_oM)$. The dotted path indicates the flow of σ under this vector field.

This theorem first appeared in Driver [47] for $h \in H(T_oM) \cap C^1([0,1], T_oM)$ and was soon extended to all $h \in H(T_oM)$ by E. Hsu [96, 97]. Other proofs may also be found in [76, 127, 146]. The proof of this theorem is rather involved and will not be given here. A sketch of the argument and more information on the technicalities involved may be found in [49].

Example 7.30. When $M = \mathbb{R}^d$, $//_s(\sigma)v_o = v_{\sigma_s}$ for all $v \in \mathbb{R}^d$ and $\sigma \in W(\mathbb{R}^d)$. Thus $X_s^h(\sigma) = (h_s)_{\sigma_s}$ and $e^{tX^h}(\sigma) = \sigma + th$ and so Theorem 7.29 becomes the classical Cameron-Martin Theorem 7.13.

Corollary 7.31 (Integration by Parts for $\mu_{W(M)}$). *For $h \in H(T_oM)$ and $F \in \mathcal{F}C^1(W(M))$ as in Eq. (7.2), let*

$$(X^h F)(\sigma) = \frac{d}{dt}|_0 F(e^{tX^h}(\sigma)) = G\left(DF, X^h\right)$$

as in Notation 7.11. Then

$$\int_{W(M)} X^h F \, d\mu_{W(M)} = \int_{W(M)} F \, z^h \, d\mu_{W(M)}$$

where

$$z^h := \int_0^1 \langle h'_s + \frac{1}{2}\operatorname{Ric}_{//_s} h'_s, db_s\rangle,$$

$$b_s(\sigma) := \Psi_s(\sigma) = \int_0^s //_r^{-1}\delta\sigma_r$$

and $\operatorname{Ric}_{//_s} \in \operatorname{End}(T_oM)$ is as in Eq. (5.60).

Proof. A special case of this Corollary 7.31 with $F(\sigma) = f(\sigma_s)$ for some $f \in C^\infty(M)$ first appeared in Bismut [21]. The result stated here was proved in [47] as an infinitesimal form of the flow Theorem 7.29. Other proofs of this corollary may be found in [2, 5, 50, 72, 73, 70, 76, 78, 96, 97, 122, 123, 127, 146]. This corollary is a special case of Theorem 7.32 below. ∎

7.4. **Divergence and Integration by Parts.** In the next theorem, it will be shown that adapted Cameron-Martin vector fields, X, are in the domain of D^* and consequently D^* is densely defined. For the purposes of this subsection, we assume that b is a T_oM-valued Brownian motion, $\Sigma = \phi(b)$ is the evolved Brownian motion on M and $//_s$ is stochastic parallel translation along Σ.

Theorem 7.32. *Let $X \in \mathcal{X}_a$ be an adapted Cameron-Martin vector field on $W(M)$ and $h := //^{-1}X$. Then $X \in \mathcal{D}(D^*)$ and*

$$X^*1 = D^*X = \int_0^1 \langle B(h), db\rangle = \int_0^1 \langle h'_s + \frac{1}{2}\operatorname{Ric}_{//_s} h_s, db_s\rangle, \tag{7.35}$$

where B is the random linear operator mapping H to $L^2(ds, T_oM)$ given by

$$[B(h)]_s := h'_s + \frac{1}{2}\operatorname{Ric}_{//_s} h_s. \tag{7.36}$$

Remark 7.33. There is a non-random constant $C < \infty$ depending only on the bound on the Ricci tensor such that $\|B\|_{H\to L^2(ds,T_oM)} \le C$.

Proof. I will give a sketch of the proof here, the interested reader may find complete details of this proof in [45]. Moreover, we will give two more proofs of this theorem, see Theorem 7.40 and Corollary 7.50 below.

We start by proving the theorem under the additional assumption that $h := //^{-1}X$ satisfies $\sup_{s\in[0,1]} |h'_s| \le K$, where K is a non-random constant.

Let b_s^t be defined as in Eq. (7.33). (Notice that b^t is **not** the flow of the vector-field Y in Eq. (7.30) but does have the property that $\frac{d}{dt}|_0 b_s^t = Y_s$.) Since C_s is skew-symmetric, e^{tC_s} is orthogonal and so by Levy's criteria, $s \to \int_0^s e^{tC_r}db_r$ is a Brownian motion. Combining this with Girsanov's theorem, $s \to b_s^t$ (for fixed t) is a Brownian motion relative to the measure $Z_t \cdot \mu$, where

$$Z_t := \exp\left(-\int_0^1 t\langle r, e^{tC}db\rangle - \frac{1}{2}t^2\int_0^1 \langle r, r\rangle ds\right). \tag{7.37}$$

For $t \in \mathbb{R}$, let $\Sigma(t,\cdot) := \phi(b^t)$ where ϕ is the stochastic development map as in Theorem 5.29. Then by Theorem 7.28, $X^h = \frac{d}{dt}|_0\Sigma(t,\cdot)$ and in particular if F is a smooth cylinder function then $X^hF = \frac{d}{dt}|_0 F(\Sigma(t,\cdot))$. So differentiating the identity,

$$\mathbb{E}\left[F(\Sigma(t,\cdot)Z_t\right] = \mathbb{E}\left[F(\Sigma)\right],$$

at $t = 0$ gives:

$$\mathbb{E}\left[XF\right] - \mathbb{E}\left[F\int_0^1 \langle r, db\rangle\right] = 0.$$

This last equation may be written alternatively as

$$\langle DF, X\rangle_{\mathcal{X}} = \mathbb{E}\left[G\left(DF, X\right)\right] = \mathbb{E}\left[F \cdot \int_0^1 \langle B(h), db\rangle\right].$$

Hence it follows that $X \in \mathcal{D}(D^*)$ and

$$D^*X = \int_0^1 \langle B(h), db\rangle.$$

This proves the theorem in the special case that h' is uniformly bounded.

Let X be a general adapted Cameron-Martin vector-field and $h := //^{-1}X$. For each $n \in \mathbb{N}$, let $h_n(s) := \int_0^s h'(r)\cdot 1_{|h'(r)|\le n}dr$ be as in Eq. (7.11). Set $X^n := //h_n$, then by the special case above we know that $X^n \in \mathcal{D}(D^*)$ and $D^*X^n = \int_0^1 \langle B(h_n), db\rangle$. It is easy to check that

$$\langle X - X^n, X - X^n\rangle_{\mathcal{X}} = \mathbb{E}\langle h - h_n, h - h_n\rangle_H \to 0 \text{ as } n \to \infty.$$

Furthermore,

$$\mathbb{E}\left|D^*(X^m - X^n)\right|^2 = \mathbb{E}\int_0^1 |B(h_m - h_n)|^2 ds \le C\mathbb{E}\langle h_m - h_n, h_m - h_n\rangle_H,$$

from which it follows that D^*X^m is convergent. Because D^* is a closed operator, it follows that $X \in \mathcal{D}(D^*)$ and

$$D^*X = \lim_{n\to\infty} D^*X^n = \lim_{n\to\infty} \int_0^1 \langle B(h_n), db\rangle = \int_0^1 \langle B(h), db\rangle.$$

■

Corollary 7.34. *The operator $D^* : \mathcal{X} \to L^2\left(W(M), \mu_{W(M)}\right)$ is densely defined. In particular D is closable. (Let $\bar{D}$ denote the closure of D.)*

Proof. Let $h \in H$, $X^h := //h$, and F and K be smooth cylinder functions. Then, by the product rule,

$$\begin{aligned}\langle DF, KX^h\rangle_{\mathcal{X}} &= \mathbb{E}[G\left(KDF, X^h\right)] = \mathbb{E}[G\left(D\left(KF\right) - FDK, X^h\right)] \\ &= \mathbb{E}[F \cdot KD^*X^h - FG\left(DK, X^h\right)].\end{aligned}$$

Therefore $KX^h \in \mathcal{D}(D^*)$ ($\mathcal{D}(D^*)$ is the domain of D^*) and

$$D^*(KX^h) = KD^*X^h - G(DK, X^h).$$

Since

$$\mathrm{span}\{KX^h | h \in H \text{ and } K \in \mathcal{F}C^\infty\} \subset \mathcal{D}(D^*)$$

is is a dense subspace of $\mathcal{X}$, D^* is densely defined. ■

Corollary 7.35. *Let h be an adapted Cameron-Martin valued process and Q_s be defined as in Eq. (6.1). Then*

$$\left(X^{Q^{\mathrm{tr}}h}\right)^* 1 = \int_0^1 \langle Q^{\mathrm{tr}}h', db\rangle. \tag{7.38}$$

Proof. Taking the transpose of Eq. (6.1) shows Q^{tr} solves,

$$\frac{d}{ds}Q^{\mathrm{tr}} + \frac{1}{2}\mathrm{Ric}_{//} Q^{\mathrm{tr}} = 0 \text{ with } Q_0^{\mathrm{tr}} = Id. \tag{7.39}$$

Therefore, from Eq. (7.35),

$$\begin{aligned}\left(X^{Q^{\mathrm{tr}}h}\right)^* 1 &= \int_0^1 \langle \left(Q^{\mathrm{tr}}h\right)' + \frac{1}{2}\mathrm{Ric}_{//} Q^{\mathrm{tr}}h, db\rangle \\ &= \int_0^1 \left\langle \left[\frac{d}{ds} + \frac{1}{2}\mathrm{Ric}_{//}\right]\left(Q^{\mathrm{tr}}h\right), db\right\rangle \\ &= \int_0^1 \langle Q^{\mathrm{tr}}h', db\rangle.\end{aligned}$$

■

Theorem 7.32 may be extended to allow for vector-fields on the paths of M which are not based. This theorem and it Corollary 7.37 will not be used in the sequel and may safely be skipped.

Theorem 7.36. *Let h be an adapted T_oM – valued process such that $h(0)$ is non-random and $h - h(0)$ is a Cameron-Martin process, $X := X^h := //h$, $\mathbb{E}_x$ denote the path space expectation for a Brownian motion starting at $x \in M$, $F : C([0,1] \to M) \to \mathbb{R}$ be a cylinder function as in Definition 7.4 and X^hF be defined as in Eq. (7.7). Then (writing $\langle df, v\rangle$ for $df\,(v)$)*

$$\mathbb{E}_o[X^hF] = \mathbb{E}_o[FD^*X^h] + \langle d(\mathbb{E}_{(\cdot)}F), h(0)_o\rangle, \tag{7.40}$$

where

$$D^*X^h := \int_0^1 \langle h_s' + \frac{1}{2}\mathrm{Ric}_{//_s} h_s, db_s\rangle := \int_0^1 \langle B(h), db\rangle,$$

as in Eq. (7.35) and $B(h)$ is defined in Eq. (7.36).

Proof. Start by choosing a smooth path α in M such that $\dot{\alpha}(0) = h(0)_o$. Let

$$\begin{aligned}C &:= \int R_{//}(h, \delta b), \\ r &= h' + \frac{1}{2}\mathrm{Ric}_{//}(h), \\ b_s^t &= \int_0^s e^{tC}db + t\int_0^s rd\lambda \text{ and} \\ Z_t &= \exp - \left\{\int_0^1 t\langle r, e^{tC}db\rangle + \frac{1}{2}t^2\int_0^1 \langle r, r\rangle ds\right\}\end{aligned}$$

be defined by the same formulas as in the proof of Theorem 7.32. Let $u_0(t)$ denote parallel translation along α, that is

$$du_0(t)/dt + \Gamma(\dot{\alpha}(t))u_0(t) = 0 \quad \text{with} \quad u_0(0) = id.$$

For $t \in \mathbb{R}$, define $\Sigma(t, \cdot)$ by

$$\Sigma(t, \delta s) = u(t,s)\delta b_s^t \quad \text{with} \quad \Sigma(t, 0) = \alpha(t)$$

and

$$u(t, \delta s) + \Gamma(u(t,s)\delta_s b_s^t)u(t,s) = 0 \quad \text{with} \quad u(t, 0) = u_o(t).$$

Appealing to a stochastic version of Theorem 4.14 (after choosing a good version of Σ) it is possible to show that $\dot{\Sigma}(0, \cdot) = X$, so the $XF = \frac{d}{dt}|_0 F\left[\Sigma(t, \cdot)\right]$. As in the proof of Theorem 7.32, b^t is a Brownian motion relative to the expectation $\mathbb{E}_t$ defined by $\mathbb{E}_t(F) := \mathbb{E}\left[Z_tF\right]$. From this it is easy to see that $\Sigma(t, \cdot)$ is a Brownian motion on M starting at $\alpha(t)$ relative to the expectation $\mathbb{E}_t$. Therefore, for all t,

$$\mathbb{E}\left[F\left(\Sigma(t, \cdot)\right)Z_t\right] = \mathbb{E}_{\alpha(t)}F$$

and differentiating this last expression at $t = 0$ gives:

$$\mathbb{E}\left[XF(\Sigma)\right] - \mathbb{E}\left[F\int_0^1 \langle r, db\rangle\right] = \langle d\mathbb{E}_{(\cdot)}F, h(0)_o\rangle.$$

The rest of the proof is identical to the previous proof. ■

As a corollary to Theorem 7.36 we get Elton Hsu's derivative formula which played a key role in the original proof of his logarithmic Sobolev inequality on $W(M)$, see Theorem 7.52 below and [98].

Corollary 7.37 (Hsu's Derivative Formula). *Let $v_o \in T_oM$. Define h to be the adapted T_oM – valued process solving the differential equation:*

$$h_s' + \frac{1}{2}\operatorname{Ric}_{//_s} h_s = 0 \quad \text{with} \quad h_0 = v_o. \tag{7.41}$$

Then

$$\langle d(\mathbb{E}_{(\cdot)}F), v_o\rangle = \mathbb{E}_o[X^h F]. \tag{7.42}$$

Proof. Apply Theorem 7.36 to X^h with h defined by (7.41). Notice that h has been constructed so that $B(h) \equiv 0$, i.e. $D^* X^h = 0$. ■

The idea for the proof used here is similar to Hsu's proof, the only question is how one describes the perturbed process $\Sigma(t,\cdot)$ in the proof of Theorem 7.36 above. It is also possible to give a much more elementary proof of Eq. (7.42) based on the ideas in Section 6, see for example [58].

7.5. **Elworthy-Li Integration by Parts Formula.** In this subsection, let $\{X_i\}_{i=0}^n \subset \Gamma(TM)$, B be a $\mathbb{R}^n$ – valued Brownian motion and $T_s^B(m)$ denote the solution to Eq. (5.1) with $\beta = B$ be as in Notation 7.25. We will further assume that $\mathbf{X}(m) : \mathbb{R}^n \to T_mM$ (as in Notation 5.4) is surjective for all $m \in M$ and let $\mathbf{X}(m)^\# = \left[\mathbf{X}(m)|_{\mathrm{Nul}(\mathbf{X}(m))^\perp}\right]^{-1}$ as in Eq. (6.5). The following Lemma is an elementary exercise in linear algebra.

Lemma 7.38. *For $m \in M$ and $v, w \in T_mM$ let*

$$\langle v, w\rangle_m := \langle \mathbf{X}(m)^\# v, \mathbf{X}(m)^\# w\rangle_{\mathbb{R}^n}.$$

Then

1. *$m \to \langle\cdot,\cdot\rangle_m$ is a smooth Riemannian metric on M.*
2. *$\mathbf{X}(m)^{\mathrm{tr}} = \mathbf{X}(m)^\#$ and in particular $\mathbf{X}(m)\mathbf{X}(m)^{\mathrm{tr}} = id_{T_mM}$ for all $m \in M$.*
3. *Every $v \in T_mM$ may be expanded as*

$$v = \sum_{j=1}^n \langle v, X_j(m)\rangle X_j(m) = \sum_{j=1}^n \langle v, \mathbf{X}(m) e_j\rangle \mathbf{X}(m) e_j \tag{7.43}$$

where $\{e_j\}_{j=1}^n$ is the standard basis for $\mathbb{R}^n$.

The proof of this lemma is left to the reader with the comment that Eq. (7.43) is proved in the same manner as item (1) in Proposition 3.48.

Theorem 7.39 (Elworthy - Li). *Suppose k_s is a T_oM valued Cameron-Martin process such that $\mathbb{E}\int_0^1 |k_s'|^2 ds < \infty$ and $F : W(M) \to \mathbb{R}$ is a bounded C^1 – function with bounded derivative on W, for example F could be a cylinder function. Then*

$$\mathbb{E}\left[\left(d_{W(M)}F\right)(Z.k.)\right] = \mathbb{E}\left[F(\Sigma)\int_0^T \langle Z_s k_s', \mathbf{X}(\Sigma_s)\, dB_s\rangle\right]$$
$$= \mathbb{E}\left[F(\Sigma)\int_0^T \langle \mathbf{X}(\Sigma_s)^{\mathrm{tr}} Z_s k_s', dB_s\rangle\right] \tag{7.44}$$

where and $Z_s = \left(T_s^B\right)_{*o}$ is the differential of $m \to T_s^B(m)$ at o.

Proof. Notice that $Z_s k_s \in T_{\Sigma_s}M$ for all s as it should be. By the reduction argument used in the proof of Theorem 7.32, it suffices to consider the case where $|k_s'| \le K$ where K is a non-random constant. Let h_s be the T_oM – valued Cameron-Martin process defined by

$$h_s := \int_0^s \mathbf{X}(\Sigma_r)^{\mathrm{tr}} Z_r k_r' dr.$$

Then by Lemma 7.38 and Theorem 7.26,

$$\partial_h T_s^B(o) = Z_s \int_0^s Z_r^{-1}\mathbf{X}(\Sigma_r) h_r' dr$$
$$= Z_s \int_0^s Z_r^{-1}\mathbf{X}(\Sigma_r)\mathbf{X}(\Sigma_r)^{\mathrm{tr}} Z_r k_r' dr = Z_s k_s.$$

In particular this implies

$$\partial_h F\left(T_{(\cdot)}^B(o)\right) = \langle dF(\Sigma), \partial_h T_s^B(o)\rangle = \langle d_{W(M)}F(\Sigma), Zk\rangle$$

and therefore by integration by parts on the flat Wiener space (Theorem 7.32) with $M = \mathbb{R}^n$) implies

$$\mathbb{E}\left[\left(d_{W(M)}F\right)(\Sigma)(Z.k.)\right] = \mathbb{E}\left[\partial_h\left[F(\Sigma)\right]\right] = \mathbb{E}\left[F(\Sigma)\int_0^T \langle h_s', dB_s\rangle\right]$$
$$= \mathbb{E}\left[F(\Sigma)\int_0^T \langle \mathbf{X}(\Sigma_s)^{\mathrm{tr}} Z_s k_s', dB_s\rangle\right].$$

■

By factoring out the redundant noise in Theorem 7.39, we get yet another proof of Corollary 7.35 which also easily gives another proof of Theorem 7.32.

Theorem 7.40 (Factoring out the redundant noise). *Assume $\mathbf{X}(m) = P(m)$ and $X_0 = 0$, k_s is a Cameron-Martin valued process adapted to the filtration, $\mathcal{F}_s^\Sigma := \sigma(\Sigma_r : r \le s)$, then*

$$\mathbb{E}\left[\left(d_{W(M)}F\right)\left(//Q_t^{\mathrm{tr}} k\right)\right] = \mathbb{E}\left[F(\Sigma)\int_0^T \langle Q_s^{\mathrm{tr}} k_s', db_s\rangle\right]$$

where Q_s solves Eq. (6.1).

Proof. By Theorems 7.39 and 5.40, we have

$$\begin{aligned}\mathbb{E}\left[\left(d_{W(M)}F\right)(//zk)\right] &= \mathbb{E}\left[F(\Sigma)\int_0^T \langle //_s z_s k_s', P(\Sigma_s)\, dB_s\rangle\right] \\ &= \mathbb{E}\left[F(\Sigma)\int_0^T \langle z_s k_s', db_s\rangle\right].\end{aligned}$$

Combining this with Theorem 5.44 implies

$$\mathbb{E}\left[\left(d_{W(M)}F\right)(//\bar{z}k)\right] = \mathbb{E}\left[F(\Sigma)\int_0^T \langle \bar{z}_s k_s', db_s\rangle\right].$$

As observed in the proof of Corollary 6.4, $\bar{z}_t = Q_t^{\mathrm{tr}}$ which completes the proof. ■

The reader interested in seeing more of these type of arguments is referred to Elworthy, Le Jan and Li [71] where these ideas are covered in much greater detail and in full generality.

7.6. **Fang's Spectral Gap Theorem and Proof.** As in the flat case we let $\mathcal{L} = D^*\bar{D}$ – an unbounded operator on $L^2\left(W(M), \mu_{W(M)}\right)$ which is a "curved" analogue of the Ornstein-Uhlenbeck operator used in Theorem 7.23. It has been shown in Driver and Röckner [56] that this operator generates a diffusion on $W(M)$. This last result also holds for pinned paths on M and free loops on $\mathbb{R}^N$, see [6].

In this section, we will give a proof of S. Fang's [79] spectral gap inequality for $\mathcal{L}$. Hsu's stronger logarithmic Sobolev inequality will be covered later in Theorem 7.52 below.

Theorem 7.41 (Fang)**.** *Let $\bar{D}$ be the closure of D and $\mathcal{L}$ be the self-adjoint operator on $L^2\left(\mu_{W(M)}\right)$ defined by $\mathcal{L} = D^*\bar{D}$. (Note, if $M = \mathbb{R}^d$ then $\mathcal{L}$ would be an infinite dimensional Ornstein-Uhlenbeck operator.) Then the null-space of $\mathcal{L}$ consists of the constant functions on $W(M)$ and $\mathcal{L}$ has a spectral gap, i.e. there is a constant $c > 0$ such that $\langle \mathcal{L}F, F\rangle_{L^2(\mu_{W(M)})} \ge c\langle F, F\rangle_{L^2(\mu_{W(M)})}$ for all $F \in \mathcal{D}(\mathcal{L})$ which are perpendicular to the constant functions.*

This theorem is the $W(M)$ analogue of Theorem 7.23. The proof of this theorem will be given at the end of this subsection. We first will need to represent F in terms of DF. (Also see Section 7.7 below.)

Lemma 7.42. *For each $F \in L^2\left(W(M), \mu_{W(M)}\right)$, there is a unique adapted Cameron-Martin vector field X on $W(M)$ such that*

$$F = \mathbb{E}F + D^*X.$$

Proof. By the martingale representation theorem (see Corollary 7.20), there is a predictable T_oM–valued process, a, (which is not in general continuous) such that

$$\mathbb{E}\int_0^1 |a_s|^2 ds < \infty,$$

and

$$F = \mathbb{E}F + \int_0^1 \langle a_s, db_s\rangle. \tag{7.45}$$

Define $h := B^{-1}(a)$, where B is as in Eq. (7.36); that is to say let h be the solution to the differential equation:

$$h_s' + \frac{1}{2}\operatorname{Ric}_{//_s} h_s = a_s \text{ with } h_0 = 0. \tag{7.46}$$

Claim: B_σ^{-1} is a bounded linear map from $L^2(ds, T_oM) \to H$ for each $\sigma \in W(M)$, and furthermore the norm of B_σ^{-1} is bounded independent of $\sigma \in W(M)$.

To prove the claim, use Duhamel's principle to write the solution to (7.46) as:

$$h_s = \int_0^s Q_s^{\mathrm{tr}}\left(Q_\tau^{\mathrm{tr}}\right)^{-1} a_\tau d\tau, \tag{7.47}$$

where Q_s is as in Eq. (6.1). Since, $W_s := Q_s^{\mathrm{tr}}\left(Q_\tau^{\mathrm{tr}}\right)^{-1}$ solves the differential equation

$$W_s' + \frac{1}{2}\operatorname{Ric}_{//_s} W_s = 0 \text{ with } W_\tau = I$$

it is easy to show from the boundedness of $\operatorname{Ric}_{//_s}$ and an application of Gronwall's inequality that

$$\left|Q_s^{\mathrm{tr}}\left(Q_\tau^{\mathrm{tr}}\right)^{-1}\right| = |W_s| \le C,$$

where C is a non-random constant independent of s and τ. Therefore,

$$\begin{aligned}\langle h, h\rangle_H &= \int_0^1 |a_s - \frac{1}{2}\operatorname{Ric}_{//_s} h_s|^2 ds \\ &\le 2\int_0^1 |a_s|^2 ds + 2\int_0^1 |\frac{1}{2}\operatorname{Ric}_{//_s} h_s|^2 ds \\ &\le 2(1 + C^2K^2)\int_0^1 |a_s|^2 ds,\end{aligned}$$

where K is a bound on the process $\frac{1}{2}\operatorname{Ric}_{//_s}$. This proves the claim.

Because of the claim, $h := B^{-1}(a)$ satisfies $\mathbb{E}\left[\langle h, h\rangle_H\right] < \infty$ and because of Eq. (7.47), h is adapted. Hence, $X := //h$ is an adapted Cameron-Martin vector field and

$$D^*X = \int_0^1 \langle B(h), db\rangle = \int_0^1 \langle a, db\rangle.$$

The existence part of the theorem now follows from this identity and Eq. (7.45).

The uniqueness assertion follows from the energy identity:

$$\mathbb{E}\left[D^*X\right]^2 = \mathbb{E}\int_0^1 |B(h)_s|^2 ds \geq C\mathbb{E}\left[\langle h, h\rangle_H\right].$$

Indeed if $D^*X = 0$, then $h = 0$ and hence $X = //h = 0$. ∎

The next goal is to find an expression for the vector-field X in the above lemma in terms of the function F itself. This will be the content of Theorem 7.45 below.

Notation 7.43. Let $L_a^2(\mu_{W(M)} : L^2(ds, T_oM))$ denote the T_oM – valued predictable processes, v_s on $W(M)$ such that $\mathbb{E}\int_0^1 |v_s|^2 ds < \infty$. Define the bounded linear operator $\bar{B} : \mathcal{X}_a \to L_a^2(\mu_{W(M)} : L^2(ds, T_oM))$ by

$$\bar{B}(X) = B(//^{-1}X) = \frac{d}{ds}\left[//_s^{-1}X_s\right] + \frac{1}{2}//_s^{-1}\operatorname{Ric} X_s.$$

Also let $\mathcal{Q} : \mathcal{X} \to \mathcal{X}$ denote the orthogonal projection of $\mathcal{X}$ onto $\mathcal{X}_a$.

Remark 7.44. Notice that $D^*X = \int_0^1 \langle \bar{B}(X), db\rangle$ for all $X \in \mathcal{X}_a$. We have seen that $\bar{B}$ has a bounded inverse, in fact $\bar{B}^{-1}(a) = //B^{-1}(a)$.

Theorem 7.45. *As above let $\bar{D}$ denote the closure of D. Also let $T : \mathcal{X} \to \mathcal{X}_a$ be the bounded linear operator defined by*

$$T(X) = (\bar{B}^*\bar{B})^{-1}\mathcal{Q}X$$

for all $X \in \mathcal{X}$. Then for all $F \in \mathcal{D}(\bar{D})$,

$$F = \mathbb{E}F + D^*T\bar{D}F. \tag{7.48}$$

It is worth pointing out that $\bar{B}^*$ is not $//B^*$ but is instead given by $\mathcal{Q}//B^*$. This is because $//B^*$ does not take adapted processes to adapted processes. This is the reason it is necessary to introduce the orthogonal projection, $\mathcal{Q}$.

Proof. Let $Y \in \mathcal{X}_a$ be given and $X \in \mathcal{X}_a$ be chosen so that $F = \mathbb{E}F + D^*X$. Then

$$\begin{aligned}\langle Y,\ \mathcal{Q}\bar{D}F\rangle_{\mathcal{X}} &= \langle Y, \bar{D}F\rangle_{\mathcal{X}} = \mathbb{E}\left[D^*Y \cdot F\right] \\ &= \mathbb{E}\left[D^*Y \cdot D^*X\right] = \mathbb{E}\left[\langle \bar{B}(Y), \bar{B}(X)\rangle_{L^2(ds)}\right] \\ &= \langle Y, \bar{B}^*\bar{B}(X)\rangle_{\mathcal{X}},\end{aligned}$$

where in going from the first to the second line we have used $\mathbb{E}\left[D^*Y\right] = 0$. From the above displayed equation it follows that $\mathcal{Q}\bar{D}F = \bar{B}^*\bar{B}(X)$ and hence $X = (\bar{B}^*\bar{B})^{-1}\mathcal{Q}\bar{D}F = T(\bar{D}F)$. ∎

7.6.1. *Proof of Theorem 7.41.* Let $F \in \mathcal{D}(\bar{D})$. By Theorem 7.45,

$$\mathbb{E}\left[F - \mathbb{E}F\right]^2 = \mathbb{E}\left[D^*T\bar{D}F\right]^2 = \mathbb{E}|\bar{B}(T\bar{D}F)|^2_{L^2(ds,T_oM)} \leq C\langle \bar{D}F, \bar{D}F\rangle_{\mathcal{X}}$$

where C is the operator norm of $\bar{B}T$. In particular if $F \in \mathcal{D}(\mathcal{L})$, then $\langle \bar{D}F, \bar{D}F\rangle_{\mathcal{X}} = \mathbb{E}[\mathcal{L}F \cdot F]$, and hence

$$\langle \mathcal{L}F, F\rangle_{L^2(\mu_{W(M)})} \geq C^{-1}\langle F - \mathbb{E}F, F - \mathbb{E}F\rangle_{L^2(\mu_{W(M)})}.$$

Therefore, if $F \in \operatorname{Nul}(\mathcal{L})$, it follows that $F = \mathbb{E}F$, i.e. F is a constant. Moreover if $F \perp 1$ (i.e. $\mathbb{E}F = 0$) then

$$\langle \mathcal{L}F, F\rangle_{L^2(\mu_{W(M)})} \geq C^{-1}\langle F, F\rangle_{L^2(\mu_{W(M)})},$$

proving Theorem 7.41 with $c = C^{-1}$.

7.7. $W(M)$ – **Martingale Representation Theorem.** In this subsection, Σ is a Brownian motion on M starting at $o \in M$, $//_s$ is stochastic parallel translation along Σ and

$$b_s = \left[\Psi(\Sigma)\right]_s = \int_0^s //_r^{-1}\delta\Sigma_r$$

is the undeveloped T_oM – valued Brownian motion associated to Σ as described before Theorem 5.29.

Lemma 7.46. *If $f \in C^\infty\left(M^{n+1}\right)$ and $i \leq n$, then*

$$\begin{aligned}&\mathbb{E}\left[//_{s_i}^{-1}\operatorname{grad}_i f\left(\Sigma_{s_1}, \dots, \Sigma_{s_n}, \Sigma_{s_{n+1}}\right)\middle|\, \mathcal{F}_{s_n}\right] \\ &= //_{s_i}^{-1}\operatorname{grad}_i(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)\left(\Sigma_{s_1}, \dots, \Sigma_{s_n}, \Sigma_{s_n}\right).\end{aligned} \tag{7.49}$$

Proof. Let us begin with the special case where $f = g \otimes h$ for some $g \in C^\infty(M^n)$ and $h \in C^\infty(M)$ where $g \otimes h\,(x_1, \dots, x_{n+1}) := g\,(x_1, \dots, x_n)\,h\,(x_{n+1})$. In this case

$$//_{s_i}^{-1}\operatorname{grad}_i f\left(\Sigma_{s_1}, \dots, \Sigma_{s_n}, \Sigma_{s_{n+1}}\right) = //_{s_i}^{-1}\operatorname{grad}_i g\left(\Sigma_{s_1}, \dots, \Sigma_{s_n}\right) \cdot h\left(\Sigma_{s_{n+1}}\right)$$

where $//_{s_i}^{-1}\operatorname{grad}_i g\,(\Sigma_{s_1}, \dots, \Sigma_{s_n})$ is $\mathcal{F}_{s_n}$ – measurable. Hence by the Markov property we have

$$\begin{aligned}&\mathbb{E}\left[//_{s_i}^{-1}\operatorname{grad}_i f\left(\Sigma_{s_1}, \dots, \Sigma_{s_n}, \Sigma_{s_{n+1}}\right)\middle|\, \mathcal{F}_{s_n}\right] \\ &= //_{s_i}^{-1}\operatorname{grad}_i g\left(\Sigma_{s_1}, \dots, \Sigma_{s_n}\right)\mathbb{E}\left[h\left(\Sigma_{s_{n+1}}\right)\middle|\, \mathcal{F}_{s_n}\right] \\ &= //_{s_i}^{-1}\operatorname{grad}_i g\left(\Sigma_{s_1}, \dots, \Sigma_{s_n}\right)(e^{(s_{n+1}-s_n)\bar{\Delta}/2}h)\left(\Sigma_{s_n}\right) \\ &= //_{s_i}^{-1}\operatorname{grad}_i(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)\left(\Sigma_{s_1}, \dots, \Sigma_{s_n}, \Sigma_{s_n}\right).\end{aligned}$$

Alternatively, as we have already seen, $M_s := (e^{(s_{n+1}-s)\bar{\Delta}/2}h)\left(\Sigma_s\right)$ is a martingale for $s \leq s_{n+1}$, and therefore,

$$\mathbb{E}\left[h\left(\Sigma_{s_{n+1}}\right)\middle|\, \mathcal{F}_{s_n}\right] = \mathbb{E}\left[M_{s_{n+1}}\middle|\, \mathcal{F}_{s_n}\right] = M_{s_n} = (e^{(s_{n+1}-s_n)\bar{\Delta}/2}h)\left(\Sigma_{s_n}\right).$$

Since Eq. (7.49) is linear in f, this proves Eq. (7.49) when f is a linear combination of functions of the form $g\otimes h$ as above.

Using a partition unity argument along with the standard convolution approximation methods; to any $f\in C^{\infty}\left(M^{n+1}\right)$ there exists a sequence of $f_k\in C^{\infty}\left(M^{n+1}\right)$ with each f_k being a linear combination of functions of the form $g\otimes h$ such that f_k along with all of its derivatives converges uniformly to f. Passing to the limit in Eq. (7.49) with f being replaced by f_k, shows that Eq. (7.49) holds for all $f\in C^{\infty}\left(M^{n+1}\right)$. ■

Recall that Q_s is the $\operatorname{End}(T_oM)$ – valued process determined in Eq. (6.1) and since

$$\frac{d}{ds}Q_s^{-1}=-Q_s^{-1}\left[\frac{d}{ds}Q_s\right]Q_s^{-1},$$

Q_s^{-1} solves the equation,

$$\frac{d}{ds}Q_s^{-1}=\frac{1}{2}\operatorname{Ric}_{//_s}Q_s^{-1}\text{ with }Q_0^{-1}=I. \tag{7.50}$$

Theorem 7.47 (Representation Formula)**.** *Suppose that F is a smooth cylinder function of the form $F(\sigma)=f(\sigma_{s_1},\dots,\sigma_{s_n})$, then*

$$F(\Sigma)=\mathbb{E}F+\int_0^1\langle a_s,db_s\rangle \tag{7.51}$$

where a_s is a bounded predictable process, a_s is zero if $s\geq s_n$ and $s\to a_s$ is continuous off the partition set, $\{s_1,\dots,s_n\}$. Moreover a_s may be expressed as

$$a_s:=Q_s^{-1}\mathbb{E}\left[\sum_{i=1}^n 1_{s\le s_i}Q_{s_i}//_{s_i}^{-1}\operatorname{grad}_i f\left(\Sigma_{s_1},\dots,\Sigma_{s_n}\right)\middle|\mathcal{F}_s\right]. \tag{7.52}$$

Proof. The proof will be by induction on n. For $n=1$ suppose $F(\Sigma)=f(\Sigma_t)$ for some $t\in(0,1]$. Integrating Eq. (5.38) from $[0,t]$ with $g=f$ implies

$$F(\Sigma)=f(\Sigma_t)=e^{t\bar{\Delta}/2}f(o)+\int_0^t\langle//_s^{-1}\operatorname{grad}\ e^{(t-s)\bar{\Delta}/2}f(\Sigma_s),db_s\rangle. \tag{7.53}$$

Since $e^{t\bar{\Delta}/2}f(o)=\mathbb{E}F$, Eq. (7.53) shows Eq. (7.51) holds with

$$a_s=1_{0\le s\le t}//_s^{-1}\operatorname{grad}\ e^{(t-s)\bar{\Delta}/2}f(\Sigma_s).$$

By Lemma 6.1, $Q_s//_s^{-1}\operatorname{grad}\ e^{(t-s)\bar{\Delta}/2}f(\Sigma_s)$ is a martingale, and hence

$$Q_s//_s^{-1}\operatorname{grad}\ e^{(t-s)\bar{\Delta}/2}f(\Sigma_s)=\mathbb{E}\left[Q_t//_t^{-1}\operatorname{grad}\ f(\Sigma_t)\middle|\mathcal{F}_s\right]$$

from which it follows that

$$a_s=1_{0\le s\le t}//_s^{-1}\operatorname{grad}\ e^{(t-s)\bar{\Delta}/2}f(\Sigma_s)=1_{0\le s\le t}Q_s^{-1}\mathbb{E}\left[Q_t//_t^{-1}\operatorname{grad}\ f(\Sigma_t)\middle|\mathcal{F}_s\right].$$

This shows that Eq. (7.52) is valid for $n=1$.

To carry out the inductive step, suppose the result holds for level n and now suppose that

$$F(\Sigma)=f\left(\Sigma_{s_1},\dots,\Sigma_{s_{n+1}}\right)$$

with $0<s_1<s_2\dots<s_{n+1}\le 1$. Let

$$(\Delta_{n+1}f)(x_1,x_2,\dots,x_{n+1})=(\Delta g)(x_{n+1})$$

where $g(x):=f(x_1,x_2,\dots,x_n,x)$. Similarly, let $\operatorname{grad}_{n+1}$ denote the gradient acting on the $(n+1)^{\text{th}}$ – variable of a function $f\in C^\infty(M^{n+1})$. Set

$$H(s,\Sigma):=(e^{(s_{n+1}-s)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_s)$$

for $s_n\le s\le s_{n+1}$. By Itô's Lemma, (see Corollary 5.18 and also Eq. (5.38),

$$d\left[H(s,\Sigma_s)\right]=\langle\operatorname{grad}_{n+1}e^{(s_{n+1}-s)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_s,//_s db_s\rangle$$

for $s_n\le s\le s_{n+1}$. Integrating this last expression from s_n to s_{n+1} yields:

$$F(\Sigma)=(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n})$$

$$+\int_{s_n}^{s_{n+1}}\langle//_s^{-1}\operatorname{grad}_{n+1}e^{(s_{n+1}-s)\bar{\Delta}_{n+1}/2}f)\left(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_s\right),db_s\rangle \tag{7.54}$$

$$=(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n})+\int_{s_n}^{s_{n+1}}\langle\alpha_s,db_s\rangle, \tag{7.55}$$

where $\alpha_s:=//_s^{-1}(\operatorname{grad}_{n+1}e^{(s_{n+1}-s)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_s)$. By the induction hypothesis, the smooth cylinder function,

$$(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n}),$$

may be written as a constant plus $\int_0^1\langle\tilde{a}_s,db_s\rangle$, where $\tilde{a}_s$ is bounded and piecewise continuous and $\tilde{a}_s\equiv 0$ if $s\ge s_n$. Thus if we let $a_s:=\tilde{a}_s+1_{s_n<s\le s_{n+1}}\alpha_s$, we have shown

$$F(\Sigma)=C+\int_0^{s_{n+1}}\langle a_s,db_s\rangle$$

for some constant C. Taking expectations of both sides of this equation then shows $C=\mathbb{E}\left[F(\Sigma)\right]$ and the proof of Eq. (7.51) is complete. So to finish the proof it only remains to verify Eq. (7.52).

Again by Lemma 6.1,

$$s\to M_s:=Q_s//_s^{-1}(\operatorname{grad}_{n+1}e^{(s_{n+1}-s)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_s)$$

is a martingale for $s\in[s_n,s_{n+1}]$ and therefore,

$$\begin{aligned}M_s&=Q_s//_s^{-1}(\operatorname{grad}_{n+1}e^{(s_{n+1}-s)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_s)\\&=\mathbb{E}\left[M_{s_{n+1}}\middle|\mathcal{F}_s\right]=\mathbb{E}\left[Q_{s_{n+1}}//_{s_{n+1}}^{-1}\left(\operatorname{grad}_{n+1}f\right)\left(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}}\right)\middle|\mathcal{F}_s\right],\end{aligned} \tag{7.56}$$

i.e.

$$\begin{aligned}&//_s^{-1}(\operatorname{grad}_{n+1}e^{(s_{n+1}-s)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_s)\\&=Q_s^{-1}\mathbb{E}\left[Q_{s_{n+1}}//_{s_{n+1}}^{-1}\left(\operatorname{grad}_{n+1}f\right)\left(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}}\right)\middle|\mathcal{F}_s\right].\end{aligned} \tag{7.57}$$

Using this identity, Eq. (7.54) may be written as

$$F(\Sigma) = g(\Sigma_{s_1},\dots,\Sigma_{s_n}) + \int_{s_n}^{s_{n+1}} \left\langle Q_s^{-1}\mathbb{E}\left[Q_{s_{n+1}}//_{s_{n+1}}^{-1}\left(\mathrm{grad}_{n+1}f\right)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}})\middle|\mathcal{F}_s\right],db_s\right\rangle. \tag{7.58}$$

where

$$g(x_1,\dots,x_n) := (e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(x_1,\dots,x_n,x_n).$$

By the induction hypothesis,

$$g(\Sigma_{s_1},\dots,\Sigma_{s_n}) = C + \int_0^1 \left\langle Q_s^{-1}\mathbb{E}\left[\sum_{i=1}^n 1_{s\le s_i}Q_{s_i}//_{s_i}^{-1}\mathrm{grad}_i g(\Sigma_{s_1},\dots,\Sigma_{s_n})\middle|\mathcal{F}_s\right],db_s\right\rangle \tag{7.59}$$

where $C = \mathbb{E}[F(\Sigma)]$ as we have already seen or alternatively, by the Markov property,

$$\begin{aligned} C &:= \mathbb{E}(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n}) \\ &= \mathbb{E}f(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}}) = \mathbb{E}[F(\Sigma)]. \end{aligned} \tag{7.60}$$

By Lemma 7.46, for $s \le s_n$ and $i < n$

$$\begin{aligned} &\mathbb{E}\left[Q_{s_i}//_{s_i}^{-1}\mathrm{grad}_i g(\Sigma_{s_1},\dots,\Sigma_{s_n})\middle|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[Q_{s_i}\mathbb{E}\left[//_{s_i}^{-1}\mathrm{grad}_i(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n})\middle|\mathcal{F}_{s_n}\right]\middle|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[Q_{s_i}//_{s_i}^{-1}\mathrm{grad}_i f(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}})\middle|\mathcal{F}_s\right]. \end{aligned} \tag{7.61}$$

While for $s \le s_n$ and $i = n$, we have:

$$\begin{aligned} \mathrm{grad}_n g(\Sigma_{s_1},\dots,\Sigma_{s_n}) &= \mathrm{grad}_n(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n}) \\ &\quad + \mathrm{grad}_{n+1}(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n}), \end{aligned}$$

$$\begin{aligned} &\mathbb{E}\left[Q_{s_n}//_{s_n}^{-1}\mathrm{grad}_n(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n})\middle|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[Q_{s_n}\mathbb{E}\left[//_{s_n}^{-1}\mathrm{grad}_n(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n})\middle|\mathcal{F}_{s_n}\right]\middle|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[Q_{s_n}//_{s_n}^{-1}\mathrm{grad}_n f(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}})\middle|\mathcal{F}_s\right] \end{aligned}$$

by Lemma 7.46 and

$$\begin{aligned} &\mathbb{E}\left[\mathbb{E}\left[Q_{s_n}//_{s_n}^{-1}\mathrm{grad}_{n+1}(e^{(s_{n+1}-s_n)\bar{\Delta}_{n+1}/2}f)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_n})\middle|\mathcal{F}_{s_n}\right]\middle|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[Q_{s_{n+1}}//_{s_{n+1}}^{-1}\left(\mathrm{grad}_{n+1}f\right)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}})\middle|\mathcal{F}_s\right] \end{aligned}$$

from Eq. (7.57) with $s = s_n$. Combining the previous three displayed equations shows,

$$\begin{aligned} &\mathbb{E}\left[Q_{s_n}//_{s_n}^{-1}\mathrm{grad}_n g(\Sigma_{s_1},\dots,\Sigma_{s_n})\middle|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[Q_{s_n}//_{s_n}^{-1}\mathrm{grad}_n f(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}})\middle|\mathcal{F}_s\right] \\ &\quad + \mathbb{E}\left[Q_{s_{n+1}}//_{s_{n+1}}^{-1}\left(\mathrm{grad}_{n+1}f\right)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}})\middle|\mathcal{F}_s\right] \end{aligned} \tag{7.62}$$

Assembling Eqs. (7.59), (7.60), (7.61) and (7.62) implies

$$\begin{aligned} &g(\Sigma_{s_1},\dots,\Sigma_{s_n}) \\ &= \mathbb{E}[F(\Sigma)] + \int_0^1 \sum_{i=1}^n \left\langle Q_s^{-1}\mathbb{E}\left[1_{s\le s_i}Q_{s_i}//_{s_i}^{-1}\mathrm{grad}_i f(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}})\middle|\mathcal{F}_s\right],db_s\right\rangle \\ &\quad + \int_0^1 \left\langle Q_s^{-1}\mathbb{E}\left[1_{s\le s_n}Q_{s_{n+1}}//_{s_{n+1}}^{-1}\left(\mathrm{grad}_{n+1}f\right)(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}})\middle|\mathcal{F}_s\right],db_s\right\rangle \end{aligned}$$

which combined with Eq. (7.58) shows

$$\begin{aligned} &F(\Sigma) = \mathbb{E}[F(\Sigma)] \\ &+ \int_0^1 \left\langle Q_s^{-1}\mathbb{E}\left[\sum_{i=1}^{n+1} 1_{s\le s_i}Q_{s_i}//_{s_i}^{-1}\mathrm{grad}_i f(\Sigma_{s_1},\dots,\Sigma_{s_n},\Sigma_{s_{n+1}})\middle|\mathcal{F}_s\right],db_s\right\rangle. \end{aligned}$$

This completes the induction argument and hence the proof. ∎

Proposition 7.48. *Equation (7.51) may also be written as*

$$F(\Sigma) = \mathbb{E}[F(\Sigma)] + \int_0^1 \left\langle \mathbb{E}\left[\xi_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r\,\mathrm{Ric}_{//_r}\,\xi_r dr\middle|\mathcal{F}_s\right],db_s\right\rangle. \tag{7.63}$$

where

$$\xi_s := //_s^{-1}\frac{d}{ds}(DF)_s.$$

Proof. Let $v_i := //_{s_i}^{-1}\,\mathrm{grad}_i\, f(\Sigma_{s_1},\dots,\Sigma_{s_n})$, so that

$$\xi_s := //_s^{-1}\frac{d}{ds}(DF)_s = \sum_{i=1}^n 1_{s<s_i}v_i,$$

and let

$$\alpha_s := \sum_{i=1}^n 1_{s\le s_i}Q_s^{-1}Q_{s_i}//_{s_i}^{-1}\mathrm{grad}_i f(\Sigma_{s_1},\dots,\Sigma_{s_n}) = \sum_{i=1}^n 1_{s\le s_i}Q_s^{-1}Q_{s_i}v_i.$$

Then the Lebesgue-Stieljtes measure associate to ξ_s is

$$d\xi_s = -\sum_{i=1}^n \delta_{s_i}(ds)\,v_i$$

and therefore

$$\alpha_s = -Q_s^{-1}\int_s^1 Q_r d\xi_r = -\int_s^1 Q_s^{-1}Q_r d\xi_r.$$

So by integration by parts we have, for $s \notin \{0, s_1, \ldots, s_n, 1\}$,

$$\begin{aligned}\alpha_s &= -\int_s^1 Q_s^{-1}Q_r d\xi_r = -\left[Q_s^{-1}Q_r\xi_r\right]|_{r=s}^{r=1} + \int_s^1 Q_s^{-1}\left[\frac{d}{dr}Q_r\right]\xi_r \\ &= \xi_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r \operatorname{Ric}_{//_r}\xi_r\end{aligned}$$

where we have used $\xi_1 = 0$. This completes the proof since from Eqs. (7.51) and (7.52),

$$F(\Sigma) = \mathbb{E}[F(\Sigma)] + \int_0^1 \langle E[\alpha_s|\mathcal{F}_s], db_s\rangle.$$

■

Corollary 7.49. *Let F be a smooth cylinder function, then there is a predictable, piecewise continuously differentiable Cameron-Martin vector field X such that $F = \mathbb{E}[F] + D^*X$.*

Proof. Just follow the proof of Lemma 7.42 using Theorem 7.47 in place of Corollary 7.20. ■

7.7.1. *The equivalence of integration by parts and the representation formula.*

Corollary 7.50. *The representation formula in Theorem 7.47 may be used to prove the integration by parts Theorem 7.32 in the case F is a cylinder function.*

Proof. Let F be a cylinder function, a_s be as in Eq. (7.52), h be an adapted Cameron-Martin process and $k_s := (Q_s^{\mathrm{tr}})^{-1}h_s$. Then, by the product rule and Eq. (7.39),

$$h_s' + \frac{1}{2}\operatorname{Ric}_{//_s}h_s = \left(\frac{d}{ds} + \frac{1}{2}\operatorname{Ric}_{//_s}\right)Q_s^{\mathrm{tr}}k_s = Q_s^{\mathrm{tr}}k_s'.$$

Hence,

$$\begin{aligned}&\mathbb{E}\left[F\int_0^1\langle h_s' + \frac{1}{2}\operatorname{Ric}_{//_s}h_s, db_s\rangle\right] \\ &= \mathbb{E}\left[\left(\mathbb{E}F + \int_0^1\langle a_s, db_s\rangle\right)\int_0^1\langle Q_s^{\mathrm{tr}}k_s', db_s\rangle\right] \\ &= \mathbb{E}\left[\int_0^1\langle Q_s^{\mathrm{tr}}k_s', a_s\rangle ds\right] \\ &= \mathbb{E}\left[\int_0^1\langle Q_s^{\mathrm{tr}}k_s', \sum_{i=1}^n 1_{s\le s_i}Q_s^{-1}Q_{s_i}//_{s_i}^{-1}\operatorname{grad}_i f(\Sigma_{s_1},\ldots,\Sigma_{s_n})\rangle ds\right] \\ &= \mathbb{E}\left[\int_0^1\langle k_s', \sum_{i=1}^n 1_{s\le s_i}Q_{s_i}//_{s_i}^{-1}\operatorname{grad}_i f(\Sigma_{s_1},\ldots,\Sigma_{s_n})\rangle ds\right] \\ &= \mathbb{E}\left[\sum_{i=1}^n\langle k_{s_i}, Q_{s_i}//_{s_i}^{-1}\operatorname{grad}_i f(\Sigma_{s_1},\ldots,\Sigma_{s_n})\rangle\right] \\ &= \mathbb{E}\left[\sum_{i=1}^n\langle //_{s_i}h_{s_i}, \operatorname{grad}_i f(\Sigma_{s_1},\ldots,\Sigma_{s_n})\rangle\right] = \mathbb{E}\left[X^hF\right].\end{aligned}$$

■

Conversely we may give a proof of Theorem 7.47 which is based on the integration by parts Theorem 7.32.

Theorem 7.51 (Representation Formula)**.** *Suppose F is a cylinder function on $W(M)$ as in Eq. (7.2) and $\xi_s := //_s^{-1}\frac{d}{ds}(DF)_s$, then*

$$F = \mathbb{E}F + \int_0^1\left\langle\mathbb{E}\left[\xi_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r\operatorname{Ric}_{//_r}\xi_r dr\middle|\mathcal{F}_s\right], db_s\right\rangle. \tag{7.64}$$

where Q_s is the solution to Eq. (6.1).

Proof. Let $h \in \mathcal{X}_a$ be a predictable adapted Cameron-Martin valued process such that $\mathbb{E}\int_0^1|h_s'|^2 ds < \infty$. By the martingale representation property in Corollary 7.20,

$$F = \mathbb{E}F + \int_0^1\langle a, db\rangle \tag{7.65}$$

for some predictable process a such that $\mathbb{E}\int_0^1|a_s|^2 ds < \infty$. Then from Corollary 7.35 and the Itô isometry property,

$$\begin{aligned}\mathbb{E}\left[X^{Q^{\mathrm{tr}}h}F\right] &= \mathbb{E}\left[F\cdot\left(X^{Q^{\mathrm{tr}}h}\right)^*1\right] = \mathbb{E}\left[F\cdot\int_0^1\langle Q^{\mathrm{tr}}h', db\rangle\right] \\ &= \mathbb{E}\left[\int_0^1\langle Q_s^{\mathrm{tr}}h_s', a_s\rangle ds\right] = \mathbb{E}\left[\int_0^1\langle h_s', Q_s a_s\rangle ds\right].\end{aligned} \tag{7.66}$$

On the other hand we may compute $\mathbb{E}\left[X^{Q^{\mathrm{tr}}h}F\right]$ as:

$$\mathbb{E}\left[X^{Q^{\mathrm{tr}}h}F\right] = \mathbb{E}\left[\langle DF, //Q^{\mathrm{tr}}h\rangle_H\right] = \mathbb{E}\int_0^1 \langle \xi_s, \frac{d}{ds}\left(Q^{\mathrm{tr}}h\right)_s\rangle ds$$
$$= \mathbb{E}\int_0^1 \left\langle \xi_s, Q_s^{\mathrm{tr}}h_s' - \frac{1}{2}\operatorname{Ric}_{//_s} Q_s^{\mathrm{tr}}h_s\right\rangle ds \tag{7.67}$$

where we have used Eq. (7.39) in the last equality. We will now rewrite the right side of Eq. (7.67) so that it has the same form as Eq. (7.66) To do this let $\rho_s := \frac{1}{2}\operatorname{Ric}_{//_s}$ and notice that

$$\int_0^1 \langle \xi_s, \rho_s Q_s^{\mathrm{tr}}h_s\rangle ds = \int_0^1 \left\langle Q_s\rho_s^*\xi_s, \left(\int_0^s h_r' dr\right)\right\rangle ds$$
$$= \int dr ds 1_{0\le r\le s\le 1}\langle Q_s\rho_s^*\xi_s, h_r'\rangle = \int_0^1 \left\langle \int_s^1 Q_r\rho_r^*\xi_r dr, h_s'\right\rangle ds$$

wherein the last equality we have interchanged the role of r and s. Using this result back in Eq. (7.67) implies

$$\mathbb{E}\left[X^{Q^{\mathrm{tr}}h}F\right] = \mathbb{E}\int_0^1 \left\langle Q_s\xi_s - \int_s^1 Q_r\rho_r^*\xi_r dr, h_s'\right\rangle ds. \tag{7.68}$$

and comparing this with Eq. (7.66) shows

$$\mathbb{E}\int_0^1 \left\langle Q_s a_s - Q_s\xi_s + \int_s^1 Q_r\rho_r^*\xi_r dr, h_s'\right\rangle ds = 0 \tag{7.69}$$

for all $h \in \mathcal{X}_a$.

Up to now we have only used $F \in \mathcal{D}(D)$ and not the fact that F is a cylinder function. We will use this hypothesis now. From the easy part of Theorem 7.47 we know that a_s satisfies the additional properties of being 1) bounded, 2) zero if $s \ge s_n$ and most importantly 3) $s \to a_s$ is continuous off the partition set, $\{s_1, \dots, s_n\}$.

Fix $\tau \in (0,1)\backslash\{s_1,\dots,s_n\}$, $v \in T_oM$ and let G be a bounded $\mathcal{F}_\tau$ – measurable function. For $n \in \mathbb{N}$ let

$$l_n(s) := \int_0^s n 1_{\tau\le r\le \tau+\frac{1}{n}} dr.$$

Replacing h in Eq. (7.69) by $h_n(s) := G\cdot l_n(s)v$ and then passing to the limit as $n \to \infty$, implies

$$0 = \lim_{n\to\infty}\mathbb{E}\int_0^1 \left\langle Q_s a_s - Q_s\xi_s + \int_s^1 Q_r\rho_r^*\xi_r dr, h_n'(s)\right\rangle ds$$
$$= \mathbb{E}\left[G\left\langle Q_\tau a_\tau - Q_\tau\xi_\tau + \int_\tau^1 Q_r\rho_r^*\xi_r dr, v\right\rangle\right]$$

and since G and v were arbitrary we conclude from this equation that

$$\mathbb{E}\left[Q_\tau\xi_\tau - \int_\tau^1 Q_r\rho_r\xi_r dr\middle|\mathcal{F}_\tau\right] = Q_\tau a_\tau.$$

Thus for all but finitely many $s \in [0,1]$,

$$a_s = Q_s^{-1}\mathbb{E}\left[Q_s\xi_s - \int_s^1 Q_r\rho_r\xi_r dr\middle|\mathcal{F}_s\right]$$
$$= \mathbb{E}\left[\xi_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r\operatorname{Ric}_{//_r}\xi_r dr\middle|\mathcal{F}_s\right].$$

Combining this with Eq. (7.65) proves Eq. (7.64). ■

7.8. **Logarithmic-Sobolev Inequality for $W(M)$.** The next theorem is the "curved" generalization of Theorem 7.24.

Theorem 7.52 (Hsu's Logarithmic Sobolev Inequality)**.** *Let M be a compact Riemannian manifold, then for all $F \in \mathcal{D}(\bar{D})$*

$$\mathbb{E}\left[F^2\log F^2\right] \le \mathbb{E}F^2\cdot\log\mathbb{E}F^2$$
$$+ 2\mathbb{E}\int_0^1 \left|//_s^{-1}(DF)_s' - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r\operatorname{Ric}_{//_r}//_r^{-1}(DF)_r' dr\right|^2 ds, \tag{7.70}$$

where $(DF)_s' := \frac{d}{ds}(DF)_s$. Moreover, there is a constant $C = C(\operatorname{Ric})$ such that

$$\mathbb{E}\left[F^2\log F^2\right] \le C\mathbb{E}\left[\langle DF, DF\rangle_{H(T_oM)}\right] + \mathbb{E}F^2\cdot\log\mathbb{E}F^2. \tag{7.71}$$

Proof. The proof we give here follows the paper of Capitaine, Hsu and Ledoux [29]. We begin in the same way as the proof of Theorem 7.24. Let $F \in \mathcal{F}C^1(W(M))$, $\varepsilon > 0$, $H_\varepsilon := F^2 + \varepsilon \in \mathcal{D}(\bar{D})$ and

$$a_s := \mathbb{E}\left[\xi_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r\operatorname{Ric}_{//_r}\xi_r dr\middle|\mathcal{F}_s\right]$$

where

$$\xi_s = //_s^{-1}\frac{d}{ds}(DH_\varepsilon)_s = 2F\cdot //_s^{-1}\frac{d}{ds}(DF)_s.$$

Then by Theorem 7.47,

$$H_\varepsilon = \mathbb{E}H_\varepsilon + \int_0^1 \langle a, db\rangle.$$

The same proof used to derive Eq. (7.23) shows, with $\phi(x) = x\ln x$,

$$\mathbb{E}\left[\phi(H_\varepsilon)\right] = \mathbb{E}\left[\phi(M_1)\right] = \phi(\mathbb{E}M_1) + \frac{1}{2}\mathbb{E}\left[\int_0^1 \frac{1}{M_s}|a_s|^2 ds\right]$$
$$= \phi(\mathbb{E}H_\varepsilon) + \frac{1}{2}\mathbb{E}\left[\int_0^1 \frac{1}{\mathbb{E}\left[H_\varepsilon|\mathcal{F}_s\right]}|a_s|^2 ds\right].$$

By the Cauchy-Schwarz inequality and the contractive properties of conditional expectations,

$$\begin{aligned}|a_s|^2 &= \left|\mathbb{E}\left[2F\left\{//_s^{-1}(DF)'_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r \operatorname{Ric}_{//_r} //_r^{-1}(DF)'_r\,dr\right\}\middle|\mathcal{F}_s\right]\right|^2 \\ &\le 4\mathbb{E}\left[F^2|\mathcal{F}_s\right]\cdot\mathbb{E}\left[\left|//_s^{-1}(DF)'_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r \operatorname{Ric}_{//_r} //_r^{-1}(DF)'_r\,dr\right|^2\middle|\mathcal{F}_s\right]\end{aligned}$$

Combining the last two equations along with Eq. (7.24) implies

$$\begin{aligned}\mathbb{E}\phi(H_\varepsilon) \le&\phi(\mathbb{E}H_\varepsilon) \\ &+ 2\mathbb{E}\int_0^1 \mathbb{E}\left[\left|//_s^{-1}(DF)'_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r \operatorname{Ric}_{//_r} //_r^{-1}(DF)'_r\,dr\right|^2\middle|\mathcal{F}_s\right]ds \\ =&\phi(\mathbb{E}H_\varepsilon) \\ &+ 2\mathbb{E}\int_0^1 \left|//_s^{-1}(DF)'_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r \operatorname{Ric}_{//_r} //_r^{-1}(DF)'_r\,dr\right|^2 ds.\end{aligned}$$

We may now let $\varepsilon \downarrow 0$ in this inequality to learn Eq. (7.70) holds for all $F \in \mathcal{F}C^1(W)$. By compactness of M, Ric_m is bounded on M and so by simple Gronwall type estimates on Q and Q^{-1}, there is a non-random constant $K < \infty$ such that

$$\left\|Q_s^{-1}Q_r \operatorname{Ric}_{//_r}\right\|_{op} \le K \text{ for all } r, s.$$

Therefore,

$$\begin{aligned}&\left|//_s^{-1}(DF)'_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r \operatorname{Ric}_{//_r} //_r^{-1}(DF)'_r\,dr\right|^2 \\ &\le \left[\left|(DF)'_s\right| + \frac{1}{2}K\int_0^1 \left|(DF)'_s\right| ds\right]^2 \\ &\le 2\left|(DF)'_s\right|^2 + \frac{1}{2}K^2\left[\int_0^1 \left|(DF)'_s\right| ds\right]^2 \\ &\le 2\left|(DF)'_s\right|^2 + \frac{1}{2}K^2\int_0^1 \left|(DF)'_s\right|^2 ds\end{aligned}$$

and hence

$$\begin{aligned}&2\mathbb{E}\int_0^1 \left|(DF)'_s - \frac{1}{2}\int_s^1 Q_s^{-1}Q_r \operatorname{Ric}_{//_r}(DF)'_r\,dr\right|^2 ds \\ &\le \left(4+K^2\right)\int_0^1 \left|(DF)'_s\right|^2 ds.\end{aligned}$$

Combining this estimate with Eq. (7.70) implies Eq. (7.71) holds with $C = \left(4+K^2\right)$. Again, since $\mathcal{F}C^1(W)$ is a core for $\bar{D}$, standard limiting arguments show that Eq. (7.70) and Eq. (7.71) are valid for all $F \in \mathcal{D}\left(\bar{D}\right)$. ■

Theorem 7.52 was first proved by Hsu [98] with an independent proof given shortly thereafter by Aida and Elworthy [4]. Hsu's original proof relied on a Markov dependence version of a standard additivity property for logarithmic Sobolev inequalities and makes key use of Corollary 7.37. On the other hand Aida and Elworthy show, using the projection construction of Brownian motion, the logarithmic Sobolev inequality on $W(M)$ is a consequence of Gross' [92] original logarithmic Sobolev inequality on the classical Wiener space $W(\mathbb{R}^N)$, see Theorem 7.24. In Aida's and Elworthy's proof, Theorem 5.43 plays an important role.

7.9. **More References.** Many people have now proved some version of integration by parts for path and loop spaces in one context or another, see for example [21, 28, 32, 26, 28, 27, 47, 48, 49, 76, 75, 78, 85, 122, 128, 146, 161, 159, 160, 163, 102]. We have followed Bismut in these notes who proved integration by parts formulas for cylinder functions depending on one time. However, as is pointed out by Leandre and Malliavin and Fang, Bismut's technique works with out any essential change for arbitrary cylinder functions. In [47, 48], the flow associated to a general class of vector fields on paths and loop spaces of a manifold were constructed. The reader is also referred to the texts [71, 100, 171] and the related articles [81, 80, 35, 77, 82, 83, 84, 34, 37, 33, 38, 36, 39, 125].

Many of the results in this section extend to pinned Wiener measure on loop spaces, see [48] for example. Loop spaces are more interesting than path spaces since they have nontrivial topology, The issue of the spectral gap and logarithmic Sobolev inequalities for general loop spaces is still an open problem. In [93], Gross has prove a logarithmic Sobolev inequality on Loop groups with an added "potential term" for a special geometry on loop groups. Here Gross uses pinned Wiener measure as the reference measure. In Driver and Lohrenz [54], it is shown that a logarithmic Sobolev inequality **without** a potential term does hold on the Loop group provided one replace pinned Wiener measure by a "heat kernel" measure. The quasi-invarariance properties of the heat kernel measure on loop groups was first established in [50, 51]. For more results on heat kernel measures on the loop groups see for example, [57, 3, 30, 31, 82, 83, 106].

The question as to when or if the potential is needed in Gross's setting for logarithmic Sobolev inequalities is still an open question, but see Gong, Röckner and Wu [89] for a positive result in this direction. Eberle [59, 60, 61, 62] has provided examples of Riemannian manifolds where the spectral gap inequality fails in the loop space setting. The reader is referred to [52, 53] and the references therein for some more perspective on the stochastic analysis on loop spaces.

8. Malliavin's Methods for Hypoelliptic Operators

In this section we will be concerned with determining smoothness properties of the Law(Σ_t) where Σ_t denotes the solution to Eq. (5.1) with $\Sigma_0 = o$ and $\beta = B$ being an $\mathbb{R}^n$ – valued Brownian motion. Unlike the previous sections in these notes, the map $\mathbf{X}(m) : \mathbb{R}^n \to T_mM$ is **not** assumed to be surjective. Equivalently put, the diffusion generator $L := \frac{1}{2}\sum_{i=1}^n X_i^2 + X_0$ is no longer assumed to be elliptic. However we will always be assuming that the vector fields $\{X_i\}_{i=0}^n$ satisfy Hörmander's restricted bracket condition at $o \in M$ as in Definition 8.1. Let $\mathcal{K}_1 := \{X_1, \dots, X_n\}$ and $\mathcal{K}_l$ be defined inductively by

$$\mathcal{K}_{l+1} = \{[X_i, K] : K \in \mathcal{K}_l\} \cup \mathcal{K}_l.$$

For example

$$\begin{aligned}\mathcal{K}_2 &= \{X_1, \dots, X_n\} \cup \{[X_j, X_i] : i, j = 1, \dots, n\} \text{ and}\\ \mathcal{K}_3 &= \{X_1, \dots, X_n\} \cup \{[X_j, X_i] : i, j = 1, \dots, n\}\\ &\quad \cup \{[X_k, [X_j, X_i]] : i, j, k = 1, \dots, n\} \text{ etc.}\end{aligned}$$

Definition 8.1. The collection of vector fields, $\{X_i\}_{i=0}^n \subset \Gamma(TM)$, satisfies **Hörmander's restricted bracket condition at** $m \in M$ if there exist $l \in \mathbb{N}$ such that

$$\mathrm{span}(\{K(m) : K \in \mathcal{K}_l\}) = T_mM.$$

Under this condition it follows from a classical theorem of Hörmander that solutions to the heat equation $\partial_t u = Lu$ are necessarily smooth. Since the fundamental solution to this equation at $o \in M$ is the law of the process Σ_t, it follows that the Law(Σ_t) is absolutely continuous relative to the volume measure λ on M and its Radon-Nikodym derivative is a smooth function on M. Malliavin, in his 1976 pioneering paper [130], gave a probabilistic proof of this fact. Malliavin's original paper was followed by an avalanche of papers carrying out and extending Malliavin's program including the fundamental works of Stroock [169, 170, 168], Kusuoka and Stroock [121, 119, 120], and Bismut [21]. See also [13, 12, 23, 55, 104, 132, 152, 136, 147, 148, 157, 158, 179] (and the references therein) along with Bell's article in this volume. The purpose of this section is to briefly explain (omitting some details) Malliavin methods.

8.1. Malliavin's Ideas in Finite Dimensions.

To understand Malliavin's methods it is best to begin with a finite dimensional analogue.

Theorem 8.2 (Malliavin's Ideas in Finite Dimensions)**.** *Let $W = \mathbb{R}^N$, μ be the Gaussian measure on W defined by*

$$d\mu(x) := (2\pi)^{-N/2} e^{-\frac{1}{2}|x|^2} dm(x).$$

Further suppose $F : W \to \mathbb{R}^d$ (think $F = \Sigma_t$) is a function satisfying:

(1) F is smooth and all of its partial derivatives are in
$$L^{\infty-}(\mu) := \cap_{1 \le p < \infty} L^p(W, \mu).$$
(2) F is a submersion or equivalently assume the "Malliavin" matrix
$$C(\omega) := DF(\omega)DF(\omega)^*$$
is invertible for all $\omega \in W$.
(3) Let
$$\Delta(\omega) := \det C(\omega) = \det(DF(\omega)DF(\omega)^*)$$
and assume $\Delta^{-1} \in L^{\infty-}(\mu)$.

Then the law $(\mu_F = F_*\mu = \mu \circ F^{-1})$ of F is absolutely continuous relative to Lebesgue measure, λ, on $\mathbb{R}^d$ and the Radon-Nikodym derivative, $\rho := d\mu_F/d\lambda$, is smooth.

Proof. For each vector field $Y \in \Gamma(T\mathbb{R}^d)$, define

$$\mathbb{Y}(\omega) = DF(\omega)^* C(\omega)^{-1} Y(F(\omega)) \tag{8.1}$$

— a smooth vector field on W such that $DF(\omega)\mathbb{Y}(\omega) = Y(F(\omega))$ or in more geometric notation,

$$F_*\mathbb{Y}(\omega) = Y(F(\omega)). \tag{8.2}$$

For the purposes of this proof, it is sufficient to restrict our attention to the case where Y is a constant vector field.

Explicit computations using the chain rule and Cramer's rule for computing $C(\omega)^{-1}$ shows that $D^k\mathbb{Y}$ may be expressed as a polynomial in Δ^{-1} and $D^\ell F$ for $\ell = 0, 1, 2 \dots, k$. In particular $D^k\mathbb{Y}$ is in $L^{\infty-}(\mu)$. Suppose $f, g : W \to \mathbb{R}$ are C^1 functions such that f, g, and their first order derivatives are in $L^{\infty-}(\mu)$. Then by a standard truncation argument and integration by parts, one shows that

$$\int_W (\mathbb{Y}f) g \, d\mu = \int_W f(\mathbb{Y}^* g) \, d\mu,$$

where

$$\mathbb{Y}^* = -\mathbb{Y} + \delta(\mathbb{Y}) \text{ and } \delta(\mathbb{Y})(\omega) := -\operatorname{div}(\mathbb{Y})(\omega) + \mathbb{Y}(\omega) \cdot \omega.$$

Suppose that $\phi \in C_c^\infty(\mathbb{R}^d)$ and $Y_i \in \mathbb{R}^d \subset \Gamma(\mathbb{R}^d)$, then from Eq. (8.2) and induction,

$$(Y_1Y_2 \cdots Y_k\phi)(F(\omega)) = (\mathbb{Y}_1\mathbb{Y}_2 \cdots \mathbb{Y}_k(\phi \circ F))(\omega)$$

and therefore,

$$\begin{aligned}\int_{\mathbb{R}^d} (Y_1Y_2 \cdots Y_k\phi) d\mu_F &= \int_W (Y_1Y_2 \cdots Y_k\phi)(F(\omega)) \, d\mu(\omega)\\ &= \int_W (\mathbb{Y}_1\mathbb{Y}_2 \cdots \mathbb{Y}_k(\phi \circ F))(\omega) \, d\mu(\omega)\\ &= \int_W \phi(F(\omega)) \cdot (\mathbb{Y}_k^*\mathbb{Y}_{k-1}^* \cdots \mathbb{Y}_1^* 1)(\omega) \, d\mu(\omega).\end{aligned} \tag{8.3}$$

By the remarks in the previous paragraph, $(\mathbb{Y}_k^*\mathbb{Y}_{k-1}^*\cdots\mathbb{Y}_1^*1)\in L^{\infty-}(\mu)$ which along with Eq. (8.3) shows

$$\left|\int_{\mathbb{R}^d}(Y_1Y_2\cdots Y_k\phi)d\mu_F\right|\le C\,\|\phi\|_{L^\infty(\mathbb{R}^d)},$$

where $C=\left\|\mathbb{Y}_k^*\mathbb{Y}_{k-1}^*\cdots\mathbb{Y}_1^*1\right\|_{L^1(\mu)}<\infty$. It now follows from Sobolev imbedding theorems or simple Fourier analysis that $\mu_F\ll\lambda$ and that $\rho:=d\mu_F/d\lambda$ is a smooth function. ∎

The remainder of Section 8 will be devoted to an infinite dimensional analogue of Theorem 8.2 (see Theorem 8.9) where $\mathbb{R}^d$ is replaced by a manifold M^d,

$$W:=\{\omega\in C([0,\infty),\mathbb{R}^n):\omega(0)=0\},$$

μ is taken to be Wiener measure on W, $B_t:W\to\mathbb{R}^n$ be defined by $B_t(\omega)=\omega_t$ and $F:=\Sigma_t:W(\mathbb{R}^n)\to M$ is a solution to Eq. (5.1) with $\Sigma_0=o\in M$ and $\beta=B$. Recall that μ is the unique measure on $\mathcal{F}:=\sigma(B_t:t\in[0,\infty))$ such that $\{B_t\}_{t\ge0}$ is a Brownian motion. I am now using t as the dominant parameter rather than s to be in better agreement with the literature on this subject.

8.2. Smoothness of Densities for Hörmander Type Diffusions .

For simplicity of the exposition, it will be assumed that M^d is a compact Riemannian manifold of dimension d. However this can and should be relaxed. For example most everything we are going to say would work if M is an imbedded submanifold in $\mathbb{R}^N$ and the vector fields $\{X_i\}_{i=0}^n$ are the restrictions of smooth vector fields on $\mathbb{R}^N$ whose partial derivatives to any order greater than 0 are all bounded.

Remark 8.3. The choice of Riemannian metric here is somewhat arbitrary and is an artifact of the method to be described below. It is the author's belief that this issue has still not been adequately addressed in the literature.

To abbreviate the notation, let

$$H=\left\{h\in W:\langle h,h\rangle_H:=\int_0^\infty\left|\dot h(t)\right|^2dt<\infty\right\}$$

and $D\Sigma_t:H\to T_{\Sigma_t}M$ be defined by $(D\Sigma_t)h:=\partial_hT_t^B(o)$ as defined Theorem 7.26. Recall from Theorem 7.26 that

$$(D\Sigma_t)h:=Z_t\int_0^tZ_\tau^{-1}\mathbf{X}(\Sigma_\tau)\dot h_\tau d\tau=//_tz_t\int_0^tz_\tau^{-1}//_\tau^{-1}\mathbf{X}(\Sigma_\tau)\dot h_\tau d\tau,\tag{8.4}$$

where $\dot h_\tau:=\frac{d}{d\tau}h_\tau$, $Z_t:=\left(T_t^B\right)_{*o}:T_oM\to T_{\Sigma_t}M$, $//_t$ is stochastic parallel translation along Σ and $z_t:=//_t^{-1}Z_t$. In the sequel, adjoints will be denote by either " $*$ " or " tr " with the former being used if an infinite dimensional space is involved and the latter if all spaces involved are finite dimensional.

Definition 8.4 (Reduced Malliavin Covariance)**.** The End (T_oM) – valued random variable,

$$\bar C_t:=\int_0^tZ_\tau^{-1}\mathbf{X}(\Sigma_\tau)\mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}\left(Z_\tau^{-1}\right)^{\mathrm{tr}}d\tau\tag{8.5}$$

$$=\int_0^tz_\tau^{-1}//_\tau^{-1}\mathbf{X}(\Sigma_\tau)\mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}//_\tau\left(z_\tau^{-1}\right)^{\mathrm{tr}}d\tau,\tag{8.6}$$

will be called the **reduced Malliavin covariance matrix.**

Theorem 8.5. *The adjoint, $(D\Sigma_t)^*:T_{\Sigma_t}M\to H$, of the map $D\Sigma_t$ is determined by*

$$\frac{d}{d\tau}\left[(D\Sigma_t)^*//_tv\right]_\tau=1_{\tau\le t}\mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}//_\tau\left(z_tz_\tau^{-1}\right)^{\mathrm{tr}}v\tag{8.7}$$

for all $v\in T_oM$. The Malliavin covariance matrix $C_t:=D\Sigma_t(D\Sigma_t)^*:T_{\Sigma_t}M\to T_{\Sigma_t}M$ is given by $C_t=Z_t\bar C_tZ_t^{\mathrm{tr}}$ or equivalently

$$C_t=D\Sigma_t(D\Sigma_t)^*=//_tz_t\bar C_tz_t^{\mathrm{tr}}//_t^{-1}.\tag{8.8}$$

Proof. Using Eq. (8.4),

$$\begin{aligned}\langle D\Sigma_th,//_tv\rangle_{T_{\Sigma_t}M}&=\left\langle Z_t\int_0^tZ_\tau^{-1}\mathbf{X}(\Sigma_\tau)\dot h_\tau d\tau,//_tv\right\rangle_{T_{\Sigma_t}M}\\&=\left\langle//_tz_t\int_0^tz_\tau^{-1}//_\tau^{-1}\mathbf{X}(\Sigma_\tau)\dot h_\tau d\tau,//_tv\right\rangle_{T_{\Sigma_t}M}\\&=\int_0^t\left\langle z_tz_\tau^{-1}//_\tau^{-1}\mathbf{X}(\Sigma_\tau)\dot h_\tau,v\right\rangle_{T_oM}d\tau\\&=\int_0^t\left\langle\dot h_\tau,\mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}//_\tau\left(z_tz_\tau^{-1}\right)^{\mathrm{tr}}v\right\rangle_{\mathbb{R}^n}d\tau\end{aligned}\tag{8.9}$$

which implies Eq. (8.7). Combining Eqs. (8.4) and (8.7), using

$$Z_\tau^{\mathrm{tr}}=(//_\tau z_\tau)^{\mathrm{tr}}=z_\tau^{\mathrm{tr}}//_\tau^{\mathrm{tr}}=z_\tau^{\mathrm{tr}}//_\tau^{-1},$$

shows

$$\begin{aligned}D\Sigma_t(D\Sigma_t)^*//_tv&=Z_t\int_0^tZ_\tau^{-1}\mathbf{X}(\Sigma_\tau)\mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}//_\tau\left(z_tz_\tau^{-1}\right)^{\mathrm{tr}}vd\tau\\&=Z_t\int_0^tZ_\tau^{-1}\mathbf{X}(\Sigma_\tau)\mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}\left(Z_\tau^{-1}\right)^{\mathrm{tr}}Z_t^{\mathrm{tr}}//_tvd\tau.\end{aligned}$$

Therefore,

$$C_t=Z_t\bar C_tZ_t^{\mathrm{tr}}=//_tz_t\bar C_tz_t^{\mathrm{tr}}//_t^{-1}$$

from which Eq. (8.8) follows. ∎

The next crucial theorem is at the heart of Malliavin's method and constitutes the deepest part of the theory. The proof of this theorem will be postponed until Section 8.4 below.

Theorem 8.6 (Non-degeneracy of $\bar{C}_t$)**.** *Let $\bar{\Delta}_t := \det\left(\bar{C}_t\right)$. If Hörmander's restricted bracket condition at $o \in M$ holds then $\bar{\Delta}_t > 0$ a.e. (i.e. $\bar{C}_t$ is invertible a.e.) and moreover $\bar{\Delta}_t^{-1} \in L^{\infty-}(\mu)$.*

Following the general strategy outlined in Theorem 8.2, given a vector field $Y \in \Gamma(TM)$ we wish to lift it via the map $\Sigma_t : W \to M$ to a vector field $\mathbb{Y}^t$ on $W := W(\mathbb{R}^n)$. According to the prescription used in Eq. (8.1) in Theorem 8.2,

$$\mathbb{Y}^t := (D\Sigma_t)^* \left(D\Sigma_t (D\Sigma_t)^*\right)^{-1} Y(\Sigma_t) = (D\Sigma_t)^* C_t^{-1} Y(\Sigma_t) \in H. \tag{8.10}$$

From Eq. (8.8)

$$C_t^{-1} = //_t \left(z_t^{\mathrm{tr}}\right)^{-1} \bar{C}_t^{-1} z_t^{-1} //_t^{-1}$$

and combining this with Eq. (8.10), using Eq. (8.7), implies

$$\begin{aligned}
\frac{d}{d\tau}\mathbb{Y}_\tau^t &= 1_{\tau\le t}\frac{d}{d\tau}\left[(D\Sigma_t) //_t \left(z_t^{\mathrm{tr}}\right)^{-1} \bar{C}_t^{-1} z_t^{-1} //_t^{-1} Y(\Sigma_t)\right]_\tau \\
&= 1_{\tau\le t}\mathbf{X}(\Sigma_\tau)^{\mathrm{tr}} //_\tau \left(z_t z_\tau^{-1}\right)^{\mathrm{tr}} \left(z_t^{\mathrm{tr}}\right)^{-1} \bar{C}_t^{-1} z_t^{-1} //_t^{-1} Y(\Sigma_t) \\
&= 1_{\tau\le t}\mathbf{X}(\Sigma_\tau)^{\mathrm{tr}} //_\tau \left(z_\tau^{-1}\right)^{\mathrm{tr}} \bar{C}_t^{-1} Z_t^{-1} Y(\Sigma_t) \\
&= 1_{\tau\le t}\mathbf{X}(\Sigma_\tau)^{\mathrm{tr}} \left(Z_\tau^{-1}\right)^{\mathrm{tr}} \bar{C}_t^{-1} Z_t^{-1} Y(\Sigma_t).
\end{aligned}$$

Hence, the formula for $\mathbb{Y}^t$ in Eq. (8.10) may be explicitly written as

$$\mathbb{Y}_s^t = \left[\int_0^{s\wedge t} \left(Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\right)^{\mathrm{tr}} d\tau\right] \bar{C}_t^{-1} Z_t^{-1} Y(\Sigma_t). \tag{8.11}$$

The reader should observe that the process $s \to \mathbb{Y}_s^t$ is non-adapted since $\bar{C}_t^{-1} Z_t^{-1} Y(\Sigma_t)$ depends on the entire path of Σ up to time t.

Theorem 8.7. *Let $Y \in \Gamma(TM)$ and $\mathbb{Y}^t$ be the non-adpated Cameron-Martin process defined in Eq. (8.11). Then $\mathbb{Y}^t$ is "Malliavin smooth," i.e. $\mathbb{Y}^t$ is H – differentiable (in the sense of Theorem 7.14) to all orders with all differentials being in $L^{\infty-}(\mu)$, see Nualart [148] for more precise definitions. Moreover if $f \in C^\infty(M)$, then $f(\Sigma_t)$ is Malliavin smooth and*

$$\langle \bar{D}[f(\Sigma_t)], \mathbb{Y}^t\rangle_H = Yf(\Sigma_t) \tag{8.12}$$

where $\bar{D}$ is the closure of the gradient operator defined in Corollary 7.16.

Proof. We only sketch the proof here and refer the reader to [147, 12, 148] with regard to some of the technical details which are omitted below. Let $\{e_i\}_{i=1}^d$ be an orthonormal basis for T_oM, then

$$\mathbb{Y}_s^t = \sum_{i=1}^d \left\langle e_i, \bar{C}_t^{-1} Z_t^{-1} Y(\Sigma_t)\right\rangle \int_0^s \left(Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\right)^{\mathrm{tr}} e_i d\tau = \sum_{i=1}^d a_i h_s^i \tag{8.13}$$

where

$$a_i := \left\langle e_i, \bar{C}_t^{-1} Z_t^{-1} Y(\Sigma_t)\right\rangle \text{ and } h_s^i := \int_0^{s\wedge t} \left(Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\right)^{\mathrm{tr}} e_i d\tau.$$

It is well known that solutions to stochastic differential equations with smooth coefficients are Malliavin smooth from which it follows that h^i, $Z_t^{-1}Y(\Sigma_t)$, and $\bar{C}_t$ are Malliavin smooth. It also follows from the general theory, under the conclusion of Theorem 8.6, that $\bar{C}_t^{-1}$ is Malliavin smooth and hence so are each the functions a_i for $i = 1, \dots d$. Therefore, $\mathbb{Y}^t = \sum_{i=1}^d a_i h^i$ is Malliavin smooth as well and in particular $\mathbb{Y}^t \in \mathcal{D}(D^*)$. It now only remains to verify Eq. (8.12).

Let h be a non-random element of H. Then from Theorems 7.14, 7.15, 7.26 and the chain rule for Wiener calculus,

$$\begin{aligned}
\mathbb{E}[f(\Sigma_t)\cdot D^*h] &= \mathbb{E}[\partial_h[f(\Sigma_t)]] = \mathbb{E}[df(D\Sigma_t h)] \\
&= \mathbb{E}\left[df\left(Z_t\int_0^t Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\dot{h}_\tau d\tau\right)\right] \\
&= \mathbb{E}\left[\left\langle \vec{\nabla} f(\Sigma_t), Z_t\int_0^t Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\dot{h}_\tau d\tau\right\rangle_{T_{\Sigma_t}M}\right] \\
&= \mathbb{E}\left[\int_0^t \left\langle \mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}\left(Z_\tau^{-1}\right)^{\mathrm{tr}} Z_t^{\mathrm{tr}}\vec{\nabla} f(\Sigma_t)\vec{\nabla} f(\Sigma_t), \dot{h}_\tau\right\rangle_{\mathbb{R}^n} d\tau\right]
\end{aligned}$$

from which we conclude that $f(\Sigma_t) \in \mathcal{D}(D^{**}) = \mathcal{D}(\bar{D})$ and

$$\left(\bar{D}[f(\Sigma_t)]\right)_s = \int_0^{s\wedge t} \mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}\left(Z_\tau^{-1}\right)^{\mathrm{tr}} Z_t^{\mathrm{tr}}\vec{\nabla} f(\Sigma_t)\, d\tau.$$

From this formula and the definition of $\mathbb{Y}^t$ it follows that

$$\begin{aligned}
&\langle \bar{D}[f(\Sigma_t)], \mathbb{Y}^t\rangle_H \\
&= \int_0^t \left\langle \mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}\left(Z_\tau^{-1}\right)^{\mathrm{tr}} Z_t^{\mathrm{tr}}\vec{\nabla} f(\Sigma_t), \mathbf{X}(\Sigma_\tau)^{\mathrm{tr}}\left(Z_\tau^{-1}\right)^{\mathrm{tr}} \bar{C}_t^{-1} Z_t^{-1} Y(\Sigma_t)\right\rangle d\tau \\
&= \left\langle \vec{\nabla} f(\Sigma_t), Z_t\left(\int_0^t Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\left(Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\right)^{\mathrm{tr}} d\tau\right)\bar{C}_t^{-1} Z_t^{-1} Y(\Sigma_t)\right\rangle \\
&= \left\langle \vec{\nabla} f(\Sigma_t), Z_t\bar{C}_t\bar{C}_t^{-1} Z_t^{-1} Y(\Sigma_t)\right\rangle = \left\langle \vec{\nabla} f(\Sigma_t), Y(\Sigma_t)\right\rangle \\
&= (Yf)(\Sigma_t).
\end{aligned}$$

■

Notation 8.8. Let $\mathbb{Y}^t$ act on Malliavin smooth functions by the formula, $\mathbb{Y}^t F := \left\langle \bar{D}F, \mathbb{Y}^t\right\rangle_H$ and let $(\mathbb{Y}^t)^*$ denote the $L^2(\mu)$ – adjoint of $\mathbb{Y}^t$.

With this notation, Theorem 8.7 asserts that

$$\mathbb{Y}^t[f(\Sigma_t)] = (Yf)(\Sigma_t). \tag{8.14}$$

Now suppose $F, G : W \to \mathbb{R}$ are Malliavin smooth functions, then

$$\begin{aligned}
\mathbb{E}\left[\mathbb{Y}^t F\cdot G + F\cdot \mathbb{Y}^t G\right] &= \mathbb{E}\left[\mathbb{Y}^t[FG]\right] = \mathbb{E}\left[\langle \bar{D}[FG], \mathbb{Y}^t\rangle_H\right] \\
&= \mathbb{E}\left[F\cdot G D^*\mathbb{Y}^t\right]
\end{aligned}$$

from which it follows that $G \in \mathcal{D}\left((\mathbb{Y}^t)^*\right)$ and

$$\left(\mathbb{Y}^t\right)^* G = -\mathbb{Y}^t G + G D^* \mathbb{Y}^t. \tag{8.15}$$

From the general theory (see [148] for example), D^*U is Malliavin smooth if U is Malliavin smooth. In particular $(\mathbb{Y}^t)^* G$ is Malliavin smooth if G is Malliavin smooth.

Theorem 8.9 (Smoothness of Densities)**.** *Assume the restricted Hörmander condition holds at $o \in M$ (see Definition 8.1) and suppose $f \in C^\infty(M)$ and $\{Y_i\}_{i=1}^k \subset \Gamma(TM)$. Then*

$$\mathbb{E}\left[(Y_1 \dots Y_k f)(\Sigma_t)\right] = \mathbb{E}\left[\mathbb{Y}_1^t \dots \mathbb{Y}_k^t \left[f(\Sigma_t)\right]\right]$$
$$= \mathbb{E}\left[\left[f(\Sigma_t)\right]\left(\mathbb{Y}_k^t\right)^* \dots \left(\mathbb{Y}_1^t\right)^* 1\right]. \tag{8.16}$$

Moreover, the law of Σ_t is smooth.

Proof. By an induction argument using Eq. (8.14),

$$\mathbb{Y}_1^t \dots \mathbb{Y}_k^t \left[f(\Sigma_t)\right] = (Y_1 \dots Y_k f)(\Sigma_t)$$

from which Eq. (8.16) is a simple consequence. As has already been observed, $(\mathbb{Y}_k^t)^* \dots (\mathbb{Y}_1^t)^* 1$ is Malliavin smooth and in particular $(\mathbb{Y}_k^t)^* \dots (\mathbb{Y}_1^t)^* 1 \in L^1(\mu)$. Therefore it follows from Eq. (8.16) that

$$\left|\mathbb{E}\left[(Y_1 \dots Y_k f)(\Sigma_t)\right]\right| \le \left\|\left(\mathbb{Y}_k^t\right)^* \dots \left(\mathbb{Y}_1^t\right)^* 1\right\|_{L^1(\mu)} \|f\|_\infty. \tag{8.17}$$

Since the argument used in the proof of Theorem 8.2 after Eq. (8.16) is local in nature, it follows from Eq. (8.17) that the $\mathrm{Law}(\Sigma_t)$ has a smooth density relative to any smooth measure on M and in particular the Riemannian volume measure. ■

8.3. **The Invertibility of $\bar{C}_t$ in the Elliptic Case.** As a warm-up to the proof of the full version of Theorem 8.6 let us first consider the special case where $\mathbf{X}(m) : \mathbb{R}^n \to T_m M$ is surjective for all $m \in M$. Since M is compact this will imply there exists and $\varepsilon > 0$ such that

$$\mathbf{X}(m)\mathbf{X}^{\mathrm{tr}}(m) \ge \varepsilon I_{T_m M} \text{ for all } m \in M.$$

Notation 8.10. We will write $f(\varepsilon) = O(\varepsilon^{\infty-})$ if, for all $p < \infty$,

$$\lim_{\varepsilon \downarrow 0} \frac{|f(\varepsilon)|}{\varepsilon^p} = 0.$$

Proposition 8.11 (Elliptic Case)**.** *Suppose there is an $\varepsilon > 0$ such that*

$$\mathbf{X}(m)\mathbf{X}^{\mathrm{tr}}(m) \ge \varepsilon I_{T_m M}$$

for all $m \in M$, then $\left[\det\left(\bar{C}_t\right)\right]^{-1} \in L^{\infty-}(\mu)$.

Proof. Let $\delta \in (0,1)$ and

$$T_\delta := \inf\left\{t > 0 : |z_t - I_{T_o M}| > \delta\right\} \tag{8.18}$$

where, as usual,

$$z_t := //_t^{-1} Z_t = //_t^{-1}\left(T_t^B\right)_{*o}.$$

Since for all $a \in T_o M$,

$$\begin{aligned}
&\langle Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\mathbf{X}^{\mathrm{tr}}(\Sigma_\tau)\left(Z_\tau^{\mathrm{tr}}\right)^{-1} a, a\rangle \\
&\quad = \left\langle \mathbf{X}(\Sigma_\tau)\mathbf{X}^{\mathrm{tr}}(\Sigma_\tau)\left(Z_\tau^{\mathrm{tr}}\right)^{-1} a, \left(Z_\tau^{\mathrm{tr}}\right)^{-1} a\right\rangle \\
&\quad \ge \varepsilon\left\langle \left(Z_\tau^{\mathrm{tr}}\right)^{-1} a, \left(Z_\tau^{\mathrm{tr}}\right)^{-1} a\right\rangle = \varepsilon\left\langle a, Z_\tau^{\mathrm{tr}}\left(Z_\tau^{\mathrm{tr}}\right)^{-1} a\right\rangle,
\end{aligned}$$

we have

$$\begin{aligned}
&Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\mathbf{X}^{\mathrm{tr}}(\Sigma_\tau)\left(Z_\tau^{\mathrm{tr}}\right)^{-1} \\
&\quad \ge \varepsilon Z_\tau^{\mathrm{tr}}\left(Z_\tau^{\mathrm{tr}}\right)^{-1} = \varepsilon z_t^{\mathrm{tr}} //_t^{\mathrm{tr}}\left(//_t^{\mathrm{tr}}\right)^{-1}\left(z_t^{\mathrm{tr}}\right)^{-1} = \varepsilon z_t^{\mathrm{tr}}\left(z_t^{\mathrm{tr}}\right)^{-1}.
\end{aligned}$$

Hence

$$\begin{aligned}
\bar{C}_t &= \int_0^t Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\mathbf{X}^{\mathrm{tr}}(\Sigma_\tau)\left(Z_\tau^{\mathrm{tr}}\right)^{-1} d\tau \\
&\ge \varepsilon \int_0^t Z_\tau^{-1}\left(Z_\tau^{\mathrm{tr}}\right)^{-1} d\tau \ge \varepsilon \int_0^{t \wedge T_\delta} z_\tau^{\mathrm{tr}}\left(z_\tau^{\mathrm{tr}}\right)^{-1} d\tau
\end{aligned}$$

and therefore,

$$\bar{\Delta}_t = \det\left(\bar{C}_t\right) \ge \varepsilon^d \det\left(\int_0^{t \wedge T_\delta} z_\tau^{\mathrm{tr}}\left(z_\tau^{\mathrm{tr}}\right)^{-1} d\tau\right).$$

By choosing $\delta > 0$ sufficiently small we may arrange that

$$\left\|z_\tau^{\mathrm{tr}}\left(z_\tau^{\mathrm{tr}}\right)^{-1} - I\right\| \le 1/2$$

for all $\tau \le t \wedge T_\delta$ in which case

$$\int_0^{t \wedge T_\delta} z_\tau^{\mathrm{tr}}\left(z_\tau^{\mathrm{tr}}\right)^{-1} d\tau \ge \frac{1}{2} t \wedge T_\delta \cdot Id$$

and hence $\bar{\Delta}_t = \det\left(\bar{C}_t\right) \ge \varepsilon^d\left(\frac{1}{2} t \wedge T_\delta\right)^d$. From this it follows, with $q = p \cdot d$, that

$$\mathbb{E}\left[\bar{\Delta}_t^{-p}\right] \le 2^q \varepsilon^{-q} \mathbb{E}\left(\left(\frac{1}{t \wedge T_\delta}\right)^q\right).$$

Now,

$$\begin{aligned}
\mathbb{E}\left(\left(\frac{1}{t \wedge T_\delta}\right)^q\right) &= \mathbb{E}\left(-\int_{t \wedge T_\delta}^\infty \frac{d}{d\tau}\tau^{-q} d\tau\right) = \mathbb{E}\left(q\int_0^\infty 1_{t \wedge T_\delta \le \tau} \cdot \tau^{-q-1} d\tau\right) \\
&= q\int_0^\infty \tau^{-q-1} \mu\left(t \wedge T_\delta \le \tau\right) d\tau
\end{aligned}$$

which will be finite for all $q > 1$ iff $\mu(t \wedge T_\delta \le \tau) = \mu(T_\delta \le \tau) = O(\tau^k)$ as $\tau \downarrow 0$ for all $k > 0$.

By Chebyschev's inequalities and Eq. (9.10) of Proposition 9.5 below,

$$\mu(T_\delta \le \tau) = \mu\left(\sup_{s \le \tau} |z_s - I| > \delta\right) \le \delta^{-q} \mathbb{E}\left[\sup_{s \le \tau} |z_s - I|^q\right] = O(\tau^{q/2}). \tag{8.19}$$

Since $q \ge 2$ was arbitrary it follows that $\mu(T_\delta \le \tau) = O(\tau^{\infty-})$ which completes the proof. ∎

8.4. Proof of Theorem 8.6.

Notation 8.12. Let $S := \{v \in T_oM : \langle v, v\rangle = 1\}$, i.e. S is the unit sphere in T_oM.

Proof. (*Proof of Theorem 8.6.*) To show $\bar{C}_t^{-1} \in L^{\infty-}(\mu)$ it suffices to shows

$$\mu(\inf_{v \in S} \langle \bar{C}_t v, v\rangle < \varepsilon) = O(\varepsilon^{\infty-}).$$

To verify this claim, notice that $\lambda_0 := \inf_{v \in S}\langle \bar{C}_t v, v\rangle$ is the smallest eigenvalue of $\bar{C}_t$. Since $\det \bar{C}_t$ is the product of the eigenvalues of $\bar{C}_t$ it follows that $\bar{\Delta}_t := \det \bar{C}_t \ge \lambda_0^d$ and so $\left\{\det \bar{C}_t < \varepsilon^d\right\} \subset \{\lambda_0 < \varepsilon\}$ and hence

$$\mu\left(\det \bar{C}_t < \varepsilon^d\right) \le \mu(\lambda_0 < \varepsilon) = O(\varepsilon^{\infty-}).$$

By replacing ε by $\varepsilon^{1/d}$ above this implies $\mu\left(\bar{\Delta}_t < \varepsilon\right) = O(\varepsilon^{\infty-})$. From this estimate it then follows that

$$\begin{aligned}\mathbb{E}\left[\bar{\Delta}_t^{-q}\right] &= \mathbb{E}\int_{\bar{\Delta}_t}^\infty q\tau^{-q-1} d\tau = q\mathbb{E}\int_0^\infty 1_{\bar{\Delta}_t \le \tau}\ \tau^{-q-1} d\tau \\ &= q\int_0^\infty \mu(\bar{\Delta}_t \le \tau)\ \tau^{-q-1} d\tau = q\int_0^\infty O(\tau^p)\ \tau^{-q-1} d\tau\end{aligned}$$

which is seen to be finite by taking $p \ge q + 1$.

More generally if T is any stopping time with $T \le t$, since $\langle \bar{C}_T v, v\rangle \le \langle \bar{C}_t v, v\rangle$ for all $v \in S$ it suffices to prove

$$\mu\left(\inf_{v \in S} \langle \bar{C}_T v, v\rangle < \varepsilon\right) = O(\varepsilon^{\infty-}). \tag{8.20}$$

According to Lemma 8.13 and Proposition 8.15 below, Eq. (8.20) holds with

$$T = T_\delta := \inf\left\{t > 0 : \max\left\{|z_t - I_{T_oM}|, \operatorname{dist}(\Sigma_t, \Sigma_0)\right\} > \delta\right\} \tag{8.21}$$

provided $\delta > 0$ is chosen sufficiently small. ∎

The rest of this section is now devoted to the proof of Lemma 8.13 and Proposition 8.15 below. In what follows we will make repeated use of the identity,

$$\langle \bar{C}_T v, v\rangle = \sum_{i=1}^n \int_0^T \left\langle Z_\tau^{-1} X_i(\Sigma_\tau), v\right\rangle^2 d\tau. \tag{8.22}$$

To prove this, let $\{e_i\}_{i=1}^n$ be the standard basis for $\mathbb{R}^n$. Then

$$\begin{aligned}Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\mathbf{X}^{\mathrm{tr}}(\Sigma_\tau)\left(Z_\tau^{\mathrm{tr}}\right)^{-1} v &= \sum_{i=1}^n Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)e_i\ \langle e_i, \mathbf{X}^{\mathrm{tr}}(\Sigma_\tau)\left(Z_\tau^{\mathrm{tr}}\right)^{-1} v\rangle \\ &= \sum_{i=1}^n \langle Z_\tau^{-1} X_i(\Sigma_\tau), v\rangle\ Z_\tau^{-1} X_i(\Sigma_\tau)\end{aligned}$$

so that

$$\left\langle Z_\tau^{-1}\mathbf{X}(\Sigma_\tau)\mathbf{X}^{\mathrm{tr}}(\Sigma_\tau)\left(Z_\tau^{\mathrm{tr}}\right)^{-1} v, v\right\rangle = \sum_{i=1}^n \left\langle Z_\tau^{-1} X_i(\Sigma_\tau), v\right\rangle^2$$

which upon integrating on τ gives Eq. (8.22).

In the proofs below, there will always be an implied sum on repeated indices.

Lemma 8.13 (Compactness Argument)**.** *Let T_δ be as in Eq. (8.21) and suppose for all $v \in S$ there exists $i \in \{1, \ldots, n\}$ and an open neighborhood $N \subset_o S$ of v such that*

$$\sup_{u \in N} \mu\left(\int_0^{T_\delta} \left\langle Z_\tau^{-1} X_i(\Sigma_\tau), u\right\rangle^2 d\tau < \varepsilon\right) = O\left(\varepsilon^{\infty-}\right), \tag{8.23}$$

then Eq. (8.20) holds provided $\delta > 0$ is sufficiently small.

Proof. By compactness of S, it follows from Eq. (8.23) that

$$\sup_{u \in S} \mu\left(\int_0^{T_\delta} \left\langle Z_\tau^{-1} X_i(\Sigma_\tau), u\right\rangle^2 d\tau < \varepsilon\right) = O\left(\varepsilon^{\infty-}\right). \tag{8.24}$$

For $w \in T_oM$, let ∂_w denote the directional derivative acting on functions $f(v)$ with $v \in T_oM$. Because for all $v, w \in \mathbb{R}^n$ with $|v| \le 1$ and $|w| \le 1$ (using Eq. (8.22)),

$$\begin{aligned}\left|\partial_w \left\langle \bar{C}_{T_\delta} v, v\right\rangle\right| &\le 2\sum_{i=1}^n \int_0^{T_\delta} \left|\left\langle Z_\tau^{-1} X_i(\Sigma_\tau), v\right\rangle \left\langle Z_\tau^{-1} X_i(\Sigma_\tau), w\right\rangle\right| d\tau \\ &\le 2\sum_{i=1}^n \int_0^{T_\delta} \left|Z_\tau^{-1} X_i(\Sigma_\tau)\right|^2_{\mathrm{Hom}(\mathbb{R}^n, T_oM)} d\tau \\ &= 2\sum_{i=1}^n \int_0^{T_\delta} \left|z_\tau^{-1} //_\tau^{-1} X_i(\Sigma_\tau)\right|^2_{\mathrm{Hom}(\mathbb{R}^n, T_oM)} d\tau,\end{aligned}$$

by choosing $\delta > 0$ in Eq. (8.21) sufficiently small we may assume there is a non-random constant $\theta < \infty$ such that

$$\sup_{|v|, |w| \le 1} \left|\partial_w \left\langle \bar{C}_{T_\delta} v, v\right\rangle\right| \le \theta < \infty.$$

With this choice of δ, if $v, w \in S$ satisfy $|v-w| < \theta/\varepsilon$ then

$$\left|\left\langle \bar{C}_{T_\delta} v, v\right\rangle - \left\langle \bar{C}_{T_\delta} w, w\right\rangle\right| < \varepsilon. \tag{8.25}$$

There exists $D < \infty$ satisfying: for any $\varepsilon > 0$, there is an open cover of S with at most $D \cdot (\theta/\varepsilon)^n$ balls of the form $B(v_j, \varepsilon/\theta)$. From Eq. (8.25), for any $v \in S$ there exists j such that $v \in B(v_j, \varepsilon/\theta) \cap S$ and

$$\left|\left\langle \bar{C}_{T_\delta} v, v\right\rangle - \left\langle \bar{C}_{T_\delta} v_j, v_j\right\rangle\right| < \varepsilon.$$

So if $\inf_{v\in S} \left\langle \bar{C}_{T_\delta} v, v\right\rangle < \varepsilon$ then $\min_j \left\langle \bar{C}_{T_\delta} v_j, v_j\right\rangle < 2\varepsilon$, i.e.

$$\left\{\inf_{v\in S} \left\langle \bar{C}_{T_\delta} v, v\right\rangle < \varepsilon\right\} \subset \left\{\min_j \left\langle \bar{C}_{T_\delta} v_j, v_j\right\rangle < 2\varepsilon\right\} \subset \bigcup_j \left\{\left\langle \bar{C}_{T_\delta} v_j, v_j\right\rangle < 2\varepsilon\right\}.$$

Therefore,

$$\begin{aligned} \mu\left(\inf_{v\in S} \left\langle \bar{C}_{T_\delta} v, v\right\rangle < \varepsilon\right) &\le \sum_j \mu\left(\left\langle \bar{C}_{T_\delta} v_j, v_j\right\rangle < 2\varepsilon\right) \\ &\le D \cdot (\theta/\varepsilon)^n \cdot \sup_{v\in S} \mu\left(\left\langle \bar{C}_{T_\delta} v, v\right\rangle < 2\varepsilon\right) \\ &\le D \cdot (\theta/\varepsilon)^n\, O(\varepsilon^{\infty-}) = O(\varepsilon^{\infty-}). \end{aligned}$$

■

The following important proposition is the Stochastic version of Theorem 4.9. It gives the first hint that Hörmander's condition in Definition 8.1 is relevant to showing $\bar{\Delta}_t^{-1} \in L^{\infty-}(\mu)$ or equivalently that $\bar{C}_t^{-1} \in L^{\infty-}(\mu)$.

Proposition 8.14 (The appearance of commutators)**.** *Let $W \in \Gamma(TM)$, then*

$$\delta\left[Z_s^{-1} W(\Sigma_s)\right] = Z_s^{-1}[X_0, W](\Sigma_s) ds + Z_s^{-1} \sum_{i=1}^n [X_i, W](\Sigma_s)\delta B_s^i. \tag{8.26}$$

This may also be written in Itô form as

$$\begin{aligned} d\left[Z_s^{-1} W(\Sigma_s)\right] &= Z_s^{-1}[X_i, W](\Sigma_s) dB_s^i \\ &\quad + \left\{Z_s^{-1}[X_0, W](\Sigma_s) + \frac{1}{2}\sum_{i=1}^n Z_s^{-1}\left(L_{X_i}^2 W\right)(\Sigma_s)\right\} ds, \end{aligned} \tag{8.27}$$

where $L_X W := [X, W]$ as in Theorem 4.9.

Proof. Write $W(\Sigma_s) = Z_s w_s$, i.e. let $w_s := Z_s^{-1} W(\Sigma_s)$. By Proposition 5.36 and Theorem 5.41,

$$\begin{aligned} \nabla_{\delta \Sigma_s} W = \delta^\nabla [W(\Sigma_s)] = \delta^\nabla [Z_s w_s] &= \left(\delta^\nabla Z_s\right) w_s + Z_s \delta w_s \\ &= (\nabla_{Z_s w_s} \mathbf{X})\, \delta B_s + (\nabla_{Z_s w_s} X_0)\, ds + Z_s \delta w_s. \end{aligned}$$

Therefore, using the fact that ∇ has zero torsion (see Proposition 3.36),

$$\begin{aligned} \delta w_s &= Z_s^{-1}\left[\nabla_{\delta\Sigma_s} W - (\nabla_{Z_s w_s}\mathbf{X})\,\delta B_s + (\nabla_{Z_s w_s} X_0)\, ds\right] \\ &= Z_s^{-1}\left[\nabla_{\mathbf{X}(\Sigma_s)\delta B_s + X_0(\Sigma_s) ds} W - \left(\nabla_{W(\Sigma_s)}\mathbf{X}\right)\delta B_s + \left(\nabla_{W(\Sigma_s)} X_0\right) ds\right] \\ &= Z_s^{-1}\left[\left(\nabla_{X_i(\Sigma_s)} W - \nabla_{W(\Sigma_s)} X_i\right)\delta B_s^i + \left(\nabla_{X_0(\Sigma_s)} W - \nabla_{W(\Sigma_s)} X_0\right) ds\right] \\ &= Z_s^{-1}\left([X_i, W](\Sigma_s)\,\delta B_s^i + [X_0, W](\Sigma_s) ds\right) \end{aligned}$$

which proves Eq. (8.26).

Applying Eq. (8.26) with W replaced by $[X_i, W]$ implies

$$d\left[Z_s^{-1}[X_i, W](\Sigma_s)\right] = Z_s^{-1}[X_j, [X_i, W]](\Sigma_s) dB_s^j + d[BV],$$

where BV denotes process of bounded variation. Hence

$$\begin{aligned} Z_s^{-1}[X_i, W](\Sigma_s)\delta B_s^i &= Z_s^{-1}[X_i, W](\Sigma_s) dB_s^i + \frac{1}{2} d\left\{Z_s^{-1}[X_i, W](\Sigma_s)\right\} dB_s^i \\ &= Z_s^{-1}[X_i, W](\Sigma_s) dB_s^i + \frac{1}{2} Z_s^{-1}[X_j, [X_i, W]](\Sigma_s) dB_s^j dB_s^i \\ &= Z_s^{-1}[X_i, W](\Sigma_s) dB_s^i + \frac{1}{2} Z_s^{-1}[X_i, [X_i, W]](\Sigma_s) ds \end{aligned}$$

which combined with Eq. (8.26) proves Eq. (8.27). ■

Proposition 8.15. *Let T_δ be as in Eq. (8.21). If Hörmander's restricted bracket condition holds at $o \in M$ and $v \in S$ is given, there exists $i \in \{1, 2, \dots, n\}$ and an open neighborhood $U \subset_o S$ of v such that*

$$\sup_{u\in U} \mu\left(\int_0^{T_\delta} \left\langle Z_\tau^{-1} X_i(\Sigma_\tau), u\right\rangle^2 d\tau \le \varepsilon\right) = O\left(\varepsilon^{\infty-}\right).$$

Proof. The proof given here will follow Norris [147]. Hörmander's condition implies there exist $l \in \mathbb{N}$ and $\beta > 0$ such that

$$\frac{1}{|\mathcal{K}_l|}\sum_{K\in\mathcal{K}_l} K(o)K(o)^{\mathrm{tr}} \ge 3\beta I$$

or equivalently put for all $v \in S$,

$$3\beta \le \frac{1}{|\mathcal{K}_l|}\sum_{K\in\mathcal{K}_l} \langle K(o), v\rangle^2 \le \max_{K\in\mathcal{K}_l} \langle K(o), v\rangle^2.$$

By choosing $\delta > 0$ in Eq. (8.21) sufficiently small we may assume that

$$\max_{K\in\mathcal{K}_l} \inf_{\tau \le T_\delta} \left\langle Z_\tau^{-1} K(\Sigma_\tau), v\right\rangle^2 \ge 2\beta \text{ for all } v \in S.$$

Fix a $v \in S$ and $K \in \mathcal{K}_l$ such that

$$\inf_{\tau\le T_\delta} \left\langle Z_\tau^{-1} K(\Sigma_\tau), v\right\rangle^2 \ge 2\beta$$

and choose an open neighborhood $U \subset S$ of v such that

$$\inf_{\tau \le T_\delta} \left\langle Z_\tau^{-1} K(\Sigma_\tau), u \right\rangle^2 \ge \beta \text{ for all } u \in U.$$

Then, using Eq. (8.19),

$$\sup_{u \in U} \mu \left(\int_0^{T_\delta} \left\langle Z_\tau^{-1} K(\Sigma_\tau), u \right\rangle^2 d\tau \le \varepsilon \right)$$

$$\le \mu \left(\int_0^{T_\delta} \beta dt \le \varepsilon \right) = \mu \left(T_\delta \le \varepsilon / \beta \right) = O\left(\varepsilon^{\infty -}\right). \tag{8.28}$$

Write $K = L_{X_{i_r}} \dots L_{X_{i_2} X_{i_1}}$ with $r \le l$. If it happens that $r = 1$ then Eq. (8.28) becomes

$$\sup_{u \in U} \mu \left(\left\langle \bar{C}_{T_\delta} u, u \right\rangle \le \varepsilon \right) \le \sup_{u \in U} \mu \left(\int_0^{T_\delta} \left\langle Z_\tau^{-1} X_{i_1}(\Sigma_\tau), u \right\rangle^2 dt \le \varepsilon \right) = O\left(\varepsilon^{\infty -}\right)$$

and we are done. So now suppose $r > 1$ and set

$$K_j = L_{X_{i_j}} \dots L_{X_{i_2}} X_{i_1} \text{ for } j = 1, 2, \dots, r$$

so that $K_r = K$. We will now show by (decreasing) induction on j that

$$\sup_{u \in U} \mu \left(\int_0^{T_\delta} \left\langle Z_\tau^{-1} K_j(\Sigma_\tau), u \right\rangle^2 dt \le \varepsilon \right) = O\left(\varepsilon^{\infty -}\right). \tag{8.29}$$

From Proposition 8.14 we have

$$d\left[Z_t^{-1} K_{j-1}(\Sigma_t)\right] = Z_t^{-1}[X_i, K_{j-1}](\Sigma_t) dB^i(t) + \left\{ Z_t^{-1}[X_0, K_{j-1}](\Sigma_t) + \frac{1}{2} Z_t^{-1} \left(L_{X_i}^2 K_{j-1} \right)(\Sigma_t) \right\} dt$$

which upon integrating on t gives

$$\left\langle Z_t^{-1} K_{j-1}(\Sigma_t), u \right\rangle = \left\langle K_{j-1}(\Sigma_0), u \right\rangle + \int_0^t \left\langle Z_\tau^{-1}[X_i, K_{j-1}](\Sigma_\tau), u \right\rangle dB_\tau^i + \int_0^t \left\langle Z_\tau^{-1}[X_0, K_{j-1}](\Sigma_\tau) + \frac{1}{2} Z_t^{-1} \left(L_{X_i}^2 K_{j-1} \right)(\Sigma_\tau), u \right\rangle d\tau.$$

Applying Proposition 9.13 of the appendix with $T = T_\delta$,

$$Y_t := \left\langle Z_t^{-1} K_{j-1}(\Sigma_t), u \right\rangle, \ y = \left\langle K_{j-1}(\Sigma_0), u \right\rangle,$$

$$M_t = \int_0^t \left\langle Z_\tau^{-1}[X_i, K_{j-1}](\Sigma_\tau), u \right\rangle dB_\tau^i \text{ and}$$

$$A_t := \int_0^t \left\langle Z_\tau^{-1}[X_0, K_{j-1}](\Sigma_\tau) + \frac{1}{2} Z_\tau^{-1} \left(L_{X_i}^2 K_{j-1} \right)(\Sigma_\tau), u \right\rangle dt$$

implies

$$\sup_{u \in U} \mu \left(\Omega_1(u) \cap \Omega_2(u) \right) = O\left(\varepsilon^{\infty -}\right), \tag{8.30}$$

where

$$\Omega_1(u) := \left\{ \int_0^{T_\delta} \left\langle Z_t^{-1} K_{j-1}(\Sigma_t), u \right\rangle^2 dt < \varepsilon^q \right\},$$

$$\Omega_2(u) := \left\{ \int_0^{T_\delta} \sum_{i=1}^n \left\langle Z_\tau^{-1}[X_i, K_{j-1}](\Sigma_\tau), u \right\rangle^2 d\tau \ge \varepsilon \right\}$$

and $q > 4$. Since

$$\sup_{u \in U} \mu \left([\Omega_2(u)]^c \right) = \sup_{u \in U} \mu \left(\int_0^{T_\delta} \sum_{i=1}^n \left\langle Z_\tau^{-1}[X_i, K_{j-1}](\Sigma_\tau), u \right\rangle^2 d\tau < \varepsilon \right) \le \sup_{u \in U} \mu \left(\int_0^{T_\delta} \left\langle Z_\tau^{-1} K_j(\Sigma_\tau), u \right\rangle^2 d\tau < \varepsilon \right)$$

we may applying the induction hypothesis to learn,

$$\sup_{u \in U} \mu \left([\Omega_2(u)]^c \right) = O\left(\varepsilon^{\infty -}\right). \tag{8.31}$$

It now follows from Eqs. (8.30) and (8.31) that

$$\begin{aligned}
\sup_{u \in U} \mu(\Omega_1(u)) &\le \sup_{u \in U} \mu(\Omega_1(u) \cap \Omega_2(u)) + \sup_{u \in U} \mu(\Omega_1(u) \cap [\Omega_2(u)]^c) \\
&\le \sup_{u \in U} \mu(\Omega_1(u) \cap \Omega_2(u)) + \sup_{u \in U} \mu([\Omega_2(u)]^c) \\
&= O\left(\varepsilon^{\infty -}\right) + O\left(\varepsilon^{\infty -}\right) = O\left(\varepsilon^{\infty -}\right),
\end{aligned}$$

which is to say

$$\sup_{u \in U} \mu \left(\int_0^{T_\delta} \left\langle Z_t^{-1} K_{j-1}(\Sigma_t), u \right\rangle^2 dt < \varepsilon^q \right) = O\left(\varepsilon^{\infty -}\right).$$

Replacing ε by $\varepsilon^{1/q}$ in the previous equation, using $O\left(\left(\varepsilon^{1/q}\right)^{\infty -}\right) = O\left(\varepsilon^{\infty -}\right)$, completes the induction argument and hence the proof. ■

8.5. **More References.** The literature on the "Malliavin calculus" is very extensive and I will not make any attempt at summarizing it here. Let me just add to references already mentioned the articles in [176, 105, 153] which carry out Malliavin's method in the geometric context of these notes. Also see [150] for another method which works if Hörmander's bracket condition holds at level 2, namely when

$$\operatorname{span}(\{K(m) : K \in \mathcal{K}_2\}) = T_m M \text{ for all } m \in M,$$

see Definition 8.1. The reader should also be aware of the deep results of Ben Arous and Leandre in [17, 18, 16, 15, 124].

9. Appendix: Martingale and SDE Estimates

In this appendix $\{B_t : t \geq 0\}$ will denote and $\mathbb{R}^n$ – valued Brownian motion, $\{\beta_t : t \geq 0\}$ will be a one dimensional Brownian motion and, unlike in the text, we will use the more standard letter P rather than μ to denote the underlying probability measure.

Notation 9.1. When M_t is a martingale and A_t is a process of bounded variation let $\langle M \rangle_t$ be the quadratic variation of M and $|A|_t$ be the total variation of A up to time t.

9.1. Estimates of Wiener Functionals Associated to SDE's.

Proposition 9.2. *Suppose $p \in [2, \infty)$, α_τ and A_τ are predictable $\mathbb{R}^d$ and $\operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^d)$ – valued processes respectively and*

$$Y_t := \int_0^t A_\tau dB_\tau + \int_0^t \alpha_\tau d\tau. \tag{9.1}$$

Then, letting $Y_t^* := \sup_{\tau \leq t} |Y_\tau|$, there exists $C_p < \infty$ such that

$$\mathbb{E}(Y_t^*)^p \leq C_p \left\{ \mathbb{E}\left(\int_0^t |A_\tau|^2 d\tau \right)^{p/2} + \mathbb{E}\left(\int_0^t |\alpha_\tau| d\tau \right)^p \right\} \tag{9.2}$$

where

$$|A|^2 = \operatorname{tr}(AA^*) = \sum_{i=1}^n (AA^*)_{ii} = \sum_{i,j} A_{ij} A_{ij} = \operatorname{tr}(A^* A).$$

Proof. We may assume the right side of Eq. (9.2) is finite for otherwise there is nothing to prove. For the moment further assume $\alpha \equiv 0$. By a standard limiting argument involving stopping times we may further assume there is a non-random constant $C < \infty$ such that

$$Y_T^* + \int_0^T |A_\tau|^2 d\tau \leq C.$$

Let $f(y) = |y|^p$ and $\hat{y} := y/|y|$ for $y \in \mathbb{R}^d$. Then, for $a, b \in \mathbb{R}^d$,

$$\partial_a f(y) = p|y|^{p-1} \hat{y} \cdot a = p|y|^{p-2} y \cdot a$$

and

$$\begin{aligned}\partial_b \partial_a f(y) &= p(p-2)|y|^{p-4}(y \cdot a)(y \cdot b) + p|y|^{p-2} b \cdot a \\ &= p|y|^{p-2}[(p-2)(\hat{y} \cdot a)(\hat{y} \cdot b) + b \cdot a].\end{aligned}$$

So by Itô's formula

$$\begin{aligned}d|Y_t|^p &= d[f(Y_t)] \\ &= p|Y_t|^{p-1} \hat{Y}_t \cdot dY_t + \frac{p}{2}|Y_t|^{p-2}\left[(p-2)\left(\hat{Y}_t \cdot dY_t\right)\left(\hat{Y}_t \cdot dY_t\right) + dY_t \cdot dY_t\right].\end{aligned}$$

Taking expectations of this formula (using Y is a martingale) then gives

$$\mathbb{E}|Y_t|^p = \frac{p}{2} \int_0^t \mathbb{E}\left(|Y|^{p-2}\left[(p-2)\left(\hat{Y} \cdot dY\right)\left(\hat{Y} \cdot dY\right) + dY \cdot dY\right]\right). \tag{9.3}$$

Using $dY = AdB$, we have

$$dY \cdot dY = Ae_i \cdot Ae_j dB^i dB^j = e_i \cdot A^* A e_i dt = \operatorname{tr}(A^* A) dt = |A|^2 dt$$

and

$$\begin{aligned}\left(\hat{Y} \cdot dY\right)^2 &= \left(\hat{Y} \cdot Ae_i\right)\left(\hat{Y} \cdot Ae_j\right) dB^i dB^j = \left(A^* \hat{Y} \cdot e_i\right)\left(A^* \hat{Y} \cdot e_i\right) dt \\ &= \left(A^* \hat{Y} \cdot A^* \hat{Y}\right) dt = \left(AA^* \hat{Y} \cdot \hat{Y}\right) dt \leq |A|^2 dt.\end{aligned}$$

Putting these results back into Eq. (9.3) implies

$$\mathbb{E}|Y_t|^p \leq \frac{p}{2}(p-1) \int_0^t \mathbb{E}\left(|Y_\tau|^{p-2}|A_\tau|^2\right) d\tau.$$

By Doob's inequality there is a constant C_p (for example $C_p = \left[\frac{p}{p-1}\right]^p$ will work) such that

$$\mathbb{E}|Y_t^*|^p \leq C_p \mathbb{E}|Y_t|^p.$$

Combining the last two displayed equations implies

$$\mathbb{E}|Y_t^*|^p \leq C \int_0^t \mathbb{E}\left(|Y_\tau|^{p-2}|A_\tau|^2\right) d\tau \leq C\mathbb{E}\left(|Y_t^*|^{p-2} \int_0^t |A_\tau|^2 d\tau\right). \tag{9.4}$$

Now applying Hölder's inequality to the result, with exponents $q = p(p-2)^{-1}$ and conjugate exponent $q' = p/2$ gives

$$\mathbb{E}|Y_t^*|^p \leq C\left[\mathbb{E}|Y_t^*|^p\right]^{\frac{p-2}{p}} \left[\mathbb{E}\left(\int_0^t |A_\tau|^2 d\tau\right)^{p/2}\right]^{2/p}$$

or equivalently, using $1 - (p-2)/p = 2/p$,

$$\left(\mathbb{E}|Y_t^*|^p\right)^{2/p} \leq C\left[\mathbb{E}\left(\int_0^t |A_\tau|^2 d\tau\right)^{p/2}\right]^{2/p}.$$

Taking the $2/p$ roots of this equation then shows

$$\mathbb{E}|Y_t^*|^p \leq C\mathbb{E}\left(\int_0^t |A_\tau|^2 d\tau\right)^{p/2}. \tag{9.5}$$

The general case now follows, since when Y is given as in Eq. (9.1) we have

$$Y_t^* \le \left(\int_0^\cdot A_\tau dB_\tau\right)_t^* + \int_0^t |\alpha_\tau|\, d\tau$$

so that

$$\begin{aligned}\|Y_t^*\|_p &\le \left\|\left(\int_0^\cdot A_\tau dB_\tau\right)_t^*\right\|_p + \left\|\int_0^t |\alpha_\tau|\, d\tau\right\|_p \\ &\le C\left[\mathbb{E}\left(\int_0^t |A_\tau|^2\, d\tau\right)^{p/2}\right]^{1/p} + \left[\mathbb{E}\left(\int_0^t |\alpha_\tau|\, d\tau\right)^p\right]^{1/p}\end{aligned}$$

and taking the p^{th} – power of this equation proves Eq. (9.2). ■

Remark 9.3. A slightly different application of Hölder's inequality to the right side of Eq. (9.4) gives

$$\begin{aligned}\mathbb{E}\,|Y_t^*|^p &\le C\left(\int_0^t \mathbb{E}\left[|Y_t^*|^{p-2}\,|A_\tau|^2\right] d\tau\right) \le C\left(\int_0^t \left[\mathbb{E}\,|Y_t^*|^p\right]^{\frac{p-2}{p}} \left[\mathbb{E}\,|A_\tau|^p\right]^{2/p} d\tau\right) \\ &= \left[\mathbb{E}\,|Y_t^*|^p\right]^{\frac{p-2}{p}} C \int_0^t \left[\mathbb{E}\,|A_\tau|^p\right]^{2/p} d\tau\end{aligned}$$

which leads to the estimate

$$\mathbb{E}\,|Y_t^*|^p \le C\left(\int_0^t \left[\mathbb{E}\,|A_\tau|^p\right]^{2/p} d\tau\right)^{p/2}.$$

Here are some applications of Proposition 9.2.

Proposition 9.4. *Let $\{X_i\}_{i=0}^n$ be a collection of smooth vector fields on $\mathbb{R}^N$ for which $D^k X_i$ is bounded for all $k \ge 1$ and suppose Σ_t denotes the solution to Eq. (5.1) with $\Sigma_0 = x \in M := \mathbb{R}^N$ and $\beta = B$. Then for all $T < \infty$ and $p \in [2, \infty)$,*

$$\mathbb{E}\,(\Sigma_T^*)^p := \mathbb{E}\left[\sup_{t\le T} |\Sigma_t|^p\right] < \infty. \tag{9.6}$$

Proof. Since

$$\begin{aligned}X_i(\Sigma_t)\delta B^i(t) &= X_i(\Sigma_t)dB^i(t) + \frac{1}{2}d\left[X_i(\Sigma_t)\right]\cdot dB^i(t) \\ &= X_i(\Sigma_t)dB^i(t) + \frac{1}{2}\left(\partial_{X_i(\Sigma_t)}X_i\right)(\Sigma_t)dt,\end{aligned}$$

the Itô form of Eq. (5.1) is

$$\delta\Sigma_t = \left[X_0(\Sigma_t) + \frac{1}{2}\left(\partial_{X_i(\Sigma_t)}X_i\right)(\Sigma_t)\right]dt + X_i(\Sigma_t)dB^i(t) \text{ with } \Sigma_0 = x,$$

or equivalently,

$$\Sigma_t = x + \int_0^t X_i(\Sigma_\tau)dB_\tau^i + \int_0^t \left[X_0(\Sigma_\tau) + \frac{1}{2}\left(\partial_{X_i(\Sigma_\tau)}X_i\right)(\Sigma_\tau)\right] d\tau.$$

By Proposition 9.2,

$$\begin{aligned}\mathbb{E}\,|\Sigma_t|^p \le \mathbb{E}\,(\Sigma_t^*)^p &\le C_p\,|x|^p + C_p\mathbb{E}\left(\int_0^t |\mathbf{X}(\Sigma_\tau)|^2\, d\tau\right)^{p/2} \\ &+ C_p\mathbb{E}\left(\int_0^t \left|X_0(\Sigma_\tau) + \frac{1}{2}\left(\partial_{X_i(\Sigma_\tau)}X_i\right)(\Sigma_\tau)\right| d\tau\right)^p.\end{aligned} \tag{9.7}$$

Using the bounds on the derivatives of X we learn

$$|\mathbf{X}(\Sigma_\tau)|^2 \le C\left(1 + |\Sigma_\tau|^2\right) \text{ and}$$
$$\left|X_0(\Sigma_\tau) + \frac{1}{2}\left(\partial_{X_i(\Sigma_\tau)}X_i\right)(\Sigma_\tau)\right| \le C\left(1 + |\Sigma_\tau|\right)$$

which combined with Eq. (9.7) gives the estimate

$$\begin{aligned}&\mathbb{E}\,|\Sigma_t|^p \le \mathbb{E}\,(\Sigma_t^*)^p \\ &\le C_p\,|x|^p + C_p\mathbb{E}\left(\int_0^t C\left(1 + |\Sigma_\tau|^2\right) d\tau\right)^{p/2} + C_p\mathbb{E}\left(\int_0^t C\left(1 + |\Sigma_\tau|\right) d\tau\right)^p.\end{aligned}$$

Now assuming $t \le T < \infty$, we have by Jensen's (or Hölder's) inequality that

$$\begin{aligned}\mathbb{E}\,|\Sigma_t|^p \le&\mathbb{E}\,(\Sigma_t^*)^p \\ \le& C\,|x|^p + Ct^{p/2}\mathbb{E}\int_0^t \left(1 + |\Sigma_\tau|^2\right)^{p/2}\frac{d\tau}{t} \\ &+ Ct^p\mathbb{E}\int_0^t \left(1 + |\Sigma_\tau|\right)^p\frac{d\tau}{t} \\ \le& C\,|x|^p + CT^{(p/2-1)}\mathbb{E}\int_0^t \left(1 + |\Sigma_\tau|^2\right)^{p/2} d\tau \\ &+ CT^{(p-1)}\mathbb{E}\int_0^t \left(1 + |\Sigma_\tau|\right)^p d\tau\end{aligned}$$

from which it follows

$$\mathbb{E}\,|\Sigma_t|^p \le \mathbb{E}\,(\Sigma_t^*)^p \le C\,|x|^p + C(T)\int_0^t \left(1 + \mathbb{E}\,|\Sigma_\tau|^p\right) d\tau. \tag{9.8}$$

An application of Gronwall's inequality now shows $\sup_{t\le T}\mathbb{E}\,|\Sigma_t|^p < \infty$ for all $p < \infty$ and feeding this back into Eq. (9.8) with $t = T$ proves Eq. (9.6). ■

Proposition 9.5. *Suppose $\{X_i\}_{i=0}^n$ is a collection of smooth vector fields on M, Σ_t solves Eq. (5.1) with $\Sigma_0 = o \in M$ and $\beta = B$, z_t is the solution to Eq.*

(5.59) (i.e. $z_t := //_t^{-1} T^B_{t*o}$) and further assume[9] *there is a constant $K < \infty$ such that $\|A(m)\|_{op} \le K < \infty$ for all $m \in M$, where $A(m) \in \mathrm{End}(T_m M)$ is defined by*

$$A(m)v := \frac{1}{2}\left[\nabla_v\left(\sum_{i=1}^n \nabla_{X_i} X_i + X_0\right) - \sum_{i=1}^n R^\nabla(v, X_i(m))X_i(m)\right]$$

and

$$\sum_{i=1}^n |\nabla_v X_i| \le K|v| \text{ for all } v \in TM.$$

Then for all $p < \infty$ and $T < \infty$,

$$\mathbb{E}\left[\sup_{t\le T}|z_t|^p\right] < \infty \tag{9.9}$$

and

$$\mathbb{E}\left[(z_\cdot - I)^{*p}_t\right] = O\left(t^{p/2}\right) \text{ as } t \downarrow 0. \tag{9.10}$$

Proof. In what follows C will denote a constant depending on K, T and p. From Theorem 5.43, we know that the integrated Itô form of Eq. (5.59) is

$$z_t = I_{T_oM} + \int_0^t //_\tau^{-1}\left(\nabla_{//_\tau z_\tau(\cdot)}\mathbf{X}\right)dB_\tau + \frac{1}{2}A_{//_\tau} z_\tau v d\tau \tag{9.11}$$

where $A_{//_t} := //_t^{-1} A(\Sigma_t)//_t$. By Proposition 9.2 and the assumed bounds on A and $\nabla.\mathbf{X}$,

$$\begin{aligned}\mathbb{E}(z_t^*)^p \le& C|I|^p + C\mathbb{E}\left(\int_0^t \sum_{i=1}^n \left|//_\tau^{-1}\left(\nabla_{//_\tau z_\tau(\cdot)}X_i\right)\right|^2 d\tau\right)^{p/2} \\ &+ C\mathbb{E}\left(\int_0^t \left|A_{//_\tau} z_\tau\right| d\tau\right)^p \\ \le& C + C\mathbb{E}\left(\int_0^t |z_\tau|^2 d\tau\right)^{p/2} + C\mathbb{E}\left(\int_0^t |z_\tau| d\tau\right)^p \\ \le& C + C\int_0^t \mathbb{E}|z_\tau|^p d\tau\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\left[(z_\cdot - I)^{*p}_t\right] &\le C\mathbb{E}\left(\int_0^t |z_\tau|^2 d\tau\right)^{p/2} + C\mathbb{E}\left(\int_0^t |z_\tau| d\tau\right)^p \\ &\le C\cdot\mathbb{E}|z_t^*|^p\cdot\left(t^{p/2} + t^p\right) \end{aligned} \tag{9.12}$$

where we have made use of Hölder's (or Jensen's) inequality. Since

$$\mathbb{E}|z_t|^p \le \mathbb{E}(z_t^*)^p \le C + C\int_0^t \mathbb{E}|z_\tau|^p d\tau, \tag{9.13}$$

Gronwall's inequality implies

$$\sup_{t\le T}\mathbb{E}\left[|z_t|^p\right] \le Ce^{CT} < \infty.$$

Feeding the last inequality back into Eq. (9.13) shows Eq. (9.9). Eq. (9.10) now follows from Eq. (9.9). and Eq. (9.12). ■

Exercise 9.6. Show under the same hypothesis of Proposition 9.5 that

$$\mathbb{E}\left[\sup_{t\le T}\left|z_t^{-1}\right|^p\right] < \infty$$

for all $p, T < \infty$. **Hint:** Show z_t^{-1} satisfies an equation similar to Eq. (9.11) with coefficients satisfying the same type of bounds.

9.2. **Martingale Estimates.**

This section follows the presentation in Norris [147].

Lemma 9.7 (Reflection Principle)**.** *Let β_t be a 1 - dimensional Brownian motion starting at 0, $a > 0$ and $T_a = \inf\{t > 0 : \beta_t = a\}$ – be first time β_t hits height a, see Figure 15. Then*

$$P(T_a < t) = 2P(\beta_t > a) = \frac{2}{\sqrt{2\pi t}}\int_a^\infty e^{-x^2/2t}dx$$

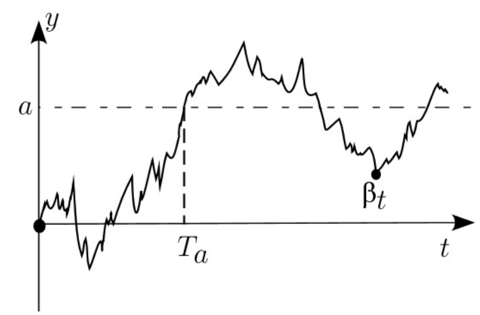

FIGURE 15. The first hitting time T_a of level a by β_t.

Proof. Since $P(\beta_t = a) = 0$,

$$\begin{aligned}P(T_a < t) &= P(T_a < t \;\&\; \beta_t > a) + P(T_a < t \;\&\; \beta_t < a) \\ &= P(\beta_t > a) + P(T_a < t \;\&\; \beta_t < a),\end{aligned}$$

[9]This will always be true when M is compact.

it suffices to prove

$$P(T_a < t \ \&\ \beta_t < a) = P(\beta_t > a).$$

To do this define a new process $\tilde{\beta}_t$ by

$$\tilde{\beta}_t = \begin{cases} \beta_t & \text{for} \quad t < T_a \\ 2a - \beta_t & \text{for} \quad t \geq T_a \end{cases}$$

(see Figure 16) and notice that $\tilde{\beta}_t$ may also be expressed as

$$\tilde{\beta}_t = \beta_{t \wedge T_a} - 1_{t \geq T_a}(\beta_t - \beta_{t\wedge T_a}) = \int_0^t (1_{\tau < T_a} - 1_{\tau \geq T_a})\, d\beta_\tau. \tag{9.14}$$

So $\tilde{\beta}_t = \beta_t$ for $t \leq T_a$ and $\tilde{\beta}_t$ is β_t reflected across the line $y = a$ for $t \geq T_a$.

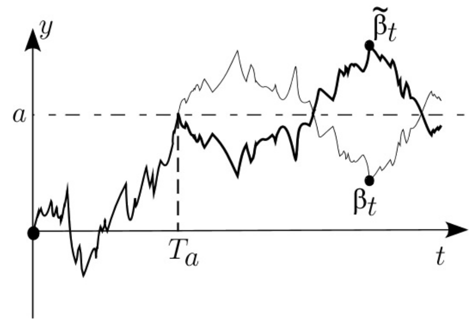

FIGURE 16. The Brownian motion β_t and its reflection $\tilde{\beta}_t$ about the line $y = a$. Note that after time T_a, the labellings of the β_t and the $\tilde{\beta}_t$ could be interchanged and the picture would still be possible. This should help alleviate the readers fears that Brownian motion has some funny asymmetry after the first hitting of level a.

From Eq. (9.14) it follows that $\tilde{\beta}_t$ is a martingale and

$$\left(d\tilde{\beta}_t\right)^2 = (1_{\tau < T_a} - 1_{\tau \geq T_a})^2\, dt = dt$$

and hence that $\tilde{\beta}_t$ is another Brownian motion. Since $\tilde{\beta}_t$ hits level a for the first time exactly when β_t hits level a,

$$T_a = \tilde{T}_a := \inf\left\{t > 0 : \tilde{\beta}_t = a\right\}$$

and $\left\{\tilde{T}_a < t\right\} = \{T_a < t\}$. Furthermore (see Figure 16),

$$\{T_a < t \ \&\ \beta_t < a\} = \left\{\tilde{T}_a < t \ \&\ \tilde{\beta}_t > a\right\} = \left\{\tilde{\beta}_t > a\right\}.$$

Therefore,

$$P(T_a < t \ \&\ \beta_t < a) = P(\tilde{\beta}_t > a) = P(\beta_t > a)$$

which completes the proof. ∎

Remark 9.8. An alternate way to get a handle on the stopping time T_a is to compute its Laplace transform. This can be done by considering the martingale

$$M_t := e^{\lambda \beta_t - \frac{1}{2}\lambda^2 t}.$$

Since M_t is bounded by $e^{\lambda a}$ for $t \in [0, T_a]$ the optional sampling theorem may be applied to show

$$e^{\lambda a} E\left[e^{-\frac{1}{2}\lambda^2 T_a}\right] = E\left[e^{\lambda a - \frac{1}{2}\lambda^2 T_a}\right] = EM_{T_a} = EM_0 = 1,$$

i.e. this implies that $E\left[e^{-\frac{1}{2}\lambda^2 T_a}\right] = e^{-\lambda a}$. This is equivalent to

$$E\left[e^{-\lambda T_a}\right] = e^{-a\sqrt{2\lambda}}.$$

From this point of view one would now have to invert the Laplace transform to get the density of the law of T_a.

Corollary 9.9. *Suppose now that $T = \inf\{t > 0 : |\beta_t| = a\}$, i.e. the first time β_t leaves the strip $(-a, a)$. Then*

$$\begin{aligned} P(T < t) &\leq 4P(\beta_t > a) = \frac{4}{\sqrt{2\pi t}} \int_a^\infty e^{-x^2/2t} dx \\ &\leq \min\left(\sqrt{\frac{8t}{\pi a^2}} e^{-a^2/2t}, 1\right). \end{aligned} \tag{9.15}$$

Notice that $P(T < t) = P(\beta_t^* \geq a)$ where $\beta_t^* = \max\{|\beta_\tau| : \tau \leq t\}$. So Eq. (9.15) may be rewritten as

$$P(\beta_t^* \geq a) \leq 4P(\beta_t > a) \leq \min\left(\sqrt{\frac{8t}{\pi a^2}} e^{-a^2/2t}, 1\right) \leq 2e^{-a^2/2t}. \tag{9.16}$$

Proof. By definition $T = T_a \wedge T_{-a}$ so that $\{T < t\} = \{T_a < t\} \cup \{T_{-a} < t\}$ and therefore

$$\begin{aligned} P(T < t) &\leq P(T_a < t) + P(T_{-a} < t) \\ &= 2P(T_a < t) = 4P(\beta_t > a) = \frac{4}{\sqrt{2\pi t}} \int_a^\infty e^{-x^2/2t} dx \\ &\leq \frac{4}{\sqrt{2\pi t}} \int_a^\infty \frac{x}{a} e^{-x^2/2t} dx = \frac{4}{\sqrt{2\pi t}} \left(-\frac{t}{a} e^{-x^2/2t}\right)\Big|_a^\infty = \sqrt{\frac{8t}{\pi a^2}} e^{-a^2/2t}. \end{aligned}$$

This proves everything but the very last inequality in Eq. (9.16). To prove this inequality first observe the elementary calculus inequality:

$$\min\left(\frac{4}{\sqrt{2\pi}y}e^{-y^2/2},1\right)\le 2e^{-y^2/2}. \tag{9.17}$$

Indeed Eq. (9.17) holds $\frac{4}{\sqrt{2\pi}y}\le 2$, i.e. if $y\ge y_0:=2/\sqrt{2\pi}$. The fact that Eq. (9.17) holds for $y\le y_0$ follows from the following trivial inequality

$$1\le 1.4552\cong 2e^{-\frac{1}{\pi}}=e^{-y_0^2/2}.$$

Finally letting $y=a/\sqrt{t}$ in Eq. (9.17) gives the last inequality in Eq. (9.16). ■

Theorem 9.10. *Let N be a continuous martingale such that $N_0=0$ and T be a stopping time. Then for all $\varepsilon,\delta>0$,*

$$P\left(\langle N\rangle_T<\varepsilon\ \&\ N_T^*\ge\delta\right)\le P(\beta_\varepsilon^*\ge\delta)\le 2e^{-\delta^2/2\varepsilon}.$$

Proof. By the Dambis, Dubins & Schwarz's theorem (see p.174 of [109]) we may write $N_t=\beta_{\langle N\rangle_t}$ where β is a Brownian motion (on a possibly "augmented" probability space). Therefore

$$\{\langle N\rangle_T<\varepsilon\ \&\ N_T^*\ge\delta\}\subset\{\beta_\varepsilon^*\ge\delta\}$$

and hence from Eq. (9.16),

$$P\left(\langle N\rangle_T<\varepsilon\ \&\ N_T^*\ge\delta\right)\le P(\beta_\varepsilon^*\ge\delta)\le 2e^{-\delta^2/2\varepsilon}.$$

■

Theorem 9.11. *Suppose that $Y_t=M_t+A_t$ where M_t is a martingale and A_t is a process of bounded variation which satisfy: $M_0=A_0=0$, $|A|_t\le ct$ and $\langle M\rangle_t\le ct$ for some constant $c<\infty$. If $T_a:=\inf\{t>0:|Y_t|=a\}$ and $t<a/2c$, then*

$$P(Y_t^*\ge a)=P(T_a\le t)\le\frac{4}{\sqrt{\pi a}}\exp\left(-\frac{a^2}{8ct}\right)$$

Proof. Since

$$Y_t^*\le M_t^*+A_t^*\le M_t^*+|A|_t\le M_t^*+ct$$

it follows that

$$\{Y_t^*\ge a\}\subset\{M_t^*\ge a/2\}\cup\{ct\ge a/2\}=\{M_t^*\ge a/2\}$$

when $t<a/2c$. Again by the Dambis, Dubins and Schwarz's theorem (see p.174 of [109]), we may write $M_t=\beta_{\langle M\rangle_t}$ where β is a Brownian motion on a possibly augmented probability space. Since

$$M_t^*=\max_{\tau\le\langle M\rangle_t}|\beta_\tau|\le\max_{\tau\le ct}|\beta_\tau|=\beta_{ct}^*$$

we learn

$$\begin{aligned}P(Y_t^*\ge a)&\le P\left(M_t^*\ge a/2\right)\le P\left(\beta_{ct}^*\ge a/2\right)\\&\le\sqrt{\frac{8ct}{\pi\left(a/2\right)^2}}e^{-(a/2)^2/2ct}=\sqrt{\frac{8ct}{\pi\left(a/2\right)^2}}e^{-(a/2)^2/2ct}\\&\le\sqrt{\frac{8c\left(a/2c\right)}{\pi\left(a/2\right)^2}}e^{-(a/2)^2/2ct}=\frac{4}{\sqrt{\pi a}}\exp\left(-\frac{a^2}{8ct}\right)\end{aligned}$$

wherein the last inequality we have used the restriction $t<a/2c$. ■

Lemma 9.12. *If $f:[0,\infty)\to\mathbb{R}$ is a locally absolutely continuous function such that $f(0)=0$, then*

$$|f(t)|\le\sqrt{2\left\|\dot f\right\|_{L^\infty([0,t])}\|f\|_{L^1([0,t])}}\ \forall\ t\ge 0.$$

Proof. By the fundamental theorem of calculus,

$$f^2(t)=2\int_0^t f(\tau)\dot f(\tau)d\tau\le 2\left\|\dot f\right\|_{L^\infty([0,t])}\|f\|_{L^1([0,t])}.$$

■

We are now ready for a key result needed in the probabilistic proof of Hörmander's theorem. Loosely speaking it states that if Y is a Brownian semi-martinagale, then it can happen **only** with small probability that the L^2 – norm of Y is small while the quadratic variation of Y is relatively large.

Proposition 9.13 (A key martingale inequality)**.** *Let T be a stopping time bounded by $t_0<\infty$, $Y=y+M+A$ where M is a continuous martingale and A is a process of bounded variation such that $M_0=A_0=0$. Further assume, on the set $\{t\le T\}$, that $\langle M\rangle_t$ and $|A|_t$ are absolutely continuous functions and there exists finite positive constants, c_1 and c_2, such that*

$$\frac{d\langle M\rangle_t}{dt}\le c_1\ \text{and}\ \frac{d\,|A|_t}{dt}\le c_2.$$

Then for all $\nu>0$ and $q>\nu+4$ there exists constants $c=c(t_0,q,\nu,c_1,c_2)>0$ and $\varepsilon_0=\varepsilon_0(t_0,q,\nu,c_1,c_2)>0$ such that

$$P\left(\int_0^T Y_t^2dt<\varepsilon^q,\ \langle Y\rangle_T=\langle M\rangle_T\ge\varepsilon\right)\le 2\exp\left(-\frac{1}{2c_1\varepsilon^\nu}\right)=O\left(\varepsilon^{-\infty}\right) \tag{9.18}$$

for all $\varepsilon\in(0,\varepsilon_0]$.

Proof. Let $q_0=\frac{q-\nu}{2}$ (so that $q_0\in(2,q/2)$), $N:=\int_0^\cdot YdM$ and

$$C_\varepsilon:=\left\{\langle N\rangle_T\le c_1\varepsilon^q,\ \ N_T^*\ge\varepsilon^{q_0}\right\}. \tag{9.19}$$

We will show shortly that for ε sufficiently small,

$$B_\varepsilon := \left\{ \int_0^T Y_t^2 dt < \varepsilon^q, \ \langle Y \rangle_T \ge \varepsilon \right\} \subset C_\varepsilon. \tag{9.20}$$

By an application of Theorem 9.10,

$$P(C_\varepsilon) \le 2 \exp\left(-\frac{\varepsilon^{2q_0}}{2c_1 \varepsilon^q} \right) = 2\exp\left(-\frac{1}{2c_1 \varepsilon^v} \right)$$

and so assuming the validity of Eq. (9.20),

$$P\left(\int_0^T Y_t^2 dt < \varepsilon^q, \ \langle Y \rangle_T \ge \varepsilon \right) \le P(C_\varepsilon) \le 2\exp\left(-\frac{1}{2c_1 \varepsilon^v} \right) \tag{9.21}$$

which proves Eq. (9.18). So to finish the proof it only remains to verify Eq. (9.20) which will be done by showing $B_\varepsilon \cap C_\varepsilon^c = \emptyset$.

For the rest of the proof, it will be assumed that we are on the set $B_\varepsilon \cap C_\varepsilon^c$. Since $\langle N \rangle_T = \int_0^T |Y_t|^2 \, d\langle M \rangle_t$, we have

$$B_\varepsilon \cap C_\varepsilon^c = \left\{ \int_0^T Y_t^2 dt < \varepsilon^q, \ \langle Y \rangle_T \ge \varepsilon, \ \int_0^T |Y_t|^2 \, d\langle M \rangle_t > c_1 \varepsilon^q, \quad N_T^* < \varepsilon^{q_0} \right\}. \tag{9.22}$$

From Lemma 9.12 with $f(t) = \langle Y \rangle_t$ and the assumption that $d\langle Y \rangle_t / dt \le c_1$,

$$\langle Y \rangle_T \le \sqrt{2 \left\| \dot{f} \right\|_{L^\infty([0,T])} \|f\|_{L^1([0,T])}} \le \sqrt{2c_1 \int_0^T \langle Y \rangle_t dt}. \tag{9.23}$$

By Itô's formula, the quadratic variation, $\langle Y \rangle_t$, of Y satisfies

$$\langle Y \rangle_t = Y_t^2 - y^2 - 2\int_0^t Y dY \le Y_t^2 + 2 \left| \int_0^t Y dY \right| \tag{9.24}$$

and on the set $\{t \le T\} \cap B_\varepsilon \cap C_\varepsilon^c$,

$$\begin{aligned} \left| \int_0^t Y dY \right| &= \left| \int_0^t Y dM + \int_0^t Y dA \right| \le |N_t| + \int_0^t |Y| \, dA \\ &\le N_T^* + c_2 \int_0^T |Y_\tau| \, d\tau \le \varepsilon^{q_0} + c_2 T^{1/2} \sqrt{\int_0^T Y_\tau^2 d\tau} \\ &\le \varepsilon^{q_0} + c_2 t_0^{1/2} \varepsilon^q. \end{aligned} \tag{9.25}$$

Combining Eqs. (9.24) and (9.25) shows, on the set $\{t \le T\} \cap B_\varepsilon \cap C_\varepsilon^c$ that

$$\langle Y \rangle_t \le Y_t^2 + 2\left[\varepsilon^{q_0} + c_2 t_0^{1/2} \varepsilon^q \right]$$

and using this in Eq. (9.23) implies

$$\begin{aligned} \langle Y \rangle_T &\le \sqrt{2c_1 \int_0^T \left(Y_t^2 + 2\left[\varepsilon^{q_0} + c_2 t_0^{1/2} \varepsilon^q \right] \right) dt} \\ &\le \sqrt{2c_1 \left[\varepsilon^q + 2\left[\varepsilon^{q_0} + c_2 t_0^{1/2} \varepsilon^q \right] t_0 \right]} = O\left(\varepsilon^{\frac{q_0}{2}} \right) = o(\varepsilon). \end{aligned} \tag{9.26}$$

Hence we may choose $\varepsilon_0 = \varepsilon_0 (c_1, c_2, t_o, q, \nu) > 0$ such that, if $\varepsilon \le \varepsilon_0$ then

$$\sqrt{2c_1 \left(\varepsilon^q + 2\varepsilon^{q_0} t_0 + 2c_2 t_0^{3/2} \varepsilon^{q/2} \right)} < \varepsilon$$

and hence on $B_\varepsilon \cap C_\varepsilon^c$ we learn $\varepsilon \le \langle Y \rangle_T < \varepsilon$ which is absurd. So we must conclude that $B_\varepsilon \cap C_\varepsilon^c = \emptyset$. ∎

References

[1] Ralph Abraham and Jerrold E. Marsden, *Foundations of mechanics*, Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., 1978, Second edition, revised and enlarged, With the assistance of Tudor Raţiu and Richard Cushman. MR **81e:**58025

[2] S. Aida, *On the irreducibility of certain dirichlet forms on loop spaces over compact homogeneous spaces*, New Trends in Stochastic Analysis (New Jersey) (K. D. Elworthy, S. Kusuoka, and I. Shigekawa, eds.), Proceedings of the 1994 Taniguchi Symposium, World Scientific, 1997, pp. 3–42.

[3] Shigeki Aida and Bruce K. Driver, *Equivalence of heat kernel measure and pinned Wiener measure on loop groups*, C. R. Acad. Sci. Paris Sér. I Math. **331** (2000), no. 9, 709–712. MR 1 797 756

[4] Shigeki Aida and David Elworthy, *Differential calculus on path and loop spaces. I. Logarithmic Sobolev inequalities on path spaces*, C. R. Acad. Sci. Paris Sér. I Math. **321** (1995), no. 1, 97–102.

[5] Helene Airault and Paul Malliavin, *Integration by parts formulas and dilatation vector fields on elliptic probability spaces*, Probab. Theory Related Fields **106** (1996), no. 4, 447–494.

[6] Sergio Albeverio, Rémi Léandre, and Michael Röckner, *Construction of a rotational invariant diffusion on the free loop space*, C. R. Acad. Sci. Paris Sér. I Math. **316** (1993), no. 3, 287–292.

[7] Y. Amit, *A multiflow approximation to diffusions*, Stochastic Process. Appl. **37** (1991), no. 2, 213–237.

[8] Lars Andersson and Bruce K. Driver, *Finite-dimensional approximations to Wiener measure and path integral formulas on manifolds*, J. Funct. Anal. **165** (1999), no. 2, 430–498. MR **2000j:**58059

[9] Louis Auslander and Robert E. MacKenzie, *Introduction to differentiable manifolds*, Dover Publications Inc., New York, 1977, Corrected reprinting. MR 57 #10717

[10] Vlad Bally, *Approximation for the solutions of stochastic differential equations. I. L^p-convergence*, Stochastics Stochastics Rep. **28** (1989), no. 3, 209–246.

[11] ———, *Approximation for the solutions of stochastic differential equations. II. Strong convergence*, Stochastics Stochastics Rep. **28** (1989), no. 4, 357–385.

[12] Denis R. Bell, *Degenerate stochastic differential equations and hypoellipticity*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 79, Longman, Harlow, 1995. MR **99e:**60124

[13] Denis R. Bell and Salah Eldin A. Mohammed, *The Malliavin calculus and stochastic delay equations*, J. Funct. Anal. **99** (1991), no. 1, 75–99. MR **92k:**60124

[14] Ya. I. Belopol′skaya and Yu. L. Dalecky, *Stochastic equations and differential geometry*, Mathematics and its Applications (Soviet Series), vol. 30, Kluwer Academic Publishers Group, Dordrecht, 1990, Translated from the Russian. MR **91b:**58271

[15] G. Ben Arous and R. Léandre, *Décroissance exponentielle du noyau de la chaleur sur la diagonale. I*, Probab. Theory Related Fields **90** (1991), no. 2, 175–202. MR **93b:**60136a

[16] ———, *Décroissance exponentielle du noyau de la chaleur sur la diagonale. II*, Probab. Theory Related Fields **90** (1991), no. 3, 377–402. MR **93b:**60136b

[17] Gérard Ben Arous, *Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale*, Ann. Inst. Fourier (Grenoble) **39** (1989), no. 1, 73–99. MR **91b:**58272

[18] ———, *Flots et séries de Taylor stochastiques*, Probab. Theory Related Fields **81** (1989), no. 1, 29–77. MR **90a:**60106

[19] Richard L. Bishop and Richard J. Crittenden, *Geometry of manifolds*, AMS Chelsea Publishing, Providence, RI, 2001, Reprint of the 1964 original. MR **2002d:**53001

[20] Jean-Michel Bismut, *Mecanique aleatoire. (french) [random mechanics]*, Springer-Verlag, Berlin-New York, 1981, Lecture Notes in Mathematics, Vol. 866.

[21] ———, *Large deviations and the Malliavin calculus*, Progress in Mathematics, vol. 45, Birkhäuser Boston Inc., Boston, Mass., 1984.

[22] G. Blum, *A note on the central limit theorem for geodesic random walks*, Bull. Austral. Math. Soc. **30** (1984), no. 2, 169–173.

[23] Nicolas Bouleau and Francis Hirsch, *Dirichlet forms and analysis on Wiener space*, de Gruyter Studies in Mathematics, vol. 14, Walter de Gruyter & Co., Berlin, 1991. MR **93e:**60107

[24] Theodor Bröcker and Tammo tom Dieck, *Representations of compact Lie groups*, Graduate Texts in Mathematics, vol. 98, Springer-Verlag, New York, 1995, Translated from the German manuscript, Corrected reprint of the 1985 translation. MR **97i:**22005

[25] R. H. Cameron and W. T. Martin, *Transformations of Wiener integrals under translations*, Ann. of Math. (2) **45** (1944), 386–396. MR 6,5f

[26] ———, *The transformation of Wiener integrals by nonlinear transformations*, Trans. Amer. Math. Soc. **66** (1949), 253–283. MR 11,116b

[27] ———, *Non-linear integral equations*, Ann. of Math. (2) **51** (1950), 629–642. MR 11,728d

[28] Robert H. Cameron, *The first variation of an indefinite Wiener integral*, Proc. Amer. Math. Soc. **2** (1951), 914–924. MR 13,659b

[29] Mireille Capitaine, Elton P. Hsu, and Michel Ledoux, *Martingale representation and a simple proof of logarithmic Sobolev inequalities on path spaces*, Electron. Comm. Probab. **2** (1997), 71–81 (electronic).

[30] Trevor R. Carson, *Logarithmic sobolev inequalities for free loop groups*, University of California at San Diego Ph.D. thesis. This may be retrieved at http://math.ucsd.edu/ driver/driver/thesis.htm, 1997.

[31] ———, *A logarithmic Sobolev inequality for the free loop group*, C. R. Acad. Sci. Paris Sér. I Math. **326** (1998), no. 2, 223–228. MR **99g:**60108

[32] Carolyn Cross, *Differentials of measure-preserving flows on path space*, University of California at San Diego Ph.D. thesis. This may be retrieved at http://math.ucsd.edu/ driver/driver/thesis.htm, 1996.

[33] A. B. Cruzeiro, S. Fang, and P. Malliavin, *A probabilistic Weitzenböck formula on Riemannian path space*, J. Anal. Math. **80** (2000), 87–100. MR **2002d:**58045

[34] A. B. Cruzeiro and P. Malliavin, *Riesz transforms, commutators, and stochastic integrals*, Harmonic analysis and partial differential equations (Chicago, IL, 1996), Chicago Lectures in Math., Univ. Chicago Press, Chicago, IL, 1999, pp. 151–162. MR **2001d:**60062

[35] Ana Bela Cruzeiro and Shizan Fang, *An L^2 estimate for Riemannian anticipative stochastic integrals*, J. Funct. Anal. **143** (1997), no. 2, 400–414. MR **98d:**60106

[36] Ana-Bela Cruzeiro and Shizan Fang, *A Weitzenböck formula for the damped Ornstein-Uhlenbeck operator in adapted differential geometry*, C. R. Acad. Sci. Paris Sér. I Math. **332** (2001), no. 5, 447–452. MR **2002a:**60090

[37] Ana-Bela Cruzeiro and Paul Malliavin, *Frame bundle of Riemannian path space and Ricci tensor in adapted differential geometry*, J. Funct. Anal. **177** (2000), no. 1, 219–253. MR **2001h:**60103

[38] ———, *A class of anticipative tangent processes on the Wiener space*, C. R. Acad. Sci. Paris Sér. I Math. **333** (2001), no. 4, 353–358. MR **2002g:**60084

[39] ———, *Stochastic calculus of variations and Harnack inequality on Riemannian path spaces*, C. R. Math. Acad. Sci. Paris **335** (2002), no. 10, 817–820. MR **2003k:**58058

[40] R. W. R. Darling, *Martingales in manifolds—definition, examples, and behaviour under maps*, Seminar on Probability, XVI, Supplement, Lecture Notes in Math., vol. 921, Springer, Berlin, 1982, pp. 217–236. MR **84j:**58133

[41] E. B. Davies, *Heat kernels and spectral theory*, Cambridge Tracts in Mathematics, vol. 92, Cambridge University Press, Cambridge, 1990. MR **92a:**35035

[42] Manfredo Perdigão do Carmo, *Riemannian geometry*, Mathematics: Theory & Applications, Birkhäuser Boston Inc., Boston, MA, 1992, Translated from the second Portuguese edition by Francis Flaherty. MR **92i:**53001

[43] Jozef Dodziuk, *Maximum principle for parabolic inequalities and the heat flow on open manifolds*, Indiana Univ. Math. J. **32** (1983), no. 5, 703–716. MR **85e:**58140

[44] H. Doss, *Connections between stochastic and ordinary integral equations*, Biological Growth and Spread (Proc. Conf., Heidelberg, 1979) (Berlin-New York), Springer, 1979, Lecture Notes in Biomath., 38, pp. 443–448.

[45] B. K. Driver, *The Lie bracket of adapted vector fields on Wiener spaces*, Appl. Math. Optim. **39** (1999), no. 2, 179–210. MR **2000b:**58063

[46] Bruce K. Driver, *Classifications of bundle connection pairs by parallel translation and lassos*, J. Funct. Anal. **83** (1989), no. 1, 185–231.

[47] ———, *A Cameron-Martin type quasi-invariance theorem for Brownian motion on a compact Riemannian manifold*, J. Funct. Anal. **110** (1992), no. 2, 272–376.

[48] ———, *A Cameron-Martin type quasi-invariance theorem for pinned Brownian motion on a compact Riemannian manifold*, Trans. Amer. Math. Soc. **342** (1994), no. 1, 375–395.

[49] ———, *Towards calculus and geometry on path spaces*, Stochastic analysis (Ithaca, NY, 1993), Proc. Sympos. Pure Math., vol. 57, Amer. Math. Soc., Providence, RI, 1995, pp. 405–422.

[50] ———, *Integration by parts and quasi-invariance for heat kernel measures on loop groups*, J. Funct. Anal. **149** (1997), no. 2, 470–547.

[51] ———, *A correction to the paper: "Integration by parts and quasi-invariance for heat kernel measures on loop groups"*, J. Funct. Anal. **155** (1998), no. 1, 297–301. MR **99a:**60054b

[52] ———, *Analysis of Wiener measure on path and loop groups*, Finite and infinite dimensional analysis in honor of Leonard Gross (New Orleans, LA, 2001), Contemp. Math., vol. 317, Amer. Math. Soc., Providence, RI, 2003, pp. 57–85. MR **2003m:**58055

[53] ———, *Heat kernels measures and infinite dimensional analysis*, To appear in Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces, Contemp. Math., vol. 338, Amer. Math. Soc., Providence, RI, 2004, p. 41 pages. MR **2003m:**58055

[54] Bruce K. Driver and Terry Lohrenz, *Logarithmic Sobolev inequalities for pinned loop groups*, J. Funct. Anal. **140** (1996), no. 2, 381–448.

[55] Bruce K. Driver and Tai Melcher, *Hypoelliptic heat kernel inequalities on the Heisenberg group*, to appear in J. Funct. Anal. (2004).

[56] Bruce K. Driver and Michael Röckner, *Construction of diffusions on path and loop spaces of compact Riemannian manifolds*, C. R. Acad. Sci. Paris Sér. I Math. **315** (1992), no. 5, 603–608.

[57] Bruce K. Driver and Vikram K. Srimurthy, *Absolute continuity of heat kernel measure with pinned wiener measure on loop groups*, Ann. Probab. **29** (2001), no. 2, 691–723.

[58] Bruce K. Driver and Anton Thalmaier, *Heat equation derivative formulas for vector bundles*, J. Funct. Anal. **183** (2001), no. 1, 42–108. MR 1 837 533

[59] Andreas Eberle, *Local Poincaré inequalities on loop spaces*, C. R. Acad. Sci. Paris Sér. I Math. **333** (2001), no. 11, 1023–1028. MR **2003c:**58025

[60] ———, *Absence of spectral gaps on a class of loop spaces*, J. Math. Pures Appl. (9) **81** (2002), no. 10, 915–955. MR **2003k:**58059

[61] ———, *Local spectral gaps on loop spaces*, J. Math. Pures Appl. (9) **82** (2003), no. 3, 313–365. MR 1 993 285

[62] ———, *Spectral gaps on discretized loop spaces*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **6** (2003), no. 2, 265–300. MR 1 991 495

[63] J. Eells and K. D. Elworthy, *Wiener integration on certain manifolds*, Problems in non-linear analysis (C.I.M.E., IV Ciclo, Varenna, 1970), Edizioni Cremonese, 1971, pp. 67–94.

[64] James Eells, *Integration on Banach manifolds*, Proceedings of the Thirteenth Biennial Seminar of the Canadian Mathematical Congress (Dalhousie Univ., Halifax, N.S., 1971), Vol. 1, Canad. Math. Congr., Montreal, Que., 1972, pp. 41–49. MR 51 #9112

[65] David Elworthy, *Geometric aspects of diffusions on manifolds*, École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87, Lecture Notes in Math., vol. 1362, Springer, Berlin, 1988, pp. 277–425. MR **90c:**58187

[66] K. D. Elworthy, *Gaussian measures on Banach spaces and manifolds*, Global analysis and its applications (Lectures, Internat. Sem. Course, Internat. Centre Theoret. Phys., Trieste, 1972), Vol. II, Internat. Atomic Energy Agency, Vienna, 1974, pp. 151–166.

[67] ———, *Measures on infinite-dimensional manifolds*, Functional integration and its applications (Proc. Internat. Conf., London, 1974), Clarendon Press, Oxford, 1975, pp. 60–68.

[68] ———, *Stochastic dynamical systems and their flows*, Stochastic Analysis (Proc. Internat. Conf., Northwestern niv.,Evanston, Ill., 1978) (New York-London), Academic Press, 1978, pp. 79–95.

[69] ———, *Stochastic differential equations on manifolds*, London Mathematical Society Lecture Note Series, vol. 70, Cambridge University Press, Cambridge, 1982. MR **84d:**58080

[70] K. D. Elworthy, Y. Le Jan, and X.-M. Li, *Integration by parts formulae for degenerate diffusion measures on path spaces and diffeomorphism groups*, C. R. Acad. Sci. Paris Sér. I Math. **323** (1996), no. 8, 921–926.

[71] K. D. Elworthy, Y. Le Jan, and Xue-Mei Li, *On the geometry of diffusion operators and stochastic flows*, Lecture Notes in Mathematics, vol. 1720, Springer-Verlag, Berlin, 1999. MR **2001f:**58072

[72] K. D. Elworthy and X.-M. Li, *Formulae for the derivatives of heat semigroups*, J. Funct. Anal. **125** (1994), no. 1, 252–286.

[73] K. D. Elworthy and Xue-Mei 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.

[74] Michel Emery, *Stochastic calculus in manifolds*, Universitext, Springer-Verlag, Berlin, 1989, With an appendix by P.-A. Meyer.

[75] O. Enchev and D. Stroock, *Integration by parts for pinned Brownian motion*, Math. Res. Lett. **2** (1995), no. 2, 161–169.

[76] O. Enchev and D. W. Stroock, *Towards a Riemannian geometry on the path space over a Riemannian manifold*, J. Funct. Anal. **134** (1995), no. 2, 392–416.

[77] S. Fang, *Stochastic anticipative calculus on the path space over a compact Riemannian manifold*, J. Math. Pures Appl. (9) **77** (1998), no. 3, 249–282. MR **99i:**60110

[78] S. Z. Fang and P. Malliavin, *Stochastic analysis on the path space of a Riemannian manifold. I. Markovian stochastic calculus*, J. Funct. Anal. **118** (1993), no. 1, 249–274.

[79] Shi Zan Fang, *Inégalité du type de Poincaré sur l'espace des chemins riemanniens*, C. R. Acad. Sci. Paris Sér. I Math. **318** (1994), no. 3, 257–260.

[80] Shizan Fang, *Rotations et quasi-invariance sur l'espace des chemins*, Potential Anal. **4** (1995), no. 1, 67–77. MR **96d:**60080

[81] ———, *Stochastic anticipative integrals on a Riemannian manifold*, J. Funct. Anal. **131** (1995), no. 1, 228–253. MR **96i:**58178

[82] ———, *Integration by parts for heat measures over loop groups*, J. Math. Pures Appl. (9) **78** (1999), no. 9, 877–894. MR 1 725 745

[83] ———, *Integration by parts formula and logarithmic Sobolev inequality on the path space over loop groups*, Ann. Probab. **27** (1999), no. 2, 664–683. MR 1 698 951

[84] ———, *Ricci tensors on some infinite-dimensional Lie algebras*, J. Funct. Anal. **161** (1999), no. 1, 132–151. MR **2000f:**58013

[85] I. B. Frenkel, *Orbital theory for affine Lie algebras*, Invent. Math. **77** (1984), no. 2, 301–352. MR **86d:**17014

[86] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine, *Riemannian geometry*, second ed., Universitext, Springer-Verlag, Berlin, 1990. MR **91j:**53001

[87] R. Gangolli, *On the construction of certain diffusions on a differenitiable manifold*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **2** (1964), 406–419.

[88] I. V. Girsanov, *On transforming a class of stochastic processes by absolutely continuous substitution of measures*, Teor. Verojatnost. i Primenen. **5** (1960), 314–330. MR 24 #A2986

[89] Fuzhou Gong, Michael Röckner, and Liming Wu, *Poincaré inequality for weighted first order Sobolev spaces on loop spaces*, J. Funct. Anal. **185** (2001), no. 2, 527–563. MR **2002j:**47074

[90] Leonard Gross, *Abstract Wiener spaces*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1, Univ. California Press, Berkeley, Calif., 1967, pp. 31–42. MR 35 #3027

[91] ———, *Potential theory on Hilbert space*, J. Functional Analysis **1** (1967), 123–181. MR 37 #3331

[92] ———, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. MR 54 #8263

[93] ———, *Logarithmic Sobolev inequalities on loop groups*, J. Funct. Anal. **102** (1991), no. 2, 268–313. MR **93b:**22037

[94] S. J. Guo, *On the mollifier approximation for solutions of stochastic differential equations*, J. Math. Kyoto Univ. **22 , no. 2** (1982), 243–254.

[95] Noel J. Hicks, *Notes on differential geometry*, Van Nostrand Mathematical Studies, No. 3, D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1965. MR 31 #3936

[96] E. P. Hsu, *Flows and quasi-invariance of the Wiener measure on path spaces*, Stochastic analysis (Ithaca, NY, 1993), Proc. Sympos. Pure Math., vol. 57, Amer. Math. Soc., Providence, RI, 1995, pp. 265–279.

[97] ———, *Quasi-invariance of the Wiener measure on the path space over a compact Riemannian manifold*, J. Funct. Anal. **134** (1995), no. 2, 417–450.

[98] Elton P. Hsu, *Inégalités de Sobolev logarithmiques sur un espace de chemins*, C. R. Acad. Sci. Paris Sér. I Math. **320** (1995), no. 8, 1009–1012.

[99] ———, *Estimates of derivatives of the heat kernel on a compact Riemannian manifold*, Proc. Amer. Math. Soc. **127** (1999), no. 12, 3739–3744. MR **2000c:**58047

[100] ———, *Quasi-invariance of the Wiener measure on path spaces: noncompact case*, J. Funct. Anal. **193** (2002), no. 2, 278–290. MR **2003i:**58069

[101] ———, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics, vol. 38, American Mathematical Society, Providence, RI, 2002. MR **2003c:**58026

[102] Y. Hu, A. S. Üstünel, and M. Zakai, *Tangent processes on Wiener space*, J. Funct. Anal. **192** (2002), no. 1, 234–270. MR **2003e:**60117

[103] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, North Holland, Amsterdam, 1981.

[104] Nobuyuki Ikeda and Shinzo Watanabe, *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam, 1989. MR **90m:**60069

[105] Peter Imkeller, *Enlargement of the Wiener filtration by a manifold valued random element via Malliavin's calculus*, Statistics and control of stochastic processes (Moscow, 1995/1996), World Sci. Publishing, River Edge, NJ, 1997, pp. 157–171. MR **99h:**60112

[106] Yuzuru Inahama, *Logarithmic Sobolev inequality on free loop groups for heat kernel measures associated with the general Sobolev spaces*, J. Funct. Anal. **179** (2001), no. 1, 170–213. MR **2001k:**60077

[107] E. Jørgensen, *The central limit problem for geodesic random walks*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **32** (1975), 1–64.

[108] H. Kaneko and S. Nakao, *A note on approximation for stochastic differential equations*, Séminaire de Probabilités, XXII, Lecture Notes in Math., vol. 1321, Springer, Berlin, 1988, pp. 155–162.

[109] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, 2:nd ed., Graduate Texts in Mathematics, no. 113, Springer Verlag, Berlin, 1991.

[110] Wilfrid S. Kendall, *Stochastic differential geometry: an introduction*, Acta Appl. Math. **9** (1987), no. 1-2, 29–60. MR **88m:**58203

[111] Wilhelm Klingenberg, *Lectures on closed geodesics*, Springer-Verlag, Berlin, 1978, Grundlehren der Mathematischen Wissenschaften, Vol. 230.

[112] Wilhelm P. A. Klingenberg, *Riemannian geometry*, de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin-New York, 1982.

[113] ———, *Riemannian geometry*, second ed., de Gruyter Studies in Mathematics, vol. 1, Walter de Gruyter & Co., Berlin, 1995.

[114] Shoshichi Kobayashi and Katsumi Nomizu, *Foundations of differential geometry. Vol. I*, John Wiley & Sons Inc., New York, 1996, Reprint of the 1963 original, A Wiley-Interscience Publication. MR **97c:**53001a

[115] ———, *Foundations of differential geometry. Vol. II*, John Wiley & Sons Inc., New York, 1996, Reprint of the 1969 original, A Wiley-Interscience Publication. MR **97c:**53001b

[116] Hiroshi Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990.

[117] T. G. Kurtz and P. Protter, *Weak limit theorems for stochastic integrals and stochastic differential equations*, Ann. Probab. **19** (1991), no. 3, 1035–1070.

[118] ———, *Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs*, Stochastic analysis, Academic Press, Boston, MA, 1991, pp. 331–346.

[119] S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus. II*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **32** (1985), no. 1, 1–76. MR **86k:**60100b

[120] ———, *Applications of the Malliavin calculus. III*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **34** (1987), no. 2, 391–442. MR **89c:**60093

[121] Shigeo Kusuoka and Daniel Stroock, *Applications of the Malliavin calculus. I*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 271–306. MR **86k:**60100a

[122] R. Léandre, *Integration by parts formulas and rotationally invariant Sobolev calculus on free loop spaces*, J. Geom. Phys. **11** (1993), no. 1-4, 517–528, Infinite-dimensional geometry in physics (Karpacz, 1992).

[123] R. Léandre and J. R. Norris, *Integration by parts Cameron–Martin formulas fo the free path space of a compact Riemannian manifold*, 1995 Warwick Univ. Preprint, 1995.

[124] Rémi Léandre, *Développement asymptotique de la densité d'une diffusion dégénérée*, Forum Math. **4** (1992), no. 1, 45–75. MR **93d:**60100

[125] Xiang Dong Li, *Existence and uniqueness of geodesics on path spaces*, J. Funct. Anal. **173** (2000), no. 1, 182–202. MR **2001f:**58074

[126] T. J. Lyons and Z. M. Qian, *Calculus for multiplicative functionals, Itô's formula and differential equations*, Itô's stochastic calculus and probability theory, Springer, Tokyo, 1996, pp. 233–250.

[127] ———, *Stochastic Jacobi fields and vector fields induced by varying area on path spaces*, Imperial College of Science, 1996.

[128] Marie-Paule Malliavin and Paul Malliavin, *An infinitesimally quasi-invariant measure on the group of diffeomorphisms of the circle*, Special functions (Okayama, 1990), ICM-90 Satell. Conf. Proc., Springer, Tokyo, 1991, pp. 234–244. MR **93h:**58027

[129] Paul Malliavin, *Geometrie differentielle stochastique*, Séminaire de Mathématiques Supérieures, Presses de l'Université de Montréal, Montreal, Que, 1978, Notes prepared by Danièle Dehen and Dominique Michel.

[130] ———, *Stochastic calculus of variation and hypoelliptic operators*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976) (New York-Chichester-Brisbane), Wiley, 1978, pp. 195—263,.

[131] ———, *Stochastic jacobi fields*, Partial Differential Equations and Geometry (Proc. Conf., Park City, Utah, 1977 (New York), Dekker, 1979, Lecture Notes in Pure and Appl. Math., 48, pp. 203–235.

[132] ———, *Stochastic analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 313, Springer-Verlag, Berlin, 1997.

[133] Gisiro Maruyama, *Notes on Wiener integrals*, Kōdai Math. Sem. Rep. **1950** (1950), 41–44. MR 12,343d

[134] E. J. McShane, *Stochastic differential equations and models of random processes*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, Univ. California Press, Berkeley, Calif., 1972, pp. 263–294.

[135] ———, *Stochastic calculus and stochastic models*, Academic Press, New York, 1974, Probability and Mathematical Statistics, Vol. 25.

[136] Tai Melcher, *Hypoelliptic heat kernel inequalities on Lie groups*, Ph.D. Thesis, 2004. See www.math.ucsd.edu/˜driver/driver/thesis.html.

[137] P.-A. Meyer, *A differential geometric formalism for the Itô calculus*, Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980), Lecture Notes in Math., vol. 851, Springer, Berlin, 1981, pp. 256–270. MR **84e:**60084

[138] Jürgen Moser, *A new technique for the construction of solutions of nonlinear differential equations*, Proc. Nat. Acad. Sci. U.S.A. **47** (1961), 1824–1831. MR 24 #A2695

[139] ———, *A rapidly convergent iteration method and non-linear differential equations. II*, Ann. Scuola Norm. Sup. Pisa (3) **20** (1966), 499–535. MR 34 #6280

[140] ———, *A rapidly convergent iteration method and non-linear partial differential equations. I*, Ann. Scuola Norm. Sup. Pisa (3) **20** (1966), 265–315. MR 33 #7667

[141] J.-M. Moulinier, *Théorème limite pour les équations différentielles stochastiques*, Bull. Sci. Math. (2) **112** (1988), no. 2, 185–209.

[142] S. Nakao and Y. Yamato, *Approximation theorem on stochastic differential equations*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976) (New York-Chichester-Brisbane), Wiley, 1978, pp. 283–296.

[143] John Nash, *The imbedding problem for Riemannian manifolds*, Ann. of Math. (2) **63** (1956), 20–63. MR 17,782b

[144] J. R. Norris, *A complete differential formalism for stochastic calculus in manifolds*, Séminaire de Probabilités, XXVI, Lecture Notes in Math., vol. 1526, Springer, Berlin, 1992, pp. 189–209. MR **94g:**58254

[145] ———, *Path integral formulae for heat kernels and their derivatives*, Probab. Theory Related Fields **94** (1993), no. 4, 525–541.

[146] ———, *Twisted sheets*, J. Funct. Anal. **132** (1995), no. 2, 273–334. MR **96f:**60094

[147] James Norris, *Simplified Malliavin calculus*, Séminaire de Probabilités, XX, 1984/85, Lecture Notes in Math., vol. 1204, Springer, Berlin, 1986, pp. 101–130.

[148] David Nualart, *The Malliavin calculus and related topics*, Probability and its Applications (New York), Springer-Verlag, New York, 1995. MR **96k:**60130

[149] Barrett O'Neill, *Semi-Riemannian geometry*, Pure and Applied Mathematics, vol. 103, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1983, With applications to relativity. MR **85f:**53002

[150] Jean Picard, *Gradient estimates for some diffusion semigroups*, Probab. Theory Related Fields **122** (2002), no. 4, 593–612. MR **2003d:**58056

[151] Mark A. Pinsky, *Stochastic Riemannian geometry*, Probabilistic analysis and related topics, Vol. 1 (New York), Academic Press, 1978, pp. 199–236.

[152] M. M. Rao, *Stochastic processes: general theory*, Mathematics and its Applications, vol. 342, Kluwer Academic Publishers, Dordrecht, 1995. MR **97c:**60092

[153] Jang Schiltz, *Time dependent Malliavin calculus on manifolds and application to nonlinear filtering*, Probab. Math. Statist. **18** (1998), no. 2, Acta Univ. Wratislav. No. 2111, 319–334. MR **2000b:**60144

[154] Laurent Schwartz, *Semi-martingales sur des variétés, et martingales conformes sur des variétés analytiques complexes*, Lecture Notes in Mathematics, vol. 780, Springer, Berlin, 1980. MR **82m:**60051

[155] ———, *Géométrie différentielle du 2ème ordre, semi-martingales et équations différentielles stochastiques sur une variété différentielle*, Seminar on Probability, XVI, Supplement, Lecture Notes in Math., vol. 921, Springer, Berlin, 1982, pp. 1–148. MR **83k:**60064

[156] ———, *Semimartingales and their stochastic calculus on manifolds*, Collection de la Chaire Aisenstadt. [Aisenstadt Chair Collection], Presses de l'Université de Montréal, Montreal, QC, 1984, Edited and with a preface by Ian Iscoe. MR **86b:**60085

[157] Ichiro Shigekawa, *Absolute continuity of probability laws of Wiener functionals*, Proc. Japan Acad. Ser. A Math. Sci. **54** (1978), no. 8, 230–233. MR **81m:**60097

[158] ———, *Derivatives of Wiener functionals and absolute continuity of induced measures*, J. Math. Kyoto Univ. **20** (1980), no. 2, 263–289. MR **83g:**60051

[159] ———, *On stochastic horizontal lifts*, Z. Wahrsch. Verw. Gebiete **59** (1982), no. 2, 211–221. MR **83i:**58102

[160] ———, *Transformations of the Brownian motion on a Riemannian symmetric space*, Z. Wahrsch. Verw. Gebiete **65** (1984), no. 4, 493–522.

[161] ———, *Transformations of the Brownian motion on the Lie group*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 409–422.

[162] Ichirō Shigekawa, *de Rham-Hodge-Kodaira's decomposition on an abstract Wiener space*, J. Math. Kyoto Univ. **26** (1986), no. 2, 191–202. MR **88h:**58009

[163] Ichiro Shigekawa, *Differential calculus on a based loop group*, New trends in stochastic analysis (Charingworth, 1994), World Sci. Publishing, River Edge, NJ, 1997, pp. 375–398. MR **99k:**60146

[164] Michael Spivak, *A comprehensive introduction to differential geometry. Vol. I*, second ed., Publish or Perish Inc., Wilmington, Del., 1979. MR **82g:**53003a

[165] Robert S. Strichartz, *Analysis of the Laplacian on the complete Riemannian manifold*, J. Funct. Anal. **52** (1983), no. 1, 48–79. MR **84m:**58138
[166] D. Stroock and S. Taniguchi, *Diffusions as integral curves, or Stratonovich without Itô*, The Dynkin Festschrift, Progr. Probab., vol. 34, Birkhäuser Boston, Boston, MA, 1994, pp. 333–369.
[167] D. W. Stroock and S. R. S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, Univ. California Press, Berkeley, Calif., 1972, pp. 333–359.
[168] Daniel W. Stroock, *The Malliavin calculus, a functional analytic approach*, J. Funct. Anal. **44** (1981), no. 2, 212–257. MR **83h:**60076
[169] ———, *The Malliavin calculus and its application to second order parabolic differential equations. I*, Math. Systems Theory **14** (1981), no. 1, 25–65. MR **84d:**60092a
[170] ———, *The Malliavin calculus and its application to second order parabolic differential equations. II*, Math. Systems Theory **14** (1981), no. 2, 141–171. MR **84d:**60092b
[171] ———, *An introduction to the analysis of paths on a Riemannian manifold*, Mathematical Surveys and Monographs, vol. 74, American Mathematical Society, Providence, RI, 2000. MR **2001m:**60187
[172] Daniel W. Stroock and James Turetsky, *Short time behavior of logarithmic derivatives of the heat kernel*, Asian J. Math. **1** (1997), no. 1, 17–33. MR **99b:**58225
[173] ———, *Upper bounds on derivatives of the logarithm of the heat kernel*, Comm. Anal. Geom. **6** (1998), no. 4, 669–685. MR **99k:**58174
[174] Daniel W. Stroock and S. R. S. Varadhan, *Diffusion processes with continuous coefficients. II*, Comm. Pure Appl. Math. **22** (1969), 479–530.
[175] H. J. Sussmann, *Limits of the Wong-Zakai type with a modified drift term*, Stochastic analysis, Academic Press, Boston, MA, 1991, pp. 475–493.
[176] Setsuo Taniguchi, *Malliavin's stochastic calculus of variations for manifold-valued Wiener functionals and its applications*, Z. Wahrsch. Verw. Gebiete **65** (1983), no. 2, 269–290. MR **85d:**58088
[177] Krystyna Twardowska, *Approximation theorems of Wong-Zakai type for stochastic differential equations in infinite dimensions*, Dissertationes Math. (Rozprawy Mat.) **325** (1993), 54. MR **94d:**60092
[178] Nolan R. Wallach, *Harmonic analysis on homogeneous spaces*, Marcel Dekker Inc., New York, 1973, Pure and Applied Mathematics, No. 19. MR 58 #16978
[179] S. Watanabe, *Lectures on stochastic differential equations and Malliavin calculus*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics, vol. 73, Published for the Tata Institute of Fundamental Research, Bombay, 1984, Notes by M. Gopalan Nair and B. Rajeev. MR **86b:**60113
[180] E. Wong and M. Zakai, *On the relation between ordinary and stochastic differential equations*, Internat. J. Engrg. Sci. **3** (1965), 213–229.
[181] ———, *On the relation between ordinary and stochastic differential equations and applications to stochastic problems in control theory*, Automatic and remote control III (Proc. Third Congr. Internat. Fed. Automat. Control (IFAC), London, 1966), Vol. 1, p. 5, Paper 3B, Inst. Mech. Engrs., London, 1967, p. 8.

Department of Mathematics, 0112, University of California at San Diego, La Jolla, CA, 92093-0112

E-mail address: `driver@math.ucsd.edu`