JOURNAL OF FUNCTIONAL ANALYSIS 110, 272–376 (1992)

A Cameron–Martin Type Quasi-invariance Theorem for Brownian Motion on a Compact Riemannian Manifold

BRUCE K. DRIVER*

Department of Mathematics, 0112, University of California, San Diego, La Jolla, California 92093-0112

Communicated by L. Gross

Received October 7, 1991

Let M be a compact manifold without boundary, o be a fixed base point in M, g be a Riemannian metric on M, and ∇ be a g-compatible covariant derivative on TM—the tangent space to M. Assume the torsion (T) of ∇ satisfies the skew symmetry condition: $g\langle T\langle X, Y\rangle, Y\rangle \equiv 0$ for all vector fields X and Y on M. (For example, take ∇ to be the Levi–Civita covariant derivative on (M, g).) Also let ν denote the Wiener measure on $W_o(M) = \{\omega \in C([0,1], M): \omega(0) = o\}$, and let $H(\omega)(s)$ denote stochastic parallel translation (relative to ν) along the path $\omega \in W_o(M)$ up to time $s \in [0,1]$. Given a C^1-function $h: [0,1] \to T_oM$, it is shown that the differential equation $\dot\sigma(t) = H(\sigma(t))h$ with initial condition $\sigma(0) = \mathrm{id}: W(M) \to W(M)$ has a solution $\sigma: \mathbb{R} \to \mathrm{Maps}(W(M), W(M))$—the measurable maps from $W(M)$ to $W(M)$. This function (σ) is a flow on $W(M)$, i.e., for all $t, \tau \in \mathbb{R}$, $\sigma(t+\tau) = \sigma(t) \circ \sigma(\tau)$ ν-a.s. Furthermore $\sigma(t)$ has the quasi-invariance property: the law $(\sigma(t)_* \nu)$ of $\sigma(t)$ with respect to the Wiener measure (ν) is equivalent to ν for all $t \in \mathbb{R}$. This result is used to prove an integration by parts formula for the h-derivative $\partial_h f$ defined by $\partial_h f(\omega) \equiv (d/dt)|_0 f(\sigma(t)(\omega))$, where f is a "C^2-cylinder" function on $W(M)$.

1. INTRODUCTION

Let H denote the Hilbert space of absolutely continuous functions functions h: $[0,1] \to \mathbb{R}^n$ such that $h(0) = 0$ and $(h, h) \equiv \int_0^1 |h'(s)|^2\, ds < \infty$. Recall the classical Cameron–Martin Theorem [CM1, CM2, Mar] which states that if μ is the standard Wiener measure on $W(\mathbb{R}^n) \equiv C([0,1], \mathbb{R}^n)$ and $h \in H$, then μ is quasi-invariant under the transformation

$$(\omega \to \omega + h): W(\mathbb{R}^n) \to W(\mathbb{R}^n). \tag{1.1}$$

Furthermore, the Radon–Nikodym derivative is given by

$$d\mu(\omega + h)/d\mu(\omega) = \exp\left(-\int_0^1 h'(s) \cdot d\omega(s) - \tfrac{1}{2}\int_0^1 |h'(s)|^2\, ds\right), \tag{1.2}$$

* This research was partially supported by NSF Grant DMS 91–01720.

0022-1236/92 $5.00

where $\int_0^1 h'(s) \cdot d\omega(s)$ is an Itô stochastic integral, see [K3, Theorem 1.2, p. 113]. The purpose of this paper is to prove the analogous result for the Wiener measure on the path space of a compact Riemannian manifold.

Let (M, g, ∇) be given, where M is a compact manifold without boundary, g is a Riemannian metric on M, and ∇ is a g-compatible covariant derivative on TM (see Section 2.) Fix a base point $o \in M$ and a C^1-function h: $[0, 1] \to T_oM$ (= tangent space of M at o) such that $h(0) = 0$. Let ν denote the Wiener measure on the path space $W(M) \equiv C([0, 1], M)$ which is concentrated on the set of based paths $W_o(M) = \{\omega \in W(M): \omega(0) = o \in M\}$. Let $H(\omega)(s)$ denote stochastic parallel translation (or horizontal lifting) along ω (relative to ν) on the interval $[0, s]$, see Section 3, and in particular Theorem 3.2. (Notice that H is really an equivalence class of processes with two processes equivalent if they are equal ν-a.s.) With the above data, define the "vector field" (X^h) (or more precisely an equivalence class of vector fields) on $W_o(M)$ by

$$X^h(\omega)(s) \equiv H(\omega)(s)\, h(s). \tag{1.3}$$

Notice that for each ω, $X^h(\omega)$ is a vector field along the curve ω. Hence, it is reasonable to interpret $X^h(\omega)$ as a tangent vector at $\omega \in W_o(M)$. See Malliavin [M1, M3] for some proposed methods of defining the tangent space $(TW(M))$ to $W(M)$ and equipping it with a Riemannian metric. Also see Remark 2.3, where it is pointed out that the map $h \to X^h(\omega)$ is an isometry with respect to the "natural" metrics on H and $T_\omega W_o(M)$.

Given such a vector field X^h, it is natural to try to construct its flow. In other words, one wants to find a function σ: $\mathbb{R} \to \text{Maps}(W(M), W(M))$ (the measurable maps from $W(M)$ to $W(M)$) which solves the initial value problem,

$$\dot\sigma(t)(\omega) = X^h(\sigma(t)(\omega)) \qquad \text{with} \quad \sigma(0)(\omega) = \omega, \tag{1.4}$$

at least for ν-almost every $\omega \in W_o(M)$. (Note that the s-variable which is taken to be the parameter for paths in $W(M)$ is now suppressed. This convention will be used whenever possible throughout this paper.)

Remark 1.1. If $M = \mathbb{R}^n$, with the usual metric and covariant derivative and $o = 0$, then $X^h(\omega)(s) = h(s)$ under the natural identification of $T\mathbb{R}^n$ with $\mathbb{R}^n \times \mathbb{R}^n$. For this case one easily solves (1.4) to find

$$\sigma(t)(\omega) = \omega + th,$$

which at $t = 1$ is the transformation (1.1) used by Cameron and Martin. The reader is referred to Example 5.1 for the more general Lie Group cases where one can still explicitly solve (1.4).

Because of the above remark it is reasonable to consider $\sigma(1)$, where $\sigma(t)$ solves (1.4), as the generalization of the transformation (1.1). Now assuming the existence of a solution (σ) to (1.4), it is natural to ask whether the map $\sigma(t)$: $W(M) \to W(M)$ leaves the Wiener measure (ν) quasi-invariant? In other words, is the law $(\nu_t \equiv \sigma(t)_* \nu)$ of $\sigma(t)$ with respect to ν equivalent to ν for all t? Recall two measures ν_t and ν are said to be equivalent it they are mutually absolutely continuous with respect to one another. Suppose for the moment $\nu_t \equiv \sigma(t)_* \nu$ is not absolutely continuous with respect to ν. In this case $X^h(\sigma(t))$ is no longer well defined, since $X^h(\omega)$ was only defined up to ν-equivalence. More precisely the equivalence class of the process $X^h(\sigma(t))$ will now depend on the particular representative chosen for X^h. This certainly renders Eq. (1.4) meaningless. Therefore, the issues of existence and quasi-invariance of solutions to (1.4) are inseparable.

Because of the difficulties discussed above, it will be beneficial to reinterpret the meaning of (1.4). The modification is as follows. First, define the coordinate functions $\sigma_o(s)$: $W(M) \to M$ by $\sigma_o(s)(\omega) = \omega(s)$ for each $s \in [0, 1]$. We now say a solution to (1.4) is a path ($\sigma(t)$ for $t \in \mathbb{R}$) of M-valued semimartingales solving

$$\dot{\sigma}(t) = H(\sigma(t)) \cdot h \qquad \text{with} \quad \sigma(0) = \sigma_o, \tag{1.5}$$

where $H(\sigma(t))$ is the stochastic parallel translation (or horizontal lift) of the M-valued semimartingale $\sigma(t)$. (We are now suppressing both the random sample path ω and the s-parameter from the notation.) Of course the process $\sigma(t)$ must be suitably differentiable in the t-variable. Recall that the notion of stochastic parallel translation along any continuous M-valued semimartingale is always well defined, see Theorem 3.2 below.

The reformulation in (1.5) of (1.4) has the advantage that it makes sense even if the flow does not satisfy the quasi-invariance property. It is now possible to summarize the main results of this paper.

THEOREM 1.1. *Let $(M, g, \nabla, W(M), H, \sigma_o)$ be as above and suppose that h: $[0, 1] \to T_o M$ is a C^1-function such that $h(0) = 0$. Then there is a unique solution σ: $\mathbb{R} \to$ "Brownian semimartingales" on M satisfying* (1.5).

See Definition 4.1 for the notion of a Brownian semimartingale. Theorem 1.1 is a consequence of Corollary 6.3. Two proofs of Corollary 6.3 are given in this paper, one in Section 6 and the other in Section 7. The condition that h is C^1 rather than an element of H is an unnatural restriction which is needed for technical reasons. I would expect the results in this paper to be true for all $h \in H$.

The next theorem is a combination of Theorem 8.1 and Theorem 8.5. In order to state the theorem, recall that a continuous M-valued process $\{X_s\}_{s \in [0, 1]}$ defined on a measurable space Ω may be thought of as a function from Ω to $W(M)$. This function is still denoted by X and is given by $X(\omega)=(s \to X_s(\omega))$. I will write X_s or $X(s)$ interchangeably depending on which notation is more convenient.

THEOREM 1.2. *Keep the assumptions and notation in Theorem* 1.1. *Further assume that* ∇ *is torsion skew symmetric* (*see Definition* 8.1), *then* $\nu_t = \sigma(t)_* \nu$ (*the law of* $\sigma(t)$) *is equivalent to* ν, *for all* t. *If* $\sigma(t)$ *is viewed as a function from* $W(M)$ *to* $W(M)$, *then* $\sigma(t)$ *solves* (1.4) *and is a flow on* $W(M)$, *i.e., for all* $t, \tau \in \mathbb{R}$, $\sigma(t) \circ \sigma(\tau) = \sigma(t+\tau)$ ν-*a.s. Furthermore, if* $\rho_t \equiv d\nu_t/d\nu$ *is the Radon–Nikodym derivative of* ν_t *with respect to* ν, *then* ρ_t^r *is* ν-*integrable for all* $r \in \mathbb{R}$. (*See* (8.16) *and Theorem* 8.5 *for a formula for* ρ_t.)

Remark 1.2. The Levi–Civita covariant derivative is an example of a torsion skew symmetric covariant derivative. See Example 8.1 for more examples.

The integration by parts formula in the next theorem is an easy consequence of Theorem 1.2. Theorem 1.3 is a combination of Theorem 9.1 and Proposition 9.1.

THEOREM 1.3.[1] *Keep the notation and hypothesis of Theorem* 1.2. *Let* $(\cdot,\cdot)$ *denote the inner product on* $L^2(W(M), d\nu)$. *For a function* f: $W(M) \to \mathbb{R}$ *let* $\partial_h f \equiv (d/dt)|_0 f(\sigma(t))$: $W(M) \to \mathbb{R}$, *if the derivative exists in probability. Then there is a function* $z(h)$: $W(M) \to \mathbb{R}$ *such that for all* "C^2-*cylinder functions*" f *and* g *on* $W(M)$ *one has the integration by parts formula*:

$$(\partial_h f, g) = (f, -\partial_h g + z(h) \cdot g).$$

Furthermore, there are constants $\varepsilon > 0$ *and* $K > 1$ *independent of* h *such that for all* $h \neq 0$ *in* C^1, $\nu(e^{\varepsilon[z(h)/\|h'\|]^2}) \leqslant K < \infty$, *where* $\|h'\|^2 \equiv \int_0^1 |h'(s)|^2\, ds$.

See Definition 9.1 for the notion of a C^2-cylinder function, and Eq. (9.2) for the explicit formula for $z(h)$ involving the Ricci curvature and the torsion of the covariant derivative ∇.

Remark 1.3. The analogues of Theorems 1.2–1.3 are valid if the Wiener measure (ν) is replaced by a pinned Wiener measure on $W(M)$, i.e., a Brownian bridge. See Driver [D3].

There have been numerous other nonlinear extensions to the classical Cameron–Martin theorem in the literature. The first such extension was by Cameron and Martin [CM3]. The later nonlinear extensions are for the most part done in the setting of Gross' [Gr2] abstract Wiener spaces, see

[1] Leandre [Le] has recently proved this integration by parts formula directly using methods of Bismut.

Gross [Gr1], Kuo [K1, K2], Ramer [Ra], and Kusuoka [Ku1–Ku4]. The order is both chronological and by degree of generality.

Results which are closer to those of this paper may be found in Albeverio and Hoegh-Krohn [AH], Shigekawa [Sh1, Sh2] Frenkel [Fr], M.-P. Malliavin [MM1], and Gross [Gr4]. All of these papers include quasi-invariant results for the Wiener measure (or the pinned Wiener measure) on the based (loops) paths of a compact Lie group. See Sections 5 and 10 for a more thorough discussion of how the results in this paper relate to the Lie group and homogeneous space cases discussed in [AH, Sh1, Sh2, Fr, MM1, Gr4]. Some other related references that the reader may wish to consult are Airault and Malliavin [AM1, AM2], Airault and Van Biesen [AV1, AV2], Epperson and Lohrenz [EL1, EL2], Getzler [Ge], Jones and Leandre [JL], P. Malliavin [M1–M3], M.-P. Malliavin and M. Malliavin [MM2–MM4], and Pressley and Segal [PS].

At this point it should be pointed out that a majority of this manuscript is devoted to Theorem 1.1—the existence of the nonlinear transformations $\sigma(t)$. Once Theorem 1.1 is proved, the quasi-invariance issue is quite easily settled with the aid of Girsanov's Theorem [Gi]—which is yet another extension of the classical Cameron–Martin theorem.

The closest result in the literature (to the author's best knowledge) relating to Theorem 1.2 is the work of Cruzeiro [Cr]. Roughly stated, Cruzeiro proves the existence of flows for a certain class of vector-fields on the standard Wiener space. She also shows that these flows satisfy the quasi-invariance property. As stated, Theorem 1.2 involves a flow on $W(M)$, rather than the path space $W(\mathbb{R}^n)$. However, by using the stochastic development of Eells and Elworthy and P. Malliavin (see, for example, [EE, Ell, Em, IW] and Section 3 below), it is possible to transfer the differential equation (1.4) or (1.5) on $W(M)$ to a differential equation on $W(\mathbb{R}^n)$. (The underlying measure on the path space $W(\mathbb{R}^n)$ will be the standard Wiener measure denoted by μ.) When this is done the "vector field" X^h on $W(M)$ becomes a vector field $\bar{X}^h$ on $W(\mathbb{R}^n)$ (defined μ-a.s.) which has the form

$$\bar{X}^h(\omega) = \int C(\omega)\, d\omega + \int R(\omega)\, ds, \tag{1.6}$$

where ω is now a path in $W(\mathbb{R}^n)$, and C and R functions on $W(\mathbb{R}^n)$ such that μ-a.s. $s \to (C(\omega)(s), R(\omega)(s))$ is an adapted continuous $\mathrm{End}(\mathbb{R}^n) \times \mathbb{R}^n$-valued process. This result is included in Theorem 5.1 and Proposition 6.1, where the reader may find explicit formulas for C and R. Roughly speaking Cruzeiro's results could be used to prove the existence of a flow generated by (1.6) provided $C \equiv 0$. But as a rule in our situation C is seldom zero. (A notable exception to this rule is when M is a commutative Lie group, i.e., $\mathbb{R}^n$ or a torus.) It should be noted, however, as in [MM1], if one knows a priori the existence of the flow for (1.6) and one has good control over the "infinitesimal density" $(z(h))$, then one can still use the technique of [Cr] to prove the quasi-invariance property. This technique is not used in this paper. We will only use the more standard Girsanov theorem.

As with Eq. (1.4), it is better to first consider $\bar{X}^h$ as inducing a flow on the space of $\mathbb{R}^n$-valued semimartingales rather than on $W(\mathbb{R}^n)$ itself. That is, one should consider the equation

$$\dot{w}(t) = \int C(w(t))\, dw(t) + \int R(w(t))\, ds \qquad \text{with} \quad w(0) = b, \tag{1.7}$$

where $w(t)$ is now a path of $\mathbb{R}^n$-valued semimartingales, and b is a standard $\mathbb{R}^n$-valued Brownian motion. (For example, $b(s)$ could be the coordinate process on $(W(\mathbb{R}^n), \mu)$, $b(s)(\omega) = \omega(s)$.) Equation (1.7) is solved in Theorem 6.1. The idea of the proof is simple. First assume that $w(t)$ is a Brownian semimartingale, that is to say, $w(t) = \int O(t)\, db + \int \alpha(t)\, ds$, where for each t in $\mathbb{R}$, $(O(t), \alpha(t))$ is a continuous adapted $\operatorname{End}(\mathbb{R}^n) \times \mathbb{R}^n$-valued process. The initial condition $w(0) = b$ implies the initial condition $(O(0), \alpha(0)) = (\mathrm{Id}, 0)$ for $(O(t), \alpha(t))$. Now insert this form of $w(t)$ into (1.7) to find the equations

$$\dot{O}(t) = C(w(t))\, O(t) \qquad \text{with} \quad O(0) = I, \tag{1.8}$$

and

$$\dot{\alpha}(t) = C(w(t))\, \alpha(t) + R(w(t)) \qquad \text{with} \quad \alpha(0) \equiv 0. \tag{1.9}$$

These equations are then solved by a modified Piccard iteration scheme.

We can now explain how the torsion skew symmetry condition enters into Theorem 1.2. First it is shown that proving quasi-invariance for $w(t)$ proves quasi-invariance for $\sigma(t)$, see Theorem 8.2. It is easy to see that in order for the law of $w(t)$ to be equivalent to μ (= Wiener measure on $W(\mathbb{R}^n)$), requires $w(t)$ to have the same quadratic covariation as b, see Lemma 8.1. But this implies that $O(t)$ must be an $O(n)$-valued process, where $O(n)$ denotes the orthogonal group on $\mathbb{R}^n$. But by (1.8), if $O(t)$ is orthogonal for all t, then the process $C(w(t))$ must be $so(n)$-valued, where $so(n)$ is the Lie algebra of $O(n)$ consisting of skew symmetric real $n \times n$-matrices. But the explicit formula for C (see (5.6)–(5.8)) shows that one can not expect $C(w(t))$ to be skew-symmetric unless the torsion tensor has the skew symmetry property in Definition 8.1.

This paper is divided into ten sections. Section 2 introduces the differential geometric notation used in the paper. This section also provides a "smooth" warmup to the stochastic calculations done later. The purely differential geometric Theorem 2.4 is also proved. (Theorem 2.4 is mainly used in the second proof of Corollary 6.3 given in Section 7.) Section 3 reviews some basic definitions and properties about manifold valued semimartingales. It is in this section where the stochastic horizontal lift and the development maps are reviewed. Both of these constructions will play a crucial role throughout the manuscript. Section 4 is devoted to deriving the basic estimates needed for proving existence for the flows defined by Eqs. (1.5) and (1.7). (On the first reading the reader should probably omit the proofs in this section.) In Section 5 and the beginning of Section 6 we show that a solution to any one of the differential equations (1.5), (1.7), or (1.8)–(1.9) can be used to construct a solution for the remaining two differential equations. The rest of Section 6 is concerned with proving existence and uniqueness to Eqs. (1.8) and (1.9). Section 7 contains an alternate proof for existence and uniqueness of solutions to (1.5). Section 8 deals with the issues of quasi-invariance and the existence of flows on $W(M)$ or $W(\mathbb{R}^n)$. Section 9 is devoted to the integration by parts formula for the h-derivative. Finally, in Section 10 (also see Section 5) we discuss some less satisfactory alternatives to Eq. (1.4).

2. Geometric Preliminaries

In the beginning of this section I will fix some notation and review some basic facts from differential geometry. The rest of the section is devoted to studying the flow equation (1.1) in the smooth category. The computations done here will be used as a guide for the stochastic case.

First, some general comments on notation. I usually use angled brackets ($\langle\ \rangle$) to enclose the arguments of a function on which the function depends linearly. For example, $F(x)\langle a, b\rangle$ would denote a function which is typically nonlinear in the x-variable but is linear in a for fixed (x, b) and linear in b for fixed (x, a). If the variables a and b are elements of a vector-bundle, then linear is to be interpreted as fiber linear. If F is a fiber bundle over a manifold M, then $\Gamma(F)$ will denote the smooth sections of F. If σ is a smooth curve in M then $\Gamma_\sigma(F)$ will denote the set of smooth sections along σ. Suppose that $f\colon M \to N$ is a differentiable map between manifolds M and N, then the differential of f will be denoted by f_*. If N is a vector space, the differential $df\langle\dot\sigma(0)\rangle \equiv (d/dt)|_0 f(\sigma(t)) \in N$ will be used frequently, where $\sigma(t)$ is a smooth curve in M. If both M and N are vector spaces, then we may define the differential f' by $f'(m)\,v \equiv (d/dt)|_0 f(m+tv) \in N$ for all m and v in M. Notice that f_* maps TM to TN, df maps TM to N, and f' maps M to $\mathrm{Hom}(M, N)$, where $\mathrm{Hom}(M, N)$ denotes the space of linear maps from M to N. Finally, if α and β are 1-forms on a manifold M, the two form $\alpha \wedge \beta$ will be identified with the alternating multilinear map on TM given by

$$\alpha \wedge \beta\langle v, w\rangle \equiv \alpha\langle v\rangle \cdot \beta\langle w\rangle - \alpha\langle w\rangle \cdot \beta\langle v\rangle$$

for all v, $w \in T_m M$ and $m \in M$. (Warning: This convention differs from Kobayashi and Nomizu [KN] by a factor of 2. This explains the factor of 2 discrepancies between formulas quoted in this paper and those in [KN].)

Throughout this paper the following data will be fixed. Let (M^n, ∇, g, o, u_o) be a smooth compact n-dimensional Riemannian manifold with metric g, a g-compatible covariant derivative ∇, a fixed base point $o \in M$, and a fixed orthogonal frame u_o above o. Recall that a covariant derivative ∇ on TM is said to be g-compatible if $\nabla g \equiv 0$, i.e., $X(g\langle Y, Z\rangle) = g\langle \nabla_X Y, Z\rangle + g\langle Y, \nabla_X Z\rangle$ for all $X, Y, Z \in \Gamma(TM)$. Also recall that an orthogonal frame at $m \in M$ is a linear isometry $u: \mathbb{R}^n \to T_m M$.

The principal bundle of orthonormal frames is denoted by $\pi: O(M) \to M$ or $O(M)$ for short, where π is the canonical fiber projection. We are mostly interested in the path spaces $W_0(\mathbb{R}^n)$, $W_o(M)$, and $W_{u_o}(O(M))$, where the following notation is being used.

Notation 2.1. If (Q, q) is a pointed manifold ($q \in Q$ given) then $W(Q) \equiv C([0,1], Q)$ is the set of continuous paths in Q and $W_q(Q) \equiv \{\omega \in W(Q): \omega(0) = q\}$ is the subset of based paths. Also let $W^\infty(Q)(W_q^\infty(Q))$ denote the set of smooth (based) paths in Q.

Given a smooth path σ in M and a smooth vector field X along σ, let $\nabla X/ds \in \Gamma_\sigma(TM)$ denote the covariant derivative $\sigma^*\nabla X$ of X, where $\sigma^*\nabla$ is the pull-back of ∇ to sections along σ. Also if $u \in \Gamma_\sigma(O(M))$, define $\nabla u/ds \in \Gamma_\sigma(E)$ by $(\nabla u/ds)(s)\cdot\xi = (\nabla/ds)(u(s)\xi)$ for all ξ in $\mathbb{R}^n$, where $E \equiv \mathrm{Hom}(\mathbb{R}^n, TM)$ is the vector-bundle over M with fiber $E_m \equiv \mathrm{Hom}(\mathbb{R}^n, T_m M)$ for each $m \in M$. Note because ∇ is g-compatible, $u(s)^{-1}(\nabla/ds)(u(s))$ is in $so(n)$—the Lie algebra of $O(n)$ consisting of skew-symmetric $n \times n$ real matrices.

The covariant derivative ∇ is equivalent to a connection on the principal bundle $O(M)$. Namely, let $\omega = \omega^\nabla$ be the connection 1-form on $O(M)$ with values in $so(n)$ defined by $\omega\langle u'(s)\rangle = u(s)^{-1}(\nabla u/ds)(s)$ where $u(s)$ is any smooth path in $O(M)$. Furthermore, this ω induces the covariant derivative ∇ on the associated bundle $TM \cong O(M) \times_{O(n)} \mathbb{R}^n$.

The following definitions are standard, see Kobayashi and Nomizu [KN].

DEFINITION 2.1. The canonical 1-form on $O(M)$ is the $\mathbb{R}^n$-valued 1-form (ϑ) on $O(M)$ given by $\vartheta\langle\xi\rangle = u^{-1}\pi_*\xi$ for all $\xi \in T_u O(M)$ and $u \in O(M)$.

DEFINITION 2.2. The standard horizontal vector fields $B\langle a\rangle(\cdot) \in \Gamma(TO(M))$ for $a \in \mathbb{R}^n$ are defined by the following: $B\langle a\rangle(u)$ is the horizontal lift of $ua \in TM$ to $T_u O(M)$ for each u in $O(M)$. So $B\langle a\rangle(u)$ is the unique element in $T_u O(M)$ such that $\pi_* B\langle a\rangle(u) = ua$ and $\omega\langle B\langle a\rangle(u)\rangle = 0$.

Remark 2.1. For $A \in so(n)$ and u in $O(M)$, let $u \cdot A \in T_u O(M)$ denote the tangent vector $u \cdot A \equiv (d/dt)|_0 \, ue^{tA}$. With this notation it is easy to check that the decomposition of a tangent vector $\xi_u \in T_u O(M)$ into vertical and horizontal components is given by $\xi_u = u \cdot \omega\langle \xi_u \rangle + B\langle \vartheta\langle \xi_u \rangle \rangle(u)$. One should also note that if $u(t)$ is a smooth path in $O(M)$ then $\omega\langle \dot{u}(t) \rangle = u(t)^{-1} (\nabla u/dt)(t)$. Hence $\dot{u}(t)$ also decomposes as $\dot{u}(t) = u \cdot (u(t)^{-1}(\nabla u/dt)(t)) + B\langle \vartheta\langle \dot{u}(t) \rangle(u(t))$.

DEFINITION 2.3. (i) The curvature tensor of ∇ is defined by

$$R\langle X, Y \rangle Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z,$$

where $X, Y, Z \in \Gamma(TM)$.

(ii) The torsion tensor of ∇ is defined by

$$T\langle X, Y \rangle = \nabla_X Y - \nabla_Y X - [X, Y],$$

where $X, Y \in \Gamma(TM)$.

(iii) The curvature form Ω of ω is the $so(n)$-valued 2-form on $O(M)$ defined by

$$\Omega\langle X, Y \rangle = d\omega\langle HX, HY \rangle \equiv (d\omega)^H \langle X, Y \rangle,$$

where $X, Y \in T_u O(M)$, and HX and HY denote the horizontal components of X and Y, respectively. (So $HX = B\langle \vartheta\langle X \rangle \rangle(u)$.)

(iii′) For all $u \in O(M)$ and $a, b \in \mathbb{R}^n$ set $\Omega_u\langle a, b \rangle \equiv \Omega\langle B\langle a \rangle(u), B\langle b \rangle(u) \rangle \in so(n)$.

(iv) The torsion form Θ of ω is the $\mathbb{R}^n$-valued 2-form on $O(M)$ defined by

$$\Theta\langle X, Y \rangle = d\vartheta^H \langle X, Y \rangle \equiv d\vartheta\langle HX, HY \rangle$$

for all $X, Y \in T_u O(M)$ and $u \in O(M)$.

(iv′) For all $u \in O(M)$ and $a, b \in \mathbb{R}^n$ set $\Theta_u\langle a, b \rangle \equiv \Theta\langle B\langle a \rangle(u), B\langle b \rangle(u) \rangle \in \mathbb{R}^n$.

The next lemma summarizes some basic properties of curvature and torsion.

LEMMA 2.1. *Using the notation of Definition* 2.3 *one has the following relations*:

(i) $\Theta = d\vartheta + \omega \wedge \vartheta$ (*first structure equation*);

(ii) $\Omega = d\omega + \omega \wedge \omega$ (*second structure equation*);

(iii) $\Omega_u\langle a, b\rangle = u^{-1}R\langle ua, ub\rangle\, u$ *for all* $u \in O(M)$, *and* a, $b \in \mathbb{R}^n$;

(iv) $\Theta_u\langle a, b\rangle = u^{-1}T\langle ua, ub\rangle$ *for all* $u \in O(M)$ *and* a, $b \in \mathbb{R}^n$.

For a proof, see [KN, Sect. III, Theorem 2.4, and Sect. III.5].

DEFINITION 2.4. (i) A path $u \in W^{\infty}(O(M))$ is said to be horizontal if $(\nabla u/ds)(s) = 0$ or equivalently $\omega\langle u'(s)\rangle = 0$ for all $s \in [0, 1]$. Let $HW^{\infty}(O(M))$ denote the set of smooth horizontal paths in $O(M)$ and $HW^{\infty}_{u_o}(O(M))$ be the curves in $HW^{\infty}(O(M))$ based at u_o.

(ii) The horizontal lift of a curve $\sigma \in W^{\infty}_o(M)$ is defined to be the unique curve $u \in HW^{\infty}_{u_o}(O(M))$ such that $\sigma(s) = \pi \circ u(s)$. Denote this path u by $H(\sigma)$, and call the resulting function H: $W^{\infty}_o(M) \to HW^{\infty}_{u_o}(O(M))$ the horizontal lift map. (Note that $H(\sigma)(s)\, u_o^{-1}$ is the parallel translation operator along $\sigma|_{[0,s]}$.)

DEFINITION 2.5. The development map is the function I: $W^{\infty}_0(\mathbb{R}^n) \to HW^{\infty}_{u_o}(O(M))$ given by $I(w) = u$, where $w \in W^{\infty}_0(\mathbb{R}^n)$ is given and $u \in HW^{\infty}_{u_o}(O(M))$ is the unique solution to the differential equation

$$u'(s) = B\langle w'(s)\rangle(u(s)) \qquad \text{with} \quad u(0) = u_o. \tag{2.1}$$

(Recall that $u_o \in O_o(M)$ is a given fixed frame.)

The stochastic counterparts of the next three theorems will be crucial for this paper.

THEOREM 2.1. *The sets* $HW^{\infty}_{u_o}(O(M))$, $W^{\infty}_o(M)$, *and* $W^{\infty}_0(\mathbb{R}^n)$ *are in one to one correspondence. In particular the development map* I: $W^{\infty}_0(\mathbb{R}^n) \to HW^{\infty}_{u_o}(O(M))$, *and the projection* π: $HW^{\infty}_{u_o}(O(M)) \to W^{\infty}_o(M)$ *are bijective, where* π *now denotes (by abuse of notation) the function* $\pi(u) = \pi \circ u$. *Furthermore, the horizontal lift map* H *is the inverse of* π, *and* $w = I^{-1}(u)$ *is given by*

$$w(s) = \int_0^s \vartheta\langle u'(s')\rangle\, ds'. \tag{2.2}$$

Proof. Let $I^{-1}(u) \equiv w$ where w is given in (2.2). Suppose that $u = I(w)$ with $w \in W^{\infty}_0(\mathbb{R}^n)$. By applying π_* to both sides of Eq. (2.1), it follows from the definition of $B\langle\cdot\rangle$ that $\pi_* u' = uw'$. Therefore $w' = u^{-1}\pi_* u' \equiv \vartheta\langle u'\rangle$, from which Eq. (2.2) follows after remembering that $w(0) = 0$. We have just shown that $I^{-1} \circ I = id$. Now if $w \equiv I^{-1}(u)$ with $u \in HW^{\infty}_{u_o}(O(M))$, then $w' = \vartheta\langle u'\rangle$ and $B\langle w'\rangle(u) = B\langle\vartheta\langle u'\rangle\rangle(u) = u'$ because u' is horizontal. Therefore, u satisfies Eq. (2.1) so that $u = I(w) = I \circ I^{-1}(u)$.

It follows trivially from the definition of H that $\pi \circ H = \text{id}$ on $W^{\infty}_o(M)$. So it only remains to show that $H \circ \pi = \text{id}$ on $HW^{\infty}_{u_o}(O(M))$. But this is also trivial, since the horizontal lift of a curve is unique once the initial frame is given. Q.E.D.

For $h \in W_0^\infty(\mathbb{R}^n)$ and $\sigma \in W_o^\infty(M)$, define $X^h(\sigma) \in \Gamma_\sigma(TM)$ by $X^h(\sigma)(s) = H(\sigma)(s) \cdot h(s)$ for all s in $[0, 1]$. One should interpret X^h as a vector field on $W_o^\infty(M)$. It is natural to flow along the vector field X^h. That is, given $\sigma_o \in W_o^\infty(M)$ the flow along X^h starting at σ_o is defined to be the solution $\sigma\colon \mathbb{R} \to W_o^\infty(M)$ to the functional differential equation

$$\dot{\sigma}(t) = X^h(\sigma) \equiv H(\sigma) \cdot h, \qquad \text{with} \quad \sigma(0) = \sigma_o. \tag{2.3}$$

Remark 2.2. At this point it would be more natural to work with H^1-paths rather than smooth paths. A H^1-path (σ) in M is an element of $W_o(M)$ such that σ has a "derivative in L^2." See Klingenberg [Kl1–Kl3] or [D2, Sect. 3] for a precise definition, and the fact that these H^1-paths form a Hilbert manifold. In this H^1-setting it would be possible to prove that X^h is in fact a smooth vector field. Hence, by standard existence theorems for ordinary differential equations on Hilbert manifolds (see, for example, Lang [L]) Eq. (2.3) will have a unique solution. Since our main interest is in the stochastic case I will not pursue this issue here. Besides, the spirit of this section is to elucidate the "algebraic" structure of these flow equations and not cloud the exposition with analysis.

Remark 2.3. Again identifying $\Gamma_\sigma^\infty(TM)$ ($\equiv$ smooth sections of TM along σ) with the tangent space to $W_o^\infty(M)$ at $\sigma \in W_o^\infty(M)$, we may define a metric $(G = G^{g, \nabla})$ on $W_o^\infty(M)$. Namely, if X and Y are two vector fields along σ, set

$$G\langle X, Y\rangle = G^{g, \nabla}\langle X, Y\rangle = \int_0^1 g\left\langle \frac{\nabla X}{ds}(s), \frac{\nabla Y}{ds}(s)\right\rangle ds. \tag{2.4}$$

Notice that $G\langle X^h(\sigma), X^h(\sigma)\rangle = \langle h, h\rangle$, where $(h, h) \equiv \int_0^1 |h'(s)|^2\, ds$. Hence the map $(h \to X^h(\sigma))\colon W_0^\infty(\mathbb{R}^n) \to T_\sigma W_o^\infty(M) \equiv \Gamma_\sigma^\infty(TM)$ is an isometry for each $\sigma \in W^\infty(M)$.

THEOREM 2.2. *Assume the same notation as above, and let $\sigma(t)$ be a solution to the flow equation* (2.3). *Let $u(t) = H(\sigma(t)) \in HW_{u_o}^\infty(O(M))$ and $w(t) \equiv I^{-1}(u(t)) \in W_0^\infty(\mathbb{R}^n)$. Then $w(t)$ and $u(t)$ satisfy*

$$\dot{u}(t)(s) = -u(t)(s) \cdot \int_0^s (\Omega_u\langle h, w'\rangle)(t, s')\, ds' + B\langle h\rangle(u)(t, s) \tag{2.5}$$

and

$$\dot{w}'(t)(s) = \int_0^s (\Omega_u\langle h, w'\rangle)(t, s')\, ds' \cdot w'(t)(s) + (\Theta_u\langle h, w'\rangle)(t, s) + h'(s), \tag{2.6}$$

where $w' = dw/ds$, $\dot{w} = dw/dt$, *etc., and* $(\Omega_u\langle h, w'\rangle)(t, s) = \Omega_{u(t,s)}\langle h(s), w'(t)(s)\rangle$ *with similar notation for the* Θ_u *and the* $B\langle h\rangle$ *terms.*

Proof. I will give two proofs of this theorem. One using the covariant derivative ∇, and the other using the connection form ω. It will be this second proof which is more easily adapted to the stochastic case. The first proof is included for the reader who is more comfortable with covariant derivatives.

First Proof. Start with the basic formula for $\dot{u}$:

$$\dot{u} = u\cdot\omega\langle\dot{u}\rangle + B\langle\vartheta\langle\dot{u}\rangle\rangle(u) = u\cdot\left(u^{-1}\frac{\nabla u}{dt}\right) + B\langle\vartheta\langle\dot{u}\rangle\rangle(u), \tag{2.7}$$

see Remark 2.1. By definition of w and Theorem 2.1, $u' = B\langle w'\rangle(u)$, and so by taking π_* of both sides of this equation one finds $\sigma' = uw'$. Since $u(t)$ is a horizontal curve in s for each fixed t, it follows by definition that $\nabla u/ds \equiv 0$. Therefore $(\nabla/ds)(\nabla/dt)\,u = [\nabla/ds, \nabla/dt]\,u = R\langle\sigma', \dot{\sigma}\rangle\,u$, and hence by Lemma 2.1, and the relations $\sigma' = uw'$ and $\dot{\sigma} = uh$, one has

$$\frac{d}{ds}\left(u^{-1}\frac{\nabla u}{dt}\right) = u^{-1}R\langle\sigma', \dot{\sigma}\rangle\,u = \Omega_u\langle w', h\rangle = -\Omega_u\langle h, w'\rangle. \tag{2.8}$$

It follows by integrating (2.8) using $\nabla u/dt = 0$ at $s = 0$ that

$$\begin{aligned}\left(u^{-1}\frac{\nabla u}{dt}\right)(t, s) &= \int_0^s \Omega_u\langle w', h\rangle(t, s')\,ds' \\ &= -\int_0^s \Omega_u\langle h, w'\rangle(t, s')\,ds'.\end{aligned} \tag{2.9}$$

Combining Eqs. (2.7) and (2.9) along with the observation that $\vartheta\langle\dot{u}\rangle \equiv u^{-1}\pi_*\dot{u} = u^{-1}\dot{\sigma} = u^{-1}uh = h$ proves Eq. (2.5).

Now to prove (2.6), first suppose that $u(t)$ is any smooth curve in $O(M)$, $\sigma \equiv \pi\circ u$, and $X \in \Gamma_\sigma(TM)$ is any vector field along σ. Then I claim $(d/dt)(u^{-1}X) = -u^{-1}\{(\nabla u/dt)\,u^{-1}X - \nabla X/dt\}$. To see this, let $\hat{u} \in O(M)$ be any horizontal lift of the curve σ, and define $g(t) \in O(n)$ by the equation $u(t) = \hat{u}(t)\,g(t)$. Then

$$\frac{d}{dt}(u^{-1}X) = \frac{d}{dt}(g^{-1}\hat{u}^{-1}X) = -g^{-1}\dot{g}g^{-1}\hat{u}^{-1}X + g^{-1}\frac{d}{dt}(\hat{u}^{-1}X).$$

Since, $\hat{u}$ is horizontal, $\nabla u/dt = \hat{u}\dot{g}$ and $(d/dt)(\hat{u}^{-1}X) = \hat{u}^{-1}(\nabla X/dt)$. These two observations and the last displayed equation yield

$$\begin{aligned}\frac{d}{dt}(u^{-1}X) &= -g^{-1}\hat{u}^{-1}\frac{\nabla u}{dt}g^{-1}\hat{u}^{-1}X + g^{-1}\hat{u}^{-1}\frac{\nabla X}{dt} \\ &= u^{-1}\left\{-\frac{\nabla u}{dt}u^{-1}X + \frac{\nabla X}{dt}\right\}.\end{aligned} \tag{2.10}$$

Applying (2.10) with $u(t)$ replaced by $u(t)(s)$ and $X(t)$ replaced by $\sigma'(t)(s)$ gives

$$\dot{w}'=\frac{d}{dt}\left[u^{-1}\sigma'\right]=-u^{-1}\cdot\frac{\nabla u}{dt}\cdot u^{-1}\sigma'+u^{-1}\frac{\nabla\sigma'}{dt}. \tag{2.11}$$

Now using the definition of torsion, Lemma 2.1, and the relation $\sigma'=uw'$, it follows that $\nabla\sigma'/dt=\nabla\dot{\sigma}/ds+T\langle\dot{\sigma},\sigma'\rangle=\nabla\dot{\sigma}/ds+u\Theta_u\langle h,w'\rangle$. Since $s\to u(t)(s)$ is a horizontal path in $O(M)$, $\nabla\dot{\sigma}/ds=(\nabla/ds)(uh)=uh'$. So the last two equations show that

$$u^{-1}\frac{\nabla\sigma'}{dt}=\Theta_u\langle h,w'\rangle+h'. \tag{2.12}$$

Equation (2.6) is now a consequence of (2.11), (2.9), and (2.12).

Second Proof. Our starting point is still Eq. (2.7). We also borrow from the above proof the equations $\vartheta\langle\dot{u}\rangle=u^{-1}\dot{\sigma}=h$ and $\vartheta\langle u'\rangle=u^{-1}\sigma'=w'$. Since u' is horizontal, $\omega\langle u'\rangle=0$. Therefore, $0=(d/dt)\,\omega\langle u'\rangle=d\omega\langle\dot{u},u'\rangle+(d/ds)\,\omega\langle\dot{u}\rangle$. By the second structure equation (Lemma 2.1) and again because u' is horizontal, $\Omega\langle\dot{u},u'\rangle=d\omega\langle\dot{u},u'\rangle$. Hence $(d/ds)\,\omega\langle\dot{u}\rangle=\Omega\langle u',\dot{u}\rangle=-\Omega_u\langle h,w'\rangle$, so that

$$\omega\langle\dot{u}\rangle(t,s)=-\int_0^s\Omega_u\langle h,w'\rangle(t,s')\,ds'.$$

This last equation is the same as (2.9), which as above yields (2.5).

Now compute $\dot{w}'=(d/dt)\,\vartheta\langle u'\rangle$:

$$\begin{aligned}\dot{w}'&=\frac{d}{dt}\vartheta\langle u'\rangle=d\vartheta\langle\dot{u},u'\rangle+\frac{d}{ds}\vartheta\langle\dot{u}\rangle=d\vartheta\langle\dot{u},u'\rangle+h'\\&=\Theta\langle\dot{u},u'\rangle-\omega\wedge\vartheta\langle\dot{u},u'\rangle+h'.\end{aligned}$$

The last equality is a consequence of the first structure equation (Lemma 2.1). Because $\omega\langle u'\rangle=0$, one has

$$\dot{w}'=\Theta\langle\dot{u},u'\rangle-\omega\langle\dot{u}\rangle\vartheta\langle u'\rangle+h'=-\omega\langle\dot{u}\rangle\,w'+\Theta_u\langle h,w'\rangle+h',$$

which combined with (2.9) again proves (2.6). Q.E.D.

Equation (2.6) may be considered as a functional differential equation for w' by defining u to be $I(w)$. (Notice that w may be recovered by integration from w' since $w(0)=0$.) Theorem 2.2 shows that given a solution (σ) to the flow equation (2.3), then $w=I^{-1}\circ H(\sigma)$ solves Eq. (2.6). It will be shown in Section 5 that (2.6) is still valid for random paths provided all w''s are replaced by δw, where δw is the Stratonovich differential of w in the s-variable. One possible method for solving (2.3) would be to solve (2.6) instead and then set $\sigma \equiv \pi \circ I(w)$. (This is what is done in Section 6.) The next theorem shows that σ defined this way solves (2.3). The stochastic analogue of this theorem appears in Section 5.

THEOREM 2.3. *Suppose that $w(t)$ is a smooth path in $W^{\infty}(\mathbb{R}^n)$ satisfying Eq.* (2.6), *where u in* (2.6) *is to be interpreted as $I(w)$. If $\sigma \equiv \pi \circ I(w)$, then σ satisfies the geometric flow equation* (2.3).

Remark 2.4. If we had been working with the Hilbert manifold of H^1-paths and had shown that the maps I and H are diffeomorphisms, then a direct proof of Theorem 2.3 would be unnecessary.

Proof. To simplify notation set $A \equiv \int_0^{\cdot} \Omega_u \langle h, w' \rangle (t, s')\, ds'$, and $u = I(w)$. We first show that u satisfies Eq. (2.5). Because of Eq. (2.7) it suffices to show that $\vartheta \langle \dot{u} \rangle = h$ and $\omega \langle \dot{u} \rangle = -A$. Set $\xi = \vartheta \langle \dot{u} \rangle$ and $E \equiv \omega \langle \dot{u} \rangle$, then I claim that the pair $(\xi - h, A + E)$ satisfies the differential equations

$$(E + A)' = \Omega_u \langle w', \xi - h \rangle \tag{2.13}$$

and

$$(\xi - h)' = (E + A)\, w' + \Theta_u \langle w', \xi - h \rangle. \tag{2.14}$$

To show this, compute E' using the second structure equation:

$$E' = \frac{d}{ds} \omega \langle \dot{u} \rangle = d\omega \langle u', \dot{u} \rangle + \frac{d}{dt} \omega \langle u' \rangle$$

$$= \Omega \langle u', \dot{u} \rangle - \omega \wedge \omega \langle u', \dot{u} \rangle + \frac{d}{dt} \omega \langle u' \rangle.$$

Because u' is horizontal, $\omega \wedge \omega \langle u', \dot{u} \rangle = 0$ and $\omega \langle u' \rangle = 0$, so that $E' = \Omega \langle u', \dot{u} \rangle = \Omega_u \langle w', \xi \rangle$. Thus $(E + A)' = \Omega_u \langle w', \xi - h \rangle$, which is Eq. (2.13).

Now compute ξ' using the first structural equation and $\vartheta \langle u' \rangle = w'$:

$$\xi' \equiv \frac{d}{ds} \vartheta \langle \dot{u} \rangle = d\vartheta \langle u', \dot{u} \rangle + \frac{d}{dt} \vartheta \langle u' \rangle$$

$$= \Theta \langle u', \dot{u} \rangle - \omega \wedge \vartheta \langle u', \dot{u} \rangle + \dot{w}'.$$

Using again $\omega \langle u' \rangle = 0$ and also that w satisfies Eq. (2.6) one finds

$$\xi' = \Theta_u \langle w', \xi \rangle + \omega \langle \dot{u} \rangle\, \vartheta \langle u' \rangle + \dot{w}'$$

$$= \Theta_u \langle w', \xi \rangle + Ew' + Aw' + \Theta_u \langle h, w' \rangle + h'.$$

This last equation easily gives Eq. (2.14).

Equations (2.13) and (2.14) are linear differential equations for the pair $(E+A, \xi-h)$ with 0 initial conditions at $s=0$. By uniqueness of solutions to linear O.D.E.'s it follows that $A+E\equiv 0$, and $\xi-h\equiv 0$, and so (2.5) holds. The theorem is now completed by applying π_* to both sides of (2.5) to get

$$\dot\sigma(t)=\pi_*\dot u(t)=u(t)h=H(\sigma(t))h,$$

which is (2.3). Q.E.D.

I will end this section with a purely differential geometric theorem on the existence of "nice" extensions of covariant derivatives for manifolds imbedded in a Euclidean space. This theorem will be used in the next section to give an extrinsic proof of the existence of stochastic horizontal lifts. It will be use again, in a more serious way, in Section 7. We will need the following lemma to prepare for Theorem 2.4.

LEMMA 2.2. *Suppose that M is an imbedded submanifold of $\mathbb{R}^N$, $i\colon M\to\mathbb{R}^N$ is the inclusion map, and g is a metric on M. Then $j\colon O(M)\to\mathbb{R}^N\times\mathrm{Hom}(\mathbb{R}^N,\mathbb{R}^N)$ defined by $j(u)\equiv(i\circ\pi(u), i'(\pi(u))\circ u)$ is an imbedding of $O(M)$ in $\mathbb{R}^N\times\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^N)$. Here $i'(m)\colon T_mM\to\mathbb{R}^N$ is given by $i'(m)\equiv pr_2\circ i_{*m}$ with pr_2 projection onto the second $\mathbb{R}^N$ under the natural identification of $T\mathbb{R}^N$ with $\mathbb{R}^N\times\mathbb{R}^N$. Explicitly, this identification is given by*

$$\left((x,v)\to v_x\equiv\frac{d}{dt}\bigg|_0(x+tv)\right)\colon\mathbb{R}^N\times\mathbb{R}^N\to T\mathbb{R}^N.$$

Proof (Sketch). Let $k\equiv N-n$, where n is the dimension of M. Given a point $m\in M$, there is an open neighborhood U of $\mathbb{R}^N$ containing m, and a C^∞-function $F\colon U\to\mathbb{R}^k$ such that $U\cap M=F^{-1}(\{0\})$ and $F'(p)\colon\mathbb{R}^N\to\mathbb{R}^k$ is surjective for all $p\in U$. Choose a smooth function $\bar g\colon U\to N\times N$-positive definite symmetric matrices such that $g\langle v_x, w_x\rangle=(\bar g(x)v, w)$ for all $v_x, w_x\in T_xM$ and $x\in U\cap M$, where $(\cdot,\cdot)$ is the usual inner product on $\mathbb{R}^N$. Define $H\colon U\times\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^N)\to\mathbb{R}^k\times\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^k)\times\mathscr{S}$, where $\mathscr{S}$ is the set of all $n\times n$ real symmetric matrices, by $H(x,A)\equiv(F(x), F'(x)A, A^{\mathrm{tr}}\bar g(x)A-I)$. Then check that $j(O(M))\cap[U\times\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^N)]=H^{-1}(\{(0,0,0)\})$, and that the differential of H is surjective for $(x,A)\in j(O(M))\cap[U\times\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^N)]$. The lemma now follows by the implicit function theorem. Q.E.D.

From now on we will identify M with $i(M)$, and $O(M)$ with $j(O(M))$.

THEOREM 2.4. *Suppose that M is an imbedded submanifold of $\mathbb{R}^N$, g is a metric on M, and ∇ is a g-compatible covariant derivative on TM. Then there is an open neighborhood $Y\subset\mathbb{R}^N$ of M, a C^∞-function $\pi\colon Y\to M$, a metric $\bar g$ on Y, and $\bar g$ compatible covariant derivative $\bar\nabla$ on TY satisfying*:

(i) $\pi|_M$ *is the identity on* M;

(ii) *if* $i\colon M \to Y$ *is the inclusion map then* $i^*\bar{g} = g$;

(iii) *suppose that* $\bar{Z}\colon \mathbb{R} \to TY$ *is a smooth path then* $\pi_*(\bar{\nabla}\bar{Z}/dt) = (\nabla/dt)(\pi_*\bar{Z})$, *in particular if* $Z\colon \mathbb{R} \to TM$ *is smooth then* $\nabla Z/dt = \bar{\nabla}Z/dt$.

Let Γ *be the* $N \times N$*-matrix valued* 1*-form on* Y *such that* $\bar{\nabla} = d + \Gamma$, *and let* $P(y) \equiv \pi'(y)\colon \mathbb{R}^N \to \mathbb{R}^N$ (*i.e.*, $P(y)a \equiv (d/dt)|_0\, \pi(y + ta) \in \mathbb{R}^N$.) *Then condition* (iii) *is equivalent to*:

(iv) $P\Gamma = dP + \pi^*\Gamma\langle\cdot\rangle\, P$, *more explicitly if* $v_y \equiv (d/dt)|_0\,(y + tv)$ *is a tangent vector in* Y, *then*

$$P(y)\,\Gamma\langle v_y\rangle = dP\langle v_y\rangle + \Gamma\langle \pi_* v_y\rangle\, P(y) = \partial_v P(y) + \Gamma\langle \pi_* v_y\rangle\, P(y). \tag{2.15}$$

(v) $P(m)$ *is orthogonal projection onto* T_mM *for all* $m \in M$.

Remark 2.5. The key consequence of condition (iii) is that if $\bar{Z}$ is a $\bar{\nabla}$-covariantly constant path in TY then $Z \equiv \pi_*\bar{Z}$ is a ∇-covariantly constant path in TM. It is this property that will be used in Section 7.

Proof. First we show that (iii) and (iv) are equivalent. Write $\bar{Z}(t) = \bar{z}(t)_{\bar{\sigma}(t)}$, where $\bar{\sigma}(t) \in Y$ and $\bar{z}(t) \in \mathbb{R}^N$, similarly $Z(t) \equiv \pi_*\bar{Z}(t) = z(t)_{\sigma(t)}$, where $z(t) = P(\bar{\sigma}(t))\,\bar{z}(t)$ and $\sigma(t) = \pi \circ \bar{\sigma}(t)$. With this notation, then $\pi_*(\bar{\nabla}\bar{Z}/dt) = (P(\bar{\sigma})[\dot{\bar{z}} + \Gamma\langle\dot{\bar{\sigma}}\rangle\,\bar{z}])_\sigma$ and $(\nabla/dt)(\pi_*\bar{Z}) = ([\dot{z} + \Gamma\langle\dot{\sigma}\rangle\, z)_\sigma$. Therefore condition (iii) is equivalent to

$$P(\bar{\sigma})(\dot{\bar{z}} + \Gamma\langle\dot{\bar{\sigma}}\rangle\,\bar{z}) = \dot{z} + \Gamma\langle\dot{\sigma}\rangle\, z \tag{2.16}$$

for all functions $\bar{z}(t)$ in $\mathbb{R}^N$. Now write $u \equiv P(\bar{\sigma})$, so that $z = u\bar{z}$ and $\dot{z} = (d/dt)(u\bar{z}) = \dot{u}\bar{z} + u\dot{\bar{z}}$. Using this expression for $\dot{z}$ and $\Gamma\langle\dot{\sigma}\rangle = (\pi^*\Gamma)\langle\dot{\bar{\sigma}}\rangle$ in (2.16), one shows easily that condition (iii) is equivalent to (iv).

To prove the existence of $\bar{\nabla}$, it is convenient to transfer the problem on Y to one on the normal bundle (E) of TM in $T\mathbb{R}^N$. Let $p\colon E \to M$ be the fiber projection, and $S \in \Gamma(E)$ be the zero section of E. By the tubular neighborhood theorem [L, Chapt. 4, Theorem 9] there is an open neighborhood Y of M in $\mathbb{R}^N$ and a diffeomorphism $\psi\colon Y \to E$ such that $p \circ \psi|_M = id_M$, and $\psi \circ i = S$. Define $\pi = p \circ \psi$, then $\pi\colon Y \to M$ satisfies condition (i).

Now suppose that $\bar{g}$ is a metric on Y, ∇ is a covariant derivative on TY, and $\bar{Z}\colon J \to TY$ is a smooth map. Let $\hat{g} \equiv (\psi^{-1})^*\bar{g}$, $\hat{\nabla} = \psi_*\bar{\nabla}\psi_*^{-1}$, and $\hat{Z} = \psi_*\bar{Z}$, then it is easy to check that condition (ii) is equivalent to

(ii′) $g = S^*\hat{g}$,

and condition (iii) is equivalent to

(iii′) $p_*((\hat{\nabla}/dt)\hat{Z}) = (\nabla/dt)(p_*\hat{Z}) = \nabla Z/dt$, where $Z \equiv p \circ \hat{Z}$.

Also the condition that $\bar{\nabla}$ is $\bar{g}$-compatible is equivalent to $\hat{\nabla}$ being $\hat{g}$-compatible. Hence, if we can find a metric $\hat{g}$, and a $\hat{g}$-compatible covariant derivative $(\hat{\nabla})$ on TE satisfying condition (ii′) and (iii′), then the corresponding $\bar{g}$ and $\bar{\nabla}$ will satisfy the conclusion of the theorem.

We now construct $\hat{\nabla}$ and $\hat{g}$. First choose a fiber metric G on E, and a G-compatible covariant derivative D on E. This covariant derivative D is equivalent to a connection on E. Explicitly, we may define the horizontal subspace $\mathscr{H}T_eE$ of T_eE by

$$\mathscr{H}T_eE \equiv \left\{ \dot{V}(0) \colon V \colon (-1, 1) \to TE \text{ smooth, } V(0) = e, \text{ and } \frac{DV}{dt} = 0 \right\}.$$

As is well known $p_* \colon \mathscr{H}T_eE \to T_{p(e)}M$ is a linear isomorphism. Let $\mathscr{H}$ also denote the horizontal lift operator: $\mathscr{H}\langle v\rangle(e) \equiv p_*|_{\mathscr{H}T_eE}^{-1}(v)$ for $v \in T_{p(e)}M$. Recall that the vertical subspace $\mathscr{V}T_eE$ of T_eE is defined to be the $\operatorname{Ker}(p_{*e})$ and is isomorphic to $E_{p(e)}$. This isomorphism is given by

$$\left(e' \to (e')_e \equiv \frac{d}{dt}\bigg|_0 (e + te') \right) \colon E_{p(e)} \to \mathscr{V}T_eE.$$

Let $\kappa \colon TE \to E$ be defined by $\kappa|_{\mathscr{H}TE} \equiv 0$, and $\kappa\langle (e')_e \rangle \equiv e'$ for all $e' \in E_{p(e)}$.

Remark 2.6. It is easy now to check that $p^*TM \cong \mathscr{H}TE$, $p^*E \cong \mathscr{V}TE$, and $TE = \mathscr{H}TE \oplus \mathscr{V}TE$ so that $TE \cong p^*TM \oplus p^*E$, where p^*TM and p^*E denote respectively the pull-backs of TM and E over M to bundles over E.

We are now in a position to define $\hat{g}$ and $\hat{\nabla}$. Set $\hat{g}\langle \xi, \eta \rangle \equiv g\langle p_*\xi, p_*\eta \rangle + G\langle \kappa\langle \xi \rangle, \kappa\langle \eta \rangle \rangle$ for all $\xi, \eta \in T_eE$ and $e \in E$. Suppose that $\bar{Z}(t)$ is a smooth curve in TE, and $e(t) \equiv p(\bar{Z}(t))$, define

$$\frac{\hat{\nabla}}{dt} \bar{Z}(t) = H \left\langle \frac{\nabla}{dt} (p_* \bar{Z}(t)) \right\rangle (e(t)) + \left(\frac{D}{dt} [\kappa \langle \bar{Z}(t) \rangle] \right)_{e(t)}.$$

It is now easy to check that $\hat{\nabla}$ is $\hat{g}$ compatible and that (iii′) holds. Condition (ii′) is also easily verified using $S_*v = \mathscr{H}\langle v \rangle (S(m))$, where $v \in T_mM$. So we are now only left to prove (v).

The fact that $P(m)$ is a projection onto T_mM for $m \in M$ follows from $\pi \circ \pi = \pi$, and $\pi = id$ on M. To show that $P(m)$ is orthogonal it suffices to show that $P(m)(T_mM^\perp) \equiv \{0\}$. Since ψ is by definition an isometry from Y to E, $T_mM^\perp = \psi_*^{-1}(\mathscr{V}T_{\psi(m)}E)$. Because $\pi \equiv p \circ \psi$ and $P(m) = \pi'(m)$ ($\pi'(m)$ is π_{*m} with base points forgotten) it follows that $P(m)(T_mM^\perp) = \{0\}$, since $\pi_*(T_mM^\perp) = p_*\psi_*\psi_*^{-1}(\mathscr{V}T_{\psi(m)}E) = p_*(\mathscr{V}T_{\psi(m)}E) = \{0\}$. Q.E.D.

Remark 2.7. The condition that $\bar{\nabla}$ is $\bar{g}$-compatible is easily seen to be equivalent to the condition

$$d\bar{g}\langle\cdot\rangle = \bar{g}\Gamma\langle\cdot\rangle + \Gamma^{tr}\langle\cdot\rangle\,\bar{g}, \tag{2.17}$$

where $\bar{g}$ is being identified with the matrix function $\tilde{g}$ satisfying $\bar{g}\langle v_x, w_x\rangle = (\tilde{g}(x)v, w)$ for all $x \in Y$ and $v, w \in \mathbb{R}^N$.

3. Stochastic Preliminaries

This section will fix the probabilistic notation and review some facts about semimartingales on manifolds. We will emphasize the non-intrinsic point of view because it facilitates the derivation of the estimates in Section 4. Throughout this section and the rest of the paper $(\Omega, \mathcal{F}, \{\mathcal{F}_s\}_{s \geq 0}, P)$ will be a filtered probability space satisfying the "usual hypothesis" (or sometimes written as the usual conditions).

Usual Hypothesis. $(\Omega, \mathcal{F}, \{\mathcal{F}_s\}_{s \geq 0}, P)$ *is said to satsfy the usual hypothesis if the* σ*-algebra* $\mathcal{F}$ *is complete with respect to the probability measure* P, *the filtration* $\{\mathcal{F}_s\}$ *is right continuous, and* $\mathcal{F}_0$ *contains all* P*-null sets.*

Given a measurable function $f: \Omega \to \mathbb{R}$, the integral $\int_\Omega f(\omega)\, dP(\omega)$ will often be denoted by $P(f)$. More generally, for any $A \in \mathcal{F}$, set $P(f; A) = \int_A f(\omega)\, dP(\omega)$. If $\mathcal{H} \subset \mathcal{F}$ is a sub-sigma algebra of $\mathcal{F}$, then $P(f \mid \mathcal{H})$ denotes the conditional expectation of f with respect to P and $\mathcal{H}$. Finally, if $(X, \mathcal{H})$ is another measurable space and $F: \Omega \to X$ is $\mathcal{F}/\mathcal{H}$-measurable, let $F_* P$ denote the probability measure on $(X, \mathcal{H})$ defined by $F_* P(A) \equiv P(F^{-1}(A))$ for all $A \in \mathcal{H}$. Recall that if $f: X \to \mathbb{R}$ is a bounded measurable function then $F_* P(f) = P(f \circ F)$.

Hopefully, the reader will not be confused by the overuse of the symbol Ω for both the curvature form and the sample space Ω. The symbol d will also have a dual meaning which is likely to be even more confusing: namely, the differential of a form or the Itô stochastic differential of a process (sorry).

Suggested references for this section are Protter [Pr] for stochastic integration theory, and Emery [Em] for stochastic calculus on manifolds. Some other references are [Bi1, Bi2, El1, IW, Me, No, RW, Sc1–Sc3] to name just a few. For the reader not familiar with stochastic calculus on manifolds perhaps Meyer's short paper [Me] is a good place to start.

We will adopt the notation in [Em], in particular $\int X\,\delta Y$ will denote the process $(s \to \int_0^s X\,\delta Y)$ where the integral is the Fisk–Stratonovich stochastic integral. In terms of Itô integrals, $\int X\,\delta Y = \int X\,dY + \frac{1}{2}[X, Y]$, where

$[X, Y]$ denotes the quadratic covariation of the semimartingales X and Y. One often writes $dX\,dY$ for the differential of $[X, Y]$. The following assumptions will be in force throughout this manuscript.

Standing Conventions. A process $\{X(s)\}$ means an adapted process. A semimartingale is by definition (in this paper) continuous. (More generally, most processes appearing in this paper will be continuous.)

DEFINITION 3.1. An M-valued semimartingale is a *continuous* M valued $(\mathscr{F}_s)$ adapted stochastic process X, such that for all f in $C^\infty(M)$, $f(X)$ is a real $(\Omega, \mathscr{F}, \{\mathscr{F}_s\}, P)$-semimartingale.

Remark 3.1. If X is an M-valued semimartingale and $\varphi: M \to Q$ is a C^∞-function then $Y \equiv \varphi(X)$ is a Q-valued semimartingale.

Suppose that $\alpha(s)$ is a T^*M semimartingale and $X(s) \equiv \bar{\pi}(\alpha(s))$ (by Remark 3.1, X is an M-valued semimartingale), where $\bar{\pi}$: $T^*M \to M$ is the canonical projection. We would like to define the Stratonovich integral $\int \alpha\langle \delta X\rangle$. To this end, use the Whitney embedding theorem (see Ref. [Au] of [Sp]) to imbed M into $\mathbb{R}^N$ (for some N) in such a way that M is a closed subset of $\mathbb{R}^N$. By the tubular neighborhood theorem (see Lang [L, Sect. IV, Theorem 9]), there exists an open subset V of $\mathbb{R}^N$ containing M and a smooth map p: $V \to M$ such that $p|_M$ is the identity on M. Set $\bar{\alpha}(s) \equiv p^*\alpha(s)$ which is a T^*V-valued semimartingale, since p^*: $T^*M \to T^*V$ is a C^∞-function. With $x = (x^1, x^2, ..., x^N)$ being the standard linear coordinates on $\mathbb{R}^N$, $\bar{\alpha}(s) = \sum_{i=1}^N \alpha_i(s)\, dx^i|_{X(s)}$, where $\alpha_i(s) \equiv \bar{\alpha}(s)\langle (\partial/\partial x^i)|_{X(s)}\rangle = \alpha(s)\langle p_*((\partial/\partial x^i)|_{X(s)})\rangle$. Because the map $(\alpha \to \alpha\langle p_*((\partial/\partial x^i)|_{\pi(\alpha)})\rangle)$: $T^*M \to \mathbb{R}$ is smooth, it follows that all of the α_i's are $\mathbb{R}$-valued semimartingales. Hence, for every T^*M valued semimartingale (α), there exists a finite collection of real semimartingales $\{\alpha_i\}$, and a finite collection of C^∞-functions $\{x^i\}$ on M such that $\alpha(s) = \sum \alpha_i(s)\, dx^i|_{X(s)}$.

DEFINITION 3.2. (i) Suppose that $\alpha(s)$ is a T^*M valued semimartingale and X is the M-valued semimartingale $X(s) \equiv \bar{\pi}(\alpha(s))$. The Stratonovich integral $\int \alpha\langle \delta X\rangle$ is the real valued semimartingale $Z \equiv \sum \int f_i(s)\, \delta(g^i(X(s)))$, where $\{f_i(s)\}$ is a finite collection of real semimartingales, and $\{g^i\}$ is a finite subset $C^\infty(M)$ such that $\alpha(s) = \sum f_i(s)\, dg^i|_{X(s)}$.

(ii) Suppose that α is a smooth 1-form on M and X is an M-valued semimartingale. The Stratonovich integral $\int \alpha\langle \delta X\rangle$ is the process $\int \tilde{\alpha}\langle \delta X\rangle$, where $\tilde{\alpha}(s) \equiv \alpha|_{X(s)}$—a T^*M-valued semimartingale such that $X(s) = \bar{\pi} \circ \tilde{\alpha}$.

Remark 3.2. The fact that $\int \alpha\langle \delta X\rangle$ is well defined is proved in [Em, Proposition 7.4, Sect. (7.7), and Exercise (7.8)]. A direct proof of this fact may be modeled on the proof of the more general Theorem 6.24 of [Em]. Because $\int \alpha\langle\delta X\rangle$ is well defined it follows that $\int \alpha\langle\delta X\rangle = \int \bar{\alpha}\langle\delta X\rangle \equiv \int \sum_i \alpha_i \delta(x^i(X))$, where we are now using the notation preceding Definition 3.2. This last expression for $\int \alpha\langle\delta X\rangle$ is the ordinary Stratonovich integral for $\int \bar{\alpha}\langle\delta X\rangle$ when X is viewed as an $\mathbb{R}^N$-valued semimartingale.

Remark 3.3. If α is a smooth 1-form on M, then again by the Whitney imbedding theorem α may be written as a finite sum $\sum f_i\, dg^i$ where the functions $\{f_i\}$ and $\{g^i\}$ are smooth functions on M. For $\alpha = \sum f_i\, dg^i$, the Stratonovich integral $\int \alpha\langle\delta X\rangle$ is given by $\sum \int f_i(X)\, \delta(g^i \circ X)$.

The following easily proved elementary properties of these stochastic integrals will be used routinely in the sequel.

PROPOSITION 3.1. *Suppose that X is an M-valued semimartingale.*

(i) *Let α be a T^*M-valued semimartingale above X ($\bar{\pi} \circ \alpha = X$), and let Z be a real-valued semimartingale, so that $\eta \equiv Z \cdot \alpha$ is also a T^*M semimartingale over X. Then $\int \eta\langle\delta X\rangle = \int Z\, \delta(\int \alpha\langle\delta X\rangle)$.*

(ii) *For $f \in C^\infty(M)$, $\int_0^s df\langle\delta X\rangle = f(X(s)) - f(X(0))$.*

(iii) *Suppose that $\varphi: M \to Q$ is a C^∞ mapping between two manifolds and that η is a 1-form on Q. Let $Y = \varphi(X)$ (a Q-valued semimartingale), then $\int (\varphi^*\eta)\langle\delta X\rangle = \int \eta\langle\delta Y\rangle$. This rule may be written informally as $\varphi_*\, \delta X = \delta(\varphi \circ X)$.*

I will now recall some facts about stochastic differential equations and stochastic parallel translation on manifolds, see [Em, Sc1–Sc3]. The emphasis will be on the special cases used later in the paper.

DEFINITION 3.3. Suppose that Q is a manifold and $X: \mathbb{R}^n \to \Gamma(TQ)$ $(a \to X\langle a\rangle(\cdot))$ is a linear map. Given an $\mathbb{R}^n$-valued semimartingale (w), a Q-valued semimartingale (q) is said to satisfy the Stratonovich stochastic differential equation

$$\delta q = X\langle\delta w\rangle(q) \tag{3.1}$$

iff for all f in $C^\infty(Q)$

$$d(f(q)) = \omega_f(q)\langle\delta w\rangle \qquad \text{with} \quad \omega_f(q)\langle a\rangle \equiv df\langle X\langle a\rangle(q)\rangle.$$

More precisely,

$$f(q(s)) - f(q(0)) = \sum_{i=1}^{n} \int_0^s (X\langle e_i\rangle(q(s'))f)\, \delta w^i(s'), \tag{3.2}$$

where $\{e_i\}_{i=1}^n$ is the standard basis for $\mathbb{R}^n$.

DEFINITION 3.4. The support of the linear map $X: \mathbb{R}^n \to \Gamma(TQ)$ is the union of the supports of the vector fields $X_i \equiv X\langle e_i \rangle$, $i \in \{1, 2, ..., n\}$. So X is said to have compact support if each X_i has compact support.

THEOREM 3.1. *Suppose that (w, Q, X) are given as above and that X has compact support. Fix a point $q_o \in Q$, then there exists a unique solution to* (3.1) *such that $q(0) = q_o$. Furthermore, if Q is an imbedded submanifold of an open subset (V) of $\mathbb{R}^N$, and $\bar{X}: \mathbb{R}^n \to \Gamma(TV)$ is a linear map extending X (that is, $X\langle a \rangle(q) = \bar{X}\langle a \rangle(q)$ for all $a \in \mathbb{R}^n$ and $q \in Q$) then q is also the unique solution to $\delta q = \bar{X}\langle \delta w \rangle(q)$ with $q(0) = q_o$.*

Remark 3.4. Theorem 3.1 is a consequence of the much more general Theorem 7.21 in [Em], which is a consequence of [Em, Theorem 6.41]. Unfortunately, the proof of Theorem 6.41 contains an error. Namely, it is assumed that the normal bundle to an imbedded submanifold is always trivial. This is not in general true. For example, it is false if the manifold is non-orientable or more generally if one of the Stiefel–Whitney classes of the tangent bundle are non-trivial (see Milnor and Stasheff [MS, Lemma 4.2]). Because of this problem and the fact that the proof of this special case is considerably simpler than the general case, a proof of Theorem 3.1 will be included. Undoubtedly, the oversight in [Em] can be fixed—probably using a modification modeled on the proof given below.

Proof. We start by first proving existence in the case that Q is an open subset of $\mathbb{R}^N$. For the moment it is not assumed that X has compact support. Write $X_i(q) = X\langle e_i \rangle(q)$ as above. By Itô's lemma, q solves $\delta q = X\langle \delta w \rangle(q)$ iff q solves the standard Stratonovich differential equation

$$dq = \sum_i X_i(q)\, \delta w^i.$$

This last equation may easily be transformed into an equivalent Itô stochastic differential equation

$$\begin{aligned} dq &= \sum X_i(q)\, dw^i + \frac{1}{2} \sum_{ij} \left(\frac{\partial}{\partial q^j} X_i \right)(q)\, d[w^i, q^j] \\ &= \sum X_i(q)\, dw^i + \frac{1}{2} \sum_{ijk} \left(\frac{\partial}{\partial q^j} X_i \right)(q)\, X_k^j(q)\, d[w^i, w^k] \qquad (3.3) \end{aligned}$$

with initial conditions $q(0) = q_o$. Therefore, by standard existence and uniqueness theorems for equations of type (3.3) (with only minor modifications due to Q being an open subset of $\mathbb{R}^N$ rather than all of $\mathbb{R}^N$), there is a unique maximal solution q to (3.3) with possible explosion at a predictable stopping time ξ. Furthermore the process q is a continuous semimartingale on the stochastic interval $[0, \xi)$, see Protter [Pr, Theorem 8, p. 199, and Theorem 38, p. 247].

Now to the general case. (Assume now that X has compact support.) By Whitney's imbedding theorem we may and do assume that Q is imbedded in $\mathbb{R}^N$. We now use the notation in Theorem 2.4 with M replaced by Q. Recall that E is the normal bundle to $Q \subset \mathbb{R}^N$, S is the zero section of E, $p: E \to Q$ is the canonical projection, Y is an open neighborhood of Q diffeomorphic to E, etc. Define a linear map $\bar{X}: \mathbb{R}^n \to \Gamma(T(E))$ by letting $\bar{X}\langle a\rangle(e) = \mathscr{H}\langle X\langle a\rangle(p(e))\rangle(e)$ be the horizontal lift of $X\langle a\rangle(p(e))$ to e for each $e \in E$. (To define this horizontal lift just choose any connection on E.) Then, for each $a \in \mathbb{R}^n$, $q \in Q$, and $e \in E$, $\bar{X}$ satisfies: (i) $p_* \bar{X}\langle a\rangle(e) = X\langle a\rangle(p(e))$ and (ii) $S_* X\langle a\rangle(q) = \bar{X}\langle a\rangle(S(q))$.

By the special case proved above there is a unique solution with possible explosion time ξ to the equation $\delta e = \bar{X}\langle \delta w\rangle(e)$, with initial condition $e(0) = e_o$, where $e_o \equiv S(q_o) \in E$. Define $q \equiv p \circ e$, and $\hat{e} \equiv S(q)$.

Claim. The semimartingale $q \equiv p \circ e$ solves $\delta q = X\langle \delta w\rangle(q)$ and $e = \hat{e} = S(q)$. Compute $\delta q = p_* \delta e = p_* \bar{X}\langle \delta w\rangle(e) = X\langle \delta w\rangle(p(e)) = X\langle \delta w\rangle(q)$ as desired where we used property (i) above. Now compute $\delta \hat{e} = S_*\, \delta q = S_* X\langle \delta w\rangle(q) = \bar{X}\langle \delta w\rangle(S(q)) = \bar{X}\langle \delta w\rangle(\hat{e})$ by property (ii) above. So by the uniqueness of solutions to stochastic differential equations it follows that $\hat{e} = e$, since $\hat{e}(0) = e(0) = S(q_o)$. This proves the claim.

Because $e = S(p(q))$, it follows that e remains in the zero section of E for all time. Let K be a compact subset of Q containing the support of X and set $\bar{K} = S(K)$. It is now clear that e must remain in the set $\bar{K}$ for all time, and therefore there can be no explosion. Thus $\xi = \infty$ a.s. This proves existence.

We now prove uniqueness and the last statement of the theorem. Suppose that q is any solution to the $\delta q = X\langle \delta w\rangle(q)$ with $q(0) = q_o$. Also suppose that Q is an imbedded submanifold of V (an open subset of $\mathbb{R}^N$) and $\bar{X}: \mathbb{R}^n \to \Gamma(TV)$ is a linear map extending X. Since, $q \in Q$ and $\bar{X}\langle a\rangle(q) = X\langle a\rangle(q)$ for all q in Q, it follows that q also satisfies $\delta q = \bar{X}\langle \delta w\rangle(q)$ with $q(0) = q_o$. But this equation can be converted as in (3.3) to a standard Itô type equation which is known to have unique solutions. So the solution q must be unique, and can be found by solving, $\delta q = \bar{X}\langle \delta w\rangle(q)$ with $q(0) = q_o$. Q.E.D.

In order to state and prove the stochastic analogues of Theorem 2.1 it is necessary to discuss stochastic horizontal lifts or equivalently parallel translation. The stochastic differential equation for parallel translation does not quite fit into the context of the above theorem. So I will now treat this special case. We will use the notation in Section 2—recall that $\pi: O(M) \to M$ was the principal $O(n)$-bundle over M of orthogonal frames, and ω was the $so(n)$-valued connection 1-form on $O(M)$ constructed from (∇).

DEFINITION 3.5. A semimartingale $u(s)$ in $O(M)$ is said to be ω-horizontal (or just horizontal) if $\int \omega\langle \delta u\rangle \equiv 0$ —abbreviated as $\omega\langle \delta u\rangle = 0$.

DEFINITION 3.6. Suppose that σ is an M-valued semimartingale, then an $O(M)$-valued semimartingale u is said to be a horizontal lift of σ if (i) $\pi \circ u = \sigma$, and (ii) u is horizontal $(\omega\langle \delta u\rangle = 0)$.

The next theorem is Proposition 8.13 of [Em] which guarantees the existence of horizontal lifts. A fairly easy proof of this theorem may be given using essentially only Theorem 2.4, Proposition 3.1, and Itô's formula. Because the notation and techniques of the proof will be needed in Section 7, I will give a proof here.

THEOREM 3.2. *Suppose that σ is an M-valued semimartingale and that u_o is a $O(M)$-valued $\mathscr{F}_0$-measurable random variable such that $\pi \circ u_o = \sigma(0)$. Then there is a unique horizontal lift (u) of σ such that $u(0) = u_o$. (In all applications u_o will be the fixed base-frame in $O(M)$.)*

Proof. First use Whitney's imbedding theorem to imbed M into a Euclidean space $(\mathbb{R}^N)$. Then choose a $(Y, \bar{g}, \pi, \bar{\nabla}, \Gamma, P)$ as in Theorem 2.4. For notational simplicity I will drop the bars from the notation. There is no danger in doing this, since Theorem 2.4 guarantees that $\bar{\nabla} = \nabla$, and $\bar{g} = g$ on the domains of ∇ and g, respectively. Recall that $\nabla = d + \Gamma$ on TY (when TY is identified with $Y \times \mathbb{R}^N$) and $P(y) = \pi'(y)$. (P throughout this proof will denote π' and not the probability measure on $(\Omega, \mathscr{F})$.) Let $pr_2\colon Y \times \mathbb{R}^N \to \mathbb{R}^N$ be projection onto the second factor, $(\cdot, \cdot)$ denote the standard inner product on both $\mathbb{R}^n$ and $\mathbb{R}^N$, and identify the metric g_y on $T_y Y$ with the positive definite matrix $g(y)$ such that $g_y\langle \alpha, \beta\rangle = (g(y)\alpha, \beta)$ for all $\alpha, \beta \in \mathbb{R}^N$. Since $M \subset \mathbb{R}^N$ is an imbedded submanifold, TM can be identified with

$$\{(m, a) \in M \times \mathbb{R}^N \mid P(m)a = 0\},$$

and the frame bundle $O(M)$ may be identified with

$$\{(m, u) \in M \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)\colon P(m)u = u,\ u^{\mathrm{tr}} g(m) u = id\},$$

see Lemma 2.4.

We now can easily compute ω in this non-intrinsic notation. For this suppose that $U(s) = (\sigma(s), u(s))$ is a smooth path in $O(M)$, then $\omega\langle U'\rangle = U^{-1}(\nabla U/ds) = u^{-1}(u' + \Gamma\langle \sigma'\rangle u)$, where σ' is the derivative as a tangent vector in M, while $u'(s) \equiv \lim_{\varepsilon \to 0} [u(s+\varepsilon) - u(s)]/\varepsilon$ is the

$\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$-valued derivative. (Notice under all of our identifications that $U' = (\sigma', u'_u) \in TM \times T(\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N))$. Therefore in general $\omega\langle (v_m, A_u)\rangle = u^{-1}(A + \Gamma\langle v_m\rangle u)$, where $(v_m, A_u) \in TO(M)$, and $u^{-1} \equiv u|^{-1}_{\mathrm{Ran}(P(m))}$. The form ω may easily be extended to a smooth form $(\bar{\omega})$ on all of $Y \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$ by setting

$$\bar{\omega}\langle (v_m, A_u)\rangle \equiv u^{-1}P(m)(A + \Gamma\langle v_m\rangle\, u) = u^{\mathrm{tr}}g(m)(A + \Gamma\langle v_m\rangle\, u).$$

Suppose that $U(s)$ is an $O(M)$-valued semimartingale. In this extrinsic notation $U(s) = (\sigma(s), u(s))$, where σ is a $\mathbb{R}^N$-valued semimartingale and u is a $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$-valued semimartingale such that a.s.:

(i) $\sigma(\cdot) \in M$;

(ii) $P(\sigma(\cdot))\, u(\cdot) = u(\cdot)$;

(iii) and $u^{\mathrm{tr}}g(\sigma)u = \mathrm{id}$.

The condition that U is horizontal ($\omega\langle \delta u\rangle = 0$) translates to

$$0 = \bar{\omega}\langle \delta(\sigma, u)\rangle = u^{-1}P(\sigma)(\delta u + \Gamma\langle \delta\sigma\rangle\, u)$$

or equivalently to

$$0 = u^{\mathrm{tr}}g(\sigma)(\delta u + \Gamma\langle \delta\sigma\rangle\, u).$$

Either one of these last two equations is equivalent to:

(iv) $P(\sigma)(\delta u + \Gamma\langle \delta\sigma\rangle\, u) = 0$.

Claim. Let $Q(m) \equiv I - P(m)$, then under conditions (i)–(iii) above $Q(\sigma)(\delta u + \Gamma\langle \delta\sigma\rangle\, u) = 0$.

One way to verify the claim is just to notice that the form $\nu\langle v_m, A_u\rangle \equiv Q(m)(A + \Gamma\langle v_m\rangle\, u)$ is identically zero on $TO(M)$, and that $Q(\sigma)(\delta u + \Gamma\langle \delta\sigma\rangle\, u) = \nu\langle \delta U\rangle$ which must be zero, since Stratonovich integrals are intrinsic objects. Alternatively, one could use Eq. (2.15) of Theorem 2.4 as follows. Because of (ii), $Q(\sigma)u = 0$ and hence

$$\delta Q(\sigma)\cdot u + Q(\sigma)\,\delta u = 0. \tag{3.4}$$

But by (2.15)

$$\delta Q(\sigma) = -\delta P(\sigma) = -dP\langle \delta\sigma\rangle = \Gamma\langle \delta\sigma\rangle\, P(\sigma) - P(\sigma)\, \Gamma\langle \delta\sigma\rangle,$$

so that

$$\delta Q(\sigma)\cdot u = \Gamma\langle \delta\sigma\rangle\, u - P(\sigma)\, \Gamma\langle \delta\sigma\rangle\, u = Q(\sigma)\, \Gamma\langle \delta\sigma\rangle\, u, \tag{3.5}$$

where (ii) was used once again. The claim clearly follows from (3.4) and (3.5).

As a consequence of the claim, condition (iv) in the presence of conditions (i)–(iii) is equivalent to

$$\text{(iv}'\text{)} \quad \delta u + \Gamma\langle \delta\sigma\rangle\, u = 0, \tag{3.6}$$

because $P(\sigma) + Q(\sigma) = \mathrm{Id}$. Therefore in order to find a horizontal lift we need only find a $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$-valued semimartingale satisfying conditions (i)–(iii), and (iv′) above with $u(0) = u_o$. At this point we have no choice but to define u as the unique solution to the linear Stratonovich differential equation (3.6) with initial condition $u(0) = u_o$.

Remark. One may expand (3.6) into the Itô form to get

$$du = -\Gamma(\sigma)\langle d\sigma\rangle\, u + \tfrac{1}{2}[\Gamma(\sigma)\langle d\sigma\rangle\, \Gamma(\sigma)\langle d\sigma\rangle - \Gamma'(\sigma)\langle d\sigma, d\sigma\rangle]\, u.$$

(If the notation is not clear see Corollary 8.3 below.) Thus if

$$Z \equiv -\int \Gamma(\sigma)\langle d\sigma\rangle + \tfrac{1}{2}\int [\Gamma(\sigma)\langle d\sigma\rangle\, \Gamma(\sigma)\langle d\sigma\rangle - \Gamma'(\sigma)\langle d\sigma, d\sigma\rangle]$$

(a $\mathrm{Hom}(\mathbb{R}^N, \mathbb{R}^N)$-valued semimartingale), then (3.6) is equivalent to the linear stochastic differential equation $du = dZu$. By Theorem 7 of [Pr, p. 197], the equation $du = dZu$ with $u(0) = u_o$ has a unique solution, and hence so does (3.6) with $u(0) = u_o$. Furthermore, this solution (u) is a semimartingale.

It now remains to show that this solution u satisfies conditions (ii) and (iii) above. This is where the choice of a nice covariant derivative in Theorem 2.4 comes to play. Recall from Remark 2.7 that the $\mathrm{Hom}(\mathbb{R}^N, \mathbb{R}^N)$-valued 1-form $\eta \equiv dg - g\Gamma - \Gamma^{\mathrm{tr}} g$ defined on TY is identically zero. (For the proof, it suffices that η vanishes on TM.) Now by assumption $u_o \in O(M)$, so $v \equiv u^{\mathrm{tr}} g(\sigma) u = \mathrm{id}$ at $s = 0$. Therefore to show codition (iii) holds ($v \equiv \mathrm{id}$), it suffices to show that $\delta v = 0$. Write g_σ for $g(\sigma)$ and compute

$$\begin{aligned}
\delta v &= \delta u^{\mathrm{tr}} g_\sigma u + u^{\mathrm{tr}}\, \delta g_\sigma u + u^{\mathrm{tr}} g_\sigma\, \delta u \\
&= -u^{\mathrm{tr}} \Gamma^{\mathrm{tr}}\langle \delta\sigma\rangle\, g_\sigma u + u^{\mathrm{tr}}\, dg\langle \delta\sigma\rangle\, u - u^{\mathrm{tr}} g_\sigma \Gamma\langle \delta\sigma\rangle\, u \\
&= u^{\mathrm{tr}}[\Gamma^{\mathrm{tr}}\langle \delta\sigma\rangle \cdot g_\sigma + dg\langle \delta\sigma\rangle - g_\sigma \Gamma\langle \delta\sigma\rangle] \cdot u \\
&= u^{\mathrm{tr}} \eta\langle \delta\sigma\rangle\, u = 0,
\end{aligned}$$

where we have used (by definition) that $dg\langle \delta\sigma\rangle = \delta(g(\sigma))$, the differential equation (3.6) for u, the differential equation for u^{tr} (the transpose of (3.6)), and the fact that $\eta \equiv 0$.

Let $w \equiv P_\sigma \cdot u$, where $P_\sigma \equiv P(\sigma)$. To show that condition (ii) holds, we must show that $w = u$. Since $w(0) = u(0) = u_o$, it suffices—by uniqueness of solutions for linear stochastic differential equations—to show that w satisfies the same differential equation as u, i.e., that $\delta w + \Gamma\langle\delta\sigma\rangle\, w = 0$. For this just compute $dw + \Gamma\langle\delta\sigma\rangle\, w$,

$$\begin{aligned} dw + \Gamma\langle\delta\sigma\rangle\, w &= \delta P_\sigma \cdot u + P_\sigma \cdot \delta u + \Gamma\langle\delta\sigma\rangle\, P_\sigma u \\ &= dP\langle\delta\sigma\rangle\, u - P_\sigma \Gamma\langle\delta\sigma\rangle\, u + \Gamma\langle\delta\sigma\rangle\, P_\sigma u \\ &= [dP - P\cdot\Gamma + \Gamma\cdot P]\langle\delta\sigma\rangle\, u = 0, \end{aligned}$$

where the last equality is a consequence of Theorem 2.4 which guarantees that the $\mathrm{Hom}(\mathbb{R}^N, \mathbb{R}^N)$-valued 1-form $[dP - P\cdot\Gamma + \Gamma\cdot P]$ vanishes on TM.
Q.E.D.

Notation 3.1. Let $\mathcal{S}M$, $H\mathcal{S}O(M)$, and $\mathcal{S}\mathbb{R}^n$ denote the space of based M-valued semimartingales starting at $o \in M$, horizontal $O(M)$-valued semimartingales starting at $u_o \in O(M)$, and $\mathbb{R}^n$-valued semimartingales starting at $0 \in \mathbb{R}^n$, respectively.

The next theorem, which is the analogue of Theorem 2.1, establishes a 1-1 correspondence between the three sets $\mathcal{S}M$, $H\mathcal{S}O(M)$, and $\mathcal{S}\mathbb{R}^n$. Recall the canonical 1-form (ϑ) on $O(M)$ is given by $\vartheta\langle\xi_u\rangle = u^{-1}\pi_*\xi_u$ for all ξ_u in $T_uO(M)$, and for $a \in \mathbb{R}^n$ the horizontal vector-field $B\langle a\rangle$ at $u \in O(M)$ is defined to be the horizontal lift of $ua \in TM$ to $T_uO(M)$.

THEOREM 3.3. *Define the maps* $H\colon \mathcal{S}M \to H\mathcal{S}O(M)$, $\pi\colon H\mathcal{S}O(M) \to \mathcal{S}M$, $I\colon \mathcal{S}\mathbb{R}^n \to H\mathcal{S}O(M)$, *and* $I^{-1}\colon H\mathcal{S}O(M) \to \mathcal{S}\mathbb{R}^n$ *as follows. Let* $H(\sigma)$ *be the horizontal lift of* $\sigma \in \mathcal{S}M$ *to* $O(M)$ *starting at* u_o, *and* $\pi(u) \equiv \pi \circ u$. *For* $b \in \mathcal{S}\mathbb{R}^n$, *let* $I(b) \equiv u$ *be the solution to the Stratonovich differential equation* $\delta u = B\langle\delta b\rangle(u)$ *starting at* $u(0) = u_o$. *Finally for* $u \in H\mathcal{S}O(M)$, *set* $b = I^{-1}(u) \equiv \int \vartheta\langle\delta u\rangle$. *Then* H *and* π *are inverses to one another as are* I *and* I^{-1}.

Proof. It is clear by the definition of H and π, and the uniqueness of horizontal lifts that $\pi \circ H = id$ on $\mathcal{S}M$ and $H \circ \pi = id$ on $H\mathcal{S}O(M)$. Now $I^{-1} \circ I(b) \equiv \int \vartheta\langle\delta u\rangle = \int \vartheta\langle B\langle\delta b\rangle(u)\rangle$, where $u \equiv I(b)$. But $\vartheta\langle B\langle a\rangle(u)\rangle = a$ for all a in $\mathbb{R}^n$ and u in $O(M)$, so that $I^{-1} \circ I(b) = \int db = b$. (If the reader is not convinced, he/she should write $\vartheta = \sum f_i\, dg^i$ and redo the argument—the proof amounts to unwinding definitions.)

Suppose that $b \equiv I^{-1}(u) \equiv \int \vartheta\langle\delta u\rangle$. We wish to show that $\delta u = B\langle\delta b\rangle(u)$. More explicitly, it must be shown for all $f \in C^\infty(O(M))$,

$$d(f(u)) = (B(u)f)\langle\delta b\rangle = (B(u)f)\langle\delta \textstyle\int \vartheta\langle\delta u\rangle\rangle = \mathscr{H}_f\langle\delta u\rangle, \tag{3.7}$$

where $(B(u)f)\langle a\rangle \equiv B\langle a\rangle(u)f$ and $\mathscr{H}_f\langle\xi_u\rangle \equiv (B\langle\vartheta\langle\xi_u\rangle\rangle(u))f$. For ξ in $TO(M)$ define $V\xi$ and $H\xi$ to be the vertical and horizontal components of ξ, respectively, then $\mathscr{H}_f\langle\xi\rangle = df\langle H\xi\rangle$. Let $\mathscr{V}_f\langle\xi\rangle \equiv df\langle V\xi\rangle = df\langle u\cdot\omega\langle\xi\rangle\rangle$. Since $\xi = H\xi + V\xi$ for all ξ in $TO(M)$, it follows that $df = \mathscr{H}_f + \mathscr{V}_f$. (Notice that $\mathscr{H}_f$ and $\mathscr{V}_f$ are both 1-forms on $O(M)$.) Therefore, for any $O(M)$-valued semimartingale we have that

$$d(f(u)) = df\langle\delta u\rangle = \mathscr{H}_f\langle\delta u\rangle + \mathscr{V}_f\langle\delta u\rangle. \tag{3.8}$$

Therefore (3.7) will be a consequence of (3.8) provided that $\mathscr{V}_f\langle\delta u\rangle = 0$ when u is a horizontal semimartingale. By choosing a basis $\{T_a\}$ for $so(n)$, we may write $\omega = \sum \omega^a T_a$ and $\mathscr{V}_f = \sum g_a\omega^a$, where $g_a(u) = df\langle u\cdot T_a\rangle$ (recall $u\cdot T \equiv (d/dt)|_0\, ue^{tT}$ for $T\in so(n)$). Therefore assuming u is horizontal,

$$\int \mathscr{V}_f\langle\delta u\rangle = \sum\int (g_a(u)\cdot\omega^a)\langle\delta u\rangle = \sum\int g_a(u)\,\delta\left(\int\omega^a\langle\delta u\rangle\right) = 0,$$

by Proposition 3.1(i), and the fact that $\omega^a\langle\delta u\rangle = 0$ for all a. Q.E.D.

Remark 3.5. Equation (3.8) is a stochastic analogue of (2.7). This may be made more explicit. Let $\hat{T}_a$ be the vertical vector fields on $O(M)$ defined by $\hat{T}_a(u) \equiv u\cdot T_a$. For $u\in\mathscr{S}O(M)$, set $w \equiv \int\vartheta\langle\delta u\rangle$ then (3.8) may be written as

$$\delta u = \sum_a \omega^a\langle\delta u\rangle\, \hat{T}_a + \sum_i \delta w_i\, B_i(u), \tag{3.9}$$

to be interpreted as $\delta(f(u)) = \sum_a (\hat{T}_a(u)f)\,\omega^a\langle\delta u\rangle + \sum_i (B_i(u)f)\,\delta w_i$ for all $f\in C^\infty(O(M))$.

To end this section, the reader is reminded of the definition of an M-valued Brownian motion starting at $o\in M$ and its relationship to the standard Brownian motion on $\mathbb{R}^n$.

DEFINITION 3.7. Let (M, g, ∇) be a Riemannian manifold equiped with a g-compatible covariant derivative ∇. The Laplacian with respect to ∇ is the second order elliptic differential operator (Δ) on $C^\infty(M)$ defined by

$$\Delta f \equiv sp(\nabla\, df) = \sum_i \nabla\, df\langle E_i, E_i\rangle = \sum_i \{E_i^2 f - df\langle\nabla_{E_i}E_i\rangle\},$$

where $\{E_i\}$ is any local orthonormal frame.

It will be useful to record the following method for computing Δf.

LEMMA 3.1. *Suppose that $f\in C^\infty(M)$ and a, $b\in\mathbb{R}^n$ then*

(i) $(B\langle a\rangle\, B\langle b\rangle f\circ\pi)(u) = (\nabla\, df)\langle ua, ub\rangle$;

(ii) $\sum_i B_i^2(f\circ\pi) = \Delta f\circ\pi$, *where* $B_i \equiv B\langle e_i\rangle$ *and the* e_i *is the* ith *standard basis element for* $\mathbb{R}^n$.

Proof. (i) $B\langle b\rangle(u) f\circ\pi = df\langle \pi_* B\langle b\rangle(u)\rangle = df\langle ua\rangle$. Thus

$$(B\langle a\rangle\, B\langle b\rangle f\circ\pi)(u) = (d/ds)|_0\, df\langle e^{sB\langle a\rangle}(u)\cdot b\rangle.$$

Let $Y(s) \equiv e^{sB\langle a\rangle}(u)\cdot b$, a tangent vector field along the curve $\sigma(s) = \pi(e^{sB\langle a\rangle}(u))$. Therefore,

$$\begin{aligned}(B\langle a\rangle\, B\langle b\rangle f\circ\pi)(u) &= \frac{d}{ds}\bigg|_0 df\langle Y(s)\rangle \\ &= \nabla\, df\langle \sigma'(0), Y(0)\rangle + df\left\langle \frac{\nabla Y}{ds}(0)\right\rangle \\ &= (\nabla\, df)(\pi_* B\langle a\rangle(u), ub\rangle = (\nabla\, df)\langle ua, ub\rangle,\end{aligned}$$

since $Y(s)$ is a parallel vector field along σ.

(ii) Let $a = b = e_i$ in (i) and sum on i. Q.E.D.

DEFINITION 3.8. An M-valued semimartingale (σ_o) is a Brownian motion iff for all $f\in C^\infty(M)$, there is a real-valued local martingale M^f such that $d(f(\sigma)) = dM^f + \frac{1}{2}\Delta f(\sigma)\, ds$.

The following theorem restates [Em, Proposition 8.26(iii)] which relates the standard Brownian motion on $\mathbb{R}^n$ to Brownian motions on M. I will only give the easy direction of the proof here. The reader is invited to give a non-intrinsic proof of the other direction using the ideas and notation in this section.

THEOREM 3.4. *Let σ be an element of $\mathscr{S}M$, then σ is a Brownian motion iff $b \equiv I^{-1}\circ H(\sigma)$ is a standard Brownian motion on $\mathbb{R}^n$.*

Proof (Easy Direction Only). Assume that b is a standard Brownian motion on $\mathbb{R}^n$, and set $u = I(b)$ and $\sigma = \pi\circ u$. By definition of u satisfying $\delta u = B\langle \delta b\rangle(u)$, for any $F\in C^\infty(O(M))$

$$\begin{aligned}\delta(F(u)) &= B\langle db\rangle(u)F + \tfrac{1}{2}\sum (B_i B_j F)(u)\, d[b^i, b^j] \\ &= B\langle db\rangle(u)F + \tfrac{1}{2}\sum (B_i^2 F)(u)\, ds.\end{aligned}$$

Now if $F = f\circ\pi$, where $f\in C^\infty(M)$ we have by Lemma 3.1(ii) that $\sum (B_i^2 F)(u) = \Delta f\circ\pi(u)$, so that the above equations becomes

$$df(\sigma) = B\langle db\rangle(u)F + \tfrac{1}{2}\Delta f(\sigma)\, ds.$$

This shows σ is an M-valued Brownian motion, since $\int B\langle db\rangle(u)F \equiv \sum_i \int (B\langle e_i\rangle(u)\, F)\, db_i$ is a martingale. Q.E.D.

4. Estimates and Differentiability

In this section and for the remainder of the paper it is assumed, as in Section 3, that $(\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s \geqslant 0}, P)$ is a filtered probability space satisfying the usual hypothesis. We further assume that this probability space supports an $\mathbb{R}^n$-valued Brownian motion $\{b(s)\}_{s \in [0,1]}$ with respect to the Filtration $\mathscr{F}_s$. For example, take $\Omega \equiv W(\mathbb{R}^n)$, $P =$ Wiener measure, $b(s)\colon \Omega \to \mathbb{R}^n$ to be given by $b(s)(\omega) = \omega(s)$, and $\mathscr{F}_s$ to be the augmentation by all P-negligible sets of the σ-algebra generated by the maps $b(s')$ for $s' \leqslant s$. Alternatively and more geometrically, we could take $\Omega = W(M)$, P to be the Wiener measure on $W_o(M) \subset W(M)$, $\mathscr{F}_s$ to be the augmentation by all P-negligible sets of the σ-algebra generated by the maps $\sigma_o(s')(\omega) = \omega(s')$ $(\omega \in W(M))$ for $s' \leqslant s$, and $b \equiv I^{-1} \circ H(\sigma_o)$. Notice that Theorem 3.4 guarantees that b is a standard $\mathbb{R}^n$-Brownian motion. (See Section 8 for a more detailed discussion of these two examples.)

It is now useful to restrict the class of semimartingales to "Brownian semimartingales." But first a word on notation and conventions. In the sequel we will be interested in processes $(X(t, s))$ indexed by $s \in [0, 1]$ and $t \in J$ or $\mathbb{R}$, where $J \equiv [-1, 1]$. These processes will usually be $C^{1,0}$ as a function of (t, s)—that is, P-a.s. the map $(t, s) \to X(t, s)$ is differentiable in the t-variable and the derivative $\dot{X}(t, s)$ is jointly continuous in (t, s). Typically for each $t \in \mathbb{R}$, the process $X(t) \equiv X(t, \cdot)$ will be a semimartingale. The following conventions on the differentials of such two-parameter processes are strictly followed in the sequel.

Standing Conventions. For each $t \in \mathbb{R}$, let $X(t)$ be a semimartingale in the suppressed s-variable, i.e., $\{X(t)(s)\}_{s \in [0,1]}$ is an $\mathscr{F}_s$-adapted semimartingale. Then $dX(t)$ $(\delta X(t))$ denotes the Itô (Stratonovich) differential of $X(t)$ with respect to the suppressed s-variable. So if $\{Y(t)\}_{t \in \mathbb{R}}$ is another one parameter family of semimartingales, then

$$\int Y(t)\, dX(t) \text{ is the process } s \to \int_0^s Y(t)(s')\, dX(t)(s')$$

and

$$\int Y(t)\, \delta X(t) \text{ is the process } s \to \int_0^s Y(t)(s')\, dX(t)(s') + \tfrac{1}{2}[Y(t), X(t)](s).$$

Definition 4.1. (a) Let V be a finite dimensional vector space and $\mathrm{Hom}(\mathbb{R}^n, V)$ be the finite dimensional vector space of linear operators from $\mathbb{R}^n$ to V. A V-valued process (w) is a Brownian semimartingale if w is a continuous $\mathscr{F}_s$-adapted process such that there exists a continuous adapted $\mathrm{Hom}(\mathbb{R}^n, V) \times V$-valued process (O, α) such that $w(s) = w(0) + \int_0^s O(s')\, db(s') + \int_0^s \alpha(s')\, ds'$. In the future we write this as $w = w(0) + \int O\, db + \int \alpha\, ds$.

Remark 4.1. Assuming that $w(0)=0$, the map $(O, \alpha) \to w = \int O\, db + \int \alpha\, ds$ is injective. Indeed, if $w = \int O\, db + \int \alpha\, ds \equiv 0$, then the finite variation part $(\int \alpha\, ds)$ of w is zero and so $\alpha = 0$. Also the quadratic variations $[\lambda \circ w, \lambda \circ w] = \int \sum_i [\lambda \circ Oe_i]^2\, ds = 0$, where $\lambda \in V^*$ (the dual space of V) and $\{e_i\}$ is the standard basis for $\mathbb{R}^n$. This shows that $\lambda \circ O \equiv 0$ for all λ and so $O \equiv 0$. (Of course this is true up to indistinguishability, a comment that will usually be omitted in the sequel.)

DEFINITION 4.1. (b) Let Q be a manifold, then a Q-valued semimartingale (X) is said to be a Brownian semimartingale iff $f \circ X$ is an $\mathbb{R}$-valued Brownian semimartingale for all $f \in C^\infty(Q)$.

The proof of the following proposition is easy and is left to the reader.

PROPOSITION 4.1. (i) *Suppose that X is a Q-valued semimartingale where Q is an imbedded submanifold of $\mathbb{R}^N$. Then X is a Brownian semimartingale in $\mathbb{R}^N$ in the sense of Definition* 4.1(a) *iff X is a Brownian semimartingale in Q in the sense of Definition* 4.1(b).

(ii) *Suppose that σ is an M-valued semimartingale, $u \equiv H(\sigma)$, and $w \equiv I^{-1}(u) = I^{-1} \circ H(\sigma)$, where I, H, and I^{-1} are as in Theorem* 3.3. *Then if any of the processes X, u, or w is a Brownian semimartingale then so are the remaining two processes.*

Before starting on the estimates, it is necessary to introduce a number of different norms. First a convention. If V is a finite dimensional vector space, then $|v|$ will denote the length of v with respect to some norm $(|\cdot|)$ on V. Since V is finite dimensional, it will in general not matter which norm is chosen, and this choice is left to the reader if the norm is not given explicitly. For the following standard vector spaces it is convenient to make the following choices of norms.

Notation 4.1. (i) If A is an $n \times m$ matrix, we put $|A| = \mathrm{tr}(A^*A)^{1/2} =$ the Hilbert–Schmidt norm.

(ii) If $a \in \mathbb{R}^n$, $|a|$ will denote the standard Euclidean length of a.

(iii) Suppose that $(B, |\cdot|)$ is a normed space and $f\colon [0, 1] \to B$, define $f_s^* = \sup\{|f(r)| : 0 \leqslant r \leqslant s\}$ and $f^* = f_1^*$. We also write $|f|_\infty$ for f_1^*.

DEFINITION 4.2. Suppose that f_s is a continuous adapted stochastic process taking values in a normed space (V). For $p \in [1, \infty)$, define $\|f\|_{S^p(s)} \equiv \|f_s^*\|_{L^p(P)}$, and $\|f\|_{S^p} \equiv \|f\|_{S^p(1)} = \|f_1^*\|_{L^p(P)}$. Also let $S^p(V)$ or just S^p stand for the space of continuous adapted processes $f_s \in V$ such that $\|f\|_{S^p} < \infty$.

For spaces of Brownian semimartingales it will be convenient to define two types of norms.

DEFINITION 4.3. Suppose that $w=\int O\,db+\int \alpha\,ds$ is a V-valued Brownian semimartingale and that $p\in[1,\infty]$. Define

$$\text{(i)}\quad \|w\|_{H^p(s)}=\left\| \left[\int_0^s |O(s')|^2\,ds'\right]^{1/2}+\int_0^s |\alpha(s')|\,ds' \right\|_{L^p(P)}, \tag{4.1}$$

and

$$\text{(ii)}\quad \|w\|_{B^p(s)}=\|O\|_{S^p(s)}+\|\alpha\|_{S^p(s)}. \tag{4.2}$$

Again set $\|w\|_{H^p}=\|w\|_{H^p(1)}$, $\|w\|_{B^p}=\|w\|_{B^p(1)}$. Also let $H^p=H^p(V)$ and $B^p=B^p(V)$ denote the set of V-valued Brownian semimartingales such that $\|w\|_{H^p}<\infty$ and $\|w\|_{B^p}<\infty$, respectively.

LEMMA 4.1 (Basic Inequalities). *Let r, r', $p\in[2,\infty]$ such that $1/p=1/r+1/r'$. Suppose that w is a V-valued Brownian semimartingale and Z is a* $\operatorname{Hom}(\mathbb{R}^n, V)$*-valued continuous adapted process (more generally just locally bounded and predictable); then:*

(i) *(Burkholder's inequality) For $p\in[2,\infty)$, there is a positive constant c_p, such that for all $s\in[0,1]$, $\|w\|_{S^p(s)}\leqslant c_p\|w\|_{H^p(s)}$;*

(ii) $\|w\|_{H^p}\leqslant\|w\|_{B^p}$;

(iii) *(Emery's inequality)* $\|\int Z\,dw\|_{H^p}\leqslant\|Z\|_{S^r}\cdot\|w\|_{H^{r'}}$;

(iv) *For all s in $[0,1]$,*

$$\left\|\int Z\,dw\right\|^p_{H^p(s)}\leqslant\|w\|^p_{B^\infty(s)}\cdot\int_0^s P(|Z(s')|^p)\,ds'$$
$$\leqslant\|w\|^p_{B^\infty(s)}\cdot\int_0^s \|Z\|^p_{S^p(s')}\,ds';$$

(v) $\|\int Z\,dw\|_{B^p}\leqslant\|Z\|_{S^r}\|w\|_{B^{r'}}$.

Now assume that Z is also a Brownian semimartingale and r, r', $p\in[2,\infty)$ such that $1/p=1/r+1/r'$; then

(iv′) $\|\int Z\,dw\|^p_{H^p(s)}\leqslant c_p\|w\|^p_{B^\infty(s)}\cdot\int_0^s\|Z\|^p_{H^p(s')}\,ds'$;

(v′) $\|\int Z\,dw\|_{B^p}\leqslant c_r\|Z\|_{H^r}\|w\|_{B^{r'}}\leqslant c_r\|Z\|_{B^r}\|w\|_{B^{r'}}$;

(vi) *there are constants $c_{r,r'}$ such that* $\|Zw\|_{B^p}\leqslant c_{r,r'}\|Z\|_{B^r}\|w\|_{B^{r'}}$;

(vii) *If Z is a process which is P-a.s. absolutely continuous with respect to ds, then* $\|Zw\|_{B^p}\leqslant\|Z'\|_{S^\infty}\|w\|_{S^p}+\|Z\|_{S^\infty}\|w\|_{B^p}$.

Proof. (i) See Stroock [St1, St2].

(ii) This one is trivial and is left to the reader.

(iii) The proof of a more general form of Emery's inequality may be found in Protter [Pr, Theorem 3, p. 191]. I will give a short proof for this special case.

By definition of $\|\cdot\|_{H^p}$, $\|\int Z\,dw\|_{H^p} \equiv \|\,|ZO|_{L^2} + |Z\alpha|_{L^1}\|_{L^p(P)}$ *where* $|f|_{L^k} \equiv [\int_0^1 |f(s)|^k\,ds]^{1/k}$. Therefore,

$$\begin{aligned}\left\| \int Z\,dw \right\|_{H^p} &\leqslant \|Z^* \cdot [\,|O|_{L^2} + |\alpha|_{L^1}]\|_{L^p(P)} \\ &\leqslant \|Z^*\|_{L^r(P)} \cdot \|\,|O|_{L^2} + |\alpha|_{L^1}\|_{L^{r'}(P)} \\ &= \|Z\|_{S^r} \cdot \|w\|_{H^{r'}},\end{aligned}$$

where Holder's inequality was used in the second inequality.

(iv)

$$\begin{aligned}\left\| \int Z\,dw \right\|_{H^p(s)} &\equiv \|\,|ZO|_{L^2([0,s)]} + |Z\alpha|_{L^1([0,s])}\|_{L^p(P)} \\ &\leqslant \|\,|ZO|_{L^2([0,s)]} + |Z\alpha|_{L^2([0,s])}\|_{L^p(P)} \\ &\leqslant \|w\|_{B^\infty(s)}\, \|\,|Z|_{L^2([0,s)]}\|_{L^p(P)}.\end{aligned}$$

Now

$$\|\,|Z|_{L^2([0,s)]}\|_{L^p(P)}^p = P\left[\int_0^s |Z(s')|^2\,ds'\right]^{p/2} \leqslant P\left[\int_0^s |Z(s')|^p\,ds'\right],$$

by Jensen's or Holder's inequality. Finally

$$P\left[\int_0^s |Z(s')|^p\,ds'\right] \leqslant \int_0^s P[Z_{s'}^*]^p\,ds' = \int_0^s \|Z\|_{S^p(s')}^p\,ds'.$$

The estimates in (iv) now easily follow from the last three displayed equations.

(v) By Holder's inequality, $\|ZO\|_{S^p} \leqslant \|Z\|_{S^r}\|O\|_{S^{r'}}$, and $\|Z\alpha\|_{S^p} \leqslant \|Z\|_{S^r}\|\alpha\|_{S^{r'}}$. Therefore

$$\left\| \int Z\,dw \right\|_{B^p} \equiv \|ZO\|_{S^p} + \|Z\alpha\|_{S^p} \leqslant \|Z\|_{S^r}[\,\|O\|_{S^{r'}} + \|\alpha\|_{S^{r'}}] = \|Z\|_{S^r}\|w\|_{B^{r'}},$$

as claimed.

The statements in (iv′) and (v′) follow immediately using (i), (ii), (iv), and (v).

To prove (vi) and (vii), write $Z = \int A\langle db\rangle + \int \gamma\,ds$, where A is a $\mathrm{Hom}(\mathbb{R}^n, \mathrm{Hom}(\mathbb{R}^n, V))$-valued process and γ is a $\mathrm{Hom}(\mathbb{R}^n, V)$-valued process. Using the definition of $\|\cdot\|_{B^p}$ and basic stochastic calculus one finds that

$$\|Zw\|_{B^p} = \|A\langle\cdot\rangle\, w + ZO\|_{S^p} + \|\gamma w + Z\alpha + A\cdot O\|_{S^p}, \tag{$*$}$$

where $A\cdot O \equiv \sum_{i=1}^{n} A\langle e_i\rangle\, Oe_i$. Using Holder's inequality on $(*)$ it is easy to deduce that

$$\begin{aligned}\|Zw\|_{B^p} &\leqslant \|Z\|_{S^r}\,\|w\|_{B^{r'}} + \|Z\|_{B^r}\,\|w\|_{S^{r'}} + \|A\|_{S^r}\,\|O\|_{S^{r'}}\\ &\leqslant c_r\,\|Z\|_{B^r}\,\|w\|_{B^{r'}} + c_{r'}\,\|Z\|_{B^r}\,\|w\|_{B^{r'}} + \|Z\|_{B^r}\,\|w\|_{B^{r'}},\end{aligned}$$

which proves (vi) with $c_{r,r'} = c_r + c_{r'} + 1$. (The actual size of $c_{r,r'}$ will depend on the choice of norms put on the spaces V, $\mathrm{Hom}(\mathbb{R}^n, V)$, and $\mathrm{Hom}(\mathbb{R}^n, \mathrm{Hom}(\mathbb{R}^n, V))$.) Assertion (vii) also easily follows from $(*)$. Indeed, we now are assuming that $A\equiv 0$ and $\gamma = Z'$ so that $(*)$ reduces to

$$\|Zw\|_{B^p} = \|ZO\|_{S^p} + \|Z'w + Z\alpha\|_{S^p},$$

which clearly implies (vii). Q.E.D.

LEMMA 4.2. *For each finite dimensional vector space V, and $p\in[1,\infty]$, the spaces $S^p(V)$ and $B^p(V)$ are Banach spaces.*

Proof. First note that as a normed space $B^p(V)\cong S^p(\mathrm{Hom}(\mathbb{R}^n, V))\oplus S^p(V)$, so it will suffice to show that $S^p(V)$ is a Banach space. Now it is clear that $S^p(V)$ is a normed space, so that only completeness remains to be verified. For this suppose that $\{f_n\}_{n=1}^{\infty}\subset S^p(V)$ with $\sum_n \|f_n\|_{S^p} \equiv \sum_n \|f_n^*\|_{L^p(P)} < \infty$. Since, $L^p(P)$ is complete, it follows that $\sum_n f_n^*$ exists in $L^p(P)$, and in particular P-a.s. $\sum_n f_n^* < \infty$. Therefore P-a.s. $f\equiv\sum_n f_n$ is a continuous function. By Holder's inequality, the monotone convergence theorem, and the fact that $f^* \leqslant \sum_n f_n^*$ P-a.s., it follows that $\|f\|_{S^p} \leqslant \sum_n \|f_n\|_{S^p} < \infty$ so that $f\in S^p(V)$. Similarly,

$$\left\| f - \sum_{n=1}^{N} f_n \right\|_{S^p} = \left\| \sum_{n=N+1}^{\infty} f_n \right\|_{S^p} \leqslant \sum_{n=N+1}^{\infty} \|f_n\|_{S^p},$$

and this last expression tends to zero as $N\to\infty$. Therefore $f = \sum_{n=1}^{\infty} f_n$ in S^p, and so $S^p(V)$ is complete. Q.E.D.

The next theorems will be used in proving existence, regularity, and differentiability of solutions to (1.5) and (1.7). But let us first record Gronwall's inequality in the form that it is used in this paper.

LEMMA 4.3 (Gronwall's Inequality). *Suppose that $\psi(s)$, $\varepsilon(s)$, and $\eta(s)$ are non-negative functions on $[0,\infty)$ such that*

$$\psi(s) \leqslant \int_0^s \eta(\tau)\,\psi(\tau)\,d\tau + \varepsilon(s), \tag{4.3}$$

then

$$\psi(s) \leqslant \int_0^s \eta(\tau)\,\varepsilon(\tau)\exp\left\{\int_0^s \eta(u)\,du\right\} d\tau + \varepsilon(s). \tag{4.4}$$

In particular if η *and* ε *are constants then* (4.4) *reduces to*

$$\psi(s) \leqslant \varepsilon e^{\eta s}. \tag{4.4$'$}$$

THEOREM 4.1. *Let* $X\colon \mathbb{R}^n \to \Gamma(T\mathbb{R}^N)$ *be a linear map with compact support as in Definition* 3.4. *For convenience write* $X(q)a$ *for* $X\langle a\rangle(q)$, *and* $X_i(q)$ *for* $X\langle e_i\rangle(q)$, *where* $\{e_i\}_{i=1}^n$ *is the standard basis for* $\mathbb{R}^n$. *Assume that* w *and* $\bar{w}$ *are two Brownian semimartingales in* $B^\infty \equiv B^\infty(\mathbb{R}^N)$ *with canonical decompositions* $w = \int O\,db + \int \alpha\,ds$ *and* $\bar{w} = \int \bar{O}\,db + \int \bar{\alpha}\,ds$. *Let* $q_o \in \mathbb{R}^N$ *be fixed, and define* q *and* $\bar{q}$ *to be the solutions to the* Itô *stochastic differential equations* $dq = X(q)\,dw$ *and* $d\bar{q} = X(\bar{q})\,d\bar{w}$, *respectively, with initial conditions* $q(0) = \bar{q}(0) = q_o$. *Then for* $p \in [2, \infty)$, *there is a constant* $K_p = K_p(\|w\|_{B^\infty}, \|\bar{w}\|_{B^\infty}, X)$ *such that* $\|q-\bar{q}\|_{H^p} \leqslant K_p\|w-\bar{w}\|_{H^p}$ *and* $\|q-\bar{q}\|_{B^p} \leqslant K_p\|w-\bar{w}\|_{B^p}$.

Proof. Since X has compact support, there is a constant C such that $|X(q)-X(\bar{q})| \leqslant C\,|q-\bar{q}|$ and $|X(q)| \leqslant C$ for all q and $\bar{q}$ in $\mathbb{R}^N$. Also because X has compact support, the processes q and $\bar{q}$ remain inside any ball containing the support of X and the initial starting point q_o. Let $Q = q - \bar{q}$, then

$$dQ = (X(q) - X(\bar{q}))\,dw + X(\bar{q})\,d(w-\bar{w}). \tag{4.5}$$

Hence,

$$\begin{aligned}
\|Q\|^p_{H^p(s)} &\leqslant \|w\|^p_{B^\infty} \cdot \int_0^s \|X(q)-X(\bar{q})\|^p_{S^p(s')}\,ds' + C\,\|w-\bar{w}\|^p_{H^p(s)}\\
&\leqslant C\,\|w\|^p_{B^\infty} \cdot \int_0^s \|Q\|^p_{S^p(s')}\,ds' + C\,\|w-\bar{w}\|^p_{H^p(s)}\\
&\leqslant C\,\|w\|^p_{B^\infty} \cdot c_p \cdot \int_0^s \|Q\|^p_{H^p(s')}\,ds' + C\,\|w-\bar{w}\|^p_{H^p(s)},
\end{aligned}$$

where Lemma 4.1 parts (iii) and (iv) were used in the first inequality, and (i) in the last. It now follows from Gronwall's Lemma (Lemma 4.3) that there is a constant K_p such that $\|q-\bar{q}\|_{H^p} \leqslant K_p\,\|w-\bar{w}\|_{H^p}$.

We may now compute $\|Q\|_{B^p}$ using (4.5) and Lemma 4.1 repeatedly:

$$\begin{aligned}
\|Q\|_{B^p} &\equiv \|(X(q)-X(\bar{q}))O + X(\bar{q})(O-\bar{O})\|_{S^p}\\
&\quad + \|(X(q)-X(\bar{q}))\alpha + X(\bar{q})(\alpha-\bar{\alpha})\|_{S^p}\\
&\leqslant C\,\|O\|_{S^\infty}\,\|Q\|_{S^p} + C\,\|O-\bar{O}\|_{S^p} + C\,\|\alpha\|_{S^\infty}\,\|Q\|_{S^p} + C\,\|\alpha-\bar{\alpha}\|_{S^p}\\
&= C\,\|w\|_{B^\infty}\,\|Q\|_{S^p} + C\,\|w-\bar{w}\|_{B^p}\\
&\leqslant C\{c_p\,\|w\|_{B^\infty}\,\|Q\|_{H^p} + \|w-\bar{w}\|_{B^p}\}\\
&\leqslant K'_p\{\|w-\bar{w}\|_{H^p} + \|w-\bar{w}\|_{B^p}\} \leqslant K_p\,\|w-\bar{w}\|_{B^p},
\end{aligned}$$

as claimed, where K_p has been increased in size appropriately. Q.E.D.

COROLLARY 4.1. *Let Q be a compact manifold which is assumed to be imbedded in $\mathbb{R}^N$ for some N. Suppose that $X: \mathbb{R}^n \to \Gamma(TQ)$ is a linear map of $\mathbb{R}^n$ to the smooth vector fields on Q. Let $w = \int O\, db + \int \alpha\, ds$ and $\bar{w} = \int \bar{O}\, db + \int \bar{\alpha}\, ds$ be two Brownian semimartingales in $B^\infty(\mathbb{R}^n)$ as above, q_o be a fixed point in Q, and q and $\bar{q}$ be the solutions to the* Fisk–Stratonovich *differential equations $dq = X(q)\,\delta w$ and $d\bar{q} = X(q)\,\delta\bar{w}$, respectively, with initial conditions $q(0) = \bar{q}(0) = q_o$. Then for $2 \leqslant p < \infty$, there is a constant $K_p = K_p(\|w\|_{B^\infty}, \|\bar{w}\|_{B^\infty}, X)$ such that $\|q - \bar{q}\|_{H^p} \leqslant K_p \|w - \bar{w}\|_{H^p}$, and $\|q - \bar{q}\|_{B^p} \leqslant K_p \cdot \|w - \bar{w}\|_{B^p}$, where the norms on Q are determined by the imbedding of Q into $\mathbb{R}^N$.*

Proof. We make use of the imbedding of Q in $\mathbb{R}^N$ to write the Fisk–Stratonovich differential equations non-intrinsically as Itô equations. First extend X to a linear map from $\bar{X}: \mathbb{R}^n \to \Gamma(T\mathbb{R}^N)$ in such a way that $\bar{X}$ has compact support. Then the equation for q may be written non-intrinsically as

$$\begin{aligned} dq &= \bar{X}(q)\, dw + \tfrac{1}{2}\bar{X}'(q)\langle dq\rangle\, dw = \bar{X}(q)\, dw + \tfrac{1}{2}\bar{X}'(q)\langle \bar{X}(q)\, dw\rangle\, dw \\ &= \bar{X}(q)\, dw + \tfrac{1}{2}\sum_{i=1}^{n} \bar{X}'(q)\langle X(q)\, Oe_i\rangle\, Oe_i\, ds \\ &= \bar{X}(q)\, dw + Y(q)[O \otimes O]\, ds, \end{aligned}$$

where $Y(q)[A \otimes B] \equiv (1/2) \sum_{i=1}^n \bar{X}'(q)\langle \bar{X}(q)\, Ae_i\rangle\, Be_i$—a smooth vector field on $\mathbb{R}^N$ with compact support. Here A and B are in $\operatorname{End}(\mathbb{R}^n)$ and $\{e_i\}$ is the standard basis for $\mathbb{R}^n$. Let $V \equiv \operatorname{End}(\mathbb{R}^n) \otimes \operatorname{End}(\mathbb{R}^n)$. The corollary follows from the above theorem with X replaced by $\hat{X}: \mathbb{R}^n \times V \to \Gamma(T\mathbb{R}^N)$ given by $\hat{X}(q)(a, v) \equiv \bar{X}(q)a + Y(q)v$ ($a \in \mathbb{R}^n$ and $v \in V$), and w and $\bar{w}$ replaced by W and $\bar{W}$ defined to be $W(s) \equiv (w(s), \int O \otimes O\, ds)$ and $\bar{W}(s) = (\bar{w}(s), \int \bar{O} \otimes \bar{O}\, ds)$, respectively. Q.E.D.

The next lemma is a version of Kolmogorov's Lemma which will be used often in the sequal, see [Pr, Theorem 53, p. 171, and Corollary, p. 173].

LEMMA 4.4 (Kolmogorov's Lemma). *Let $p > 1$ and V be a finite dimensional vector space. Suppose $f: J \equiv [-1, 1] \to S^p(V)$ is a K-Lipschitz function, i.e., $\|f(t_1) - f(t_2)\|_{S^p} \leqslant K\, |t_1 - t_2|$ for all $t_1, t_2 \in J \equiv [-1, 1]$. Then there is a version of f such that P-a.s. $(t \to f(t)): J \to W(V) \equiv C([0, 1], V)$ is continuous. In particular, there is a version of f such that $((t, s) \to f(t)(s)): J \times [0, 1] \to V$ is P-a.s. continuous.*

LEMMA 4.5. *Let $p > 1$, and suppose that $q: J \to S^p(\mathbb{R}^N)$ is an $\|\cdot\|_{S^p}$-differentiable function and the derivative $(\dot{q})$ is K-Lipschitz on J. Then there is a version of q such that P-a.s. the function $(t, s) \to q(t)(s)$ is $C^{1,0}$.*

Proof. First notice that the hypothesis implies $\|\dot{q}\|_{S^p}$ is bounded on J. Therefore, using the fundamental theorem of calculus in the Banach space $S^p(\mathbb{R}^N)$ one finds

$$\begin{aligned}\|q(t_1)-q(t_2)\|_{S^p} &\leqslant \left\| \int_{t_2}^{t_1} \dot{q}(t)\,dt \right\|_{S^p}\\ &\leqslant \left| \int_{t_2}^{t_1} \|\dot{q}(t)\|_{S^p}\,dt \right| \leqslant C\,|t_1-t_2|.\end{aligned}$$

So by Kolmogorov's Lemma, we may choose a version of q and $\dot{q}$ such that P-a.s. the function $(t\to\dot{q}(t))$ and $(t\to q(t))$ is sup-norm continuous in $W(\mathbb{R}^N)$. Assume that $t\geqslant 0$ ($t\leqslant 0$ may be handled similarly) and let $\Pi=\{0=t_0<t_1<t_2<\cdots<t_k=t\}$ denote a partition of the interval $[0,t]$, and set $I_\Pi(\dot{q})\equiv\sum_{i=0}^{k-1}\dot{q}(t_i)(t_{i+1}-t_i)$. Again by the fundamental theorem of calculus and the definition of the Riemann integral,

$$q(t)=q(0)+\lim_{\operatorname{mesh}(\Pi)\to 0} I_\Pi(\dot{q}) \qquad \text{in} \quad S^p(\mathbb{R}^N).$$

By the definition of $\|\cdot\|_{S^p}$, this implies that

$$\lim_{\operatorname{mesh}(\Pi)\to 0} |q(t)-q(0)-I_\Pi(\dot{q})|_\infty=0 \qquad \text{in} \quad L^p(P),$$

where $|\cdot|_\infty$ denotes the sup-norm on $W(\mathbb{R}^N)$. Therefore, by choosing an appropriate sequence of partitions Π_k with $\operatorname{mesh}(\Pi_k)\to 0$ as $k\to\infty$, we can assume that

$$\lim_{k\to\infty} |q(t)-q(0)-I_{\Pi_k}(\dot{q})|_\infty=0 \qquad P\text{-a.s.}$$

In particular this implies off a fixed null set independent of s,

$$\lim_{k\to\infty} |q(t)(s)-q(0)(s)-I_{\Pi_k}(\dot{q})(s)|=0.$$

Therefore, by the definition of the $\mathbb{R}$-valued Riemann integral we find, off a fixed null set, that

$$q(t)(s)=q(0)(s)+\int_0^t \dot{q}(\tau)(s)\,d\tau.$$

But, this clearly implies that q is P-a.s. $C^{1,0}$. Q.E.D.

THEOREM 4.2. *Let $w(t)\in B^\infty(\mathbb{R}^n)$ for each t in J and assume $C_\infty\equiv\sup_{t\in J}\|w(t)\|_{B^\infty}<\infty$. Assume for each $p\in[2,\infty)$:*

(i) *the map $t\to w(t)$ is continuously differentiable into B^p;*

(ii) *there is a constant* K' *such that* $\|w(t)-w(t')\|_{B^p} \leqslant K' |t-t'|$ *and* $\|\dot{w}(t)-\dot{w}(t')\|_{B^p} \leqslant K' |t-t'|$ *for all* $t, t' \in J$.

Suppose that $X\colon \mathbb{R}^n \to \Gamma(T\mathbb{R}^N)$ *is a linear map with compact support,* $q_o \in \mathbb{R}^N$ *is a fixed point, and for each* t *let* $q(t)$ *be the solution to the Itô stochastic differential equation*

$$dq(t) = X(q(t))\, dw(t) \qquad \text{with} \quad q(t)(0) = q_o. \tag{4.6}$$

Then there is a version of $q(t)$ *for each* t, *such that a.s. the process* $(t \to q(t))\colon J \to W(\mathbb{R}^N)$ *is continuous. Furthermore this version of* q *is a.s.* $C^{1,0}$, *the map* $t \to q(t)$ *is differentiable into* B^p *for each* $p \in [2, \infty)$, *and* $\dot{q}(t)$ *is a Brownian semimartingale satisfying the stochastic differential equation*

$$d\dot{q}(t) = X(q(t))\, d\dot{w}(t) + X'(q(t))\langle \dot{q}(t) \rangle\, dw(t). \tag{4.7}$$

Finally for each $p \in [2, \infty)$, *there are constants* $K_p = K_p(C_\infty, K', X)$ *such that*

$$\|q(t)-q(t')\|_{B^p} \leqslant K_p\, |t-t'|, \tag{4.8}$$

and

$$\|\dot{q}(t)-\dot{q}(t')\|_{B^p} \leqslant K_p\, |t-t'|. \tag{4.9}$$

Proof. In this proof K_p will denote a generic constant depending only on p, C_∞, K', and X. The value of K_p will vary from place to place.

According to Theorem 4.1, there is a constant K_p such that $\|q(t)-q(t')\|_{B^p} \leqslant K_p \|w(t)-w(t')\|_{B^p} \leqslant K_p |t-t'|$ which proves (4.8). By Burkholder's inequality (Lemma 4.1), $\|q(t)-q(t')\|_{S^p} \leqslant c_p \|q(t)-q(t')\|_{B^p}$, so that $\|q(t)-q(t')\|_{S^p} \leqslant K_p |t-t'|$. Therefore by Kolomogorov's Lemma (Lemma 4.4), there is a version of $q(t)$ such that a.s. $(t \to q(t))\colon J \to W(\mathbb{R}^N)$ is continuous, and hence the map $(t, s) \to q(t)(s)$ is also continuous P-a.s. We now take $q(t)$ to be such a continuous version.

Let $\dot{q}(t)$ denote the solution to the stochastic differential equation

$$d\dot{q}(t) = X(q(t))\, d\dot{w}(t) + X'(q(t))\langle \dot{q}(t) \rangle\, dw(t) \qquad \text{with} \quad \dot{q}(0) = 0, \tag{4.10}$$

where $q(t)$ is the solution to (4.6). (I do not claim yet that $\dot{q}$ is the derivative of q.)

Remark 4.2. Note that Eq. (4.10) has global solutions. To see this fix t and let $W(t) \equiv \int X'(q(t))\langle \cdot \rangle\, dw(t)$, so that $W(t)$ is a $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$-valued Brownian semimartingale. Also set $J(t) \equiv \int X(q(t))\, d\dot{w}(t)$. With this notation Eq. (4.10) may be rewritten as

$$\dot{q}(t) = J(t) + \int dW(t)\, \dot{q}(t) \tag{4.11}$$

to which one may directly apply [Pr, Theorem 7, p. 197].

We now proceed by proving the following assertions:

(i) both $\sup_{t\in J}\|\dot q(t)\|_{H^p}$ and $\sup_{t\in J}\|\dot q(t)\|_{B^p}$ are finite for all $p\in[2,\infty)$;

(ii) there exists K_p such that $\|\dot q(t_1)-\dot q(t_2)\|_{H^p}\leqslant K_p\,|t_1-t_2|$ for $t_1, t_2\in J$;

(iii) there exists K_p such that $\|\dot q(t_1)-\dot q(t_2)\|_{B^p}\leqslant K_p\,|t_1-t_2|$ for $t_1, t_2\in J$;

(iv) $\dot q$ is the derivative of q in the $\|\cdot\|_{H^p}$ norm;

(v) $\dot q$ is the derivative of q in the $\|\cdot\|_{B^p}$ norm.

Step (i). By (4.10) and Lemma 4.1(iii) and (iv′) one finds that

$$\|\dot q(t)\|^p_{H^p(s)}\leqslant C\,\|\dot w(t)\|^p_{H^p(s)}+C\,\|w(t)\|^p_{B^\infty}\cdot\int_0^s\|\dot q(t)\|^p_{H^p(s')}\,ds',$$

where C is a constant depending on p, and the sup-norm of X and its first derivatives. If we knew $\|\dot q(t)\|_{H^p(s)}$ were finite for each s, then it would follow from Gronwall's inequality that $\|\dot q(t)\|_{H^p}\leqslant K(C\,\|\dot w(t)\|_{H^p}, C\,\|w(t)\|_{B^\infty})$, where K is a function increasing in its arguments. This technicality is easily overcome by replacing $\dot q(t)$ by $\dot q(t)^\sigma$ where σ is the first exit time of the process $\dot q(t)$ from a large ball. (It then follows from (4.10) that $\|\dot q(t)^\sigma\|_{H^p(s)}$ is also bounded.) By the same argument above it follows

$$\begin{aligned}\|\dot q(t)^\sigma\|^p_{H^p(s)}&\leqslant C\,\|\dot w(t)^\sigma\|^p_{H^p(s)}+C\,\|w(t)^\sigma\|^p_{B^\infty}\cdot\int_0^s\|\dot q(t)^\sigma\|^p_{H^p(s')}\,ds'\\&\leqslant C\,\|\dot w(t)\|^p_{H^p(s)}+C\,\|w(t)\|^p_{B^\infty}\cdot\int_0^s\|\dot q(t)^\sigma\|^p_{H^p(s')}\,ds',\end{aligned}$$

for which Gronwall's inequality yields $\|\dot q(t)^\sigma\|_{H^p}\leqslant K(C\,\|\dot w(t)\|_{H^p}, C\,\|w(t)\|_{B^\infty})$, independent of the stopping time σ. Finally, letting the size of the ball tend to infinity, it follows that $\|\dot q(t)\|_{H^p}\leqslant K(\|w(t)\|_{H^p}, \|w(t)\|_{B^\infty})$ which is bounded since $\|\dot w(t)\|_{H^p}\leqslant\|\dot w(t)\|_{B^p}$ and $\|w(t)\|_{B^\infty}$ are bounded by hypothesis.

Now it is easy to estimate $\|\dot q(t)\|_{B^p}$ from (4.10) using Lemma 4.1 to find

$$\|\dot q\|_{B^p}\leqslant C\{\|\dot w\|_{B^p}+\|\dot q\|_{S^r}\,\|w\|_{B^{r'}}\}\leqslant C\{\|\dot w\|_{B^p}+c_r\,\|\dot q\|_{H^r}\,\|w\|_{B^{r'}}\},$$

where $1/r+1/r'=1/p$. This shows that $\|\dot q(t)\|_{B^p}$ is bounded, since it has already been shown that $\|\dot q\|_{H^r}$ is bounded and by hypothesis $\|\dot w\|_{B^p}$ and $\|w\|_{B^{r'}}$ remain bounded.

Step (ii). In order to simplify notation, write q_i for $q(t_i)$, $\dot q_i$ for $\dot q(t_i)$, w_i for $w(t_i)$, and $\dot w_i$ for $\dot w(t_i)$ for $i=1$ and 2. Using the differential equation (4.10) for $\dot q(t)$, we have

$$
\begin{aligned}
d(\dot{q}_2-\dot{q}_1)=&\,[X(q_1)-X(q_2)]\,d\dot{w}_1\ +X(q_2)\,d[\dot{w}_1-\dot{w}_2]\\
&+[X'(q_1)\langle\dot{q}_1\rangle-X'(q_2)\langle\dot{q}_2\rangle]\,dw_1\\
&+X'(q_2)\langle\dot{q}_2\rangle\,d[w_1-w_2].
\end{aligned}
\tag{4.12}
$$

From (4.12) and Lemma 4.1 it follows that

$$
\begin{aligned}
\|\dot{q}_2-\dot{q}_1\|_{H^p(s)}\leqslant&\ \|X(q_1)-X(q_2)\|_{S^r(s)}\,\|\dot{w}_1\|_{H^{r'}(s)}\\
&+\|X(q_2)\|_{S^r(s)}\,\|\dot{w}_1-\dot{w}_2\|_{S^{r'}(s)}\\
&+\left\|\int[X'(q_1)\langle\dot{q}_1\rangle-X'(q_2)\langle\dot{q}_2\rangle]\,dw_1\right\|_{H^p(s)}\\
&+\|X'(q_2)\langle\dot{q}_2\rangle\|_{S^r(s)}\,\|[w_1-w_2]\|_{H^{r'}(s)},
\end{aligned}
\tag{4.13}
$$

where $1/r+1/r'=1/p$. Using the given estimates on $w(t)$, $\dot{w}(t)$, the fact that X is globally Lipschitz, $|X|$ and $|X'|$ are uniformly bounded, the estimate that $\|\dot{q}\|_{H^r}\leqslant C<\infty$, Eq. (4.8), Lemma 4.1, and Theorem 4.1, it follows that the first, second, and fourth terms on the right hand side of (4.13) are bounded by a constant times $|t_1-t_2|$. Thus

$$
\|\dot{q}_2-\dot{q}_1\|_{H^p(s)}\leqslant K_p\,|t_2-t_1|+\left\|\int[X'(q_1)\langle\dot{q}_1\rangle-X'(q_2)\langle\dot{q}_2\rangle]\,dw_1\right\|_{H^p(s)},
$$

and so using Lemma 4.1(iv)

$$
\begin{aligned}
\|\dot{q}_2-\dot{q}_1\|^p_{H^p(s)}\leqslant&\ K\,|t_2-t_1|^p\\
&+K\int_0^s\|X'(q_1)\langle\dot{q}_1\rangle-X'(q_2)\langle\dot{q}_2\rangle\|^p_{S^p(s')}\,ds'.
\end{aligned}
\tag{4.14}
$$

Because X is C^∞ with compact support it follows that

$$
\begin{aligned}
|X'(q_1)\langle\dot{q}_1\rangle-X'(q_2)\langle\dot{q}_2\rangle|&\leqslant|X'(q_1)\langle\dot{q}_1-\dot{q}_2\rangle|+|[X'(q_1)-X'(q_2)]\ \langle\dot{q}_2\rangle|\\
&\leqslant C\,|\dot{q}_1-\dot{q}_2|+C\,|q_1-q_2|\,|\dot{q}_2|.
\end{aligned}
\tag{4.15}
$$

By (4.14), (4.15), and the inequality

$$
\begin{aligned}
\|\,|\dot{q}_2|\cdot|q_1-q_2|\,\|_{S^p}&\leqslant\|\dot{q}_2\|_{S^r}\,\|q_1-q_2\|_{S^{r'}}\\
&\leqslant c_r\,\|\dot{q}_2\|_{H^r}\,\|q_1-q_2\|_{S^{r'}}\leqslant K_r\,|t_1-t_2|,
\end{aligned}
$$

where $1/r+1/r'=1/p$ one finds

$$
\begin{aligned}
\|\dot{q}_2-\dot{q}_1\|^p_{H^p(s)}&\leqslant K\,|t_2-t_1|^p+C\int_0^s\|\dot{q}_1-\dot{q}_2\|^p_{S^p(s')}\,ds'\\
&\leqslant K\,|t_2-t_1|^p+c_p\cdot C\int_0^s\|\dot{q}_1-\dot{q}_2\|^p_{H^p(s')}\,ds'.
\end{aligned}
\tag{4.16}
$$

Gronwall's inequality applied to (4.16) shows $\|\dot{q}_2-\dot{q}_1\|_{H^p}^p \leqslant K_p\,|t_2-t_1|$, which completes step (ii).

Step (iii). By (4.12) and Lemma 4.1,

$$\begin{aligned}\|\dot{q}_2-\dot{q}_1\|_{B^p} \leqslant{} & \|X(q_1)-X(q_2)\|_{S^r}\,\|\dot{w}_1\|_{B^{r'}} \\ & + \|X(q_2)\|_{S^r}\,\|\dot{w}_1-\dot{w}_2\|_{B^{r'}} \\ & + \|X'(q_1)\langle \dot{q}_1\rangle - X'(q_2)\langle \dot{q}_2\rangle\|_{S^r}\,\|w_1\|_{B^{r'}} \\ & + \|X'(q_2)\langle \dot{q}_2\rangle\|_{S^r}\,\|w_1-w_2\|_{B^{r'}},\end{aligned} \tag{4.17}$$

where again $1/r+1/r'=1/p$. Using similar arguments as above it is easy to see that the first, second, and fourth terms on the right hand side of (4.17) may be estimated by a constant times $|t_2-t_1|$. The third term in (4.17) may be estimated with the help of Lemma 4.1 as

$$\begin{aligned}&\|X'(q_1)\langle \dot{q}_1\rangle - X'(q_2)\langle \dot{q}_2\rangle\|_{S^r}\,\|w_1\|_{B^{r'}} \\ &\quad \leqslant C\,\|\,|\dot{q}_2-\dot{q}_1|+|\dot{q}_2|\cdot|q_1-q_2|\,\|_{S^r} \\ &\quad \leqslant Cc_r\,\|\dot{q}_2-\dot{q}_1\|_{H^r}+\|\dot{q}_2\|_{S^k}\cdot\|q_1-q_2\|_{S^{k'}} \\ &\quad \leqslant K\,|t_2-t_1|+c_k\cdot c_{k'}\,\|\dot{q}_2\|_{H^k}\,\|q_1-q_2\|_{H^{k'}} \\ &\quad \leqslant K_p\,|t_2-t_1|,\end{aligned}$$

where $1/k+1/k'=1/r$. Hence, putting all the estimates together shows that $\|\dot{q}_2-\dot{q}_1\|_{B^p}\leqslant K_p\,|t_2-t_1|$, which is the third assertion. We may now apply Kolmogorov's Lemma to conclude that the process $\dot{q}$ has a version such that a.s. $(t\to\dot{q}(t))\colon J\to W(\mathbb{R}^N)$ is continuous. We now assume that such a continuous version of $\dot{q}$ has been chosen.

Step (iv). Fix $t\in J$, $\varDelta\neq 0$, and set $Q=[q(t+\varDelta)-q(t)]/\varDelta$. For notational simplicity set $q=q(t)$, $w=w(t)$, $\bar{q}=q(t+\varDelta)$, $\bar{w}=w(t+\varDelta)$, $\dot{w}_\varDelta\equiv[w(t+\varDelta)-w(t)]/\varDelta$. Also for any function $f\colon\mathbb{R}^n\to V$, where V is a finite dimensional vector space, set $\bar{f}'(a,b)=\int_0^1 f'(a+tb)\,dt$, for all $a,b\in\mathbb{R}^n$. This function $\bar{f}'(a,b)$ satisfies $\bar{f}'(a,0)=f'(a)$, and $f(a+b)=f(a)+\bar{f}'(a,b)b$, for all a, b in $\mathbb{R}^n$.

Claim. $\|Q\|_{H^p}$ is bounded independent of t and $\varDelta$.
To see this consider

$$\begin{aligned}dQ &= (X(\bar{q})\,d\bar{w}-X(q)\,dw)/\varDelta \\ &= \frac{1}{\varDelta}\{[X(q)+\bar{X}'(q,\varDelta\cdot Q)\langle\varDelta\cdot Q\rangle]\,d\bar{w}-X(q)\,dw\} \\ &= X(q)[d\bar{w}-dw]/\varDelta+\bar{X}'(q,\varDelta\cdot Q)\langle Q\rangle\,d\bar{w} \\ &= X(q)\,d\dot{w}_\varDelta+\bar{X}'(q,\varDelta\cdot Q)\langle Q\rangle\,d\bar{w}.\end{aligned} \tag{4.18}$$

Recall that X has compact support so that $q(t)$ remains in the ball containing the support of X and the initial point q_o. Therefore, $\Delta \cdot Q = q - \bar{q}$ is also bounded for all Δ. From these last comments, (4.18), and Lemma 4.1 one can estimate

$$\begin{aligned} \|Q\|^p_{H^p(s)} &\leqslant C\left[\|\dot{w}_\Delta\|^p_{H^p} + C_\infty \int_0^s \|Q\|^p_{S^p(s')}\, ds'\right] \\ &\leqslant K_p\left[\|\dot{w}_\Delta\|^p_{B^p} + \int_0^s \|Q\|^p_{H^p(s')}\, ds'\right]. \end{aligned} \tag{4.19}$$

From (4.19) and Gronwall's inequality it follows that $\|Q\|_{H^p} \leqslant K_p \|\dot{w}_\Delta\|_{B^p}$. So it suffices to show that $\|\dot{w}_\Delta\|_{B^p}$ is bounded independent of t and Δ. But by the fundamental theorem of calculus in the Banach space $B^p(\mathbb{R}^n)$ we get

$$\begin{aligned} \|\dot{w}_\Delta - \dot{w}(t)\|_{B^p} &= \left\| \frac{1}{\Delta}\int_t^{t+\Delta} [\dot{w}(\tau) - \dot{w}(t)]\, d\tau \right\|_{B^p} \\ &\leqslant \left| \frac{1}{\Delta}\int_t^{t+\Delta} \|\dot{w}(\tau) - \dot{w}(t)\|_{B^p}\, d\tau \right| \\ &\leqslant \left| \Delta^{-1} \cdot K' \int_t^{t+\Delta} |\tau - t|\, d\tau \right| \leqslant K' |\Delta|/2, \end{aligned} \tag{4.20}$$

where hypothesis (ii) of the theorem was used to get the last inequality. Since K' is independent of t and Δ, (4.20) implies $\|\dot{w}_\Delta\|_{B^p}$ is uniformly bounded which proves the claim.

We now may finish the proof of step (iv). Set $\varepsilon = Q - \dot{q}(t)$, then by (4.10) and (4.18)

$$\begin{aligned} d\varepsilon &= X(q)\, d[\dot{w}_\Delta - \dot{w}] + \bar{X}'(q, \Delta \cdot Q)\langle Q \rangle\, d\bar{w} - X'(q)\langle \dot{q} \rangle\, dw \\ &= X(q)\, d[\dot{w}_\Delta - \dot{w}] + \bar{X}'(q, \Delta \cdot Q)\langle \dot{q} + \varepsilon \rangle\, d\bar{w} - X'(q)\langle \dot{q} \rangle\, dw \\ &= X(q)\, d[\dot{w}_\Delta - \dot{w}] + [\bar{X}'(q, \Delta \cdot Q)\langle \dot{q} \rangle - X'(q)\langle \dot{q} \rangle]\, d\bar{w} \\ &\quad + X'(q)\langle \dot{q} \rangle\, d[w - \bar{w}] + \bar{X}'(q, \Delta \cdot Q)\langle \varepsilon \rangle\, d\bar{w}. \end{aligned} \tag{4.21}$$

The second term "divided" by $d\bar{w}$ in the last line of (4.21) may be rewritten as

$$\begin{aligned} [\bar{X}'(q, \Delta Q)\langle \dot{q} \rangle - X'(q)\langle \dot{q} \rangle] &= \int_0^1 [X'(q + r\,\Delta Q)\langle \dot{q} \rangle - X'(q)\langle \dot{q} \rangle]\, dr \\ &= \int_0^1 dr \int_0^1 du\, X''(q + ur\,\Delta Q)\langle r\,\Delta Q, \dot{q} \rangle \\ &\equiv \bar{X}''(q, \Delta Q)\langle \Delta Q, \dot{q} \rangle. \end{aligned}$$

With this (4.21) becomes

$$d\varepsilon = X(q)\, d[\dot w_\Delta - \dot w] + \bar X''(q, \Delta Q)\langle \Delta Q, \dot q\rangle\, d\bar w$$
$$+ X'(q)\langle \dot q\rangle\, d[w - \bar w] + \bar X'(q, \Delta Q)\langle \varepsilon\rangle\, d\bar w. \tag{4.22}$$

Using Lemma 4.1, the claim that $\|Q\|_{H^p}$ is bounded, parts (i)–(iii), and (4.20), it follows from (4.22) that

$$\|\varepsilon\|^p_{H^p(s)} \leqslant C\Bigg[\|\dot w_\Delta - \dot w\|^p_{B^p} + \|\,|\Delta\cdot Q|\,|\dot q|\,\|^p_{S^p}$$
$$+ (\|\dot q\|_{S^r}\,\|w - \bar w\|_{S^{r'}})^p + C_\infty \int_0^s \|\varepsilon\|^p_{S^p(s')}\, ds'\Bigg]$$
$$\leqslant K_p\,|\Delta|^p\,[1 + (\|Q\|_{S^r}\cdot\|\dot q\|_{S^{r'}})^p + \|\dot q\|_{S^r}] + K_p \int_0^s \|\varepsilon\|^p_{S^p(s')}\, ds'$$
$$\leqslant K_p\left[|\Delta|^p + \int_0^s \|\varepsilon\|^p_{H^p(s')}\, ds'\right], \tag{4.23}$$

where $1/r + 1/r' = 1/p$. By (4.23) and Gronwall's inequality, it follows that

$$\|\varepsilon\|_{H^p} \leqslant K_p\,|\Delta|, \tag{4.24}$$

i.e., $\|[q(t+\Delta) - q(t)]/\Delta - \dot q(t)\|_{H^p} \leqslant K_p\,|\Delta|$. Hence, for all $p \in [2, \infty)$, $q(t)$ is H^p-differentiable with the derivative $\dot q(t)$ solving the stochastic differential equation (4.10). Because $\|\cdot\|_{S^p} \leqslant c_p\,\|\cdot\|_{H^p}$, it follows from Lemma 4.5 (using (4.9)) that our version of q is already $C^{1,0}$.

Step (v). By (4.22), (4.20), and Lemma 4.1, it is easy to estimate

$$\|\varepsilon\|_{B^p} \leqslant CK'\Delta/2 + C\Delta\,\|\dot q\|_{S^r}\,\|\bar w\|_{B^{r'}}$$
$$+ C\,\|\dot q\|_{S^r}\,\|w - \bar w\|_{B^{r'}} + C\,\|\varepsilon\|_{S^r}\,\|\bar w\|_{B^{r'}}$$
$$\leqslant CK'\Delta/2 + Cc_r\,\|\dot q\|_{H^r}\,\{\Delta\,\|\bar w\|_{B^{r'}} + \|w - \bar w\|_{B^{r'}}\}$$
$$+ Cc_r\,\|\varepsilon\|_{H^r}\,\|\bar w\|_{B^{r'}}, \tag{4.25}$$

where $\dot q \equiv \dot q(t)$. Hence, hypothesis (ii) of the theorem, the fact that $\|\dot q\|_{H^r}$ is bounded, and (4.24)–(4.25) imply that $\|\varepsilon\|_{B^p} \leqslant K_p\cdot\Delta$—which proves that q is also B^p-differentiable with $\dot q(t)$ solving (4.10). Q.E.D.

COROLLARY 4.2. *Keep the hyposthesis of Theorem 4.2. Then all the conclusions of theorem 4.2 remain valid if the Itô differential equation (4.6) is replaced by the Stratonovich differential equation*

$$dq(t) = X(q(t))\,\delta w(t) \qquad \text{with} \quad q(0) = q_o. \tag{4.6'}$$

For the proof it will be useful to have the following:

LEMMA 4.6. *Suppose that V and W are finite dimensional vector spaces. Let $F: \mathbb{R} \to \bigcap_{p \geqslant 2} S^p(V)$, $G: \mathbb{R} \to \bigcap_{p \geqslant 2} S^p(W)$, and $R: V \to \mathbb{R}$ be given functions, and assume that R is smooth with compact support. Let $r(t) \equiv R \circ F(t) = R(F(t))$.*

(i) *If F and G are S^p-continuous for all $p \in [2, \infty)$ then so are $F \otimes G$ and r.*

(ii) *If F and G are S^p-Lipschitz for all $p \in [2, \infty)$ then so are $F \otimes G$ and r.*

(iii) *If F and G are S^p-differentiable for all $p \in [2, \infty)$ then so are $F \otimes G$ and r. The derivatives are given by $(d/dt)(F \otimes G) = \dot{F} \otimes G + F \otimes \dot{G}$, and $\dot{r} = R'(F)\langle \dot{F} \rangle$.*

(iv) *Furthermore, if $\dot{F}$ and $\dot{G}$ are S^p-continuous (Lipschitz) for all $p \in [2, \infty)$, then $(d/dt)/(F \otimes G)$ and $\dot{r}$ are also S^p-continuous (Lipschitz) for all $p \in [2, \infty)$.*

Proof. (i) and (ii). To simplify notation, let $\dot{r}_i = \dot{r}(t_i)$, $\dot{F}_i = \dot{F}(t_i)$, $G_i = G(t_i)$, and $F_i = F(t_i)$ for $i = 1$ and 2. Then by Holder's inequality,

$$\begin{aligned}\|F \otimes G(t_1) - F \otimes G(t_2)\|_{S^p} &\leqslant \|(F_1 - F_2 \otimes G_1)\|_{S^p} + \|F_2 \otimes (G_1 - G_2)\|_{S^p} \\ &\leqslant C\, \|F_1 - F_2\|_{S^r}\, \|G_1\|_{S^{r'}} + C\, \|F_2\|_{S^r}\, \|G_1 - G_2\|_{S^{r'}},\end{aligned}$$

where $1/r + 1/r' = 1/p$, and C is a constant such that $|A \otimes B| \leqslant C\, |A| \cdot |B|$. This inequality clearly proves the assertions in (i) and (ii) involving $F \otimes G$. The assertions in (i) and (ii) involving $r(t)$ are trivial, because

$$\|r_1 - r_2\|_{S^p} \leqslant K\, \|F_1 - F_2\|_{S^p},$$

where K is a Lipschitz constant for R.

(iii) Let $\varepsilon(h) \equiv \|[F \otimes G(t+h) - F \otimes G(t)]/h - \dot{F} \otimes G(t) - F \otimes \dot{G}(t)\|_{S^p}$, then as in the ordinary proof of the product rule one finds

$$\begin{aligned}\varepsilon(h) \leqslant\ & \|[F(t) \otimes \{G(t+h) - G(t)]/h - \dot{G}(t)\}\|_{S^p} \\ &+ \|\{[F(t+h) - F(t)]/h - \dot{F}(t)\} \otimes G(t)\|_{S^p} \\ &+ \left\| \frac{1}{h} [F(t+h) - F(t)] \otimes [G(t+h) - G(t)] \right\|_{S^p}.\end{aligned}$$

Let r and r' be such that $1/p = 1/r + 1/r'$, it then follows from Holder's inequality that

$$\begin{aligned}\varepsilon(h) \leqslant\ & C\, \|F(t)\|_{S^r}\, \|G(t+h) - G(t)]/h - \dot{G}(t)\|_{S^{r'}} \\ &+ C\, \|[F(t+h) - F(t)]/h - \dot{F}(t)\|_{S^r}\, \|G(t)\|_{S^{r'}} \\ &+ C \left\| \frac{1}{h} [F(t+h) - F(t)] \right\|_{S^r} \|[G(t+h) - G(t)]\|_{S^{r'}},\end{aligned}$$

which tends to zero as $h \to 0$, proving $(d/dt)(F \otimes G) = \dot{F} \otimes G + F \otimes \dot{G}$.

To see that $r(t)$ is differentiable, use Taylor's theorem to conclude for all $x, y \in V$ that

$$|R(y) - R(x) - R'(x)(y-x)| \leqslant K\,|x-y|^2,$$

where K is a bound on the second derivatives of R. Insert $y = F(t+h)$ and $x = F(t)$ and divide by h in this last inequality and then take the S^p-norm of both sides of the result to conclude

$$\frac{1}{h}\,\|r(t+h) - r(t) - R'(F(t))(F(t+h) - F(t))\|_{S^p}$$

$$\leqslant K\,\|F(t+h) - F(t)\|^2_{S^{2p}}/h.$$

Since F is differentiable the right member of this last inequality tends to zero as h tends to zero. This clearly concludes the proof of (iii) because $\|(1/h)\,R'(F(t))(F(t+h) - F(t)) - R'(F(t))\,\dot{F}(t)\|_{S^p}$ is bounded by $K\,\|(F(t+h) - F(t))/h - \dot{F}(t)\|_{S^p}$ which tends to zero as $h \to 0$, where K is now a bound on R'.

(iv) The assertions in (iv) follow from (i)–(iii). To apply (i)–(iii) for the $\dot{r}(t)$ case, left $F(t) \to R'(F(t))$ and $G(t) \to \dot{F}(t)$ and notice that the assertions involving $F \otimes G$ hold for $B\langle F, G\rangle$, where $B\colon V \times W \to Z$ is any bounded vector-valued bi-linear form. Q.E.D.

Proof of Corollary 4.2. As in the proof of Corollary 4.1, Eq. (4.6′) may be written as the Itô equation

$$dq = X(q)\,dw + Y(q)[O \otimes O]\,ds, \tag{4.26}$$

where $Y(q)[A \otimes B] \equiv (1/2)\sum_{i=1}^n X'(q)\langle X(q)\,Ae_i\rangle\,Be_i$ is a smooth vector field on $\mathbb{R}^N$ with compact support. Again A and B are in $\mathrm{End}(\mathbb{R}^n)$ and $\{e_i\}$ is the standard basis for $\mathbb{R}^n$. As before let $V \equiv \mathrm{End}(\mathbb{R}^n) \otimes \mathrm{End}(\mathbb{R}^n)$, $\hat{X}\colon \mathbb{R}^n \times V \to \Gamma(T\mathbb{R}^N)$ be given by $\hat{X}(q)\langle (a, v)\rangle \equiv X(q)a + Y(q)v$ ($a \in \mathbb{R}^n$ and $v \in V$), and $W(t)(s) \equiv (w(t)(s), \int_0^s O(t)(s') \otimes O(t)(s')\,ds')$. Then Eq. (4.26) can be written in Itô form as

$$dq(t) = \hat{X}(q(t))\,dW(t) \qquad \text{with} \quad q(t)(0) = q_o, \tag{4.27}$$

where $W(t)$ is an $\mathbb{R}^n \times V$-valued Brownian semimartingale in $B^\infty(\mathbb{R}^n \times V)$. To finish the proof one needs only to replace $\mathbb{R}^n$ by $\mathbb{R}^n \times V$ in Theorem 4.2, and to verify that $W(t)$ still satisfies hypotheses (i)–(iii) of Theorem 4.2.

Now if $W = (w, \int v\,ds)$ where $w = \int O\,db + \int \alpha\,ds$ is an $\mathbb{R}^n$-valued Brownian semimartingale and v is a continuous adapted V-valued process, then $dW = (O, 0)\,db + \int (\alpha, v)\,ds$. From this it follows that

$$\begin{aligned}\|W\|_{B^p} = \|O\|_{S^p} + \|(\alpha, v)\|_{S^p} &\leqslant \|O\|_{S^p} + \|\alpha\|_{S^p} + \|v\|_{S^p} \\ &= \|w\|_{B^p} + \|v\|_{S^p}.\end{aligned} \tag{4.28}$$

So by (4.28), in order to verify the corresponding hypothesis (i)–(iii) of Theorem 4.2, it suffices to show with $v(t) \equiv O(t) \otimes O(t)$ that:

(i) $t \to v(t)$ is differentiable in $S^p(V)$ with $\dot{v}(t) = \dot{O}(t) \otimes O(t) + O(t) \otimes \dot{O}(t)$;

(ii) $v(t)$ and $\dot{v}(t)$ are S^p-Lipschitz on J;

(iii) $\|v(t)\|_{S^\infty}$ is bounded for $t \in J$.

Now (iii) is obvious, since $\|v(t)\|_{S^\infty} \leqslant C \cdot \|O(t)\|^2_{B^\infty} \leqslant C \cdot \|w(t)\|^2_{B^\infty}$, and (i) and (ii) follow from Lemma 4.6. Q.E.D.

LEMMA 4.7. *Suppose that $F(t)$ and $G(t)$ are Brownian semimartingales for each t, and for each $p \in [2, \infty)$ the functions F and G are B^p-differentiable. Then the path of Brownian semimartingales $Z(t) \equiv \int F(t)\, \delta G(t)$ is also B^p-differentiable for all $p \in [2, \infty)$ and the derivative process $\dot{Z}(t)$ is given by $\dot{Z}(t) = \int \dot{F}(t)\, \delta G(t) + \int F(t)\, \delta \dot{G}(t)$. Furthermore if $\dot{F}$ and $\dot{G}$ are B^p-continuous (Lipschitz) for $p \in [2, \infty)$, then so is $\dot{Z}$. (This lemma still holds true if $F(t)$ and $G(t)$ are vector-valued processes, in which case the multiplication should be replaced by the tensor product or some bilinear form.)*

Proof. Write $dF = A\, db + \alpha\, ds$, and $dG = C\, db + \gamma\, ds$, so that

$$dZ = F\, dG + \tfrac{1}{2} d[F, G] = FC\, db + \{F\gamma + \tfrac{1}{2} A \cdot C\}\, ds,$$

where $A \cdot C \equiv \sum_i (Ae_i)(Ce_i)$ and $\{e_i\}_{i=1}^n$ is the standard basis for $\mathbb{R}^n$. By Lemma 4.1 the B^p norm is stronger than the S^p norm, so the process F is S^p-differentiable for all p. Hence by Lemma 4.6, FC and $\{F\gamma + (1/2)\, A \cdot C\}$ are S^p differentiable for all $p \in [2, \infty)$ with the derivatives given by the product rule. Therefore by the definition of the B^p-norm, Z is B^p-differentiable for all p and the differential $d\dot{Z}$ of $\dot{Z}$ is given by

$$d\dot{Z} = [\dot{F}C + F\dot{C}]\, db + \{\dot{F}\gamma + F\dot{\gamma} + \tfrac{1}{2}[\dot{A} \cdot C + A \cdot \dot{C}]\}\, ds.$$

This last expression is easily seen to be the same as $\dot{F}\, \delta G + F\, \delta \dot{G}$ as claimed in the lemma. The continuity (Lipschitz) assertion for $\dot{Z}$ also follows from Lemma 4.6 and the explicit formula for $d\dot{Z}$ given above. Q.E.D.

COROLLARY 4.3. *Let $q(t)$ be an $\mathbb{R}^N$-valued Brownian semimartingale such that $t \to q(t)$ is B^p-differentiable for all $p \in [2, \infty)$, and let ω be a smooth 1-form on $\mathbb{R}^N$ with compact support. Then the path of Brownian semimartingales $Z(t) \equiv \int \omega\langle \delta q(t) \rangle$ is B^p-differentiable for all $p \in [2, \infty)$ with derivative*

$$\begin{aligned} \dot{Z}(t) &= \int d\omega\langle \dot{q}(t), \delta q(t) \rangle + \int d(\omega\langle \dot{q}(t) \rangle) \\ &= \int d\omega\langle \dot{q}(t), \delta q(t) \rangle + \omega\langle \dot{q}(t) \rangle|_0^{\cdot}. \end{aligned}$$

(Informally this may be written $(d/dt)(\omega\langle\delta q\rangle)=(d\omega)\langle\dot q,\delta q\rangle+\delta(\omega\langle\dot q\rangle)$.) *Furthermore, if* $\dot q$ *is* B^p*-continuous (Lipschitz) for all* $p\in[2,\infty)$ *then so is* $\dot Z$.

The proof of this corollary will be given after the following lemma.

LEMMA 4.8. *Let* $q(t)$ *be an* $\mathbb{R}^N$*-valued Brownian semimartingale such that* $q(t)$ *is* B^p*-differentiable for all* $p\in[2,\infty)$ *and let* $f\colon\mathbb{R}^N\to\mathbb{R}$ *be a smooth function with compact support. Then* $F(t)\equiv f(q(t))$ *is a Brownian semimartingale which is* B^p*-differentiable for all* $p\in[2,\infty)$ *and* $\dot F(t)=df\langle\dot q(t)\rangle\equiv f'(q(t))\,\dot q(t)$. *Furthermore, if* $\dot q$ *is* B^p*-continuous (Lipschitz) for all* $p\in[2,\infty)$ *then so is* $\dot F$.

Proof. Write $dq(t)=A(t)\,db+\alpha(t)\,ds$, so that A and α are S^p-differentiable. Then by Itô's Lemma,

$$dF(t)=f'(q(t))\,A(t)\,db+[f'(q(t))\,\alpha(t)+Y(q(t))\langle A(t)\otimes A(t)\rangle]\,ds, \tag{4.29}$$

where $Y(q)\langle A\otimes B\rangle=(1/2)\sum_i f''(q(t))\langle Ae_i, Be_j\rangle$ and $\{e_i\}$ is the standard basis for $\mathbb{R}^n$. From Lemma 4.1 we know that $\|\cdot\|_{S^p}\leqslant c_p\,\|\cdot\|_{B^p}$, and hence $q(t)$ is S^p-differentiable for all $p\in[2,\infty)$. Repeated application of Lemma 4.6 shows that $f'(q(t))\,A(t)$, and $[f'(q(t))\,\alpha(t)+Y(q(t))\langle A(t)\otimes A(t)\rangle]$ are S^p-differentiable for $p\in[2,\infty)$. So by the definition of the B^p-norm, it follows that F is B^p-differentiable for all $p\in[2,\infty)$. The fact that the B^p-derivative of F is given by $f'(q(t))\langle\dot q(t)\rangle$ follows from the fact that this is the correct formula for the weaker S^p-derivative, see Lemma 4.6. The continuity and Lipschitz assertion of the lemma are proved in a similar way. Q.E.D.

Proof of Corollary 4.3. Without loss of generality we may assume that $\omega=f\,dg$, where f is a C^∞-function on $\mathbb{R}^N$ with compact support, and g is linear. Let $\iota_v(d\omega)\equiv(d\omega)\langle v,\cdot\rangle$, that is, ι_v is interior multiplication by $v\in T\mathbb{R}^N$. Now $d\omega=df\wedge dg$, and hence $d\omega\langle\dot q,\delta q\rangle\equiv(\iota_{\dot q}\,d\omega)\langle\delta q\rangle=df\langle\dot q\rangle\,\delta(g\circ q)-dg\langle\dot q\rangle\,\delta(f\circ q)$. Set $F(t)\equiv f(q(t))$ and $G(t)\equiv g(q(t))$, then by assumption (using g is linear) G is B^p-differentiable for all $p\geqslant 2$, and by Lemma 4.8, so is $F(t)$ with $\dot F(t)=df\langle\dot q(t)\rangle$. Thus

$$d\omega\langle\dot q,\delta q\rangle\equiv(\iota_{\dot q}\,d\omega)\langle\delta q\rangle=\dot F\,\delta G-\dot G\,\delta F. \tag{4.30}$$

Now by definition of $Z(t)=\int\omega\langle\delta q(t)\rangle$, $Z(t)=\int F(t)\,\delta G(t)$. Therefore, by Lemma 4.7, Z is continuously differentiable and $\dot Z(t)=\int\dot F(t)\,\delta G(t)+\int F(t)\,\delta\dot G(t)$. So from this expression for $\dot Z$ and (4.30) one has

$$\begin{aligned}\dot Z(t)-\int d\omega\langle\dot q(t),\delta q(t)\rangle&=\int[F(t)\,\delta\dot G(t)+\dot G(t)\,\delta F(t)]=\int\delta(F(t)\,\dot G(t))\\&=\int\delta[f(q(t))\,d\lambda\langle\dot q(t)\rangle]=\omega\langle\dot q(t)\rangle|_0^{\cdot}\end{aligned}$$

as claimed. The continuity and Lipschitz assertions directly follow from Lemma 4.7. Q.E.D.

5. Geometric and Non-geometric Flow Equations

We now assume that we have the following data:

(i) $(\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\geqslant 0}, b, P)$ is a filtered probability space satisfying the usual hypothesis and b is an $\mathbb{R}^n$-valued Brownian motion with respect to the filtration $\{\mathscr{F}_s\}$ as in Section 3;

(ii) $(M^n, \nabla, g, o, u_o, h)$ is a smooth compact n-dimensional Riemannian manifold with metric g, a g-compatible covariant derivative ∇, a fixed base point $o \in M$, a fixed orthogonal frame $u_o \in O_o(M)$, and a C^1-function h: $[0, 1] \to \mathbb{R}^n$ such that $h(0) = 0$.

Definition 5.1. The *geometric flow equation* (associated to h) is the differential equation

$$\dot{\sigma}(t) = H(\sigma(t)) \cdot h, \tag{5.1}$$

where σ: $\mathbb{R} \to \mathscr{S}M$ is a path of semimartingales, and $H(\cdot)$ is the horizontal lift operator in Theorem 3.3. We assume here that P-a.s. the function $((t, s) \to \sigma(t)(s))$: $\mathbb{R} \times [0, 1] \to M$ is $C^{1,0}$.

Remark 5.1. Since h is C^1 and $H(\sigma(t))$ is a semimartingale, $H(\sigma(t)) \cdot h$ is also a semimartingale. Therefore, if σ solves (5.1) then necessarily $\dot{\sigma}(t)$ is a TM-valued semimartingale. It seems that the most general possible h one might allow (for general manifolds) is a semimartingale. In this paper h is assumed to be a C^1 deterministic function.

Example 5.1. (a) Take $M = \mathbb{R}^n$, $o = 0$, and $\nabla =$ to the usual covariant derivative on $T\mathbb{R}^n$. Then upon identifying $T\mathbb{R}^n$ with $\mathbb{R}^n \times \mathbb{R}^n$, $H(\cdot) = \mathrm{id}$. Therefore the solution to (5.1) is

$$\sigma(t) = \sigma(0) + t \cdot h. \tag{5.2}$$

That is, (5.1) just generates translations by h.

(b) Take $M = G$ to be a Lie group, $o = e = \mathrm{id}$ in G, u_o: $\mathbb{R}^n \to \mathfrak{g} = \mathrm{Lie}(G)$ (the Lie algebra of G) to be a fixed frame, and let ∇ be the covariant derivative for which the left invariant vector fields are covariantly constant. In this case $H(\sigma(t)) = L_{\sigma(t)*}$, where $L_{\sigma(t)}$: $G \to G$ is left translation by $\sigma(t)$—$L_{\sigma(t)} g = \sigma(t) g$. Again the solution to (5.1) may be found explicitly,

$$\sigma(t) = \sigma(0)\, e^{t\bar{h}}, \tag{5.3}$$

where $\bar{h}(s) \equiv u_o h(s) \in \mathfrak{g}$.

(c) This is the same as example (b), but now take ∇ to be the covariant derivative for which the right invariant vector fields are parallel. Then in this case the solution to (5.1) is

$$\sigma(t) = e^{t\bar{h}}\sigma(0). \tag{5.4}$$

The transformations in Example 5.1 have been highly studied when $\sigma(0)$ is a Brownian motion and more recently when $\sigma(0)$ is a Brownian Bridge process. In particular, one is interested in whether the law of $\sigma(t)$ is equivalent to the law of $\sigma(0)$. Example 5.1(a) is the domain of the classical Cameron–Martin formula, see [CM1, CM2, Mar] and also [Gr1, Gi, K1–K3, Ra, Ku1–Ku4]. The quasi-invariance for the flow (5.1) in Examples 5.1(b) and (c) is discussed in [AH, Sh1, Sh2, Fr, MM1, Gr4].

In an effort to convince the reader that the solution to Eq. (5.1) is the correct generalizations to the formulas in (5.2)–(5.4), let us briefly discuss two other possible alternatives. A more thorough discussion can be found in Section 10. One alternative to (5.1) which coincides with Examples 5.1(a)–(c) is to use the exponential function. Explicitly, define $\sigma(t) = \exp(tH(\sigma(0))\, h)$, where $\exp: TM \to M$ is the geodesic flow associated with the covariant derivative ∇. But this "shifting" procedure suffers from two serious problems. The first is that in general the map $T_t(\sigma_o) \equiv \exp(tH(\sigma_o)\, h)$ is not a flow on the space of semimartingales. The second is that in most cases the $\mathrm{Law}(T_t(\sigma_o))$ will not be equivalent to the $\mathrm{Law}(\sigma_o)$ when σ_o is a Brownian motion on M. See Section 10 for more details.

A second possible curve shifting technique is to use the flow of a given s-dependent vector field. Explicitly, let $X: [0, 1] \to \Gamma(TM)$ be an s-dependent vector field on M such that $X(0) = 0$. Now for any vector field $Y \in \Gamma(TM)$, let e^{tY} denote the flow on M generated by Y. With this notation, define the shift $T_t(\sigma_o)$ of σ_o by $T_t(\sigma_o)(s) \equiv e^{tX(s)}(\sigma_o(s))$. This prescription again reproduces (5.2)–(5.4) after an appropriate choice of vector field X depending on h. This procedure is considered in [MM1] in the special case that M is a homogeneous space. In this case T_t is always a flow on $W(M)$, but in general T_t will not leave Brownian motions on M quasi-invariant. In Section 10 it is shown that in order for T_t to leave the Wiener measure quasi-invariant the vector field $X(s)$ must be a Killing vector field for each s. In other words, $e^{tX(s)}$ should act isometrically on M. Of course the generic manifold does not admit any non-trivial Killing vector fields, and this shifting technique is then useless.

Notation 5.1. Suppose Q is an imbedded submanifold of some Euclidean space $\mathbb{R}^N$ and that $q_o \in Q$ is a fixed base point. Let $\mathcal{S}^\infty Q$-denote the space of Brownian semimartingales in Q that start at q_o which are also in $B^\infty(\mathbb{R}^N)$. (If $Q = \mathbb{R}^n$, M, or $O(M)$, then $q_o = 0, o, u_o$, respectively.) The space $\mathcal{S}^\infty Q$ is equipped with the topology of convergence in the B^p-norm for all $p \in [2. \infty)$. (Note $p \neq \infty$ here.) So a function $q: \mathbb{R} \to \mathcal{S}^{\infty}Q$ is said to be continuous if it is B^p-norm continuous for all $p \in [2, \infty)$. We say q is C^1 if: (i) for each $T > 0$, $\sup_{|t| \leqslant T} \|q(t)\|_{B^{\infty}} < \infty$, and (ii) $\dot{q}$ exists and is continuous in the B^p-norm for each $p \in [2, \infty)$. It is not required that $\dot{q}$ be in $B^{\infty}(\mathbb{R}^N)$.

Remark 5.2. If $q: J \to \mathcal{S}^{\infty}Q$ is C^1, then it follows that q is B^p-Lipschitz for all $p \in [2, \infty)$. Therefore by Kolmogorov's lemma (Lemma 4.4), we can choose a version of q such that P-a.s. $(t, s) \to q(t)(s)$ is continuous. In the sequel, such a version will always be chosen.

We now restrict our attention to the spaces $\mathcal{S}^{\infty}M$, and $\mathcal{S}^{\infty}O(M)$, where an imbedding of M into $\mathbb{R}^N$ has been chosen and fixed once and for all. By Lemma 2.2, the imbedding of M into $\mathbb{R}^N$ induces an imbedding of $O(M)$ into $\mathbb{R}^N \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$. The next proposition is a regularity result for solutions to (5.1).

PROPOSITION 5.1. *Suppose that $\sigma: \mathbb{R} \to \mathcal{S}^{\infty}M$ is a C^1-solution to* (5.1), *then in fact $\dot{\sigma}$ is B^p-Lipschitz for all $p \in [2, \infty)$.*

Proof. Since σ is C^1, σ is B^p-Lipschitz for all $p \in [2, \infty)$. Therefore, by Corollary 4.1 and our method of constructing $H(\sigma)$ in Theorem 3.2, it follows that $t \to H(\sigma(t))$ is B^p-Lipschitz for all $p \geqslant 2$ (see Lemma 7.3(i) below). The lemma now follows, because $\dot{\sigma}(t) = H(\sigma(t))h$ and Lemma 4.1(vi) gives

$$\|\dot{\sigma}(t_1) - \dot{\sigma}(t_2)\|_{B^p} \leqslant c_{r,r'} \|H(\sigma(t_1)) - H(\sigma(t_2))\|_{B^r} \|h\|_{B^{r'}},$$

where $1/r + 1/r' = 1/p$. Q.E.D.

The following theorem is the stochastic analogue of Theorem 2.2.

THEOREM 5.1. *Let $\sigma: J \to \mathcal{S}^{\infty}M$ be a C^1 function satisfying* (5.1), *and set $w(t) \equiv I^{-1} \circ H(\sigma(t))$. Then $w: J \to \mathcal{S}^{\infty}(\mathbb{R}^n)$ is a C^1-function and satisfies*

$$\dot{w}(t) = \int C(w(t))\, \delta w(t) + h, \tag{5.5}$$

where for any Brownian semimartingale (w),

$$C(w) \equiv A(w) + T(w) \tag{5.6}$$

with

$$A(w) \equiv \int \Omega_u \langle h, \delta w \rangle, \tag{5.7}$$

and

$$T(w) \equiv \Theta_u \langle h, \cdot \rangle, \tag{5.8}$$

where $u \equiv \pi \circ I(w)$.

Proof. We will closely imitate the proof of Theorem 2.2. First notice by Theorem 3.3 that $u = H(\sigma)$. Therefore by the method of construction of $H(\sigma)$ in in Theorem 3.2 and by Corollary 4.2, it follows that $u(t)$ is C^1 in $\mathscr{S}^\infty O(M)$ (see the proof of Lemma 7.3(i) below for more details). By Corollary 4.3 it follows that $w(t) \equiv I^{-1}(u(t)) = \int \vartheta\langle \delta u(t)\rangle$ is C^1 in $\mathscr{S}^\infty \mathbb{R}^n$. In fact because of Proposition 5.1, we know that $\dot{w}$ and $\dot{u}$ are in fact B^p-Lipschitz for all $p \geqslant 2$—but we will not need this property. By definition u is horizontal ($\omega\langle \delta u\rangle = 0$), thus by Corollary 4.3,

$$0 = \frac{d}{dt}\,\omega\langle \delta u\rangle = d\omega\langle \dot{u}, \delta u\rangle + \delta(\omega\langle \dot{u}\rangle) = \Omega\langle \dot{u}, \delta u\rangle + \delta(\omega\langle \dot{u}\rangle), \tag{5.9}$$

where the second equality is a result of the second structure equation (Lemma 2.1) and the fact that $\omega \wedge \omega\langle \dot{u}, \delta u\rangle = [\omega\langle \dot{u}\rangle, \omega\langle \delta u\rangle] = 0$. Now $\Omega\langle \dot{u}, \delta u\rangle = \Omega\langle \dot{u}, B\langle \delta w\rangle(u)\rangle = \Omega_u\langle h, \delta w\rangle$, since $\Omega\langle \dot{u}, \cdot\rangle$ only depends on the horizontal component of $\dot{u}$ which is equal to $B\langle h\rangle(u)$. (Note: $\pi_* \dot{u} = \dot{\sigma} = uh = \pi_* B\langle h\rangle(u)$, see Section 2 for the notation.) So (5.9) may be rewritten as

$$\omega\langle \dot{u}(t)\rangle = -\int \Omega_{u(t)}\langle h, \delta w(t)\rangle = -A(w(t)), \tag{5.10}$$

because $\dot{u}(t)(0) = 0 \in T_{u_0}O(M)$ and hence $\omega\langle \dot{u}(t)(0)\rangle = 0$.

Now again use Corollary 4.3 to differentiate the equation $w = I^{-1}(u) = \int \vartheta\langle \delta u\rangle$ with respect to t to get

$$\dot{w}(t)(s) = \int_0^s d\vartheta\langle \dot{u}(t), \delta u(t)\rangle + \vartheta\langle \dot{u}(t)\rangle|_0^s. \tag{5.11}$$

Using the first structure equation (Lemma 2.1) and arguments like those used to go from (5.9) to (5.10), (5.11) may be rewritten as

$$\dot{w}(t)(s) = \int_0^s \Theta_{u(t)}\langle h, \delta w(t)\rangle - \int_0^s \omega\langle \dot{u}(t)\rangle\, \delta w(t) + h(s). \tag{5.12}$$

The theorem follows from Eqs. (5.6)–(5.8), (5.10), and (5.12). Q.E.D.

The next theorem is the converse of Theorem 5.1 and is the stochastic analogue of Theorem 2.3.

THEOREM 5.2. *Suppose that $w\colon \mathbb{R} \to \mathscr{S}^\infty \mathbb{R}^n$ is a C^1-function that satisfies* (5.5). *If $\sigma \equiv \pi \circ I(w)$, then $\sigma\colon \mathbb{R} \to \mathscr{S}^\infty M$ is a C^1-function satisfying* (5.1).

Proof. Let $u(t) \equiv I(w(t))$, so that $\sigma(t) = \pi \circ u(t)$. By Proposition 6.3 below, $\dot{w}(t)$ is automatically B^p-Lipschitz for all $p \in [2, \infty)$, and hence by Corollary 4.2, $u\colon \mathbb{R} \to \mathscr{S}^\infty O(M)$ is C^1. By Lemma 4.8, $\sigma\colon \mathbb{R} \to \mathscr{S}^\infty M$ is also C^1. Define $\xi = \vartheta\langle \dot{u}\rangle$, $E = \omega\langle \dot{u}\rangle$, and $A \equiv A(w) = \int \Omega_u\langle h, \delta w\rangle$, where for notational simplicity the t-variable is also being suppressed. Our goal is to show that $\xi = h$, since then $\dot{\sigma} = \pi_* \dot{u} = u\vartheta\langle \dot{u} \rangle = uh = H(\sigma)h$ as desired. To prove $\xi = h$, we will follow closely Theorem 2.2 and show that the semimartingales $(\xi - h, A + E)$ satisfies a linear stochastic differential equation. Because $(\xi - h, A + E) = 0$ if $s = 0$, it will follow that $(\xi - h, A + E) \equiv 0$.

Start by computing the differential of ξ using Corollary 4.3,

$$\begin{aligned} d\xi = d(\vartheta\langle \dot{u} \rangle) &= d\vartheta\langle \delta u, \dot{u} \rangle + \frac{d}{dt} \langle \vartheta\langle \delta u \rangle) \\ &= d\vartheta\langle \delta u, \dot{u} \rangle + \frac{d}{dt} (dw), \end{aligned} \tag{5.13}$$

where $\vartheta\langle \delta u \rangle = d(I^{-1}(u)) = dw$ was used in the last equality. Using the first structure equation $(\Theta = d\vartheta + \omega \wedge \vartheta)$ and computing as in the proof of Theorem 5.1 we find

$$\begin{aligned} d\xi &= d\dot{w} + d\vartheta\langle \delta u, \dot{u} \rangle = d\dot{w} + \Theta\langle \delta u, \dot{u} \rangle - \omega \wedge \vartheta\langle \delta u, \dot{u} \rangle \\ &= d\dot{w} + \Theta\langle \delta u, \dot{u} \rangle + \omega\langle \dot{u} \rangle\, \vartheta\langle \delta u \rangle = d\dot{w} + \Theta_u\langle \delta w, \xi \rangle + E\, \delta w, \end{aligned}$$

since $\omega\langle \delta u \rangle = 0$. Combining the formula for $d\xi$ with that of $d\dot{w}$ from (5.5), $(d\dot{w} = A\, \delta w + \Theta_u\langle h, \delta w \rangle + dh)$ yields

$$d(\xi - h) = (A + E)\, \delta w + \Theta_u\langle \delta w, \xi - h \rangle. \tag{5.14}$$

Again using Corollary 4.3 and the fact that $\omega\langle \delta u \rangle = 0$, compute

$$0 = \frac{d}{dt} \omega\langle \delta u \rangle = d\omega\langle \dot{u}, \delta u \rangle + d(\omega\langle \dot{u} \rangle) = \Omega\langle \dot{u}, \delta u \rangle + dE. \tag{5.15}$$

Because $\vartheta\langle \dot{u} \rangle = \xi$, and $\vartheta\langle \delta u \rangle = \delta w$, (5.15) may be rewritten as $0 = \Omega_u\langle \xi, \delta w \rangle + dE$, which when added with $\Omega_u\langle h, \delta w \rangle - \delta A = 0$ yields

$$d(A + E) = \Omega_u\langle h - \xi, \delta w \rangle. \tag{5.16}$$

For each fixed t, (5.14) and (5.16) are a coupled pair of linear stochastic differential equations for $(\xi(t) - h)$ and $(A(t) + E(t))$ with initial conditions $\xi(t) - h = 0$, and $A(t) + E(t) = 0$ at $s = 0$. By uniqueness of solutions to such equations we see that $\xi(t) - h = 0$ and $A(t) + E(t) = 0$ P-a.s. This proves the theorem. Q.E.D.

6. Existence and Uniqueness for the Non-geometric Flow

In this section an existence and uniqueness theorem for the "non-geometric flow equation" (5.5) of Theorem 5.1 will be proved. Hence, by

Theorems 5.1 and 5.2 this will also prove existence and uniqueness for the "geometric flow equation" (5.1). The following section is devoted to proving directly existence and uniqueness for the geometric flow equation (5.1).

The first step in the proof is to reformulate the non-geometric flow equations (5.5), see Proposition 6.1 below. For the purposes of this section it is important to remember that a C^1-function $w: J \to \mathcal{S}^\infty \mathbb{R}^n$ is equivalent to giving a pair of processes $(O, \alpha): J \to S^\infty(\mathrm{End}(\mathbb{R}^n) \times \mathbb{R}^n)$, such that (O, α) is S^p continuously differentiable for all $p \geqslant 2$ and $\sup_{t \in J} \{ \|O(t)\|_{S^\infty} + \|\alpha(t)\|_{S^\infty} \} < \infty$. Of course w is related to (O, α) by $w(t) = \int O(t)\, db + \int \alpha(t)\, ds$.

Definition 6.1. Suppose that $f: O(M) \to V$ ($V = a$ vector space) is a C^∞-function, let $f': O(M) \to \mathrm{End}(\mathbb{R}^n, V)$ be defined by $f'(u)\langle a \rangle = df\langle B\langle a \rangle(u) \rangle$ for all $u \in O(M)$ and $a \in \mathbb{R}^n$. We will call f' the horizontal derivative of f.

Remark 6.1. The two main examples of interest are $f(u) = \Omega_u\langle \cdot, \cdot \rangle$, and $f(u) = \Theta_u\langle \cdot, \cdot \rangle$, in which case $f'(u)\langle a \rangle$ will be denoted by $\Omega'_u\langle a, \cdot, \cdot \rangle$ and $\Theta'_u\langle a, \cdot, \cdot \rangle$, respectively.

Definition 6.2. Let A and $B \in \mathrm{End}(\mathbb{R}^n) \equiv \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^n)$, $a \in \mathbb{R}^n$, and $u \in O(M)$. Define

$$\text{(i)} \qquad \bar{\Omega}_u\langle A, a, B \rangle \equiv \sum_{i=1}^n \Omega'_u\langle Ae_i, a, Be_i \rangle, \tag{6.1}$$

$$\text{(ii)} \qquad \bar{\Theta}_u\langle A, a, B \rangle = \sum_{i=1}^n \Theta'_u\langle Ae_i, a, Be_i \rangle, \text{ and} \tag{6.2}$$

$$\text{(iii)} \qquad \mathrm{Ric}_u\langle a, A, B \rangle = \sum_{i=1}^n \Omega_u\langle a, Ae_i \rangle Be_i, \tag{6.3}$$

where $\{e_i\}$ is the standard basis for $\mathbb{R}^n$.

Remark 6.2. If O is an orthogonal matrix and $A = B = O$, then the formulas (6.1)–(6.3) are independent of O. The reason for the terminology Ric_u in the (6.3) is that $\mathrm{Ric}_u\langle a, O, O \rangle = \sum_{i=1}^n \Omega_u\langle a, e_i \rangle e_i$ is essentially the Ricci-tensor when O is orthogonal.

Proposition 6.1. *Suppose that $w: J \to \mathcal{S}^\infty \mathbb{R}^n$, and $(O, \alpha): J \to \mathcal{S}^\infty \mathbb{R}^n$ is defined by*

$$w(t) = \int O(t)\, db + \int \alpha(t)\, ds. \tag{6.4}$$

Then Eq. (5.5) *is equivalent to the pair of differential equations*

$$\dot{O}(t) = C(w(t))\, O(t) \tag{6.5}$$

and

$$\dot{\alpha}(t) = C(w(t))\, \alpha(t) + R(w(t)), \tag{6.6}$$

where $C(w)$ *is defined as in Theorem* 5.1 *and* $R(w)$ *is defined to be*

$$R(w) = \tfrac{1}{2}\{\mathrm{Ric}_u\langle h, O, O\rangle + \bar{\Theta}_u\langle O, h, O\rangle\} + h', \tag{6.7}$$

where w *is the Brownian semimartingale* $w = \int O\, db + \int \alpha\, ds$. *The derivatives in* (6.5) *and* (6.6) *are to be taken in the* S^p*-topologies for all* $p \in [2, \infty)$.

Proof. Insert the expression (6.4) for $w(t)$ into (5.5) to get

$$\dot{O}\, db + \dot{\alpha}\, ds = C(w)[O\, db + \alpha\, ds] + \tfrac{1}{2}\, dC(w)\, dw + h'\, ds, \tag{6.8}$$

where the t is now also being suppressed from the notation. By (5.7), (5.8), Itô's Lemma, and $\delta u = B\langle \delta w\rangle(u)$,

$$\begin{aligned} dA(w) &= \Omega_u\langle h, \delta w\rangle = \Omega_u\langle h, dw\rangle + \tfrac{1}{2}\Omega'_u\langle dw, h, dw\rangle \\ &= \Omega_u\langle h, O\, db\rangle + \{\Omega_u\langle h, \alpha\rangle + \tfrac{1}{2}\bar{\Omega}_u\langle O, h, O\rangle\}\, ds, \end{aligned} \tag{6.9}$$

and

$$\begin{aligned} dT(w) &= d(\Theta_u\langle h, \cdot\rangle) = \Theta'_u\langle \delta w, h, \cdot\rangle + \Theta_u\langle dh, \cdot\rangle \\ &= \Theta'_u\langle dw, h, \cdot\rangle + \tfrac{1}{2}\Theta''_u\langle dw, dw, h, \cdot\rangle + \Theta_u\langle dh, \cdot\rangle \\ &= \Theta'_u\langle O\, db, h, \cdot\rangle + \{\Theta'_u\langle \alpha, h, \cdot\rangle \\ &\quad + \Theta_u\langle h', \cdot\rangle + \tfrac{1}{2}\bar{\Theta}''_u\langle O, O, h, \cdot\rangle\}\, ds, \end{aligned} \tag{6.10}$$

where $\bar{\Theta}''_u\langle O, O, h, \cdot\rangle \equiv \sum_i \Theta''_u\langle Oe_i, Oe_i, h, \cdot\rangle$, and Θ'' is the horizontal derivative of Θ'.

Thus using $d[C(w)] = dA(w) + dT(w)$, (6.9), and (6.10) one finds

$$\begin{aligned} dC(w)\, dw &= \Omega_u\langle h, dw\rangle\, dw + \Theta'_u\langle dw, h, dw\rangle \\ &= \mathrm{Ric}_u\,\langle h, O, O\rangle\, ds + \bar{\Theta}_u\langle O, h, O\rangle\, ds. \end{aligned} \tag{6.11}$$

Noting that the B^p-norm on $w = \int O\, db + \int \alpha\, ds$ is equivalent to the S^p-norm on (O, α), the lemma is proved upon inserting (6.11) into (6.8) and then comparing the db and ds terms on both sides of the result. Q.E.D.

Remark 6.3. Let $C = \sup_{u \in O(M)} |\Theta_u\langle\cdot, \cdot\rangle|$. Notice that $C(w) = A(w) + T(w)$, where $A(w)(s) \in so(n)$ and $T(w)(s) \in \mathrm{End}(\mathbb{R}^n)$ such that $|T(w)(s)| \leqslant$

$C\,|h|_\infty$ independent of w, s, and the random sample point. Lemma 6.1 below and this observation imply that any solution to (6.5) and (6.6) has the property that $O(t)(s)$ is uniformly bounded provided $O(0)(s)$ was uniformly bounded. With $O(t)$ uniformly bounded, it follows from the definition of $R(w(t))$ that $R(w(t))(s)$ remains uniformly bounded. This in turn will imply (by Lemma 6.1) that α must remain bounded. (See Corollary 6.1.)

LEMMA 6.1. *Let $V=\mathbb{R}^n$ or $\mathrm{End}(\mathbb{R}^n)$, and assume that $A\colon J\to so(n)$, $T\colon J\to \mathrm{End}(\mathbb{R}^n)$, and $R\colon J\to V$ are all continuous maps, where $J=[-1,1]$. Let $\alpha_o\in V$ be a fixed vector and define $\alpha\colon J\to V$ to be the unique solution to the ordinary differential equation*

$$\dot\alpha(t)=A(t)\,\alpha(t)+T(t)\,\alpha(t)+R(t)\qquad \textit{with}\quad \alpha(0)=\alpha_o. \tag{6.12}$$

Then

(i)
$$|\alpha|_\infty\leqslant(\,|\alpha_o|+|R|_\infty)\,e^{|T|_\infty}, \tag{6.13}$$

where $|T|_\infty=\sup_{t\in J}|T(t)|$, and $|R|_\infty=\sup_{t\in J}|R(t)|$. (Notice this estimate is independent of the size of A—intuitively A supplies only a rotation term.)

(ii) *Now also suppose that $\bar A\colon J\to so(n)$, $\bar T\colon J\to\mathrm{End}(\mathbb{R}^n)$, and $\bar R\colon J\to V$ are another collection of continuous maps and that $\bar\alpha$ denotes the solution to (6.12) with A, T, and R replaced by $\bar A$, $\bar B$, $\bar R$. Then there exists a constant $K=K(\,|\alpha_o|,\,|T|_\infty,\,|R|_\infty,\,|\bar T|_\infty,\,|\bar R|_\infty)$, such that*

$$|\alpha(t)-\bar\alpha(t)|\leqslant K\left|\int_0^t\{\,|A(\tau)-\bar A(\tau)|+|T(\tau)-\bar T(\tau)|\right.$$
$$\left.+\,|R(\tau)-\bar R(\tau)|\,\}\,d\tau\,\right|. \tag{6.14}$$

(Notice that the constant K is independent of A and $\bar A$. This is crucial to later applications.)

Proof. (i) For simplicity assume that $t\geqslant 0$—the case $t\leqslant 0$ is similar. Let $U\colon J\to O(n)$ be the unique solution to

$$\dot U(t)=A(t)\,U(t)\qquad\text{with}\quad U(0)=\mathrm{id}\in O(n) \tag{6.15}$$

and set $Z(t)\equiv U(t)^{-1}\,\alpha(t)$, then

$$\dot Z=U^{-1}TUZ+U^{-1}R\qquad\text{and}\quad Z(0)=\alpha_o. \tag{6.16}$$

Because $U(t)$ is orthogonal, we have that $|\alpha(t)|=|Z(t)|$, so it suffices to show that Z satisfies the estimate (6.13). From (6.16) one easily finds

$$|Z(t)|\leqslant|\alpha_o|+\int_0^t[\,|T(\tau)|\cdot|Z(\tau)|+|R(\tau)|\,]\,d\tau.$$

This last equation and Gronwall's inequality (Lemma 4.3) yields (6.13).

(ii) In this argument K will denote a constant depending on $(|\alpha_o|, |T|_\infty, |R|_\infty, |\bar{T}|_\infty, |\bar{R}|_\infty)$. The different K's in a string of inequalities may vary from place to place.

Let U and Z be defined as above and define $\bar{U}$ and $\bar{Z}$ analogously with (A, T, R) replaced by $(\bar{A}, \bar{T}, \bar{R})$. Then

$$\begin{aligned}|\alpha-\bar{\alpha}| = |UZ-\bar{U}\cdot\bar{Z}| &\leqslant |U|\;|Z-\bar{Z}| + |U-\bar{U}|\;|\bar{Z}| \\ &\leqslant |Z-\bar{Z}| + K\,|U-\bar{U}|, \qquad (6.17)\end{aligned}$$

where we have used part (i) to replace $|\bar{Z}|$ by K. Estimate the U-term:

$$\begin{aligned}|U(t)-\bar{U}(t)| &= |U(t)^{-1}\,\bar{U}(t) - I| \\ &\leqslant \int_0^t \left|\frac{d}{d\tau}\,[U(\tau)^{-1}\,\bar{U}(\tau)]\right| d\tau \\ &= \int_0^t |U(\tau)^{-1}\,[\bar{A}(\tau)-A(\tau)]\,\bar{U}(\tau)|\,d\tau \\ &= \int_0^t |\bar{A}(\tau)-A(\tau)|\,d\tau. \qquad (6.18)\end{aligned}$$

Now to the Z term. From (6.16) we see that

$$\begin{aligned}|\dot{Z}-\dot{\bar{Z}}| &= |U^{-1}[TUZ+R] - \bar{U}^{-1}[\bar{T}\bar{U}\bar{Z}+\bar{R}]| \\ &\leqslant |U^{-1}-\bar{U}^{-1}|\;|TUZ+R| + |TUZ-\bar{T}\bar{U}\bar{Z}| + |R-\bar{R}| \\ &\leqslant K\,|U-\bar{U}| + |T-\bar{T}|\;|UZ| + |\bar{T}|\;|UZ-\bar{U}\bar{Z}| + |R-\bar{R}| \\ &\leqslant K\,|U-\bar{U}| + K\,|T-\bar{T}| + K\,|UZ-\bar{U}\bar{Z}| + |R-\bar{R}| \\ &\leqslant K\,[\,|U-\bar{U}| + |T-\bar{T}| + |U-\bar{U}|\;|\bar{Z}| + |Z-\bar{Z}|\,] + |R-\bar{R}| \\ &\leqslant K\,[\,|U-\bar{U}| + |T-\bar{T}| + |Z-\bar{Z}| + |R-\bar{R}|\,]. \qquad (6.19)\end{aligned}$$

Upon integration of this inequality from 0 to t, and an application of Gronwall's inequality it is seen that

$$\begin{aligned}|Z(t)-\bar{Z}(t)| &\leqslant K\int_0^t [\,|U(\tau)-\bar{U}(\tau)| \\ &\qquad + |T(\tau)-\bar{T}(\tau)| + |R(\tau)-\bar{R}(\tau)|\,]\,d\tau \\ &\leqslant K\int_0^t [\,|A(\tau)-\bar{A}(\tau)| \\ &\qquad + |T(\tau)-\bar{T}(\tau)| + |R(\tau)-\bar{R}(\tau)|\,]\,d\tau, \qquad (6.20)\end{aligned}$$

where (6.18) was used in going from the first to the second inequality. The result follows by combining (6.17), (6.18), and (6.20). Q.E.D.

DEFINITION 6.3. A 1-parameter family of $\mathrm{End}(\mathbb{R}^n) \times \mathbb{R}^n$-valued processes $(O(t), \alpha(t))$ solves (6.5) and (6.6) *pointwise* if the following conditions are verified.

(i) P-a.s. the function $(t, s) \to (O(t)(s), \alpha(t)(s))$ is $C^{1,0}$.

(ii) Let $w(t) \equiv \int O(t)\, db + \int \alpha(t)\, ds$. There exists versions of $u(t)$ of $I(w(t))$, and $A(t)$ of $A(w(t))$ such that P-a.s. the maps $(t, s) \to u(t)(s)$ and $(t, s) \to A(t)(s)$ are $C^{1,0}$.

Let $R(t)$ be the version of $R(w(t))$ found by inserting $O(t)$ and $u(t)$ into Eq. (6.6). Let $T(t) \equiv \Theta_{u(t)}\langle h, \cdot\rangle$ (a version of $T(w(t))$), and $C(t) \equiv A(t) + T(t)$ (a version of $C(w(t))$). Notice, with these choices, P-a.s. the map $(t, s) \to (R(t)(s), T(t)(s), C(t)(s))$ is $C^{1,0}$.

(iii) There is a fixed set $\Omega_o \subset \Omega$ of full measure such that on Ω_o, $(O(t)(s), \alpha(t)(s))$ verifies pointwise (6.5) and (6.6) with $C(w(t))$ and $R(w(t))$ replaced by $C(t)$ and $R(t)$.

COROLLARY 6.1. *Let $(O(t), \alpha(t))$ be a 1-parameter family of $\mathrm{End}(\mathbb{R}^n) \times \mathbb{R}^n$-valued processes which solves (6.5) and (6.6) pointwise, and set $w(t) = \int O(t)\, db + \int \alpha(t)\, ds$. If $\|w(0)\|_{B^\infty} = \|O(0)\|_{S^\infty} + \|\alpha(0)\|_{S^\infty} < \infty$, then $\|O(t)\|_{S^\infty} \leqslant \|O(0)\|_{S^\infty} e^{C|h|_\infty}$, and $\|w(t)\|_{B^\infty} = \|O(t)\|_{S^\infty} + \|\alpha(t)\|_{S^\infty} \leqslant C_o$ for all $t \in J$, where $C_o = C_o(|h'|_\infty, \|w(0)\|_{B^\infty})$ and $C \equiv \sup_{u \in O(M)} |\Theta_u\langle\cdot, \cdot\rangle|$.*

Remark 6.4. Throughout the paper I will make repeated use of the following measure theoretic fact. Let $f\colon U \times \Omega \to V$ be a measurable map where U is a subset of $\mathbb{R}^k$ and V is a normed vector space. Suppose that for each $\omega \in \Omega$ the map $x \to f(x, \omega)$ is continuous, and there is a constant C_o such that for each $x \in U$, $P(\{\omega\colon |f(x, \omega)| \leqslant C_o\}) = 1$, then $P(\{\omega\colon \sup_{x \in U} |f(x, \omega)| \leqslant C_o\}) = 1$. Indeed, if D is a countable dense subset of U, then it is clear that $P(\Omega_D) = 1$, where $\Omega_D \equiv \{\omega\colon \sup_{x \in D} |f(x, \omega)| \leqslant C_o\}$. But by continuity, for all $\omega \in \Omega_D$ one has $\sup_{x \in U} |f(x, \omega)| \leqslant C_0$. Since, $\Omega_D \subset \Omega_U \equiv \{\omega\colon \sup_{x \in U} |f(x, \omega)| \leqslant C_o\}$ and our σ-algebra is P-complete, it follows that Ω_U is measurable and that $P(\Omega_U) = 1$.

Proof of Corollary 6.1. For $t \in J$, let $A(t)$, $T(t)$, $R(t)$, and $C(t)$ be as in Definition 6.3. Notice that $A(t) \in so(n)$, and on Ω_o, $|T(t)|_\infty \leqslant C\,|h|_\infty$. (I will omit the statement "on Ω_o" in the future.) So by Lemma 6.1(i), $|O(t)|_\infty \leqslant |O(0)|_\infty\, e^{C|h|_\infty}$, from which it follows that

$$\|O(t)\|_{S^\infty} \leqslant \|O(0)\|_{S^\infty}\, e^{C|h|_\infty}.$$

Also it is easy to show there is a constant K such that

$$|R(t)|_\infty \leqslant K\,|O(t)|_\infty^2\,|h|_\infty + |h'|_\infty \leqslant K\,\|O(0)\|_{S^\infty}^2\,e^{2C\,|h|_\infty}\,|h|_\infty + |h'|_\infty,$$

where in the second inequality we used the estimate on O just proved. Therefore we may apply Lemma 6.1(i) to (6.6) to get

$$\|\alpha(t)\|_{S^\infty} \leqslant (\,\|\alpha(0)\|_{S^\infty} + K\,\|O(0)\|_{S^\infty}^2\,e^{2C\,|h|_\infty}\,|h|_\infty + |h'|_\infty)\,e^{C\,|h|_\infty}.$$

The lemma now easily follows with

$$C_o \equiv (2\,\|w(0)\|_{B^\infty} + K\,\|w(0)\|_{B^\infty}^2\,|h'|_\infty + |h'|_\infty)\,e^{3C\,|h'|_\infty} \tag{6.21}$$

because $|h|_\infty \leqslant |h'|_\infty$, and $\|w(t)\|_{S^\infty} \equiv \|O(t)\|_{B^\infty} + \|\alpha(t)\|_{B^\infty}$. Q.E.D.

After proving some basic estimates (Proposition 6.2) on $A(w)$, $T(w)$, and $R(w)$, we will see that the $w\colon J \to \mathscr{S}^\infty\mathbb{R}^n$ in the above lemma is in fact C^1 and $\dot{w}$ is B^p-Lipschitz for $p \in [2, \infty)$. Furthermore, this w satisfies (5.5), or equivalently (6.5) and (6.6) with derivatives taken in the S^p-topologies.

PROPOSITION 6.2 (Basic Estimates). *Let $0 < C_o < \infty$ be a fixed constant and suppose that w and $\bar{w}$ are in $\mathscr{S}^\infty\mathbb{R}^n$ with $\|w\|_{B^\infty} \leqslant C_o$ and $\|\bar{w}\|_{B^\infty} \leqslant C_o$. Also assume that $|h'|_\infty \leqslant C_o$. Then for all $p \in [2, \infty)$ there exists a constant $K = K(p, C_o)$ such that:*

(i) $\|A(w) - A(\bar{w})\|_{H^p} \leqslant K\,\|w - \bar{w}\|_{H^p}$;

(ii) $\|A(w) - A(\bar{w})\|_{S^p} \leqslant c_p K\,\|w - \bar{w}\|_{H^p} \leqslant c_p K\,\|w - \bar{w}\|_{B^p}$;

(iii) $\|T(w) - T(\bar{w})\|_{H^p} \leqslant K\,\|w - \bar{w}\|_{H^p}$;

(iv) $\|T(w) - T(\bar{w})\|_{S^p} \leqslant c_p K\,\|w - \bar{w}\|_{H^p} \leqslant c_p K\,\|w - \bar{w}\|_{B^p}$;

(v) $\|R(w) - R(\bar{w})\|_{S^p} \leqslant K\,\|w - \bar{w}\|_{B^p}$;

(vi) $\|C(w) - C(\bar{w})\|_{H^p} \leqslant K\,\|w - \bar{w}\|_{H^p}$;

(vii) $\|C(w) - C(\bar{w})\|_{S^p} \leqslant c_p K\,\|w - \bar{w}\|_{H^p} \leqslant c_p K\,\|w - \bar{w}\|_{B^p}$.

Remark 6.5. If O and $\bar{O}$ are restricted to be orthogonal matricies, then estimate (v) above can be improved to an H^p-estimate as in (i) and (iii). We will not need this improved version of (v).

Proof. Throughout this proof K will denote a constant depending only on C_o, p, and the underlying geometrical data. Let $u = I(w)$, $\bar{u} = I(\bar{w})$, $A = A(w)$, and $\bar{A} = A(\bar{w})$. By Corollary 4.1 there is a constant K such that $\|u - \bar{u}\|_{H^p} \leqslant K\,\|w - \bar{w}\|_{H^p}$ and $\|u - \bar{u}\|_{B^p} \leqslant K\,\|w - \bar{w}\|_{B^p}$. These inequalities will be used frequently along with the obvious inequalities that $\|O - \bar{O}\|_{S^p} \leqslant \|w - \bar{w}\|_{B^p}$, $\|\alpha - \bar{\alpha}\|_{S^p} \leqslant \|w - \bar{w}\|_{B^p}$, and $|h|_\infty \leqslant |h'|_\infty \leqslant C_o$.

From (6.9)

$$dA = \Omega_u\langle h, dw\rangle + \tfrac{1}{2}\bar{\Omega}_u\langle O, h, O\rangle\,ds$$

with a similar equation holding for $d\bar{A}$. Thus

$$d(A-\bar{A})=\Omega_u\langle h, dw\rangle-\Omega_{\bar{u}}\langle h, d\bar{w}\rangle+\tfrac{1}{2}\{\bar{\Omega}_u\langle O, h, O\rangle-\bar{\Omega}_{\bar{u}}\langle \bar{O}, h, \bar{O}\rangle\}\, ds.$$

Therefore

$$\begin{aligned}\|A-\bar{A}\|_{H^p}\leqslant &\left\| \int \{\Omega_u\langle h, dw\rangle-\Omega_{\bar{u}}\langle h, dw\rangle\}\right\|_{H^p}\\ &+\left\|\int \Omega_{\bar{u}}\langle h, d(w-\bar{w})\rangle\right\|_{H^p}\\ &+\tfrac{1}{2}\left\|\int \{\bar{\Omega}_u\langle O, h, O\rangle-\bar{\Omega}_{\bar{u}}\langle \bar{O}, h, \bar{O}\rangle\}\, ds'\right\|_{H^p}. \qquad (6.22)\end{aligned}$$

The three terms on the right side of (6.22) will be estimated separately.

The first term is estimated using Emery's and Burkholder's inequalities (Lemma 4.1(i) and (iii)),

$$\begin{aligned}\left\|\int \{\Omega_u\langle h, dw\rangle-\Omega_{\bar{u}}\langle h, dw\rangle\}\right\|_{H^p} &\leqslant K\,\|w\|_{H^\infty}\,\|\Omega_u\langle h, \cdot\rangle-\Omega_{\bar{u}}\langle h, \;\rangle\|_{S^p}\\ &\leqslant K\,\|w\|_{H^\infty}\,|h|_\infty\,\|u-\bar{u}\|_{S^p}\\ &\leqslant c_p K\,\|w\|_{H^\infty}\,|h|_\infty\,\|w-\bar{w}\|_{H^p},\end{aligned}$$

where the first K is dependent on the norms and the second K includes the Lipschitz constant for the function $u\to\Omega_u\langle\cdot,\cdot\rangle$. The second term is easily estimated using Emery's inequality:

$$\left\|\int \Omega_{\bar{u}}\langle h, d(w-\bar{w})\rangle\}\right\|_{H^p}\leqslant \|\Omega_{\bar{u}}\langle h, \cdot\rangle\|_{S^\infty}\,\|w-\bar{w}\|_{H^p}\leqslant K\,|h|_\infty\,\|w-\bar{w}\|_{H^p}.$$

Now for the third term, by elementary estimates it is easily seen P-a.s. that

$$\begin{aligned}|\bar{\Omega}_u\langle O, h, O\rangle-\bar{\Omega}_{\bar{u}}\langle \bar{O}, h, \bar{O}\rangle| &\leqslant |\bar{\Omega}_u-\bar{\Omega}_{\bar{u}}|\;|h|_\infty\, C_o^2+|\bar{\Omega}_{\bar{u}}|\;|h|\;2C_o\,|O-\bar{O}|\\ &\leqslant K\,|u-\bar{u}|+K\,|O-\bar{O}|,\end{aligned}$$

from which it follows that twice the third-term of (6.22) can be estimated by

$$\begin{aligned}K\,\|u-\bar{u}\|_{S^p}+K\left\|\int_0^1 |O-\bar{O}|\,ds\right\|_{L^p} &\leqslant Kc_p\,\|u-\bar{u}\|_{H^p}+K\,\|w-\bar{w}\|_{H^p}\\ &\leqslant K\,\|w-\bar{w}\|_{H^p}.\end{aligned}$$

This finishes the proof of (i). Part (ii) follows from (i) with an application of Burkholder's inequality. I will omit the proofs of (iii) and (iv), since they are similar to the proofs of (i) and (ii) with Eq. (6.10) used in place of (6.9).

By elementary estimates one has P-a.s. that

$$|R(w)-R(\bar{w})|_\infty \leqslant K\,|u-\bar{u}|_\infty + K\,|O-\bar{O}|_\infty,$$

which implies

$$\|R(w)-R(\bar{w})\|_{S^p} \leqslant K\,\|u-\bar{u}\|_{S^p} + K\,\|O-\bar{O}\|_{S^p} \leqslant K\,\|w-\bar{w}\|_{B^p}.$$

This proves (v). Finally, (vi) and (vii) are a direct consequence of (i)–(iv) and the definition $C(w) \equiv A(w)+T(w)$. Q.E.D.

COROLLARY 6.2 (Regularity). *Keeping the same notation and hypothesis as Corollary* 6.1, *the function* $w\colon J \to \mathcal{S}^\infty\mathbb{R}^n$ *is in fact* C^1 *with* $\dot{w}$ B^p*-Lipschitz for all* $p \in [2,\infty)$. *Furthermore, this* w *satisfies* (5.5), *or equivalently* (6.5) *and* (6.6) *where the derivatives are now taken in the* S^p*-topologies.*

Proof. First notice that

$$\begin{aligned}\|C(w)\|_{S^p} &\leqslant \|C(w)-C(0)\|_{S^p} + \|C(0)\|_{S^p}\\ &\leqslant K\,\|w\|_{H^p} + K\,|h|_\infty \leqslant K[\,|h|_\infty + \|w\|_{B^\infty}] \leqslant KC_o < \infty,\end{aligned}$$

provided $\|w\|_{B^\infty} \leqslant C_o$. In particular $\|C(t)\|_{S^p} = \|C(w(t))\|_{S^p} \leqslant K_p < \infty$ for all $t \in J$. Since $|O(t)-O(t')| \leqslant \int_t^{t'} |C(w(\tau))|\,|O(\tau)|\,d\tau \leqslant C_o \int_t^{t'} |C(w(\tau))|\,d\tau$ for all $t'>t$, it follows that $\|O(t)-O(t')\|_{S^p} \leqslant C_o \int_t^{t'} \|C(w(\tau))\|_{S^p}\,d\tau \leqslant K\,|t-t'|$. Similarly, since $R(w)$ is bounded when O and h are bounded, it follows that $\|\alpha(t)-\alpha(t')\|_{S^p} \leqslant K\,|t-t'|$. Therefore, $\|w(t)-w(t')\|_{B^p} \leqslant K\,|t-t'|$ for all $t, t' \in J$. That is, w is B^p-Lipschitz. By Proposition 6.2(vii) it follows that $C(t)=C(w(t))$ is S^p-Lipschitz.

Now let $h>0$, and set $\varepsilon(h) = \|[O(t+h)-O(t)]/h - \dot{O}(t)\|_{S^p}$, where $\dot{O}(t)$ is the pointwise derivative. Using the fundamental theorem of calculus pointwise, we learn

$$[O(t+h)-O(t)]/h - \dot{O}(t) = \frac{1}{h}\int_t^{t+h} [C(\tau)\,O(\tau) - C(t)\,O(t)]\,d\tau.$$

Consequently

$$\begin{aligned}\varepsilon(h) &\leqslant \frac{1}{h}\int_t^{t+h} \|C(\tau)\,O(\tau)-C(t)\,O(t)\|_{S^p}\,d\tau\\ &\leqslant \frac{1}{h}\int_t^{t+h} [K\,\|O(\tau)-O(t)\|_{S^{r}} + K\,\|C(\tau)-C(t)\|_{S^p}]\,d\tau\\ &\leqslant \frac{1}{h}\int_t^{t+h} K\,|\tau-t|\,d\tau \leqslant Kh,\end{aligned}$$

where Holder's inequality (with $1/r + 1/r' = 1/p$) was used in the second inequality along with the boundedness of $C(\tau)$ in the S^r-norms. The argument also works for $h<0$, so that $\varepsilon(h) \leqslant K|h|$. This shows that $O(t)$ is S^p-differentiable with derivative given by $C(t)\,O(t)$. A similar computation would show that $\alpha(t)$ is also S^p-differentiable with derivative given by $C(t)\,\alpha(t) + R(t)$. Therefore, $w(t)$ is B^p-differentiable and satisfies (5.5).

To show that $\dot{w}(t)$ is B^p-Lipschitz, we start with the easy estimate,

$$|\dot{O}(t) - \dot{O}(t')| \leqslant |C(t) - C(t')|\; C_o + |C(t')|\; |O(t) - O(t')|,$$

where C_o is the constant in Corollary 6.1.

From this estimate and Holder's inequality one finds

$$\|\dot{O}(t) - \dot{O}(t')\|_{S^p} \leqslant C_o\, \|C(t) - C(t')\|_{S^p} + \|C(t')\|_{S^r}\, \|O(t) - O(t')|\, \|_{S^{r'}},$$

where $1/p = 1/r + 1/r'$. Since for each $p \in [2, \infty)$, w is B^p-differentiable, it follows w is B^p-Lipschitz. Hence by Proposition 6.2, $C(t) = C(w(t))$ is S^p-Lipschitz for all $p \in [2, \infty)$. These comments and the above displayed estimate clearly imply that $\dot{O}$ is S^p-Lipschitz. The proof that $\dot{\alpha}$ is S^p-Lipschitz is similar. Q.E.D.

The next proposition along with Corollary 6.2 shows that the pointwise notion and "S^p-notion" (in Proposition 6.1) of the solution to (6.5) and (6.6) agree.

PROPOSITION 6.3. *Suppose that* $w: J \to \mathscr{S}^\infty \mathbb{R}^n$ $(w(t) = \int O(t)\, db + \int \alpha(t)\, ds)$ *is a function such that* $(O(t), \alpha(t))$ *satisfy* (6.5) *and* (6.6) *with derivatives taken in the* S^p*-topologies* $(p \in [2, \infty))$ *as in Proposition* 6.1. *Further assume* $\sup_{t \in J} \|w(t)\|_{B^\infty} < \infty$, *then* $w: J \to \mathscr{S}^\infty \mathbb{R}^n$ *is* C^1 *and the derivative* $\dot{w}$ *is* B^p*-Lipschitz for* $p \in [2, \infty)$. *Furthermore, it is possible to choose a version of* $(O(t), \alpha(t))$ *such that* $(O(t), \alpha(t))$ *solves* (6.5) *and* (6.6) *in the pointwise sense of Definition* 8.3.

Proof. Let $C(t)$ and $R(t)$ be versions of $C(w(t))$ and $R(w(t))$, respectively. Since w is B^p-differentiable, w is B^p-continuous on J. Therefore by Proposition 6.2, $t \to C(t)$ and $t \to R(t)$ is S^p-continuous for all $p \in [2, \infty)$. It now follows (with a Holder's inequality argument) from (6.5) and (6.6) that $\dot{O}(t)$ and $\dot{\alpha}(t)$ are S^p-continuous for all p. Thus $w: J \to \mathscr{S}^\infty \mathbb{R}^n$ is C^1, and hence w is B^p-Lipschitz. So by Proposition 6.2, $C(t)$ and $R(t)$ are S^p-Lipschitz for all p. Going back to (6.5) and (6.6), we can now conclude by a Holder's inequality argument that $\dot{O}(t)$ and $\dot{\alpha}(t)$ are in fact S^p-Lipschitz—i.e., $\dot{w}(t)$ is B^p-Lipschitz.

Because $\dot{w}$ is B^p-Lipschitz, Lemma 4.5 asserts the existence of $C^{1,0}$ versions of $O(t)$ and $\alpha(t)$ which will still be denoted by $O(t)$ and $\alpha(t)$, respectively. Also using Lemma 4.4, Lemma 4.5, and Theorem 4.2, we may choose version $A(t)$ and $u(t)$ of $A(w(t))$ and $I(w(t))$, respectively, such that $(t, s) \to A(t)(s)$ is continuous, and $(t, s) \to u(t)(s)$ is $C^{1,0}$. Now let $C(t)$, $R(t)$, and $T(t)$ denote the versions of $C(w(t))$, $R(w(t))$, and $T(w(t))$ described in Definition 6.3. It is easy to see that the pointwise notion of the derivative of $(O(t), \alpha(t))$ and the S^p-notion of the derivative agree P-a.s. Therefore from (6.5), we know for each t that P-a.s. $\dot{O}(t) = C(t)\, O(t)$, where the derivative is now taken pointwise. Since both sides of this last equation are continuous processes, it follows (in the standard way) that there is a fixed subset $\Omega_o \subset \Omega$ of full measure such that on Ω_o, $\dot{O}(t)(s) = C(t)(s)\, O(t)(s)$ for all $(t, s) \in J \times [0, 1]$. A similar argument shows that α satisfies (6.6) pointwise. Q.E.D.

We now come to the first proof of existence and uniqueness of solutions to equations (6.5) and (6.6).

THEOREM 6.1. *Let h: $[0, 1] \to \mathbb{R}^n$ be a fixed C^1-function with $h(0) = 0$. Suppose that $w_o = \int O_o\, db + \int \alpha_o\, ds$ is a Brownian semimartingale in $\mathcal{S}^\infty \mathbb{R}^n$. Then in the class of all differentiable functions w: $\mathbb{R} \to \mathcal{S}^\infty \mathbb{R}^n$ such that $\sup_{|t| \leqslant T} \|w(t)\|_{B^\infty} < \infty$ for all $T > 0$, there is a unique member (w) satisfying (5.5) and $w(0) = w_o$. (Equivalently if $(O(t), \alpha(t))$ is defined by $w(t) = \int O(t)\, db + \int \alpha(t)\, ds$, then there exists a unique solution $(O(t), \alpha(t))$ to (6.5) and (6.6) with $O(0) = O_o$ and $\alpha(0) = \alpha_o$ such that $\sup_{|t| \leqslant T} [\|O(t)\|_{S^\infty} + \|\alpha(t)\|_{S^\infty}] < \infty$ for all $T > \infty$.) This solution (w) has the property that $\dot{w}$ is B^p-Lipschitz for all $p \in [2, \infty)$. Furthermore, if T^h_t: $\mathcal{S}^\infty \mathbb{R}^n \to \mathcal{S}^\infty \mathbb{R}^n$ is defined by $T^h_t(w_o) = w(t)$, then T^h is a flow on $\mathcal{S}^\infty \mathbb{R}^n$ in the sense that $T^h_t \circ T^h_\tau(w_o)$ and $T^h_{t+\tau}(w_o)$ are indistinguishable.*

Proof. First let me make some initial comments and reductions. The theorem will be proved with $\mathbb{R}$ replaced by $J = [-1, 1]$. It should be clear to the reader that any other compact interval would work just as well. From existence on compact intervals and uniqueness, it is easy to conclude the existence of a solution w on all of $\mathbb{R}$. The fact that T^h_t is a flow on $\mathcal{S}^\infty \mathbb{R}^n$ is a direct consequence of uniqueness of solutions by the usual O.D.E. proof. Finally we have already seen that any solution w: $J \to \mathcal{S}^\infty \mathbb{R}^n$ to (5.5) is necessarily C^1 and $\dot{w}$ is B^p-Lipschitz, see Corollary 6.1, Corollary 6.2, and Proposition 6.3. So it suffices to consider only C^1-functions w: $J \to \mathcal{S}^\infty \mathbb{R}^n$. (It is readily verified by a scaling argument that $J = [-1, 1]$ in Corollary 6.1, Corollary 6.2, and Proposition 6.3 can be replaced by $[-T, T]$ provided h is replaced by $T \cdot h$ in all of the estimates.)

Let $C \equiv \sup_{u \in O(M)} |\Theta_u\langle \cdot, \cdot \rangle|$, and C_o be the constant in Corollary 6.1 defined in (6.21). Let X denote the set of C^1-functions w: $J \to \mathcal{S}^\infty \mathbb{R}^n$ $(w(t) = \int O(t)\, db + \int \alpha(t)\, ds)$ such that $w(0) = w_o$, $\|O(t)\|_{S^\infty} \leqslant \|O(0)\|_{S^\infty}\, e^{C\, |h|_\infty}$ and $\|w(t)\|_{B^\infty} \leqslant C_o$ for $t \in J$. By Corollary 6.1, Corollary 6.2, and Proposition 6.3, any solution to the differential equation (5.5) with initial conditions w_o must be in X. We may and do assume that versions for O and α have been chosen to be jointly continuous in (t, s), see Lemma 4.5. We now define a function $L: X \to X$ as follows. For $w(t) = \int O(t)\, db + \int \alpha(t)\, ds$, define O_1 and α_1 as the solutions to the ordinary differential equations

$$\dot{O}_1(t) = C(w(t))\, O_1(t) \qquad \text{with} \quad O_1(0) = O_o \tag{6.23}$$

and

$$\dot{\alpha}_1(t) = C(w(t))\, \alpha_1(t) + R(w(t)) \qquad \text{with} \quad \alpha_1(0) = \alpha_o, \tag{6.24}$$

where we make the convention that versions of $I(w(t))$, $T(w(t))$, $C(w(t))$, and $R(w(t))$ have been chosen to be P-a.s. jointly continuous in (t, s). These equations may be solved for each fixed sample point and for each fixed s. Because of continuous dependence (of solutions to ordinary differential equations) on parameters it follows that $(t, s) \to (O_1(t, s), \alpha(t, s))$ is $C^{1,0}$ P-a.s. Set $L(w)(t) \equiv w_1(t) = \int O_1(t)\, db + \int \alpha_1(t)\, ds$. With no essential modification, it follows by the same arguments in Corollaries 6.1 and 6.2 that $w_1 = L(w)$ is back in X. In fact, even more is true. By the same proof as in Corollary 6.2, there is a constant $K = K_p$ independent of $w \in X$ such that $w_1 \equiv L(w)$ is B^p-Lipschitz with Lipschitz constant K—i.e., $\|L(w)(t) - L(w)(t')\|_{B^p} \leqslant K_p\, |t - t'|$ for $t, t' \in J$, and $w \in X$.

The key feature of L is that if $w: J \to \mathscr{S}^\infty \mathbb{R}^n$ is a solution to (6.5)–(6.6) with $w(0) = w_o$, then w is a fixed point for L in the sense that $P(\{L(w)(t) = w(t)$ for $t \in J\}) = 1$. To verify this statement, recall from the proof of Proposition 6.3 that if $w(t) = \int O(t)\, db + \int \alpha(t)\, ds$ solves (5.5) where $(O(t), \alpha(t))$ are chosen to be P-a.s. jointly continuous in (t, s), then $(O(t), \alpha(t))$ solves (6.5) and (6.6) in the pointwise sense. It is now clear that solving (6.23) and (6.24) will yield $(O_1(t), \alpha_1(t)) = (O(t), \alpha(t))$ P-a.s. We have used the fact that any two versions of $C(w(t))$ or $R(w(t))$ which are P-a.s. jointly continuous in (t, s) must be identically equal on a set of full P-measure.

The strategy of the proof is to choose a $w_0 \in X$, define $w_n = L^{(n)}(w_0)$, where $L^{(n)}$ means L composed with itself n-times, and then show that $w = \lim_{n \to \infty} w_n$ exists and solves the differential equation (5.5). To show uniqueness of the solution and existence of the limit w, we will show that $L^{(n)}$ is a contraction in B^p for some sufficiently large n. This will follow from the next claim:

Claim. *There is a constant $K = K(p)$, independent of w_1 and w_2 in X such that*

$$\|L(w_1)(t) - L(w_2)(t)\|_{B^p} \leqslant K \left| \int_0^t \|w_1(\tau) - w_2(\tau)\|_{B^p}\, d\tau \right|, \tag{6.25}$$

for all $t \in J$.

Proof of Claim. For simplicity assume that $t \geqslant 0$. Let $\bar{w}_i(t) \equiv L(w_i)(t) = \int \bar{O}_i(t)\, db + \int \bar{\alpha}_i(t)\, ds$ for $i = 1, 2$. By Lemma 6.1(ii) and the fact that $|T(w)|_\infty \leqslant C\, |h|_\infty$ for any Brownian semimartingale w, it follows that there is a constant K independent of w_1 and w_2 in X such that P-a.s.

$$|\bar{O}_1(t)(s) - \bar{O}_2(t)(s)| \leqslant K \int_0^t [\, |A(w_1(\tau)(s) - A(w_2(\tau))(s)| + |T(w_1(\tau))(s) - T(w_2(\tau))(s)|\,]\, d\tau.$$

Consequently letting K vary from place to place,

$$\begin{aligned} \|\bar{O}_1(t) - \bar{O}_2(t)\|_{S^p} &\leqslant K \int_0^t [\, \|A(w_1(\tau)) - A(w_2(\tau))\|_{S^p} + \|T(w_1(\tau)) - T(w_2(\tau))\|_{S^p}]\, d\tau \\ &\leqslant K \int_0^t \|w_1(\tau) - w_2(\tau)\|_{B^p}\, d\tau, \end{aligned} \tag{6.26}$$

where Proposition 6.2 was used in the last inequality. The last K in (6.26) now depends on p and C_o. Similarly one finds by Lemma 6.1 and Proposition 6.2 that

$$\|\bar{\alpha}_1(t) - \bar{\alpha}_2(t)\|_{S^p} \leqslant K \int_0^t \|w_1(\tau) - w_2(\tau)\|_{B^p}\, d\tau. \tag{6.27}$$

In the application of Lemma 6.1 we have used Remark 6.4 to guarantee the existence of a constant K_o independent of $w \in X$ such that P-a.s. $\sup_{t \in J} |R(w(t))|_\infty \leqslant K_o$. Clearly (6.25) is a consequence of (6.26) and (6.27) proving the claim.

Iterating (6.25) leads to

$$\|L^{(n)}(w_1)(t) - L^{(n)}(w_2)(t)\|_{B^p} \leqslant \gamma K^n\, |t|^n/n!, \tag{6.28}$$

where $\gamma \equiv \sup_{t \in J} \|w_2(t) - w_1(t)\|_{B^p} \leqslant 2C_o$. This immediately proves uniqueness of solutions. Indeed if w_1 and w_2 are two solutions, then since $w_i = L^{(n)}(w_i)$ for all n, it follows from (6.28) that $\|w_1(t) - w_2(t)\|_{B^p} \leqslant 2C_0 K^n\, |t|^n/n!$, which tends to 0 as $n \to \infty$. To prove existence, choose $w_0 \in X$, for example, take $w_0(t) \equiv w_o$ for all t. Define $w_n \equiv L^{(n)}(w_0)$, then (6.28) shows that

$$\begin{aligned} \|w_{n+1}(t) - w_n(t)\|_{B^p} &= \|L^{(n)}(w_1)(t) - L^{(n)}(w_0)(t)\|_{B^p} \\ &\leqslant 2C_o K^n\, |t|^n/n!. \end{aligned}$$

This last inequality shows that w_n is B^p-Cauchy uniformly in t, since $\sum_{n=0}^\infty 2C_o K^n\, |t|^n/n! < \infty$. Thus $w(t) \equiv B^p\text{-}\lim_{n \to \infty} w_n(t)$ exists uniformly in

t and is a B^p-continuous function. In fact, since each w_n is B^p-Lipschitz with the Lipschitz constant independent of n, w is also B^p-Lipschitz. It is also clear, by passing to a subsequence to get uniform in s almost sure convergence of $O_n(t)$ and $\alpha_n(t)$, that $\|w(t)\|_{B^\infty} \leqslant C_o$ for all $t \in J$.

To finish the proof it suffices by Proposition 6.3 to show $w: J \to \mathscr{S}^\infty \mathbb{R}^n$ is differentiable in the B^p-topologies and that $w(t) \equiv \int O(t)\, db + \int \alpha(t)\, ds$ satisfies (5.5). Or equivalently $O(t)$, and $\alpha(t)$ are S^p-differentiable and satisfy (6.5) and (6.6). For this fix a $p \geqslant 2$, the function $\tau \to C(w_n(\tau))\, O_{n+1}(\tau)$ is S^p-continuous and hence S^p-Riemannian integrable. Using the same argument as in Lemma 4.5 and the definition of L, it is easy to show that

$$O_{n+1}(t) - O_o = \int_0^t C(w_n(\tau))\, O_{n+1}(\tau)\, d\tau \qquad (P\text{-a.s.}), \tag{6.29}$$

where the integral is a Riemannian integral in S^p. We now estimate the difference between the right member of (6.29) and $\int_0^t C(w(\tau))\, O(\tau)\, d\tau$,

$$\begin{aligned}
&\left\| \int_0^t [C(w_n(\tau))\, O_{n+1}(\tau) - C(w(\tau))\, O(\tau)]\, d\tau \right\|_{S^p} \\
&\quad \leqslant \int_0^t \|C(w_n(\tau))\, O_{n+1}(\tau) - C(w(\tau))\, O(\tau)\|_{S^p}\, d\tau \\
&\quad \leqslant \int_0^t \|C(w(\tau))\|_{S^r}\, \|O_{n+1}(\tau) - O(\tau)\|_{S^{r'}}\, d\tau \\
&\qquad + C_o \int_0^t \|C(w_n(\tau)) - C(w(\tau))\|_{S^p}\, d\tau \\
&\quad \leqslant K \int_0^t \|w_{n+1}(\tau) - w(\tau)\|_{B^{r'}}\, d\tau,
\end{aligned}$$

where $1/p = 1/r + 1/r'$ and $K = K(C_o, \sup_{t \in J} \|C(w(t))\|_{B^p})$. Since $\|w_{n+1}(\tau) - w(\tau)\|_{B^{r'}} \to 0$ uniformly in τ as $n \to \infty$, this last inequality shows that the right-hand side of (6.29) converges to $\int_0^t C(w(\tau))\, O(\tau)\, d\tau$ in the S^p-norm. Since the left-hand side of (6.29) converges to $O(t) - O_o$ in the S^p norm, it follows that

$$O(t) = O_o + \int_0^t C(w(\tau))\, O(\tau)\, d\tau. \tag{6.30}$$

So by the fundamental theorem of calculus it follows that O is S^p differentiable with derivative $\dot{O}(t) = C(w(t))\, O(t)$. A similar argument shows that α is S^p-differentiable and that α satisfies (6.6). Q.E.D.

COROLLARY 6.3. *Let* $h: [0, 1] \to \mathbb{R}^n$ *be a* C^1*-function such that* $h(0) = 0$, *and suppose that* σ_o *is a* B^∞*-Brownian semimartingale with values in* M *such that* $\sigma_o(0) = o$. *Then there is a unique* C^1*-function* $\sigma: \mathbb{R} \to \mathcal{S}^\infty M$, *such that* (5.1) *holds* ($\dot{\sigma}(t) = H(\sigma(t))h$) *and* $\sigma(0) = \sigma_o$. *Furthermore, the function* $\bar{T}_t^h: \mathcal{S}^\infty M \to \mathcal{S}^\infty M$ *defined by* $\bar{T}_t^h(\sigma_o) = \sigma(t)$ *is a flow on* $\mathcal{S}^\infty M$.

Proof. Existence follows from Theorems 5.2 and 6.1. The uniqueness assertion follows from Theorems 5.1 and 6.1. The property that $\bar{T}_t^h$ is a flow again follows from uniqueness in the usual way. Q.E.D.

Remark. In Proposition 7.1 below it will be shown that any differentiable function $\sigma: \mathbb{R} \to \mathcal{S}^\infty M$ solving (5.1) is automatically C^1 with $\dot{\sigma}$ B^p-Lipschitz for all $p \in [2, \infty)$.

7. Existence and Uniqueness for the Geometric Flow Equations

The purpose of this section is to give a more "direct" proof of Corollary 6.3. This section may be skipped without loss of continuity. The reason for including this section is that the techniques used may be useful in the future.

The idea of the proof is to imbed M into $\mathbb{R}^N$ for some N, and use a standard Piccard iteration scheme to solve the equation

$$\dot{\sigma}(t) = H(\sigma(t))h \tag{7.1}$$

as an equation in $\mathbb{R}^N$. Then with the aid of Theorem 2.4, it will be shown that the solution $\sigma(t)$ found this way actually takes values in M.

In this section, it will be assumed that M is an imbedded submanifold of $\mathbb{R}^N$ for some N. We also suppose that (Y, g, π, Γ, P) has been chosen as in the proof of Theorem 3.2. Recall that Y is an open neighborhood in $\mathbb{R}^N$ containing M, $\pi: Y \to M$ is a C^∞-map such that $\pi|_M = id$, g is a metric on Y which extends the metric on M, Γ is a connection 1-form on $\mathbb{R}^N$ such that $\nabla = d + \Gamma$ on TM, and $P(x) = \pi'(x)$. (*Note well*: in this section π is a function on Y and not $O(M)$.) Also write ∇ for the covariant derivative on TY defined by $d + \Gamma$. We assume that (Y, g, π, Γ, P) has all the properties guaranteed by Theorem 2.3. It may further be assumed, by shrinking Y to a relatively compact subset of Y if necessary, that the function $\Gamma(x) \equiv \Gamma|_{T_x Y}$ for $x \in Y$ is bounded along with all of its derivatives and that the metric g is comparable to the Euclidean metric on $\mathbb{R}^N$. The metric g is said to be comparable with the Euclidean metric if there is a constant $\varepsilon > 0$ such that $\varepsilon^2 |v|^2 \leqslant g\langle v_x, v_x\rangle \leqslant \varepsilon^{-2} |v|^2/n$ for all $v \in \mathbb{R}^N$ and x in Y. Recall that v_x denotes the tangent vector $(d/dt)|_0 (x + tv)$.

For any semimartingale σ with values in Y starting at $o \in M$, let $u = H(\sigma)$

denote the semimartingale with values in End($\mathbb{R}^n$, $\mathbb{R}^N$) found by solving the stochastic differential equation

$$\delta u + \Gamma(\sigma)\langle \delta\sigma \rangle u = 0 \qquad \text{with} \quad u(0) = u_o. \tag{7.2}$$

If σ is a semimartingale in M, it follows from the proof of Theorem 3.2 that $H(\sigma)$ is the horizontal lift of σ to $O(M)$, written in non-intrinsic form. Moreover because Γ is g-compatible, the linear maps $u(s)\colon \mathbb{R}^n \to \mathbb{R}^N$ ($s \in [0, 1]$) solving (7.2) are isometries with respect to the Euclidean inner product on $\mathbb{R}^n$ and the inner product $g_{\sigma(s)} \equiv g|_{T_{\sigma(s)}Y}$ on $\mathbb{R}^N$, where σ is now any semimartingale in Y starting at $o \in M$. Therefore, because g is comparable to the usual metric on $\mathbb{R}^N$, it follows that

$$|u|^2 \equiv \sum_i |ue_i|^2 \leqslant \frac{1}{n} \varepsilon^{-2} \sum_i g_\sigma \langle ue_i, ue_i \rangle = \varepsilon^{-2},$$

where $\{e_i\}$ is the standard basis for $\mathbb{R}^n$. This shows that the semimartingale $u = H(\sigma)$ remains uniformly bounded independent of σ with the bound $|u| \leqslant \varepsilon^{-1}$.

LEMMA 7.1. *Let σ be any continuous semimartingale with values in Y starting at $o \in M$, then $P(\sigma)\, H(\sigma) = H(\pi \circ \sigma)$. (Recall that $P(\sigma) = \pi' \circ \sigma$, and this P does not denote the probability measure P.)*

Proof. Define $\bar{\sigma} = \pi \circ \sigma$, $u \equiv H(\sigma)$, and $\bar{u} = P(\sigma)u$. It suffices to show

$$\delta\bar{u} + \Gamma(\bar{\sigma})\langle \delta\bar{\sigma} \rangle \bar{u} = 0, \tag{7.3}$$

since $\bar{u}(0) = P(o)\, u_o = u_o$. By Itô's lemma,

$$\delta\bar{u} = P'(\sigma)\langle \delta\sigma \rangle u + P(\sigma)\, \delta u = dP\langle \delta\sigma \rangle u - P(\sigma)\, \Gamma(\sigma)\langle \delta\sigma \rangle u.$$

But by Theorem 2.3(iv), $dP - P\Gamma = -\pi^*\Gamma\langle \cdot \rangle P$, which combined with the last equation gives

$$\delta\bar{u} = -\pi^*\Gamma\langle \delta\sigma \rangle P(\sigma)\, u = -\Gamma\langle \pi_*\, \delta\sigma \rangle \bar{u} = -\Gamma\langle \delta\bar{\sigma} \rangle \bar{u}.$$

This last equation is the same as (7.3). Q.E.D.

LEMMA 7.2. *Suppose that $\sigma\colon \mathbb{R} \to \mathscr{S}^\infty Y$ $(\sigma(t) = o + \int O(t)\, db + \int \alpha(t)\, ds)$ is a differentiable function (in the B^p-topologies). Then $\sigma(t)$ is a solution to (7.1) iff $(O(t), \alpha(t))$ solves*

$$\dot{O}(t) = C(\sigma(t))\, O(t), \tag{7.4}$$

and

$$\dot{\alpha}(t) = C(\sigma(t))\, \alpha(t) + R(\sigma(t)), \tag{7.5}$$

where $C(\sigma(t))$ *and* $R(\sigma(t))$ *are defined as follows. For any* $\sigma = o + \int O\,db + \int \alpha\,ds$ *in* $\mathscr{S}^{\infty}Y$ *set*

$$C(\sigma) \equiv -\Gamma(\sigma)\langle\cdot\rangle\, H(\sigma)\, h \tag{7.6}$$

and

$$R(\sigma) \equiv \tfrac{1}{2}\sum_i \left[\Gamma(\sigma)\langle Oe_i\rangle\, \Gamma(\sigma)\langle Oe_i\rangle - \Gamma'(\sigma)\langle Oe_i, Oe_i\rangle\right] H(\sigma)h + H(\sigma)h'. \tag{7.7}$$

(*The derivatives in* (7.4) *and* (7.5) *are to be interpreted in the* S^p-*sense for all* $p \in [2, \infty)$.)

Proof. For notational simplicity, write $u(t)$ for $H(\sigma(t))$ and suppress t from the notation. Then if σ solves (7.1) one has

$$\begin{aligned} d\dot\sigma &= \dot O\,db + \dot\alpha\,ds = d(uh) = \delta u\cdot h + u\,\delta h = -\Gamma(\sigma)\langle\delta\sigma\rangle\, uh + uh'\,ds \\ &= -\Gamma(\sigma)\langle d\sigma\rangle\, uh - \tfrac{1}{2}\Gamma'(\sigma)\langle d\sigma, d\sigma\rangle\, uh \\ &\quad + \tfrac{1}{2}\Gamma(\sigma)\langle d\sigma\rangle\, \Gamma(\sigma)\langle d\sigma\rangle\, uh + uh'\,ds \\ &= -\Gamma(\sigma)\langle O\,db\rangle\, uh + \Big\{ -\Gamma(\sigma)\langle\alpha\rangle\, uh + \tfrac{1}{2}\sum_i \left[\Gamma(\sigma)\langle Oe_i\rangle\, \Gamma(\sigma)\langle Oe_i\rangle\right. \\ &\quad \left. - \tfrac{1}{2}\Gamma'(\sigma)\langle Oe_i, Oe_i\rangle\right] uh + uh' \Big\}\, ds. \\ &= C(\sigma)\, O\,db + \left[C(\sigma)\,\alpha + R(\sigma)\right] ds. \end{aligned}$$

Equating the db-terms and the ds-terms on both sides of this last equation yields (7.4) and (7.5). Q.E.D.

Equations (7.4) and (7.5) are the analogues of (6.5) and (6.6) of the last section. Equations (7.4) and (7.5) have the disadvantage of being non-intrinsic; however, they are analytically simpler than (6.5) and (6.6). This is because the coefficient $C(\sigma)$ is bounded independent of σ. Indeed, $|C(\sigma)| \leqslant M\varepsilon^{-1}\,|h|_{\infty}$, where M is a bound on Γ.

LEMMA 7.3. *Suppose that* σ *and* $\bar\sigma$ *are* $\mathscr{S}^{\infty}Y$-*Brownian semimartingales, and that* $p \in [2,\infty)$ *then there is a constant* κ *independent of* σ, h, *and* p *and constants* $K_p = K(p, \|\sigma - o\|_{B^\infty}, \|\bar\sigma - o\|_{B^\infty})$ *for* $p \in [2,\infty)$ *such that*

(i) $\|H(\sigma) - H(\bar\sigma)\|_{B^p} \leqslant K_p\, \|\sigma - \bar\sigma\|_{B^p}$ *for* $p \in [2, \infty)$; (7.8)

(ii) $\|H(\sigma) - u_o\|_{B^p} \leqslant \kappa[1 + \|\sigma - o\|^2_{B^p}]$; *and* (7.9)

(iii) $\|H(\sigma)h\|_{B^p} \leqslant \kappa\, |h'|_{\infty} \cdot [1 + \|\sigma - o\|^2_{B^p}]$. (7.10)

In (ii) *and* (iii) *above* $p = \infty$ *is permissible.*

Proof. Let $U(\sigma)=(\sigma, H(\sigma))$, then U solves the Statonovich differential equation $\delta U(\sigma)=F(U(\sigma))\,\delta\sigma$, where $F(x, u)a=(a, -\Gamma(x)\langle a\rangle u)$ for $x\in Y$, $a\in\mathbb{R}^N$, and $u\in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^N)$. Equation (7.8) is now seen to be a direct consequence of Corollary 4.1 applied to the equation for U. Notice that the only way compactness entered Theorem 4.1 and Corollary 4.1 was to guarantee that the vector field X and all of its derivatives were Lipschitz. But our Γ satisfies this condition and hence so does the function F.

To prove Eq. (7.9) first express (7.2) in Itô form,

$$du=-\Gamma(\sigma)\langle d\sigma\rangle\, u+\tfrac{1}{2}\sum_{i=1}^{n}\{(\Gamma(\sigma)\langle Oe_i\rangle)^2-\Gamma'(\sigma)\langle Oe_i, Oe_i\rangle\}\, ds,$$

where $\{e_i\}$ is the standard basis for $\mathbb{R}^n$. Use the boundedness of $u=H(\sigma)$, Γ, and Γ' along with Lemma 4.1 to conclude

$$\|u-u_o\|_{B^p}\leqslant C\,\|\sigma-o\|_{B^p}+C\,\|O\|_{S^p}^2,$$

where C is a constant independent of σ. Equation (7.9) is an easy consequence of this inequality, since $\|O\|_{S^p}\leqslant\|\sigma-o\|_{B^p}$ and $x(1+x)\leqslant$ const.$(1+x^2)$ for all $x\geqslant 0$.

To prove (7.10), write $u=H(\sigma)$ as $u_o+\int A\,db+\int \beta\,ds$ and compute

$$d(uh)=A\,db\,h+[\beta h+uh']\,ds,$$

from which it follows

$$\|uh\|_{B^p}\leqslant|h'|_\infty\,[\,\|H(\sigma)-u_o\|_{B^p}+\varepsilon^{-1}],$$

since $|H(\sigma)|_\infty\leqslant\varepsilon^{-1}$, $|h|_\infty\leqslant|h'|_\infty$. Equation (7.10) is an easy consequence of this last displayed equation and (7.9). Q.E.D.

The next proposition is the analogue of Proposition 6.3. Now that $C(\sigma)$ is bounded independent of σ, the B^∞-boundedness assumption that was used in Proposition 6.3 is no longer necessary. The following jazzed up version of Gronwall's inequality will be used in place of Lemma 6.1(i).

LEMMA 7.4. *Let $(V, \|\cdot\|)$ be a normed linear space. Assume that $f\colon(-a, a)\to V$ is a differentiable function and there are constants $\varepsilon\geqslant 0$ and $\kappa\geqslant 0$ such that $\|\dot f(t)\|\leqslant\kappa\,\|f(t)\|+\varepsilon$ for $t\in I\equiv(-a, a)$. Then for all $t\in I$*

$$\|f(t)\|\leqslant e^{\kappa|t|}[\,\|f(0)\|+\varepsilon\,|t|\,].$$

Proof. Because f is differentiable, f is continuous and so for each $T\in(0, a)$, $M_T\equiv\sup_{|t|\leqslant T}\|f(t)\|$ is finite. Let V^* denote the continuous dual of V and let $\|\cdot\|$ also denote the induced norm on V^*. Choose $\lambda\in V^*$ such that $\|\lambda\|=1$ and set $g(t)=\lambda(f(t))$. Then $g\colon I\to\mathbb{R}$ is differentiable with $\dot{g}(t) = \lambda(\dot{f}(t))$ and hence $|\dot{g}(t)| \leqslant \|\dot{f}(t)\| \leqslant M_T$ for $|t| \leqslant T$. Apply Theorem 8.21 of Rudin [Ru] to g to find

$$\lambda(f(t)) \equiv g(t) = g(0) + \int_0^t \dot{g}(\tau)\, d\tau.$$

Therefore,

$$\begin{aligned} |\lambda(f(t))| &\leqslant |g(0)| + \left| \int_0^t |\dot{g}(\tau)|\, d\tau \right| \leqslant \|f(0)\| + \left| \int_0^t \|\dot{f}(\tau)\|\, d\tau \right| \\ &\leqslant \|f(0)\| + \kappa \left| \int_0^t \|f(\tau)\|\, d\tau \right| + \varepsilon\, |t|, \end{aligned}$$

from which it follows, by taking the supremum over $\lambda \in V^*$ with $\|\lambda\| = 1$, that

$$\|f(t)\| \leqslant \|f(0)\| + \varepsilon\, |t| + \kappa \left| \int_0^t \|f(\tau)\|\, d\tau \right|.$$

(Notice that $t \to \|\dot{f}(t)\|$ is Borel measurable, since it is the pointwise limit of the Borel functions $F_n(t) = n\, \|f(t + 1/n) - f(t)\| \cdot 1_{\{t:\, |t| < a - 1/n\}}$.) The lemma now easily follows from Gronwall's inequality, Lemma 4.3. Q.E.D.

Proposition 7.1 (Regularity). *Suppose that* $\sigma\colon \mathbb{R} \to \mathscr{S}^\infty Y$ $(\sigma(t) = o + \int O(t)\, db + \int \alpha(t)\, ds)$ *is a differentiable function which satisfies* (7.1), *then* σ *is* C^1 *(in the* B^p*-topology) and* $t \to \dot{\sigma}(t)$ *is* B^p*-Lipschitz for all* $p \in [2, \infty)$. *Furthermore, it is possible to choose versions of* $O(t)$, $\alpha(t)$, $C(t) = C(\sigma(t))$, *and* $R(t) = R(\sigma(t))$ *such that* (7.4) *and* (7.5) *hold pointwise P-a.s.*

Proof. Recall that $|C(\sigma)| \leqslant \kappa\, |h|_\infty$, where κ depends only on g and Γ. So by (7.4), $\|\dot{O}(t)\|_{S^p} \leqslant \kappa\, |h|_\infty\, \|O(t)\|_{S^p}$ for all t and $p \in [2, \infty)$. Thus by Lemma 7.4, $\|O(t)\|_{S^p} \leqslant \|O(0)\|_{S^p}\, e^{\kappa t\, |h|_\infty}$ for all $t \in \mathbb{R}$ and $p \in [2, \infty)$. Letting p tend to infinity in this last estimate yields

$$\|O(t)\|_{S^\infty} \leqslant \|O(0)\|_{S^\infty}\, e^{\kappa t\, |h|_\infty}. \tag{7.11}$$

It is now easy to conclude from (7.7) that there exists a constant C_o independent of $\sigma(t)$ and h such that

$$\begin{aligned} \|R(\sigma(t))\|_{S^\infty} &\leqslant C_o\, |h'|_\infty\, [1 + \|O(t)\|^2_{S^\infty}] \\ &\leqslant C_o\, |h'|_\infty\, [1 + \|O(0)\|^2_{S^\infty}\, e^{2\kappa\, |t|\, |h|_\infty}]. \end{aligned} \tag{7.12}$$

Now apply Lemma 7.4 with $p < \infty$ using (7.5) and (7.12) to find, after letting $p \to \infty$, the estimate

$$\begin{aligned} \|\alpha(t)\|_{S^\infty} \leqslant e^{\kappa\, |t|\, |h|_\infty} \{ &\|\alpha(0)\|_{S^\infty} \\ &+ C_o\, |h'|_\infty\, |t|\, [1 + \|O(0)\|^2_{S^\infty}\, e^{2\kappa\, |t|\, |h|_\infty}] \}. \end{aligned} \tag{7.13}$$

From (7.11) and (7.13) one can show at the expense of increasing κ that

$$\|\sigma(t)-o\|_{B^\infty} \leqslant \kappa e^{\kappa\,|t|\,|h'|_\infty}\,[1+\|\sigma_o-o\|^2_{B^\infty}]. \tag{7.14}$$

In particular this shows that $\sup_{t\in J}\|\sigma(t)-o\|_{B^\infty}<\infty$ for any solution to (7.1), where J is any compact interval. For definiteness take $J=[-1,1]$.

Knowing that $Z\equiv\sup_{t\in J}\|\sigma(t)-o\|_{B^\infty}<\infty$ allows us to apply Lemma 7.3(i). By Lemma 7.3, Burkholder's inequality, and the boundedness of h', Γ, Γ', $H(\sigma)$ and $O(t)$, it can be shown that there is a constant $K_p=K(Z,p,|h'|_\infty)$ such that

$$\|C(\sigma(t))-C(\sigma(\tau))\|_{S^p}\leqslant K_p\,\|\sigma(t)-\sigma(\tau)\|_{B^p} \tag{7.15}$$

and

$$\|R(\sigma(t))-R(\sigma(\tau))\|_{S^p}\leqslant K_p\,\|\sigma(t)-\sigma(\tau)\|_{B^p} \tag{7.16}$$

for t, $\tau\in J$. (See also Lemma 4.6.) Since σ is B^p-differentiable, (7.15) and (7.16) show that $C(t)\equiv C(\sigma(t))$ and $R(t)\equiv R(\sigma(t))$ are also B^p-continuous. The remainder of the proof may be completed using the same techniques as those in the proof of Proposition 6.3 with (6.5) and (6.6) replaced by (7.4) and (7.5). Q.E.D.

We are now ready to solve (7.1) by the method of Piccard iterates. For a C^1-function $\sigma\colon[-T,T]\to\mathscr{S}^\infty Y$, let $L(\sigma)\colon[-T,T]\to\mathscr{S}^\infty\mathbb{R}^N$ be the C^1-function defined by

$$L(\sigma)(t)=\sigma_o+\int_0^t H(\sigma(\tau))\,h\,d\tau, \tag{7.17}$$

where now all integrals are to be interpreted as Riemann integrals in B^p for all $p\in[2,\infty)$. The next lemma collects a number of estimates involving the function L.

LEMMA 7.5. *Let $T>0$, κ be as in Lemma 7.3, $p\in[2,\infty)$, $\beta\equiv\kappa\,|h'|_\infty$, and for $\sigma\colon[-T,T]\to\mathscr{S}^\infty Y$ set*

$$K_p(\sigma)\equiv\beta\cdot[1+\sup_{|t|\leqslant T}\|\sigma(t)-o\|^2_{B^p}]. \tag{7.18}$$

Suppose that σ, σ_1, and σ_2 are C^1-functions from $[-T,T]\to\mathscr{S}^\infty Y$, then L satisfies the following inequalities:

(i) $\|L(\sigma)(t)-\sigma_o\|_{S^p}\leqslant\varepsilon^{-1}\,|h|_\infty\,|t|$ ($p=\infty$ *is permissible*); (7.19)

(ii) $\|L(\sigma)(t)-L(\sigma)(\tau)\|_{B^p}\leqslant K_p(\sigma)\,|t-\tau|$ *for all* $t,\tau\in[-T,T]$; (7.20)

(iii) $$\|L(\sigma)(t)-\sigma_o\|_{B^p} \leqslant K_p(\sigma)T; \tag{7.21}$$

(iv) $$\|L(\sigma_1)(t)-L(\sigma_2)(t)\|_{B^p} \leqslant C_p \left| \int_0^t \|\sigma_1(\tau)-\sigma_2(\tau)\|_{B^p}\, d\tau. \right|, \tag{7.22}$$

where $C_p = C_p(\sup_{|t| \leqslant T} \|\sigma_1(t)-o\|_{B^\infty}, \sup_{|t| \leqslant T} \|\sigma_2(t)-o\|_{B^\infty})$.

Proof. Since the B^p-norm dominates the S^p-norm for $p<\infty$, it is permissible to estimate $\|L(\sigma)(t)-\sigma_o\|_{S^p}$ by $|\int_0^t \|H(\sigma(\tau))\,h\|_{S^p}\, d\tau|$. This immediately implies (7.19) for $p<\infty$, since $\|H(\sigma(\tau))\,h\|_{S^p} \leqslant \varepsilon^{-1}\,|h|_\infty$. So by passing to the limit $p \to \infty$, (7.19) holds for $p=\infty$. The second estimate follows from (7.10), the definition of K_p in (7.18), and the estimate

$$\|L(\sigma)(t)-L(\sigma)(\tau)\|_{B^p} \leqslant \left| \int_\tau^t \|H(\sigma(u))\,h\|_{B^p}\, du \right|.$$

The third estimate is a consequence of (7.10), the definition of K_p in (7.18) and the inequality

$$\|L(\sigma)(t)-\sigma_o\|_{B^p} \leqslant \left| \int_0^t \|H(\sigma(\tau))\,h\|_{B^p}\, d\tau \right|.$$

For (iv), one estimates

$$\begin{aligned}
\|L(\sigma_1)(t)-L(\sigma_2)(t)\|_{B^p} &\leqslant \left| \int_0^t \|\,[H(\sigma_1(\tau))-H(\sigma_2(\tau))]\,h\|_{B^p}\, d\tau \right| \\
&\leqslant 2c_p\,|h'|_\infty \left| \int_0^t \|H(\sigma_1(\tau))-H(\sigma_2(\tau))\|_{B^p}\, d\tau \right| \\
&\leqslant C_p \left| \int_0^t \|\sigma_1(\tau)-\sigma_2(\tau)\|_{B^p}\, d\tau \right|,
\end{aligned} \tag{7.23}$$

where the second inequality is a consequence of Lemma 4.1(vii), and the last inequality is a consequence of (7.8). Q.E.D.

PROPOSITION 7.2 (Local Existence). *Let* $h\colon [0,1] \to \mathbb{R}^n$ *be a* C^1-*function such that* $h(0)=0$, σ_o *be a Brownian semimartingale in* M *such that* $\sigma_o(0)=o$, *and* $\|\sigma_o - o\|_{B^\infty} < \infty$. *Then there is a constant* $T=T(\sigma_o, h)$ *such that there exists a* C^1-*function* $\sigma\colon [-T,T] \to \mathscr{S}^\infty Y$ *solving* (7.1) *with* $\sigma(0)=\sigma_o$.

Proof. For $T>0$ let X_T be the set of C^1-functions $\sigma\colon [-T,T] \to \mathscr{S}^\infty Y$ such that $\sigma(0)=\sigma_o$, and $\sup_{|t| \leqslant T} \|\sigma(t)-\sigma_o\|_{B^\infty} \leqslant 1$. By (7.19) with $p=\infty$ and the compactness of M we see that $L(\sigma)(t) \in \mathscr{S}^\infty Y$ provided that

$|t| \leqslant \varepsilon\rho/|h|_\infty$, where ρ is the distance of M to the compliment of Y. By the triangle inequality $\|\sigma(t) - o\|_{B^\infty} \leqslant 1 + \|\sigma_o - o\|_{B^\infty}$ for all $\sigma \in X_T$. Hence for

$$K \equiv \beta[1 + (1 + \|\sigma_o - o\|_{B^\infty})^2],$$

then $\sup_{\sigma \in X_T} K_\infty(\sigma) \leqslant K < \infty$, where β and $K_\infty(\sigma)$ are defined in Lemma 7.5. Consequently by (7.21), $\|L(\sigma)(t) - \sigma_o\|_{B^\infty} \leqslant K|t|$ and hence $\|L(\sigma)(t) - \sigma_o\|_{B^\infty} \leqslant 1$ provided $|t| < 1/K$. As a result of these comments L maps X_T back into X_T provided T is less than $1/K$ and $\varepsilon\rho/|h|_\infty$. Also notice that the constant C_p which occurs in (7.22) may be chosen to be independent of σ_1 and $\sigma_2 \in X_T$.

To summarize the above paragraph, for T sufficiently small L maps X_T back into X_T and L satisfies, for all $p \in [2, \infty)$, (7.20) and (7.22) with constants C_p and K_p which can be chosen to be independent of σ, σ_1, and σ_2 in X_T.

We now fix a $T < \min\{1/K, \varepsilon\rho/|h|_\infty\}$. Define $\sigma_0(t) = \sigma_o$ for $t \in [-T, T]$, so that $\sigma_0 \in X_T$ and let $\sigma_n(t) = L^{(n)}(\sigma_0)(t)$, where $L^{(n)}$ is the nth iterate of L. Because of (7.22), it follows (as in the proof of Theorem 6.1) that σ_n converges uniformly in the B^p-norms for $p \in [2, \infty)$ to a Lipschitz function $\sigma\colon [-T, T] \to \mathscr{S}^\infty Y$. It is also clear from (7.22) that $L(\sigma) = \sigma$, i.e.,

$$\sigma(t) = \sigma_o + \int_0^t H(\sigma(\tau))\, h\, d\tau.$$

But this last equation shows $\sigma\colon [-T, T] \to \mathscr{S}^\infty Y$ is a C^1-solution to (7.1). Q.E.D.

It is now easy to give the second proof of Corollary 6.3.

Proof of Corollary 6.3. First we start with the uniqueness assertion of Corollary 6.3. Suppose that σ and $\hat{\sigma}\colon [-T, T] \to \mathscr{S}^\infty Y$ are two necessarily C^1-solutions of (7.1) with initial condition σ_o. According to Lemma 7.4, there is a constant C_p depending on σ_1 and σ_2 such that (7.22) holds. Now using $L(\sigma_1) = \sigma_1$ and $L(\sigma_2) = \sigma_2$, iteration of the inequality (7.22) shows $\sigma_1 = \sigma_2$ just as in the proof of uniqueness in Theorem 6.1. (See the argument starting at Eq. (6.28).)

Suppose that T is as in Proposition 7.2, so there exists a necessarily unique C^1-solution $\sigma\colon [-T, T] \to \mathscr{S}^\infty Y$ to (7.1) with values in Y. By Proposition 7.1, $\dot{\sigma}$ is B^p-Lipschitz for all $p \in [2, \infty)$. By Lemma 4.8, $\bar{\sigma} \equiv \pi \circ \sigma\colon [-T, T] \to \mathscr{S}^\infty M$ is a C^1-function for which $\dot{\bar{\sigma}}$ is also B^p-Lipschitz for $p \in [2, \infty)$. (Recall in this section π maps Y to M and not $O(M)$ to M.) We will now see that $\bar{\sigma}$ also satisfies (7.1). Since, $\bar{\sigma}$ is differentiable, it suffices to identify the derivative of $\bar{\sigma}$ with $H(\bar{\sigma})h$. With the aid of Kolmogorov's Lemma we may assume that versions of σ, $\bar{\sigma}$, $H(\sigma)$, and

$H(\bar{\sigma})$ have been chosen to be P-a.s. $C^{1,0}$ as functions of (t, s). For these versions we have

$$\dot{\bar{\sigma}}(t) = P(\sigma(t))\,\dot{\sigma}(t) = P(\sigma(t))\,H(\sigma(t))h = H(\bar{\sigma}(t))h,$$

where the last equality is a consequence of Lemma 7.1. Thus $\bar{\sigma}$ is also a solution to (7.1) with initial condition $\bar{\sigma}(0) = \pi \circ \sigma(0) = \pi \circ \sigma_o = \sigma_o$, where $\pi \circ \sigma_o = \sigma_o$ because σ_o is in M and $\pi|_M = id$ on M. (In fact by uniqueness, $\sigma = \bar{\sigma}$.) So we have shown for $T = T(\sigma_o, h) > 0$, there exists a unique solution $\sigma\colon [-T, T] \to \mathscr{S}^\infty M$ solving (7.1) and $\sigma(0) = \sigma_o$.

Up to now only intervals $[-T, T]$ centered about $t = 0$ have been considered, but it is clear because Eq. (7.1) is autonomous that we may equally well center the interval about any other t_o in $\mathbb{R}$. So as is standard in ordinary differential equations, it is possible to construct (using uniqueness and local existence) a "maximal" solution (σ) to (7.1) with $\sigma(0) = \sigma_o$. By standard arguments, this maximal solution (σ) will be defined on all of $\mathbb{R}$ provided σ does not blow up in finite time. But (7.14) clearly rules out any finite time blow up. Q.E.D.

8. Quasi-invariance of the Flow

One purpose of this section is to show the solution to (5.1) with initial condition $\sigma(0) = \sigma_o$ equal to a Brownian motion has the "quasi-invariance" property, i.e., the law of $\sigma(t)$ remains equivalent to the law of σ_o. Recall two measures P and Q are said to be equivalent if they are mutually absolutely continuous with respect to one another. The other purpose of this section is to use (5.1) to construct a flow on $W(\mathbb{R}^n)$ rather than a flow on the space of Brownian semimartingales $\mathscr{S}^\infty \mathbb{R}^n$. As mentioned in the Introduction, these two issues are closely related. In order to solve both of these problems it is necessary to assume that covariant derivative (∇) has the following skew symmetry condition.

Definition 8.1. The covariant derivative ∇ is said to be *torsion skew symmetric* (TSS) if for each $m \in M$ and $v \in T_m M$ the map $(w \to T\langle v, w\rangle)\colon T_m M \to T_m M$ is skew symmetric with respect to the metric $g_m \equiv g|_{T_m M}$. An equivalent condition is that for each frame $u \in O(M)$ and $a \in \mathbb{R}^n$, the map $(b \to \Theta_u\langle a, b\rangle)\colon \mathbb{R}^n \to \mathbb{R}^n$ should be a skew-symmetric linear transformation.

Example 8.1. (a) If ∇ is the Levi–Civita connection, then ∇ is TSS because the torsion is zero.

(b) Suppose G^n is a compact Lie group with Lie algebra $\mathfrak{g}$, and choose an Ad_G-invariant inner product $((\cdot, \cdot))$ on $\mathfrak{g}$. Define a metric (η) on

G by the formula $\eta\langle L_{g*}a, L_{g*}b\rangle \equiv (a, b)$, where $a, b \in \mathfrak{g}$, $g \in G$, and L_g is left multiplication by $g \in G$. Clearly $L_g: G \to G$ acts isometrically, and because $(\cdot, \cdot)$ is Ad-invariant, it also follows that the right multiplications $\{R_g\}_{g \in G}$ act isometrically. If ∇ is the left covariant derivative on TG (see Example 5.1(b)) then ∇ is TSS. The reason is that upon identifying $\mathbb{R}^n$ with $\mathfrak{g}$ (so that a frame u at g is now an isometry from $\mathfrak{g}$ to T_gG) one finds for $u \in O_g(G)$ and $a, b \in \mathfrak{g}$ that $\Theta_u\langle a, b\rangle = -O^{-1}[Oa, Ob]$, where O is the orthogonal transformation on $\mathfrak{g}$ such that $u = L_{g*}O$. Therefore, $\Theta_u\langle a, \cdot\rangle = -O^{-1}\, ad_{Oa}O$ which is skew-adjoint on $\mathfrak{g}$ for all $a \in \mathfrak{g}$ because of the Ad_G-invariance of $(\cdot, \cdot)$.

(c) Keeping the same notation as above but now take ∇ to be the right covariant derivative on TG (See Example 5.1(c)), then ∇ is TSS. Here $\Theta_u\langle a, \cdot\rangle = O^{-1}\,\mathrm{ad}_{Oa}O$, where now O is the orthogonal transformation such that $u = R_{g*}O$. (The change in sign comes from the definition of the Lie-Bracket in terms of left invariant vector fields rather than right invariant vector fields. The function $(g \to g^{-1}): G \to G$ transforms right invariant vector fields to left invariant vector fields.)

(d) Suppose that (G, η) is as in part (b) and that H is a closed subgroup of G. Let $M = G/H = \{gH: g \in G\}$ be the homogeneous space of right cosets. For $g \in G$ set $\bar{g} \equiv gH$, and let $p: G \to M$ be the canonical projection $p(g) = \bar{g}$. Then $p: G \to M$ is a principal bundle with structure group H. Let $\mathfrak{h}$ be the Lie algebra of H, and $\mathfrak{h}^\perp$ be the orthogonal compliment of $\mathfrak{h}$ in $\mathfrak{g}$ relative to $(\cdot, \cdot)$. Then for each $g \in G$, the map $\rho(g) \equiv p_* L_{g*}|_{\mathfrak{h}^\perp}: \mathfrak{h}^\perp \to T_{\bar{g}}M$ is an isomorphism. M can be made into a Riemannian manifold by requiring $\rho(g)$ to be an isometry for each $g \in G$. (See [KN, pp. 154–155].) Similar to the Lie group case, it is convenient to identify $O_{\bar{g}}(M)$ with the set of isometries u from $\mathfrak{h}^\perp$ to $T_{\bar{g}}M$. Given the above data, there is a natural connection $\bar{\omega}$ on $p: G \to M$ defined by $\bar{\omega}\langle \xi_g\rangle = (L_{g*}^{-1}\xi_g)_{\mathfrak{h}}$, where $a_{\mathfrak{h}}$ denotes the orthogonal projection of $a \in \mathfrak{g}$ onto $\mathfrak{h}$. The $\mathfrak{h}$-valued 1-form $(\bar{\omega})$ is easily seen to be a connection 1-form using the fact that Ad_h leaves both $\mathfrak{h}$ and $\mathfrak{h}^\perp$ invariant for all $h \in H$. Let $O(\mathfrak{h}^\perp)$ denote the orthogonal transformation on $\mathfrak{h}^\perp$ and $\mathrm{Ad}: G \to O(\mathfrak{g})$ denote the adjoint representation of G. Then $\mathrm{Ad}(H)$ leaves $\mathfrak{h}^\perp$ invariant, so $\mathrm{Ad}: H \to O(\mathfrak{h}^\perp)$ defined by $\mathrm{Ad}(h)|_{\mathfrak{h}^\perp}$ is an orthogonal representation of H. Taking $O_g(M)$ to be the set of isometries (u) from $\mathfrak{h}^\perp$ to T_gM,' the map $\rho: G \to O(M)$ defined above is a principal bundle morphism covering the identity of M such that $\rho(gh) = \rho(g)\,\mathrm{Ad}(h)$.

The morphism ρ and connection $\bar{\omega}$ on $p: G \to M$ induce in a standard way a unique connection (ω) on $O(M)$, such that $\bar{\omega} = p^*\omega$, see [KN, Proposition 6.1, p. 79]. The connection (ω) has the property that a path $u(t)$ in $O(M)$ over $\sigma(t)$ in M is horizontal iff $u(t) = \rho(g(t))O$ where $O \in O(\mathfrak{h}^\perp)$ and $g(t)$ is an $\bar{\omega}$-horizontal path in G. The reader may now verify that for all $g \in G$ and $a, b \in \mathfrak{k}^{\perp}$ that $\Theta_{\rho(g)}\langle a, b\rangle = -[a, b]_{\mathfrak{k}^{\perp}}$, where $[a, b]_{\mathfrak{k}^{\perp}}$ denotes the orthogonal projection of $[a, b]$ onto $\mathfrak{k}^{\perp}$. It is now easily seen that $(b \to \Theta_{\rho(g)}\langle a, b\rangle): \mathfrak{k}^{\perp} \to \mathfrak{k}^{\perp}$ is skew symmetric.

The next proposition explores the relationship between a TSS covariant derivative (∇) and the Levi–Civita covariant derivative ($\bar{\nabla}$).

PROPOSITION 8.1. *Let (M, g) be a Riemannian manifold and suppose that ∇ is a metric compatible TSS covariant derivative on M. Let $\bar{\nabla}$ denote the Levi–Civita covariant derivative on M. Then for $v \in TM$*

(i) $\nabla_v = \bar{\nabla}_v + \frac{1}{2} T\langle v, \cdot\rangle$,
where T is the torsion tensor of ∇, and (8.1)

(ii) *the Laplacian constructed using ∇ is the same as the Laplacian constructed using the Levi–Civita covariant derivative $\bar{\nabla}$.*

Remark. This proposition shows that the notion of a ∇-Brownian motion and a $\bar{\nabla}$-Brownian motion agree. However, if $\nabla \neq \bar{\nabla}$, the horizontal lift operators H^{∇} and $H^{\bar{\nabla}}$ will be different. These observations seem to play a crucial role in Gross' paper on logarithmic Sobolev inequalities on loop groups [Gr4].

Proof. (i) Let X, Y, Z be in T_mM, and let $A\langle X\rangle$ be the operator on T_mM such that $\nabla_X = \bar{\nabla}_X + A\langle X\rangle$. Because $\bar{\nabla}$ is torsion free, it follows that $T\langle X, Y\rangle = A\langle X\rangle Y - A\langle Y\rangle X$. From the metric compatibility of ∇ and $\bar{\nabla}$ one learns that $A\langle X\rangle$ is skew adjoint on T_mM. Using these two properties it is easy to show

$$2g\langle A\langle X\rangle Y, Z\rangle = g\langle T\langle X, Y\rangle, Z\rangle - g\langle T\langle Y, Z\rangle, X\rangle + g\langle T\langle Z, X\rangle, Y\rangle. \tag{8.2}$$

To verify (8.2) just expand the right member of (8.2) in terms of the A's and simplify. Because of the skew symmetry properties, $T\langle Y, Z\rangle = -T\langle Z, Y\rangle$ and $Z \to T\langle Y, Z\rangle$ is skew symmetric, $0 = -g\langle T\langle Y, Z\rangle, X\rangle + g\langle T\langle Z, X\rangle, Y\rangle$. Therefore (8.2) simplifies to

$$2g\langle A\langle X\rangle Y, Z\rangle = g\langle T\langle X, Y\rangle, Z\rangle.$$

Since Z is arbitrary, this proves part (i).

(ii) Let $\{E_i\}_{i=1}^n$ be a local orthonormal frame near $m \in M$ and f be a C^∞-function on M. Then by definition the Laplacian (Δ) constructed from ∇,

$$\Delta f = \sum (\nabla\, df)\langle E_i, E_i\rangle = \sum (E_i^2 f - df\langle \nabla_{E_i} E_i\rangle).$$

By part (i), $\nabla_{E_i} E_i = \bar{\nabla}_{E_i} E_i + (1/2)\, T\langle E_i, E_i\rangle = \bar{\nabla}_{E_i} E_i$, so that the above displayed equation may be written as

$$\Delta f = \sum (E_i^2 f - df\langle \bar{\nabla}_{E_i} E_i\rangle),$$

which is the definition of the Levi–Civita Laplacian. Q.E.D.

The next two lemmas will enable us to determine when solutions to (5.1) have the quasi-invariance property. The second of the two lemmas is a corollary of Girsanov's theorem along with Novikov's criterion. This lemma will be used to prove quasi-invariance of the flow (5.1) when ∇ is TSS.

LEMMA 8.1. *Let $w = \int O\, db + \int \alpha\, ds$ be a Brownian semimartingale, where O is an $n \times n$ matrix valued continuous predictable process and α is an $\mathbb{R}^n$-valued predictable process. If the laws w and b are equivalent, then the process O is $O(n)$-valued P-a.s.*

Proof. Let $\mu_w = w_* P$ and $\mu = b_* P$ (Wiener measure) denote the laws of w and b, respectively, on $W(\mathbb{R}^n)$. Recall that if Q is a manifold and $\{X(s)\}_{s \in [0,1]}$ is a Q-valued continuous process, then X may also be viewed as a function from Ω to $W(Q)$ by setting $X(\omega)(s) = X(s, \omega)$. Also recall that $X_* P$ denotes the measure on $W(Q)$ such that $X_* P(f) = P(f \circ X)$ for all bounded measurable functions $f: W(Q) \to \mathbb{R}$.

Let $\xi_s: W(\mathbb{R}^n) \to \mathbb{R}^n$ denote the coordinate map $\xi_s(x) = x(s)$ where $x \in W(\mathbb{R}^n)$. (I will write ξ_s or $\xi(s)$ interchangeably.) Since $\{\xi_s\}_{s \in [0,1]}$ is a Brownian motion with respect to μ, the quadratic covariation process $[\xi^i, \xi^j](s) = \delta_{ij} s$ when computed relative to μ. Since μ_w is equivalent to μ, $[\xi^i, \xi^j](s) = \delta_{ij} s$ still holds relative to μ_w. From this it follows (see Corollary 8.1 below) that $[w^i, w^j](s) = \delta_{ij} s$ P-a.s. Now, $[w^i, w^j](s)$ can also be computed directly to find

$$[w^i, w^j](s) = \int_0^s [O(s')\, O^{\mathrm{tr}}(s')]_{ij}\, ds'.$$

Thus one finds that the P-a.s. $\int_0^s [O(s')\, O^{\mathrm{tr}}(s')]_{ij}\, ds' = \delta_{ij} s$. Taking the derivative of this last expression implies P-a.s. that $[O(s)\, O^{\mathrm{tr}}(s)]_{ij} = \delta_{ij}$. That is, $O(s)\, O^{\mathrm{tr}}(s) = I$. Q.E.D.

LEMMA 8.2 (Girsanov's Theorem). *Let $w = \int O\, db + \int \alpha\, ds$ be a Brownian semimartingale such that (O, α) is a predictable $O(n) \times \mathbb{R}^n$-valued process. Assume there is a non-random constant $C > 0$ such that $P(\int_0^1 |\alpha(s)|^2\, ds \leq C) = 1$, then $\mu = b_* P$ and $\mu_w = w_* P$ are equivalent.*

Proof. I will follow closely the proof in Protter [Pr, Theorem 21, p. 111]. First define the square integrable martingale $M_s \equiv \int_0^s \alpha \cdot O\, db$ and set

$$Z_s \equiv \exp(-M_s - \tfrac{1}{2}[M, M]_s) = \exp\left\{ -\int_0^s \alpha \cdot O\, db - \tfrac{1}{2} \int_0^s |\alpha|^2\, ds' \right\}.$$

It is standard and easy to verify that $dZ_s = -Z_s\, dM_s$, so that Z is a local martingale. Since $[M, M]_1 = \int_0^1 |\alpha(s)|^2\, ds \leqslant C$, $P[\exp(\frac{1}{2}[M, M]_1)] < e^{C/2} < \infty$. Therefore by Novikov's criterion (see [RY, Proposition 1.15, p. 308]), Z is actually a martingale and in particular $P(Z_s) = P(Z_0) = P(1) = 1$ for all s.

Define $Q = Z_1 \cdot P$, i.e., Q is the probability measure on Ω such that $dQ/dP = Z_1$. Since $Z_1 > 0$ P-a.s., the measures P and Q are equivalent. Let β be the P-martingale $\beta \equiv \int O\, db$, and notice that $[\beta^i, \beta^j](s) = \delta^{ij} s$. By Girsanov's theorem (see [Pr, Theorem 20, p. 109]), the process

$$\beta^i - \int Z^{-1} \cdot d[Z, \beta^i] = \beta^i + \int d[M, d\beta^i] = \beta^i + \int \alpha^i\, ds = w^i$$

is a Q-local martingale for each $i \in \{1, 2, ..., n\}$. Since the measure P and Q are equivalent, quadratic covariations computed with respect to P or Q give the same answer. Therefore $[w^i, w^j] = \delta_{ij} s$ relative to Q. We may now use Levy's theorem [Pr, Theorem 38, Chap. II] to conclude that w is a Q-Brownian motion.

We now know that $w_* Q = \mu = b_* P$. Thus given any bounded measurable function $f\colon W(\mathbb{R}^n) \to \mathbb{R}$, it follows that

$$P(f(b)) = Q(f(w)) \equiv P(Z_1 f(w)). \tag{8.3}$$

Suppose that $f\colon W(\mathbb{R}^n) \to \mathbb{R}$ is a non-negative measurable function. By (8.3) and the fact that $Z_1 > 0$ P-a.s. it is trivial to verify that the following statements are equivalent: (i) $\mu_w(f) = 0$, (ii) $P(f(w)) = 0$, (iii) $P(Z_1 f(w)) = 0$, (iv) $P(f(b)) = 0$, and (v) $\mu(f) = 0$. The equivalence of statements (i) and (v) above clearly implies that μ_w and μ are equivalent. Q.E.D.

Remark 8.1. It will be useful to note that $P(Z_s^r) < \infty$ for all $r \in \mathbb{R}$. To see this first notice that

$$Z_s^r = \exp\left\{ -r \int_0^s \alpha \cdot O\, db - \frac{r^2}{2} \int_0^s |\alpha|^2\, ds' \right\}$$
$$\times \exp\left\{ \frac{r^2 - r}{2} \int_0^s |\alpha|^2\, ds' \right\} \equiv U_s \cdot V_s. \tag{8.4}$$

By Novikov's criterion U_s is still a martingale and in particular $P(U_s)=1$. The second term (V_s) in (8.4) is bounded by $\exp(Cs\,|r^2-r|/2)$, and hence

$$P(Z_s^r) \leqslant \exp(Cs\,|r^2-r|/2) < \infty. \tag{8.5}$$

COROLLARY 8.1. *Keep the same assumptions and notation as in Lemma* 8.2. *Let $\rho = d\mu_w/d\mu$ be the Radon–Nikodym derivative of μ_w with respect to μ, and let $\mathscr{H}$ be the σ-field generated by the random variables $w(s)$ for s in* $[0, 1]$. *Then $1/\rho(w) = P(Z_1 \mid \mathscr{H})$. Furthermore for each $r \in \mathbb{R}$, ρ^r is μ-integrable.* (*Warning*: *the analogous formula in the proof of the Corollary on p.* 112 *of* [Pr] *is missing the above conditional expectation and a proper interpretation.*)

Proof. Suppose that f is a bounded measurable function on $W(\mathbb{R}^n)$. Then $\mu_w(f) = \mu(\rho f) = P(\rho(b) f(b)) = P(Z_1 \rho(w) f(w))$ by (8.3). Because $\rho(w) f(w)$ is $\mathscr{H}$-measurable, $\mu_w(f) = P(P(Z_1 \mid \mathscr{H})\, \rho(w) f(w))$. On the other hand by definition of μ_w, $\mu_w(f) = P(f(w))$. Hence $P(f(w)) = P(P(Z_1 \mid \mathscr{H})\, \rho(w) f(w))$, and this holds for all bounded measurable functions f: $W(\mathbb{R}^n) \to \mathbb{R}$. Thus $P(Z_1 \mid \mathscr{H})\, \rho(w) = 1$ P-a.s.

Because $\mu(\rho^r) = P(\rho^r(b)) = P(Z_1 \rho^r(w)) = P(P(Z_1 \mid \mathscr{H})\, \rho^r(w)) = P(P(Z_1 \mid \mathscr{H})^{r-1})$, the assertion that ρ^r is μ-integrable for all r is equivalent to the assertion that $P(Z_1 \mid \mathscr{H})^r$ is P-integrable for all real r. This last assertion follows from the following purely measure theoretic lemma and Remark 8.1. Q.E.D.

LEMMA 8.3. *Let $(\Omega, \mathscr{F}, P)$ be a probability space. Suppose that $\mathscr{H} \subset \mathscr{F}$ is a sub-sigma field of $\mathscr{F}$, and $Z\colon \Omega \to \mathbb{R}_+$ is an $\mathscr{F}$-measurable function for which Z^r is integrable for all $r \in \mathbb{R}$, then U^r is integrable for all $r \in \mathbb{R}$, where $U \equiv P(Z \mid \mathscr{H})$.*

Proof. If $r \geqslant 1$, then $U^r \leqslant P(Z^r \mid \mathscr{H})$ P-a.s. by Jensen's inequality, so that $P(U^r) \leqslant P(Z^r) < \infty$. It now follows from Holder's inequality that U^r is integrable for all $r \geqslant 0$.

Now suppose that $r > 0$, then

$$P(U^{-r}) = \int_0^\infty r u^{-(r+1)} P(U < u)\, du, \tag{$*$}$$

which is verified by the computation

$$\begin{aligned} P(U^{-r}) &= P\left(\int_0^{U^{-1}} r y^{r-1}\, dy\right) = P\left(\int_0^\infty 1_{\{y < U^{-1}\}} \cdot r y^{r-1}\, dy\right) \\ &= \int_0^\infty r y^{r-1} P(U^{-1} > y)\, dy = \int_0^\infty r y^{r-1} P(U < y^{-1})\, dy \\ &= \int_0^\infty r u^{-(r+1)} P(U < u)\, du. \end{aligned}$$

Since $u^{-(r+1)}$ is integrable for u near infinity and $P(U<u) \leqslant 1$, in order to show $P(U^{-r}) < \infty$ for all $r>0$ it suffices (by $(*)$) to show that for all $k>0$ there is a constant C_k such that $P(U<u) \leqslant C_k u^k$.

By Chebyshev's inequality for all $\delta>0$, $P(Z<\delta) = P(Z^{-1} > \delta^{-1}) \leqslant \delta^k c_k$ where $c_k = P(Z^{-k})$. Hence,

$$\begin{aligned} P(U<u) &= P(U<u, Z<\delta) + P(U<u, Z \geqslant \delta) \\ &\leqslant c_k \delta^k + P(P(Z \geqslant \delta \mid \mathscr{H}); U<u). \end{aligned}$$

Now again by Chebyshev's inequality,

$$P(Z \geqslant \delta \mid \mathscr{H}) \leqslant \delta^{-1} P(Z \mid \mathscr{H}) \equiv \delta^{-1} U \qquad P\text{-a.s.}$$

Combining the last two displayed equations yields

$$P(U<u) \leqslant c_k \delta^k + \delta^{-1} P(U; U<u) \leqslant c_k \delta^k + \delta^{-1} u P(U<u).$$

Now set $\delta = 2u$, and solve for $P(U<u)$ to find

$$P(U<u) \leqslant c_k 2^k u^k/(1-1/2) = c_k 2^{k+1} u^k \equiv C_k u^k.$$

This is the desired estimate and the lemma is proved. Q.E.D.

THEOREM 8.1. *Suppose that h: $[0,1] \to \mathbb{R}^n$ is a C^1-function such that $h(0)=0$. Also assume that the covariant derivative ∇ is torsion skew symmetric (TSS). Let $w_o = \int O_o\, db + \int \alpha_o\, ds$ be an $\mathbb{R}^n$-valued Brownian semimartingale such that O_o is an $O(n)$-value process and $\|\alpha_o\|_{S^\infty} < \infty$. Let w: $\mathbb{R} \to \mathscr{S}^\infty \mathbb{R}^n$ be the solution to (5.5) given in Theorem 6.1. Then the law of $w(t)$ $(\mu_{w(t)} \equiv w(t)_* P)$ is equivalent to μ—Wiener measure on $W(\mathbb{R}^n)$. Furthermore, if ρ is the Radon–Nikodym derivative $\rho \equiv d(w(t)_* P)/d\mu$, then ρ^r is μ-integrable for all $r \in \mathbb{R}$.*

Proof. We may restrict t to a compact interval which for definiteness is taken to be $J=[-1,1]$. Write $w(t) = \int O(t)\, db + \int \alpha(t)\, ds$, then the pair (O, α) satisfies Eqs. (6.5)–(6.6) with $O(0) = O_o$ and $\alpha(0) = \alpha_o$. Because ∇ is TSS the process $T(w(t)) = \Theta_{u(t)}\langle h, \cdot \rangle$ (where $u(t) = I(w(t))$, see Eq. (5.7)) is $so(n)$ valued. Therefore the process $C(w(t)) \equiv A(w(t)) + T(w(t))$ is $so(n)$-valued because $A(w(t))$ is always skew symmetric. Because of Proposition 6.3 and Corollary 6.1, we may find a version of (O, α) such that the solution (O) to (6.5) with $O(0) = O_o$ is an $O(n)$-valued process and $|\alpha(t)|_\infty$ is bounded by a non-random constant C independent of $t \in J$. More precisely, by an application of Lemma 6.1 to (6.6) using the estimates (which are easily derived from (6.7) and (5.8))

$$|R(w(t))(s)| \leqslant C[\,|h(s)| + |h'(s)|\,],$$

and

$$|T(w(t))(s)| = |\Theta_{u(t)(s)}\langle h(s), \cdot\rangle| \leqslant C\,|h(s)|,$$

we find that $|\alpha(t)(s)| \leqslant Ce^{C\,|h(s)|} \cdot [\,|\alpha_o(s)| + |h(s)| + |h'(s)|\,]$ for some constant C. In particular this implies there is a constant C' such that

$$\int_0^s |\alpha(t)(s')|^2\,ds' \leqslant C'e^{C\,|h|_\infty} \int_0^1 [\,|\alpha_o(s)|^2 + |h'(s')|^2]\,ds'. \qquad (*)$$

Lemma 8.2 may now be used to conclude that the laws of $w(t)$ and μ are equivalent for all t. Corollary 8.1 shows that ρ^r is μ-integrable for all $r \in \mathbb{R}$.
Q.E.D.

Up until now, Eqs. (5.1) and (5.5) have been used to produce a flow on the space of Brownian semimartingales. Now that we know that the flow to (5.5) in the space of Brownian semimartingales has the quasi-invariance property, it makes sense to try to consider (5.1) and (5.5) as flow equations on $W(M)$ and $W(\mathbb{R}^n)$, respectively. In order to do this, it is necessary to make a digression into the properties of stochastic integrals and differential equations as functionals on path spaces. The discussion of the existence of a flow on $W(M)$ or $W(\mathbb{R}^n)$ will begin just before Definition 8.3 below.

Notation 8.1. Let V be a finite dimensional manifold, $W(V) = C([0, 1], V)$, and for $0 \leqslant s \leqslant 1$ let $\mathscr{H}_s(V)$ be the σ-algebra on $W(V)$ generated by the coordinate functions $\{\xi_{s'} : 0 \leqslant s' \leqslant s\}$, where $\xi_s(\omega) = \omega(s) \in V$ for all $\omega \in W(V)$. For $s \geqslant 1$ set $\mathscr{H}_s(V) \equiv \mathscr{H}_1(V)$—$\mathscr{H}_1(V)$ will also be denoted simply by $\mathscr{H}(V)$. If Q is a measure on $\mathscr{H}(V)$, let $\bar{\mathscr{H}}^Q$ denote the completion of $\mathscr{H}(V)$ with respect to Q. The extension of Q to $\bar{\mathscr{H}}^Q$ will still be called Q. Let $\mathscr{N}(Q) \equiv \{A \in \bar{\mathscr{H}}^Q : Q(A) = 0\}$ be the null sets of Q. The completion of the filtration $\{\mathscr{H}_s(V)\}$ with respect to Q is the filtration $\{\bar{\mathscr{H}}_s^Q\}$ where $\bar{\mathscr{H}}_s^Q \equiv \sigma(\mathscr{H}_s(V) \cup \mathscr{N}(Q))$—the σ-algebra generated by $\mathscr{H}_s(V)$ and all Q-negligible sets. Finally let $\bar{\mathscr{H}}_{s+}^Q \equiv \bigcap_{\varepsilon > 0} \bar{\mathscr{H}}_{s+\varepsilon}^Q$, so that $\{\bar{\mathscr{H}}_{s+}^Q\}_{s \geqslant 0}$ is a right continuous complete filtration (with respect to Q) on $W(V)$, i.e., $(W(V), \{\bar{\mathscr{H}}_{s+}^Q\}_{s \geqslant 0}, \bar{\mathscr{H}}^Q, Q)$ satisfies the usual hypothesis (see Section 3).

Remark 8.2. Clearly if Q' is another measure on $\mathscr{H}(V)$ which is equivalent to Q then $\bar{\mathscr{H}}_s^Q = \bar{\mathscr{H}}_s^{Q'}$ for all s.

Suppose that $(\Omega, \{\mathscr{F}_s\}, \mathscr{F}, P)$ is a filtered probability space satisfying the usual hypothesis and $\{X_s\}_{s \in [0,1]}$ is a continuous adapted V-valued process on Ω. (So we may view X as the function from $\Omega \to W(V)$ given by $\omega \in \Omega \to (s \to X_s(\omega)) \in W(V)$.) The condition that X_s is $\mathscr{F}_s$-adapted is equivalent to the function X being $\mathscr{F}_s/\mathscr{H}_s(V)$ measurable for all $s \in [0, 1]$. The proof of the next lemma is easy and is left to the reader.

LEMMA 8.4. *Assume the notation in the above paragraph. Let* $Q \equiv X_* P$ *be the law of* X *on* $W(V)$, *then* X *is* $\mathscr{F}_s/\bar{\mathscr{H}}^Q_{s+}$*-measurable for all* $s \in [0, 1]$.

Keep the same notation as in Notation 8.1 except now suppose that V is a finite dimensional vector space. Also assume that Q is a measure for which the coordinate functions $\{\xi_s\}_{s \in [0,1]}$ form a semimartingale on $(W(V), \{\bar{\mathscr{H}}^Q_{s+}\}_{s \geq 0}, \bar{\mathscr{H}}^Q, Q)$. Suppose that Z is another finite dimensional vector space and that $\{A_s\}$ is a $\mathrm{Hom}(V, Z)$-valued $\bar{\mathscr{H}}^Q_{s+}$-adapted continuous process on $W(V)$. Let φ be a fixed continuous version of $\int A\, d\xi$, so that φ may be viewed as a function from $W(V) \to W(Z)$ which is $\bar{\mathscr{H}}^Q_{s+}/\mathscr{H}_s(Z)$ measurable for all $s \in [0, 1]$. The next proposition is a special case of [RW, Lemma 10.1, p. 125] when Q is the standard Wiener measure.

PROPOSITION 8.2. *Assume the setup in the above paragraph—so* φ *is a fixed version of* $\int A\, d\xi$. *Suppose* $\{X_s\}_{s \in [0,1]}$ *is a* V*-valued semimartingale on a filtered probability space* $(\Omega, \{\mathscr{F}_s\}, P)$ *satisfying the usual hypothesis and* $X_* P$ *and* Q *are equivalent on* $\mathscr{H}(V)$. *Then* $\varphi \circ X$ *is a version of* $\int A(X)\, dX$ *or written more suggestively*: $(\int A(\xi)\, d\xi) \circ X = \int A(X)\, dX$.

Proof. First notice by Lemma 8.4 that X is $\mathscr{F}_s/\bar{\mathscr{H}}^Q_{s+}$ measurable for each s, so that $A(X)$ is $\mathscr{F}_s/\mathscr{H}_s(\mathrm{Hom}(V, Z))$ measurable for each s—i.e., $A(X)$ is an $\mathscr{F}_s$-adapted continuous $\mathrm{Hom}(V, Z)$-valued process. Hence, the stochastic integral $\int A(X)\, dX$ is well defined. Similarly $\varphi \circ X\colon \Omega \to W(Z)$ is $\mathscr{F}_s/\mathscr{H}_s(Z)$ measurable for all s. In order to identify $\varphi \circ X$ with $\int A(X)\, dX$, we apply [Pr, Theorem 21, p. 57] to learn for each $\varepsilon > 0$ that

$$Q\left(\sup_s \left| \varphi_s - \sum_{i=1}^{K} A_{s_i}(\xi_{s \wedge s_{i+1}} - \xi_{s \wedge s_i}) \right| > \varepsilon\right) \to 0 \qquad \text{as} \quad K \to \infty, \tag{8.6}$$

where $s_i = s_i^K \equiv s \cdot i/K$ so that $0 = s_0 < s_1 < s_2 < \cdots < s_K = s$ is a partition of $[0, s]$ for each K. Since $X_* P$ is equivalent to Q, we may replace Q in (8.6) by $X_* P$, and use the fact that $\xi_s(X) = X_s$ to find

$$P\left(\sup_s \left| (\varphi \circ X)_s - \sum_{i=1}^{K} (A \circ X)_{s_i} (X_{s \wedge s_{i+1}} - X_{s \wedge s_i}) \right| > \varepsilon\right) \to 0 \qquad \text{as} \quad K \to \infty. \tag{8.7}$$

But again by [Pr, Theorem 21, p. 57] we know that

$$P\left(\sup_s \left| \int_0^s A(X)\, dX - \sum_{i=1}^{K} (A \circ X)_{s_i} (X_{s \wedge s_{i+1}} - X_{s \wedge s_i}) \right| > \varepsilon\right) \to 0 \qquad \text{as} \quad K \to \infty. \tag{8.8}$$

The proposition now follows from (8.7) and (8.8). Q.E.D.

COROLLARY 8.2. *Keep the same notation and assumptions as in Proposition* 8.2 *and for definiteness take* $V = \mathbb{R}^N$. *For* $i, j \in \{1, \ldots, N\}$ *let* ψ^{ij} *be a fixed version of* $[\xi^i, \xi^j]$. *So that* $\psi^{ij}\colon W(\mathbb{R}^N) \to W(\mathbb{R})$, *and each* ψ^{ij} *is* $\mathscr{H}^Q_{s+}/\mathscr{H}_s(\mathbb{R})$ *measurable for all* $s \in [0, 1]$. *Then* $\psi^{ij} \circ X$ *is a version of* $[X^i, X^j]$, *i.e.*, $[\xi^i, \xi^j] \circ X = [X^i, X^j]$.

Proof. By definition of the quadratic covariation ($[\xi^i, \xi^j]$), ψ^{ij} is a version of $\xi^i\xi^j - \xi^i_0 \cdot \xi^j_0 - \int \xi^i \, d\xi^j - \int \xi^j \, d\xi^i$. Let φ^{ij} be a fixed version of $\int \xi^i \, d\xi^j$ for each i and j. Then ψ^{ij} is indistinguishable from $\kappa \equiv \xi^i\xi^j - \xi^i_0 \cdot \xi^j_0 - \varphi^{ij} - \varphi^{ji}$. With the aid of Proposition 8.2 we find that $\kappa \circ X$ is a version of $X^iX^j - X^i_0 \cdot X^j_0 - \int X^i \, dX^j - \int X^j \, dX^i$ which is a version of $[X^i, X^j]$. But since $X_* P$ is equivalent to Q it easily follows that $\kappa \circ X$ and $\psi^{ij} \circ X$ are P-indistinguishable. Therefore $\psi^{ij} \circ X$ is also a version of $[X^i, X^j]$. Q.E.D.

COROLLARY 8.3. *Keep the same assumptions and notation as in Corollary* 8.2. *Let* $\mathrm{Bi}(\mathbb{R}^N)$ *denote the set of bi-linear forms on* $\mathbb{R}^N$ *taking values in some fixed finite dimensional vector space. For each* $\mathscr{H}^Q_{s+}$*-adapted* $\mathrm{Bi}(\mathbb{R}^N)$*-valued process* (G) *and any* $\mathbb{R}^N$*-valued semimartingale* (Y), *set*

$$\int G(Y)\langle dY, dY\rangle \equiv \sum_{ij} \int G_{ij}(Y) \, d[Y^i, Y^j],$$

where $G_{ij}(\omega) = G(\omega)\langle e_i, e_j\rangle$ *with* $\{e_i\}$ *the standard basis for* $\mathbb{R}^N$ *and* $\omega \in W(\mathbb{R}^N)$. *Suppose that* ψ *is any fixed version of* $\int G(\xi)\langle d\xi, d\xi\rangle$. *Then* $\psi \circ X$ *is a version of* $\int G(X)\langle dX, dX\rangle$, *i.e.*, $(\int G(\xi)\langle d\xi, d\xi\rangle) \circ X = \int G(X)\langle dX, dX\rangle$. *Recall we are assuming that* $X_* P$ *is equivalent to* Q. (*Actually it does not matter that* G *is adapted for this lemma.*)

Proof. Let ψ^{ij} be fixed versions of $[\xi^i, \xi^j]$ for each i and j. Then ψ is Q-indistinguishable from $\sum_{ij} \int G_{ij} \, d\psi^{ij}$. We can now use the Riemann sum approximation argument as in Proposition 8.2 and the fact that $\psi^{ij} \circ X$ and $[X^i, X^j]$ are P-indistinguishable to conclude that $\psi \circ X$ is indeed a version of $\int G(X)\langle dX, dX\rangle$. (In this case the Riemann sums approximating $\int G_{ij}(\xi) \, d\psi^{ij}$ will converge uniformly Q-a.s.) Q.E.D.

COROLLARY 8.4. *Keep the same assumptions and notation as in Corollary* 8.2. *Now assume that* $\alpha = \sum_i \alpha_i \, dx^i$ *is a one-form on* $\mathbb{R}^N$, *and* φ *is any fixed version of* $\int \alpha\langle \delta\xi\rangle \equiv \sum_i \int \alpha_i(\xi) \, \delta\xi^i$. *Then* $\varphi \circ X$ *is a version of* $\int \alpha\langle \delta X\rangle = \sum_i \int \alpha_i(X) \, \delta X^i$, *i.e.*, $(\int \alpha\langle \delta\xi\rangle) \circ X = \int \alpha\langle \delta X\rangle$.

Proof. For any $\mathbb{R}^N$-valued semimartingale Y, by definition of the Stratonovich integral and Itô's lemma, $\int \alpha\langle \delta Y\rangle = \sum_i \int \alpha_i(Y) \, dY^i + \frac{1}{2}\sum_{i,j} \int (\partial_j \alpha_i)(Y) \, d[Y^i, Y^j]$, where $\partial_j = \partial/\partial x^j$. Therefore φ is a version of

$$\sum_i \int \alpha_i(\xi) \, d\xi^i + \tfrac{1}{2} \sum_{ij} \int \partial_j \alpha_i(\xi) \, d[\xi^i, \xi^j].$$

So it follows from Proposition 8.2 and Corollary 8.3 that $\varphi \circ X$ is a version of

$$\sum_i \int \alpha_i(X)\, dX^i + \tfrac{1}{2} \sum_{ij} \int \partial_j \alpha_i(X)\, d[X^i, X^j]$$

which is a version of $\int \alpha\langle \delta X\rangle$. Q.E.D.

The next proposition is a special case of [RW, Theorem 10.4, p. 126] (see also [IW, Theorem 3.1, p. 178]) when Q is the standard Wiener measure on $W(\mathbb{R}^N)$.

PROPOSITION 8.3. *Keep the same setup as in Proposition* 8.2 *with* $V = \mathbb{R}^N$ *for definiteness. Suppose that* $F\colon \mathbb{R}^N \times \mathbb{R}^K \to \operatorname{Hom}(\mathbb{R}^N, \mathbb{R}^K)$ *is a* C^∞*-function and that* $q\colon W(\mathbb{R}^N) \to W(\mathbb{R}^K)$ *is a* $\mathscr{H}^Q_{s+}/\mathscr{H}_s(\mathbb{R}^K)$ *measurable function for* $s \in [0, 1]$ *which solves the Stratonovich differential equation*

$$\delta q = F(\xi, q)\, \delta\xi \qquad \text{with} \quad q(0) = q_o, \tag{8.9}$$

where q_o *is a fixed point in* $\mathbb{R}^N$. *Then* $\bar{q} \equiv q \circ X$ *is an* $(\Omega, \{\mathscr{F}_s\}, P)$*-semimartingale solving the Stratonovich differential equation* $\delta\bar{q} = F(X, \bar{q})\, \delta X$ *with* $\bar{q}(0) = q_o$. (*Again by Lemma* 8.4 *all processes are appropriately adapted for the statement to make sense.*)

Proof. First notice that for any $\mathbb{R}^N$-valued semimartingale (Y), the Stratonovich differential equation $\delta\bar{q} = F(Y, \bar{q})\, \delta Y$ is equivalent to the Itô differential equation $d\bar{q} = F(Y, \bar{q})\, dY + G(Y, \bar{q})\langle dY, dY\rangle$, where $G(Y, \bar{q})\langle a, b\rangle \equiv \frac{1}{2}(d/dt)|_0\, [F(Y + ta, \bar{q} + tF(Y, \bar{q})\, a)\, b]$. So by assumption $q\colon W(\mathbb{R}^N) \to W(\mathbb{R}^K)$ satisfies

$$q = q_o + \int F(\xi, q)\, d\xi + \int G(\xi, q)\langle d\xi, d\xi\rangle.$$

By Proposition 8.2 and Corollary 8.3, $\bar{q} \equiv q \circ X$ is P-indistinguishable from

$$q_o + \int (F(\xi, q) \circ X)\, dX + \int (G(\xi, q) \circ X)\langle dX, dX\rangle.$$

Since $F(\xi, q) \circ X = F(X, \bar{q})$ and $G(\xi, q) \circ X = G(X, \bar{q})$, $\bar{q}$ satisfies

$$\bar{q} = q_o + \int F(X, \bar{q})\, dX + \int G(X, \bar{q})\langle dX, dX\rangle = q_o + \int F(X, \bar{q})\, \delta X. \qquad \text{Q.E.D.}$$

From now on let μ denote the standard Wiener measure on $W(\mathbb{R}^n)$, and ν denote the Wiener measure on $W(M)$ such that $\nu(W_o(M)) = 1$. Now for each $s \in [0, 1]$, let $b(s)$ $(\bar{\sigma}_o(s))$ denote the coordinate function on $W(\mathbb{R}^n)$

(on $W(M)$) given by $\omega \to \omega(s)$ for $\omega \in W(\mathbb{R}^n)$ $(\omega \in W(M))$. With these definitions, $(W(\mathbb{R}^n), \{\bar{\mathscr{H}}^{\mu}_{s+}\}, \mu)$ and $(W(M), \{\bar{\mathscr{H}}^{\nu}_{s+}\}, \nu)$ are filtered probability spaces which satisfy the usual hypothesis and $\bar{b}$ is a Brownian motion on the first and $\bar{\sigma}_o$ is a Brownian motion on the second. It is well known that $\bar{\mathscr{H}}^{\mu}_{s+} = \bar{\mathscr{H}}^{\mu}_{s}$, and $\bar{\mathscr{H}}^{\nu}_{s+} = \bar{\mathscr{H}}^{\nu}_{s}$, but we won't need this here. We also define $\gamma \equiv H(\bar{\sigma}_o)_* \nu = I(\bar{b})_* \mu$ a probability measure on the "interpolating" path space $W(O(M))$. (See Theorem 3.3 for the definitions of I and H.) Let $\bar{u}(s)$ denote the coordinate function $\bar{u}(s)(\omega) = \omega(s)$ on $W(O(M))$. Therefore $(W(O(M)), \{\bar{\mathscr{H}}^{\gamma}_{s+}\}, \gamma)$ is also a filtered probability space on which $\{\bar{u}(s)_{s \in [0,1]}$ is a semimartingale because of the following lemma.

LEMMA 8.5. *Let $(\Omega, \{\mathscr{F}_s\}, P)$ be a filtered probability space satisfying the usual hypothesis, and V be a finite dimensional manifold. Assume that $X_s\colon \Omega \to V$ is a V-valued semimartingale on $(\Omega, \{\mathscr{F}_s\}, P)$. Let $\gamma \equiv X_* P$ be the law of X on $W(M)$ and $\xi_s\colon W(V) \to V$ be the coordinate function $\xi_s(\omega) = \omega(s)$ for each $s \in [0, 1]$. Then $\{\xi_s\}$ is a V-valued semimartingale defined on $(W(V), \{\bar{\mathscr{H}}^{\gamma}_{s+}\}, \gamma)$.*

Proof. Let f be a C^∞-function on M. We must show that $f \circ \xi$ is a real semimartingale on $(W(V), \{\bar{\mathscr{H}}^{\gamma}_{s+}\}, \gamma)$. But this follows easily using the "good integrator" definition [Pr, Definition, p. 44] of a semimartingale and fact that $f \circ X$ is a semimartingale. The key points to note are: (i) (using the notation in [Pr, pp. 43–44]) for each simple $\{\bar{\mathscr{H}}^{\gamma}_{s+}\}$-predictable function H and $\varepsilon > 0$

$$\gamma(|I_{f \circ \xi}(H)| > \varepsilon) = P(|I_{f \circ X}(H \circ X)| > \varepsilon),$$

and (ii) by Lemma 8.4, $X\colon \Omega \to W(V)$ is $\mathscr{F}_s/\bar{\mathscr{H}}^{\gamma}_{s+}$ measurable for each s the process $H \circ X$ is $\{\mathscr{F}_s\}$ predictable. The reader can now easily finish the proof. Q.E.D.

Remark 8.3. One description of the measure ν is the law $([\pi \circ I(\bar{b})]_* \mu)$ of $\pi(I(\bar{b}))$ with respect to μ. Another description of ν is the measure concentrated on $W_o(M)$ such that $\{\bar{\sigma}_o(s)\}_{s \in [0,1]}$ is a Markov process with transition kernel given by the heat operator $e^{t\Delta/2}$. Here Δ is the Laplacian in Definition 3.7 which according to Proposition 8.1 is the same as the Levi–Civita Laplacian because ∇ is TSS in this section.

DEFINITION 8.2. Let $\bar{I}\colon W(\mathbb{R}^n) \to W(O(M))$ be a fixed version of $I(\bar{b})$, $\hat{I}\colon W(O(M)) \to W(\mathbb{R}^n)$ be a fixed version of the stochastic integral

$$\int \vartheta\langle \delta \bar{u} \rangle,$$

and $\bar{H}\colon W(M) \to W(O(M))$ be a fixed version of $H(\bar{\sigma}_o)$. (See Theorem 3.3 for the definition of I and H, and Definition 2.1 for the definition of the canonical 1-form ϑ.) Let $\Psi \equiv \pi \circ \bar{I}$ where $\pi: O(M) \to M$ is the canonical projection and let $\hat{\Psi} \equiv \hat{I} \circ \bar{H}$.

Remark 8.4. Because of Lemma 8.4, for each $s \in [0, 1]$, $\bar{I}$ is $\bar{\mathscr{H}}^{\mu}_{s+}/\bar{\mathscr{H}}^{\gamma}_{s+}$ measurable, $\hat{I}$ is $\bar{\mathscr{H}}^{\gamma}_{s+}/\bar{\mathscr{H}}^{\mu}_{s+}$ measurable, and $\bar{H}$ is $\bar{\mathscr{H}}^{\nu}_{s+}/\bar{\mathscr{H}}^{\gamma}_{s+}$ measurable. Consequently for each $s \in [0, 1]$, $\Psi: W(\mathbb{R}^n) \to W(M)$ is $\bar{\mathscr{H}}^{\mu}_{s+}/\bar{\mathscr{H}}^{\nu}_{s+}$-measurable and $\hat{\Psi}: W(M) \to W(\mathbb{R}^n)$ is $\bar{\mathscr{H}}^{\nu}_{s+}/\bar{\mathscr{H}}^{\mu}_{s+}$-measurable.

THEOREM 8.2. *Let $(\Omega, \{\mathscr{F}_s\}, P)$ be a filtered probability space satisfying the usual hypothesis.*

(i) *Suppose that $\{b(s): \Omega \to \mathbb{R}^n\}$ is an $\mathscr{F}_s$-semimartingale such that $b_* P$ is equivalent to μ, then $I(b)$ and $\bar{I} \circ b$ are P-indistinguishable.*

(ii) *Suppose that $\{\sigma_o(s): \Omega \to M\}$ is an $\mathscr{F}_s$-semimartingale such that $\sigma_{o*} P$ is equivalent to ν, then $H(\sigma_o)$ is P-indistinguishable from $\bar{H} \circ \sigma_o$.*

(iii) *Suppose that $\{u(s): \Omega \to O(M)\}$ is an $\mathscr{F}_s$-semimartingale such that $u_* P$ is equivalent to γ, then u is horizontal and $\hat{I} \circ u$ is P-indistinguishable from $I^{-1}(u)$.*

(iv) *$\hat{\Psi} \circ \Psi: W(\mathbb{R}^n) \to W(\mathbb{R}^n)$ is μ-indistinguishable from the identity map $(\bar{b})$ on $W(\mathbb{R}^n)$.*

(v) *$\Psi \circ \hat{\Psi}: W(M) \to W(M)$ is ν-indistinguishable from the identity map $(\bar{\sigma}_o)$ on $W(M)$.*

Proof. (i) By Theorem 3.1 we may consider $I(b)$ as the solution to a Stratonovich differential equation having the form in Proposition 8.3 where F depends only on q and not ξ. So (i) follows from Proposition 8.3.

(ii) In the proof of Theorem 3.2 it was shown that $H(\sigma_o)$ may be considered as a solution to a Stratonovich differential equation having the form in Proposition 8.3, see Eq. (3.6). So (ii) follows from Proposition 8.3, and the pathwise uniqueness of solutions to (3.6). (To apply Proposition 8.3 let $\xi = \sigma_o$, $q = u$, and $F(\sigma_o, u) = \Gamma(\sigma_o)\langle \cdot \rangle u$.)

(iii) Let $\alpha = \omega$ in Corollary 8.4, in which case we may take $\varphi \equiv 0$ as a version of $\int \omega\langle \delta\bar{u} \rangle$. So by Corollary 8.4 (see the remark below), $0 = \varphi \circ u$ is a version of $\int \omega\langle \delta u \rangle$, which shows that u is indeed horizontal. To finish the proof of (iii), apply Corollary 8.4 again, now with $\alpha = \vartheta$, and $\varphi = \hat{I}$.

Remark. Notice that M may be imbedded in $\mathbb{R}^N$ for some N which induces an imbedding of $O(M)$ into $\mathbb{R}^N \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$ as in Lemma 2.2. The form ω may be extended to a form $(\bar{\omega})$ on $T(\mathbb{R}^N \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N))$. We may now view u as an $\mathbb{R}^N \times \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$-valued process and by definition $\int \omega\langle \delta u \rangle = \int \bar{\omega}\langle \delta u \rangle$ for any $O(M)$-valued semimartingale u. It should now be clear to the reader that our first application of Corollary 8.4 in (iii) above was valid. Similar comments are also needed for the second application of Corollary 8.4 above. Such comments in the future will be left to the reader to fill in.

(iv) We will apply parts (i)–(iii) above with $(\Omega, \{\mathscr{F}_s\}, P) = (W(\mathbb{R}^n), \{\mathscr{H}^{\mu}_{s+}\}, \mu)$. By (i), $\bar{I} = \bar{I} \circ \bar{b}$ is μ-indistinguishable from $I(\bar{b})$, so that $\Psi = \pi \circ \bar{I}$ is μ-indistinguishable from $\pi \circ I(\bar{b})$. By (ii), $\bar{H} \circ \Psi$ is μ-indistinguishable from $H \circ \pi \circ I(\bar{b})$ which by Theorem 3.3 is μ-indistinguishable from $I(\bar{b})$. So $\hat{I} \circ \bar{H} \circ \Psi$ is μ-indistinguishable from $\hat{I} \circ I(\bar{b})$ which by (iii) is μ-indistinguishable from $I^{-1} \circ I(\bar{b})$. Because of Theorem 3.3, $I^{-1} \circ I(\bar{b})$ and $\bar{b}$ are μ-indistinguishable, and hence $\hat{\Psi} \circ \Psi = \hat{I} \circ \bar{H} \circ \Psi$ is μ-indistinguishable from $\bar{b}$. This proves (iv) because $\bar{b}$: $W(\mathbb{R}^n) \to W(\mathbb{R}^n)$ is the identity map.

(v) The proof of (v) is similar to (iv), except one now applies (i)–(iii) with $(\Omega, \{\mathscr{F}_s\}, P) = (W(M), \{\mathscr{H}^{\nu}_{s+}\}, \nu)$. Q.E.D.

We are now ready to discuss the existence of the flows on $W(M)$ and $W(\mathbb{R}^n)$ generated by (5.1) and (5.5), respectively. First some more notation. Let $\bar{w}(t) = \int \bar{O}(t)\, d\bar{b} + \int \bar{\alpha}(t)\, ds$ be the solution to (5.5) with $\bar{w}(0) = \bar{b}$ given in Theorem 6.1. (Here the underlying filtered probability space is $(W(\mathbb{R}^n), \{\mathscr{H}^{\mu}_{s+}\}, \mu)$ with reference Brownian motion $\{\bar{b}(s)\}$.) So $\bar{O}(t)$ and $\bar{\alpha}(t)$ solve equations (6.5) and (6.6) with $\bar{O}(0) = id \in O(n)$, and $\bar{\alpha}(0) = 0$. We can and do assume that μ-a.s. the function $(t, s) \to (\bar{w}(t)(s), \bar{O}(t)(s), \bar{\alpha}(t)(s))$ is $C^{1,0}$, see Proposition 6.3 and Lemma 4.5.

DEFINITION 8.3. Using the above notation, to each $t \in \mathbb{R}$ and to each C^1-function h: $[0, 1] \to \mathbb{R}^n$ such that $h(0) = 0$, define $S^h(t)$: $W(\mathbb{R}^n) \to W(\mathbb{R}^n)$, $O^h(t)$: $W(\mathbb{R}^n) \to W(O(n))$, and $\alpha^h(t)$: $W(\mathbb{R}^n) \to W(\mathbb{R}^n)$ by $S^h(t) = \bar{w}(t)$, $O^h(t) = \bar{O}(t)$, and $\alpha^h(t) = \bar{\alpha}(t)$.

Remark 8.5. Notice for each $s \in [0, 1]$ and $t \in \mathbb{R}$ that $S^h(t)$ and $\alpha^h(t)$ are $\mathscr{H}^{\mu}_{s+}/\mathscr{H}_s(\mathbb{R}^n)$ measurable, and that $O^h(t)$ is $\mathscr{H}^{\mu}_{s+}/\mathscr{H}_s(O(n))$ measurable. In fact, by Theorem 8.1, Remark 8.2, and Lemma 8.4 it follows that $S^h(t)$ is also $\mathscr{H}^{\mu}_{s+}/\mathscr{H}^{\mu}_{s+}$ measurable for each $s \in [0, 1]$, and in particular $S^h(t) \circ S^h(\tau)$ is still $\mathscr{H}^{\mu}_{s+}/\mathscr{H}^{\mu}_{s+}$ measurable for all $s \in [0, 1]$. (Recall that in this section ∇ is always assumed to be TSS.)

The next theorem shows that $S^h(t)$ is the "universal" solution to equation (5.5) when the initial Brownian semimartingale $w_o = \int O_o\, db + \int \alpha_o\, ds$ has the property that O_o is an orthogonal process and α_o is a bounded process.

THEOREM 8.3. *Suppose that h: $[0, 1] \to \mathbb{R}^n$ is a C^1-function such that $h(0) = 0$. Let $(\Omega, \{\mathscr{F}_s\}, \{b(s)\}, P)$ be a filtered probability space satisfying the usual conditions equipped with an $\mathbb{R}^n$-valued Brownian motion $b(s)$. Assume (as in Theorem 8.1) that $w_o = \int O_o\, db + \int \alpha_o\, ds$ is an $\mathbb{R}^n$-valued*

Brownian semimartingale such that O_o is an $O(n)$-valued process and $\|\alpha_o\|_{S^\infty} < \infty$. Let $w: \mathbb{R} \to \mathscr{S}^\infty \mathbb{R}^n$ be the solution to (5.5) given in Theorem 6.1. Then $w(t)$ is P-indistinguishable from $S^h(t) \circ w_o$.

Proof. First notice by Lemma 8.2 that $w(t)_* P$ and $S^h(t)_* \mu$ are equivalent to μ for each t. By Lemma 8.4, $w_o: \Omega \to W(\mathbb{R}^n)$ is $\mathscr{F}_s/\bar{\mathscr{H}}^\mu_{s+}$-measurable for all $s \in [0, 1]$ and therefore so is $S^h(t) \circ w_o$. By Proposition 8.3, $S^h(t) \circ w_o$ is P-indistinguishable from

$$\int O^h(t) \circ w_o \, dw_o + \int \alpha^h(t) \circ w_o \, ds$$

$$= \int O^h(t) \circ w_o \cdot O_o \, db + \int [O^h(t) \circ w_o \cdot \alpha_o + \alpha^h(t) \circ w_o] \, ds.$$

So

$$S^h(t) \circ w_o = \int O(t) \, db + \int \alpha(t) \, ds \qquad P\text{-a.s.},$$

where $O(t) \equiv O^h(t) \circ w_o \cdot O_o$, and $\alpha(t) \equiv O^h(t) \circ w_o \cdot \alpha_o + \alpha^h(t) \circ w_o$. Because $O^h(0) \equiv \mathrm{id} \in O(n)$, and $\alpha^h(0) \equiv 0 \in \mathbb{R}^n$, it follows that $O(0) = O_o$ and $\alpha(0) = \alpha_o$. So in order to show $w(t) = S^h(t) \circ w_o$, it suffices to show by the uniqueness assertion in Theorem 6.1 that $O(t)$ and $\alpha(t)$ are $S^p(P)$-continuously differentiable (for $p \in [2, \infty)$) solutions of (6.5) and (6.6).

I assert that $t \to O^h(t) \circ w_o$, and $t \to \alpha^h(t) \circ w_o$ are $S^p(P)$-continuously differentiable for all $p \geqslant 2$ with derivatives given by $\dot{O}^h(t) \circ w_o$, and $\dot{\alpha}^h(t) \circ w_o$. It is then clear that $O(t)$ and $\alpha(t)$ are also $S^p(P)$-continuously differentiable with

$$\dot{O}(t) \equiv \dot{O}^h(t) \circ w_o \cdot O_o \tag{8.10}$$

and

$$\dot{\alpha}(t) \equiv \dot{O}^h(t) \circ w_o \cdot \alpha_o + \dot{\alpha}^h(t) \circ w_o. \tag{8.11}$$

The first assertion is an easy consequence of the following claim and the fact that $t \to O^h(t)$ and $t \to \alpha^h(t)$ are $S^p(\mu)$ continuously differentiable functions for all $p \geqslant 2$.

Claim. If $X: W(\mathbb{R}^n) \to W(\mathbb{R}^N)$ is a $\bar{\mathscr{H}}^\mu_1/\mathscr{H}_1(\mathbb{R}^N)$-measurable process and $p \in [2, \infty)$, then for each $r \in (p, \infty)$, there is a constant $C = C(r, p)$ independent of X such that

$$\|X \circ w_o\|_{S^p(P)} \leqslant C \, \|X\|_{S^r(\mu)}.$$

To prove the claim, let $\rho \equiv d(w_{o*}P)/d\mu$ which exists by Girsanov's Theorem (Lemma 8.2). Then compute $\|X \circ w_o\|_{S^p(P)}$ using Holder's inequality,

$$\begin{aligned}\|X \circ w_o\|_{S^p(P)} &= \|X\|_{S^p(w_{o*}P)} = \|X\|_{S^p(\rho\mu)} \\ &= \|X^* \rho^{1/p}\|_{L^p(\mu)} \leqslant \|\rho^{1/p}\|_{L^{r'}(\mu)} \|X\|_{S^r(s)},\end{aligned}$$

where $1/r' = 1/p - 1/r$. Set $C \equiv \|\rho^{1/p}\|_{L^{r'}(\mu)}$ which is finite by Corollary 8.1. This proves the claim and the above assertions.

Set $\tilde{w}(t) \equiv S^h(t) \circ w_o = \int O(t)\, db + \int \alpha(t)\, ds$. The proof will be complete if we can show $O(t) = C(\tilde{w}(t))\, O(t)$ and $\dot{\alpha}(t) = C(\tilde{w}(t))\, \alpha(t) + R(\tilde{w}(t))$. To this end we first define two functions $\bar{C}: W(\mathbb{R}^n) \to W(so(n))$ and $\bar{R}: W(\mathbb{R}^n)$ which "implement" C and R. To motivate the definitions of these functions let $X = \int O\, db + \int \alpha\, ds$ be any $\mathbb{R}^n$-valued Brownian semimartingale such that O is an orthogonal process and α is bounded. Then $A(X)$ in (5.7) is given by

$$\begin{aligned}A(X) &= \int \Omega_{I(X)}\langle h, dX\rangle + \tfrac{1}{2} \int \bar{\Omega}_{I(X)}\langle O, h, O\rangle\, ds \\ &= \int \Omega_{I(X)}\langle h, dX\rangle + \tfrac{1}{2} \int \bar{\Omega}_{I(X)}\langle id, h, id\rangle\, ds,\end{aligned}$$

where id is the $n \times n$ identity matrix as follows, see Eq. (6.9), and Remark 6.2. (See Definition 6.2 for the definition of $\bar{\Omega}$.) Therefore, using Lemma 8.2 and Theorem 8.2(ii),

$$C(X) \equiv A(X) + T(X) = A(X) + \Theta_{I(X)}\langle h, \cdot\rangle$$

is given by

$$C(X) = \int \Omega_{I \circ X}\langle h, dX\rangle + \tfrac{1}{2} \int \bar{\Omega}_{I \circ X}\langle id, h, id\rangle\, ds + \Theta_{I \circ X}\langle h, \cdot\rangle.$$

With this as motivation define $\bar{C}$ as a fixed version of

$$\int \Omega_I\langle h, d\bar{b}\rangle + \tfrac{1}{2} \int \bar{\Omega}_I\langle id, h, id\rangle\, ds + \Theta_I\langle h, \cdot\rangle. \tag{8.12}$$

It then follows by Proposition 8.2 and Lemma 8.2 that $\bar{C} \circ X$ is indistinguishable from $C(X)$. Similarly, if we define (see (6.7) and Remark 6.2) $\bar{R}: W(\mathbb{R}^n) \to W(\mathbb{R}^n)$ by

$$\begin{aligned}\bar{R} &\equiv \tfrac{1}{2}\{\mathrm{Ric}_I\langle h, \mathrm{id}, \mathrm{id}\rangle + \bar{\Theta}_I\langle \mathrm{id}, h, \mathrm{id}\rangle\} + h' \\ &= \tfrac{1}{2}\{\mathrm{Ric}_I\langle h\rangle + \bar{\Theta}_I\langle \mathrm{id}, h, \mathrm{id}\rangle\} + h',\end{aligned} \tag{8.13}$$

then by Theorem 8.2(ii) and Lemma 8.2, $\bar{R} \circ X = R(X)$.

The rest of the proof is now a simple verification. By definition, we know that $\dot{O}^h(t) = \bar{C} \circ S^h(t)\, O^h(t)$ and hence

$$\dot{O}^h(t) \circ w_o = \bar{C} \circ S^h(t) \circ w_o \cdot O^h(t) \circ w_o = \bar{C} \circ \tilde{w}(t) \cdot O^h(t) \circ w_o. \tag{8.14}$$

From (8.14) and (8.10) we find

$$\dot{O}(t) \equiv \bar{C} \circ \tilde{w}(t) \cdot O(t) = C(\tilde{w})\, O(t),$$

as desired. Similarly by definition, $\dot{\alpha}^h(t) = \bar{C} \circ S^h(t) \cdot \alpha^h(t) + \bar{R} \circ S^h(t)$, and hence

$$\dot{\alpha}^h(t) \circ w_o = \bar{C} \circ \tilde{w}(t) \cdot \alpha^h(t) \circ \tilde{w}(t) + \bar{R} \circ \tilde{w}(t). \tag{8.15}$$

By insertion of (8.14) and (8.15) into (8.11) and using the definition of $\alpha(t)$ it follows that

$$\dot{\alpha}(t) = \bar{C} \circ \tilde{w}(t) \cdot \alpha(t) + \bar{R} \circ \tilde{w}(t) = C(\tilde{w}(t)) \cdot \alpha(t) + R(\tilde{w}(t)). \qquad \text{Q.E.D.}$$

We can now easily prove our first version of the Cameron–Martin type theorem. Because, as will be shown, $S^h(t)$ is a flow on $W(\mathbb{R}^n)$ it is possible to get a rather explicit Cameron–Martin type formula for the Radon–Nikodym derivatives $d(S^h(t)_* \mu)/d\mu$. Recall again that $\mathbf{V}$ is assumed to be TSS in this section.

THEOREM 8.4. *Suppose that $h\colon [0, 1] \to \mathbb{R}^n$ is a C^1-function such that $h(0) = 0$, and S^h, α^h, and O^h are as in Definition 8.3. Then S^h is a flow on $W(\mathbb{R}^n)$ which leaves the Wiener measure (μ) quasi-invariant. More explicitly,*

(i) *for all $t, \tau \in \mathbb{R}$, $S^h(t+\tau) = S^h(t) \circ S^h(\tau)$ μ-a.s., and*

(ii) *$d(S^h(t)_* \mu)/d\mu = Z(h, t)$, where*

$$Z(h, t) \equiv \exp\left\{ -\int_0^1 \alpha^h(-t) \cdot O^h(-t)\, d\bar{b} - \tfrac{1}{2} \int_0^1 |\alpha^h(-t)(s)|^2\, ds \right\}. \tag{8.16}$$

Proof. (i) Because of Theorem 8.1, we may apply Theorem 8.3 with $(\Omega, \{\mathscr{F}_s\}, P) = (W(\mathbb{R}^n), \{\mathscr{H}^{\mu}_{s+}\}, \mu)$ and $w_o \equiv S^h(\tau)$ to learn $w_t \equiv S^h(t) \circ S^h(\tau)$ solves (5.5) with $w(0) = S^h(\tau)$. But the function $\tilde{w}(t) \equiv S^h(t+\tau)$ also solves (5.5) with the same initial condition $(S^h(\tau))$, and hence by the uniqueness assertion of Theorem 6.1, $\tilde{w}(t) = w(t)$, i.e., $S^h(t+\tau) = S^h(\tau) \circ S^h(\tau)$.

(ii) From Eq. (8.3) of Lemma 8.2 with $P = \mu$, $b = \bar{b}$, and $w = S^h(t)$ it follows that

$$\mu(f) = \mu(f(\bar{b})) = \mu(Z(h, -t)\, f(S^h(t))) \tag{8.17}$$

for all bounded measurable functions $f: W(\mathbb{R}^n) \to \mathbb{R}$. Replace f by $f \circ S^h(-t)$ and then t by $-t$ in (8.17) to find that

$$\mu(f(S^h(t))) = \mu(Z(h, t) f), \tag{8.18}$$

where part (i) was used to conclude that $f \circ S^h(t) \circ S^h(-t) = f$ μ-a.s. Equation (8.18) implies that $d(S^h(t)_* \mu)/d\mu = Z(h, t)$. Q.E.D.

I will leave it as an exercise to the reader to verify that in the case $M = \mathbb{R}^n$ with the usual covariant derivative that (8.16) reproduces (1.2). One word of caution: $d\mu(\omega + h)/d\mu(\omega)$ in (1.2) is equal to $d(S^{-h}(1)_* \mu(\omega))/d\mu(\omega)$ not $d(S^h(1)_* \mu(\omega))/d\mu(\omega)$.

Remark 8.6. At this point the t in all of the above notation is unnecessary. The reader can easily verify that $S^{\tau h}(t)$ and $S^h(t\tau)$ both satisfy (5.5) with h replaced by τh. Therefore, $S^{\tau h}(1)$ is μ-indistinguishable from $S^h(\tau)$. For this reason, we introduce the following notation.

Notation 8.2. Let $S(h) \equiv S^h(1)$: $W(\mathbb{R}^n) \to W(\mathbb{R}^n)$, $O(h) \equiv O^h(1)$: $W(\mathbb{R}^n) \to W(O(n))$, $\alpha(h) = \alpha^h(1)$: $W(\mathbb{R}^n) \to W(\mathbb{R}^n)$, and $Z(h) = Z(h, 1)$: $W(\mathbb{R}^n) \to (0, \infty)$.

We end this section by transferring Theorem 8.4 from $W(\mathbb{R}^n)$ to $W(M)$.

THEOREM 8.5. *Recall that $\bar{\sigma}_o(s)$: $W(M) \to M$ was defined by $\bar{\sigma}_o(s)(\omega) = \omega(s)$ and $\{\bar{\sigma}_o(s)\}$ is an M-valued Brownian motion on the probability space $(W(M), \{\mathscr{H}^v_{s+}\}, v)$. Let $\{b(s) \equiv \hat{\Psi} \circ \bar{\sigma}_o(s) = \hat{\Psi}_s\}_{s \in [0,1]}$ be the fixed reference $\mathbb{R}^n$-valued Brownian motion on this probability space, and $h \in C^1([0, 1], \mathbb{R}^n)$ be a given function such that $h(0) = 0$. Also let $\bar{\sigma}(t)$ denote the solution to* (5.1), *$(d/dt)\,\bar{\sigma}(t) = H(\bar{\sigma}(t))h$ with $\bar{\sigma}(0) = \bar{\sigma}_o$ whose existence is guaranteed by Corollary* 6.3 *with $(\Omega, \{\mathscr{F}_s\}, P) = (W(M), \{\mathscr{H}^v_{s+}\}, v)$ and $\sigma_o = \bar{\sigma}_o$. Then*

(i) *$\bar{\sigma}(t) = \Psi \circ S^h(t) \circ \hat{\Psi}$ are v-indistinguishable,*

(ii) *$\bar{\sigma}(t)$ is a flow on $W(M)$ which leaves v quasi-invariant, and*

(iii) *$d(\bar{\sigma}(t)_* v)/dv = Z(th) \circ \hat{\Psi}$.*

Proof. Let $\sigma(t) \equiv \Psi \circ S^h(t)$, and notice that $\sigma(t)$ is a $\{\bar{b}(s)\}_{s \in [0,1]}$-Brownian semimartingale defined on $(W(\mathbb{R}^n), \{\mathscr{H}^\mu_{s+}\}, \mu)$. Recall that $\bar{b}(s)(\omega) = \omega(s)$ for $\omega \in W(\mathbb{R}^n)$, or as a function from $W(\mathbb{R}^n)$ to $W(\mathbb{R}^n)$, $\bar{b}$ is the identity map. By Theorem 8.2, $\sigma(t)$ is indistinguishable from $\pi \circ I(S^h(t))$, and so by Theorem 5.2, $\sigma(t)$ solves the geometric flow equation (5.1),

$$\dot{\sigma}(t) = H(\sigma(t)) \cdot h = \bar{H} \circ \sigma(t) \cdot h. \tag{8.19}$$

The derivative is taken in the $B^p(\mu)$-topologies, where the μ signifies that the reference probability space and Brownian motion is $(W(\mathbb{R}^n), \{\mathscr{H}^\mu_{s+}\}, \{\bar{b}(s)\}, \mu)$. Now right compose both sides of equation (8.19) with $\hat{\Psi}$ to find

$$\dot{\sigma}(t) \circ \hat{\Psi} = \bar{H} \circ \sigma(t) \circ \hat{\Psi} \cdot h. \tag{8.20}$$

Set $\tilde{\sigma}(t) \equiv \sigma(t) \circ \hat{\Psi} = \Psi \circ S^h(t) \circ \hat{\Psi}$, then because of Lemma 8.6 below $\tilde{\sigma}(t)$ is a $(W(M), \{\mathscr{H}^{\nu}_{s+}\}, \{b(s)\}, \nu)$ Brownian semimartingale for each t which is $B^p(\nu)$-continuously differentiable for all $p \in [2, \infty)$. Furthermore, the $B^p(\nu)$ derivative of $\tilde{\sigma}(t)$ is given by $(d/dt)\,\tilde{\sigma}(t) = \dot{\sigma}(t) \circ \hat{\Psi}$. Combining these last remarks with (8.20) shows that $\tilde{\sigma}(t)$ solves the same geometric flow equation as $\bar{\sigma}(t)$:

$$\frac{d}{dt}\,\tilde{\sigma}(t) = \bar{H} \circ \tilde{\sigma}(t) \cdot h.$$

Because $\tilde{\sigma}(0) = \Psi \circ S^h(0) \circ \hat{\Psi} = \Psi \circ \hat{\Psi}$ which is indistinguishable from $\bar{\sigma}_o$, it follows by the uniqueness assertion of Corollary 6.3 that $\bar{\sigma}(t)$ and $\tilde{\sigma}(t) = \Psi \circ S^h(t) \circ \hat{\Psi}$ are indistinguishable.

Parts (ii) and (iii) of the theorem are a trivial consequence of Theorems 8.2 and 8.4 and part (i) just proved. Q.E.D.

LEMMA 8.6. *Keep the same notation as in Theorem* 8.5. *Let* $X = \int A\, d\bar{b} + \int \alpha\, ds$ *be a* $(W(\mathbb{R}^n), \{\mathscr{H}^{\mu}_{s+}\}, \{\bar{b}(s)\}, \mu)$*-Brownian semimartingale, then*

$$X \circ \hat{\Psi} = \int A \circ \hat{\Psi}\, db + \int \alpha \circ \hat{\Psi}\, ds \tag{8.21}$$

and hence $X \circ \hat{\Psi}$ *is a* $(W(M), \{\mathscr{H}^{\nu}_{s+}\}, \{b(s)\}, \nu)$*-Brownian semimartingale. Furthermore,* $\|X \circ \hat{\Psi}\|_{B^p(\nu)} = \|X\|_{B^p(\mu)}$.

Proof. Equation (8.21) follows from Proposition 8.2 and the fact that $\bar{b} \circ \hat{\Psi} = \hat{\Psi} = \hat{\Psi} \circ \bar{\sigma}_o = b$. To see this last statement just compute

$$\begin{aligned}\|X \circ \hat{\Psi}\|_{B^p(\nu)} &\equiv \|A \circ \hat{\Psi}\|_{S^p(\nu)} + \|\alpha \circ \hat{\Psi}\|_{S^p(\nu)} = \|A\|_{S^p(\hat{\Psi}_* \nu)} + \|\alpha\|_{S^p(\hat{\Psi}_* \nu)} \\ &= \|A\|_{S^p(\mu)} + \|\alpha\|_{S^p(\mu)} \equiv \|X\|_{B^p(\mu)},\end{aligned}$$

where we have used Theorem 3.4 to conclude that $\hat{\Psi}_* \nu = \mu$. Q.E.D.

Remark 8.7. One might think that the notion of solution to Eqs. (5.1) or (5.5) depends on the particular choice of a reference Brownian motion. However, this does not seem to be the case in the above path space setting. The reason is that every $\mathbb{R}^n$-valued Brownian motion (B) on $(W(\mathbb{R}^n), \{\mathscr{H}^{\mu}_{s+}\}, \mu)$ necessarily has the form $B = \int O\, d\bar{b}$, where O is an $O(n)$-valued predictable process, see, for example, [Pr, Theorem 42, p. 155]. Also because $\Psi\colon (W(\mathbb{R}^n), \{\mathscr{H}^{\mu}_{s+}\}, \mu) \to (W(M), \{\mathscr{H}^{\nu}_{s+}\}, \nu)$ is a measure theoretic isomorphism, it follows that any $(W(M), \{\mathscr{H}^{\nu}_{s+}\}, \nu)$-Brownian motion (B) must be of the form $B = \int O\, db$ where O is an $\mathscr{H}^{\nu}_{s+}$-predictable $O(n)$-valued process. In particular, this shows, for the path spaces $W(M)$ and $W(\mathbb{R}^n)$, that the B^p-norms are independent of the choice of reference Brownian motion.

9. Integration by Parts

In this section we are primarily interested in the filtered probability space $(W(M), \{\bar{\mathscr{H}}^{\nu}_{s+}\}, \nu)$. To simplify notation let $\mathscr{F}_s \equiv \bar{\mathscr{H}}^{\nu}_{s+}$ and let b be the $\mathbb{R}^n$-valued Brownian motion on $(W(M), \{\mathscr{F}_s\}, \nu)$ given by $b \equiv \hat{\Psi} = \hat{\Psi} \circ \bar{\sigma}_o$. Again, as in Section 8, the covariant derivative ∇ on TM is assumed to be torsion skew symmetric (TSS), so that all the results of Section 8 are valid.

The purpose of this section is to use the results of Section 8 to derive an integration by parts formula for the "H-derivative." For the prototype of this sort of result see Cameron [Ca]. It was L. Gross who first emphasized the importance of and systematically studied the pointwise H-derivative in the abstract Wiener space setting. A history of the H-derivative may be found in Gross' paper [Gr3]. Before introducing the H-derivative appropriate to our context, we introduce the reproducing kernel Hilbert space H.

Notation 9.1. Let H denote the Hilbert space of functions $h\colon [0, 1] \to \mathbb{R}^n$ such that h is absolutely continuous, $h(0) = 0$, and $\int_0^1 |h'(s)|^2\, ds < \infty$.

In this section we fill fix a function $\bar{H}\colon W(M) \to W(O(M))$ as in Definition 8.2 for which $H(X) = \bar{H} \circ X$ for all M-valued semimartingales X with law(X) equivalent to ν. This should help avoid any possible confusion between the horizontal lift operator H and the Hilbert space H.

Definition 9.1. Let D denote the set of C^2-cylinder functions on $W(M)$. That is, $f \in D$ iff there is a positive integer k, a C^2-function $F\colon M^k \to \mathbb{R}$, and points $s_1, s_2, ..., s_k$ in $[0, 1]$ such that $f(\omega) = F(\omega(s_1), \omega(s_2), ..., \omega(s_k))$ for $\omega \in W(M)$. For any $h \in H$, the h derivative $(\partial_h f)$ of $f \in D$ is defined ν-a.s. to be

$$\partial_h f(\omega) = \sum_{i=1}^{k} f_i(\omega)\langle \bar{H}(\omega)(s_i)\, h(s_i)\rangle, \tag{9.1}$$

where $f_i(\omega)\langle v\rangle \equiv v(F(\omega(s_1), \omega(s_2), ..., \omega(s_{i-1}), \cdot, \omega(s_{i+1}), ..., \omega(s_k))$ for $v \in T_{\omega(s_i)}M$ and $\omega \in W(M)$. So f_i is the differential of F with respect to the ith variable evaluated at $(\omega(s_1), \omega(s_2), ..., \omega(s_k))$. (It will be shown in the course of the proof of Theorem 9.1 below that $\partial_h f$ is well defined ν-a.s., independent of the way f is represented.)

Theorem 9.1. *Let $(\cdot, \cdot)$ denote the $L^2(W(M), \mathscr{F}_1, \nu)$ inner product. For $h \in H$ define*

$$z(h) \equiv \int_0^1 [(\mathrm{Ric}_{\bar{H}}\langle h\rangle + \hat{\Theta}_{\bar{H}}\langle h\rangle)/2 + h'] \cdot db, \tag{9.2}$$

where $\mathrm{Ric}_{\bar{H}}\langle h\rangle \equiv \mathrm{Ric}_{\bar{H}}\langle I, h, I\rangle = \sum_i \Omega_{\bar{H}}\langle h, e_i\rangle e_i$, *and* $\hat{\Theta}_{\bar{H}}\langle h\rangle \equiv \Theta_{\bar{H}}\langle I, h, I\rangle = \sum_i \Theta'_{\bar{H}}\langle e_i, h, e_i\rangle$. *See Definition* 6.2 *for more on this notation. Then with respect to this inner product,* ∂_h^* *is a densely defined operator, the domain* $D(\partial_h^*)$ *contains* D, *and for* $f \in D$

$$\partial_h^* f = -\partial_h f + z(h) f. \tag{9.3}$$

Furthermore for each $h \in H$ *and* $p \in [1, \infty)$ *there is a constant* c_p *such that*

$$\|z(h)\|_{L^p(\nu)} \leqslant c_p \left[\int_0^1 |h'(s)|^2\, ds\right]^{1/2}. \tag{9.4}$$

This theorem immediately implies the following corollary.

COROLLARY 9.1. *For each* $h \in H$, *the densely defined operator* ∂_h *on* $L^2(W(M), d\nu)$ *is closable.*

Proof of Theorem 9.1. The estimate in (9.4) is easily proved by Burkholder's inequality, the existence of a constant C bounding $\hat{\Theta}_u$ and Ric_u independent of $u \in O(M)$ (by compactness), and the Sobolev inequality

$$|h(s)| \leqslant \left[\int_0^1 |h'(s')|^2\, ds'\right]^{1/2}. \tag{S}$$

So it only remains to prove (9.3).

For the moment assume that $h \in H \cap C^1$, and let $\bar{\sigma}(t)$ be as in the statement of Theorem 8.5. So $\bar{\sigma}(t)$ is a version of $\Psi \circ S^h(t) \circ \hat{\Psi}$ which is $C^{1,0}$ ν-a.s. and satisfies

$$\frac{d}{dt}\bar{\sigma}(t) = \bar{H} \circ \bar{\sigma}(t) \cdot h \qquad \text{with} \quad \bar{\sigma}(0) = \bar{\sigma}_o,$$

where the derivative is relative to the $B^p(\nu)$-norms. To simplify notation, we now drop the bars and write $\sigma(t)$ for $\bar{\sigma}(t)$. (Recall that $\sigma_o = \bar{\sigma}_o$ is the process on $W(M)$, such that as a function from $W(M)$ to $W(M)$ σ_o is the identity map.) Suppose $f \in D$ (D as in Definition 9.1), then trivially $\partial_h f = (d/dt)|_0 f(\sigma(t))$ ν-a.s., which incidentally shows that $\partial_h f$ is well defined independent of the possible choices for $k, s_1, s_2, ..., s_k \in [0, 1]$, and $F\colon M^k \to \mathbb{R}$ such that $f(\omega) = F(\omega(s_1), \omega(s_2), ..., \omega(s_k))$. By Theorem 8.5,

$$\nu(f \circ \sigma(t)) = \nu(f \cdot Z(th) \circ \hat{\Psi}) \tag{9.5}$$

holds for all t. In Lemma 9.1 below it is shown that $(d/dt)|_0 f \circ \sigma(t) = \partial_h f$ in $L^p(\nu)$ for all $p \in [1, \infty)$. It follows from Lemma 8.6, Proposition 8.2, Theorem 8.2, and Lemma 9.2 below that $(d/dt)|_0\, Z(th) \circ \hat{\Psi} = z(h)$ in $L^p(\nu)$

for all $p \in [1, \infty)$. Therefore, after differentiating (9.5) at $t=0$, one finds that

$$\nu(\partial_h f) = \nu(f \cdot z(h)). \tag{9.6}$$

Now replace f in (9.6) by $f \cdot g$ where g is also in D and use $\partial_h(f \cdot g) = \partial_h f \cdot g + f \cdot \partial_h g$ to find

$$(\partial_h f, g) = (f, -\partial_h g - \dot{Z}(0)\, g) = (f, -\partial_h g + z(h)\, g), \tag{9.7}$$

which proves (9.3) for $h \in H \cap C^1$.

For general $h \in H$, choose a sequence of functions $h_n \in H \cap C^1$ such that $\int_0^1 |h'_n - h'|^2\, ds \to 0$ as $n \to \infty$, and hence by (S), $|h - h_n|_\infty \to 0$ as $n \to \infty$. Now it is easy to show $\|\partial_{h_n} f - \partial_h f\|_{L^\infty(\nu)} \leqslant C\, |h - h_n|_\infty$ which tends to zero as $n \to \infty$, where $\partial_h f$ is given by (9.1). Since $\partial_{h_n} f$ is well defined ν-a.s. this shows that $\partial_h f$ is also well defined ν-a.s. By these comments and (9.4), one easily verifies that (9.7) holds for all $h \in H$ by replacing h in (9.7) by h_n and passing to the limit $n \to \infty$. Q.E.D.

LEMMA 9.1. *Let $h \in H \cap C^1$, $\sigma(t)$, and $f \in D$ be as above. Then for all $p \in [1, \infty)$, $k(t) \equiv f \circ \sigma(t)$ is L^p-differentiable at $t = 0$ with $\dot{k} = \partial_h f$.*

Proof. By the fundamental theorem of calculus (pointwise)

$$[k(t) - k(0)]/t - \dot{k}(0) = \frac{1}{t} \int_0^t [\dot{k}(\tau) - \dot{k}(0)]\, d\tau.$$

Taking the L^p norms of both sides of this equations yields

$$\|[k(t) - k(0)]/t - \dot{k}(0)\|_{L^p} \leqslant \left| \frac{1}{t} \int_0^t \|[\dot{k}(\tau) - \dot{k}(0)]\|_{L^p}\, d\tau \right|. \tag{9.8}$$

Now

$$\begin{aligned} \dot{k}(\tau) - \dot{k}(0) = \sum_{i=1}^{k} & [f_i(\sigma(\tau)) \langle \bar{H} \circ \sigma(\tau)(s_i)\, h(s_i) \rangle \\ & - \langle f_i(\sigma_o) \langle \bar{H} \circ \sigma_o(s_i)\, h(s_i) \rangle], \end{aligned}$$

from which it is easy to get the estimate

$$|\dot{k}(\tau) - \dot{k}(0)|_\infty \leqslant C\, |\bar{H} \circ \sigma(\tau) - \bar{H} \circ \sigma_o|_\infty + C\, |\sigma(\tau) - \sigma_o|_\infty, \tag{9.9}$$

where $C = C(F, h)$ depends on the sup-norm of F and its derivatives up to order two and the sup-norm of h. Take L^p norms of both sides of (9.9) to get

$$\begin{aligned} \|\dot{k}(\tau) - \dot{k}(0)\|_{S^p} &\leqslant C' [\|\bar{H} \circ \sigma(\tau) - \bar{H} \circ \sigma_o\|_{S^p} + \|\sigma(\tau) - \sigma_o\|_{S^p}] \\ &\leqslant C_p\, \|\sigma(\tau) - \sigma_o\|_{B^p}, \end{aligned} \tag{9.10}$$

where Lemmas 4.1 and 7.3 were used to get the last inequality. The constant C_p now depends on p. Now σ is B^p-continuously differentiable and hence Lipschitz. Therefore, combining (9.8), (9.10), and this last comment shows there is a constant C_p' such that

$$\| [k(t)-k(0)]/t-\dot{k}(0)\|_{L^p} \leqslant C_p' \, |t|.$$

This shows k is B^p-differentiable at $t=0$. This proves the lemma, since in the proof of Theorem 9.1 it has already been noted that the pointwise derivative of k at $t=0$ is $\partial_h f$. Q.E.D.

LEMMA 9.2. *The function $Z(th)$ is $L^p(\mu)$ differentiable for all $p \in [1, \infty)$ and*

$$\frac{d}{dt}\bigg|_0 Z(th) \equiv \int_0^1 [\mathrm{Ric}_I\langle h\rangle + \hat{\Theta}_I\langle h\rangle)/2 + h'] \cdot d\bar{b} \equiv \bar{z}(h).$$

Proof. To simplify notation let $O(t) = O^h(t)$, and $\alpha(t) = \alpha^h(t)$, where O^h and α^h are given in Definition 8.3. Set $D(t)(s) \equiv -\int_0^s \alpha(t) \cdot O(t)\, db - (1/2)\int_0^s |\alpha(t)(s')|^2\, ds'$ so that $Z(th) = e^{D(-t)(1)}$. Let $Y(t)(s) = e^{D(t)(s)}$, clearly it suffices to prove that $Y(t)$ is $S^p(\mu)$ differentiable at $t=0$ and that $\dot{Y}(t)(1) = -\bar{z}(h)$.

Using Lemma 4.6 and the regularity properties of O and α (see Corollary 6.1, Corollary 6.2, and Proposition 6.3), it is easily seen that $D\colon J \to \mathscr{S}^\infty\mathbb{R}$ is a C^1-function, and that $\dot{D}(t)$ is $B^p(\mu)$-Lipschitz for all $p \in [2, \infty)$. By Lemma 4.5 we may and do assume that a version of D has been chosen such that $(t, s) \to D(t)(s)$ is $C^{1,0}$. Hence, pointwise $\dot{Y}(t)(s) = Y(t)(s)\, \dot{D}(t)(s)$, where μ-a.s.

$$\dot{D}(t)(s) = -\int_0^s \frac{d}{dt}[O(t)^{\mathrm{tr}}\alpha(t)] \cdot d\bar{b} - \frac{1}{2}\int_0^s \frac{d}{dt}|\alpha(t)(s')|^2\, ds'. \tag{9.11}$$

By the same techniques used in the proof of Lemma 9.1 one has

$$\| [Y(t) - Y(0)]/t - \dot{Y}(0)\|_{S^p} \leqslant \left| \frac{1}{t}\int_0^t \|\dot{Y}(\tau) - \dot{Y}(0)\|_{S^p}\, d\tau \right|. \tag{9.12}$$

This last integrand is easily estimated with the aid of Holder's inequality, and Lemma 4.1 as

$$\begin{aligned}
\|\dot{Y}(\tau) - \dot{Y}(0)\|_{S^p} &\leqslant \|(Y(\tau)-1)\,\dot{D}(\tau)\|_{S^p} + \|\dot{D}(\tau) - \dot{D}(0)\|_{S^p} \\
&\leqslant \|(Y(\tau)-1)\|_{S^q}\, \|\dot{D}(\tau)\|_{S^{q'}} + \|\dot{D}(\tau) - \dot{D}(0)\|_{S^p}, \\
&\leqslant C\, \|(Y(\tau)-1)\|_{S^q}\, \|\dot{D}(\tau)\|_{B^{q'}} + C\, \|\dot{D}(\tau) - \dot{D}(0)\|_{B^p} \\
&\leqslant C\, [\, \|(Y(\tau)-1)\|_{S^q} + |\tau|\,],
\end{aligned} \tag{9.13}$$

where the constant C may increase from line to line, and $1/p = 1/q + 1/q'$. The fact that $\dot{D}$ is $B^p(\mu)$-Lipschitz was used to get the last inequality. Using the fundamental theorem of calculus, Holder's inequality, and Burkholder's inequality one finds that

$$\begin{aligned}\|(Y(\tau)-1)\|_{S^q} &\leqslant \left| \int_0^\tau \|Y(\tau')\|_{S^r} \|\dot{D}(\tau')\|_{S^{r'}} d\tau' \right| \\ &\leqslant c_{r'} \left| \int_0^\tau \|Y(\tau')\|_{S^r} \|\dot{D}(\tau')\|_{B^{r'}} d\tau' \right|, \qquad (9.14)\end{aligned}$$

where $1/q = 1/r + 1/r'$. By Remark 8.1 and a standard martingale inequality (see Theorem 6.10, pp. 33–34, of [IW]) there exists a constant C such that $\|Y(\tau')\|_{S^r} \leqslant C$ for $\tau' \in J$. Because $\dot{D}(\tau')$ is $B^{r'}$ continuous there is a constant C such that $\|\dot{D}(\tau')\|_{B^{r'}} \leqslant C$ for $\tau' \in J$. These last two comments combined with (9.14) imply the existence of a constant C such that $\|(Y(\tau)-1)\|_{S^q} \leqslant C\,|\tau|$. Combining this estimate with (9.12) and (9.13) shows there is a constant C such that

$$\|[Y(t)-Y(0)]/t - \dot{Y}(0)\|_{S^p} \leqslant C\,|t|,$$

which shows that Y is B^p-differentiable at $t=0$. We also know that $\dot{Y}(0) = e^{D(0)}\dot{D}(0) = \dot{D}(0)$ because $D(0)=0$, since $\alpha(0) \equiv 0$. From (9.11), using (6.6), (6.7), and the initial conditions $\alpha(0)=0$, and $O(0) = \mathrm{id} \in O(n)$, it follows that

$$\begin{aligned}\dot{Y}(0)(1) = \dot{D}(0)(1) &= -\int_0^1 \dot{\alpha}(0)\cdot d\bar{b} = -\int_0^1 R(\bar{b})\, d\bar{b} \\ &= -\int_0^1 [(\mathrm{Ric}_I\langle h\rangle + \hat{\Theta}_I\langle h\rangle)/2 + h']\, db =: -\bar{z}(h). \quad \text{Q.E.D.}\end{aligned}$$

To conclude this section it will be shown that the "infinitesimal" density $z(h)$ is "highly" integrable. To simplify notation, for each $h \in H$ set $\|h\| \equiv (\int_0^1 |h'(s)|^2\, ds)^{1/2}$.

PROPOSITION 9.1. *There exists constants $\delta > 0$, and $K > 1$ such that for each $h \in H$, $\nu(\exp(\delta[z(h)/\|h\|]^2)) \leqslant K$, where $z(h)$ is defined in* (9.2).

Proof. Let N be the martingale $N(s) \equiv \int_0^s [(\mathrm{Ric}_{\tilde{H}} < h\rangle + \hat{\Theta}_{\tilde{H}}\langle h\rangle)/2 + h'] \cdot db$. It is easy to see that there is a constant $C > 0$ such that $[N, N](1) \leqslant C\,\|h\|^2$. Now apply Lemma 9.3 below with $\delta = \varepsilon/C$. Q.E.D.

LEMMA 9.3. *There exists constants $\varepsilon > 0$, and $K > 1$, such that for each continuous local martingale N (on some filtered probability space $(\Omega, \{\mathscr{F}_s\}, P)$) the following estimate holds*:

$$P(\exp(\varepsilon N_1^2/\|[N, N]_1\|_{L^\infty(P)}) \leqslant K. \qquad (9.15)$$

Proof. It is clear that if $C \equiv \| [N, N]_1 \|_{L^\infty(P)} = \infty$ then (9.15) holds since $K > 1$ by assumption. So we may assume that $C < \infty$. By the Dambis, Dubins–Schwarz Theorem (see [RY, Theorem 1.7, p. 171, or RW, Chap. IV, Sect. 34]), on a possibly "enriched" probability space $(\tilde{\Omega}, \{\tilde{\mathscr{F}}_s\}, \tilde{P})$ there is local martingale $\tilde{N}$ and a Brownian motion $\tilde{B}$ such that the laws of N and $\tilde{N}$ are the same, the laws of $[N, N]$ and $[\tilde{N}, \tilde{N}]$ are the same, and $\tilde{N}(s) = \tilde{B}([\tilde{N}, \tilde{N}](s))$. Therefore,

$$\begin{aligned} P(\exp(\varepsilon N_1^2/C)) &= \tilde{P}(\exp(\varepsilon \tilde{N}_1^2/C)) = \tilde{P}(\exp(\varepsilon \tilde{B}([\tilde{N}, \tilde{N}](1)))^2/C) \\ &\leqslant \tilde{P}(\exp(\varepsilon (\tilde{B}_C^*)^2/C)), \end{aligned} \tag{9.16}$$

where $\tilde{B}_C^* = \sup_{s \leqslant C} |\tilde{B}_s|$. Now $\tilde{B}(\cdot)$ has the same law as $C^{-1/2}\tilde{B}(C\cdot)$, and hence $\tilde{B}_C^*$ has the same law as $C^{1/2}\tilde{B}_1^*$. So (9.16) may be written as

$$P(\exp(\varepsilon N_1^2/C)) \leqslant \tilde{P}(\exp(\varepsilon (\tilde{B}_1^*)^2)). \tag{9.17}$$

But by Fernique's Theorem (see [K3, pp. 159–160] or [IW, p. 402]) there is a constant $\varepsilon > 0$, such that $K \equiv \tilde{P}(\exp(\varepsilon (\tilde{B}_1^*)^2)) < \infty$. (Notice that only the law of $\tilde{B}$ enters here so that ε and K are independent of the particular realization of the continuous Brownian motion $\tilde{B}$.) Q.E.D.

10. Final Remarks

In this final section I will briefly discuss the two alternative methods for "shifting" an M-valued semimartingale that were introduced after Example 5.1. In each of these strategies the existence of the shifted process is not at issue. However, in general, these alternative shifting strategies will not have the desirable quasi-invariance properties. Since the results of this section are negative in nature, I will only sketch the arguments involved. For the rest of this section it will *always* be assumed that the covariant derivative (∇) is torsion skew symmetric (TSS).

Let $h\colon [0, 1] \to \mathbb{R}^n$ be a C^1-function such that $h(0) = 0$ and let σ_o be an M-valued semimartingale starting at o. As described after Example 5.1, one might try to define $\sigma(t)$ by $\sigma(t)(s) = \exp(tH(\sigma_o)(s)\, h(s))$, where exp is the geodesic flow with respect to the covariant derivative ∇. Notice that $\sigma(t)$ is a semimartingale if σ_o is a semimartingale because of Itô's lemma. Let $T_t(\sigma_o) \equiv \sigma(t) = \exp(tH(\sigma_o)h)$, so that T_t transforms semimartingales on M to semimartingales on M. In general T_t is not a flow—i.e., $T_t \circ T_\tau \neq T_{t+\tau}$.

Remark 10.1 For the Riemannian manifolds in Example 5.1, T_t is actually a flow. The reason is because in each of these examples the curvature is zero so that the map T_t is the flow generated by (5.1)—see Remark 10.2. However, as soon as ∇ has curvature, the map T_t will no longer in general be a flow. This happens even on a non-commutative compact Lie group with the Levi–Civita connection, which is the average of the left and right connections.

The other main flaw of the map T_t is that it does not have the quasi-invariance property. I will now explain why this property fails when the curvature is not zero. As has been done throughout this paper, we will pull T_t back to a mapping on $\mathbb{R}^n$-valued semimartingales where it is easier to decide quasi-invariance questions. To this end set $w(t) \equiv I^{-1} \circ H(\sigma(t))$.

To simplify notation let $u(t) \equiv H(\sigma(t))$, $u_o \equiv H(\sigma_o)$, and $v(t)(s) = e^{tB\langle h(s)\rangle}(u_o(s))$, where $B\langle h(s)\rangle$ is the standard horizontal vector field in Definition 2.2. Since $\sigma = \pi \circ v = \pi \circ u$, it makes sense to define the $O(n)$-valued semimartingale $g(t)(s)$ by $g(t)(s) \equiv u(t)(s)^{-1}\, v(t)(s)$. So $v(t)(s) = u(t)(s)\, g(t)(s)$, which is just the decomposition of the non-horizontal $O(M)$-valued semimartingale (v) into a horizontal piece (u) and a "vertical" piece (g). It is convenient to define another $\mathbb{R}^n$-valued semimartingale by $x(t)(s) = \int_0^s \vartheta\langle \delta v(t)\rangle$. The two processes w and x are related by g, namely $w = \int g\, \delta x$. Now suppose that σ_o is a Brownian motion on M, so that $b \equiv w(0) = x(0)$ is a Brownian motion in $\mathbb{R}^n$. Because σ_o is a Brownian semimartingale it follows that u, v, w, and x are all Brownian semimartingales.

We can now understand why, in the case of nonzero curvature, the laws of $w(t)$ and b are not in general equivalent. The idea is to use Lemma 8.1 along with the non-orthogonality (to be shown) of the process $O(t)$, where $O(t)$ and $\alpha(t)$ are processes such that $w(t) \equiv \int O(t)\, db + \int \alpha(t)\, ds$. Because the process g is orthogonal it will suffice to show that $Q(t) \equiv g^{-1}(t)\, O(t)$ is not orthogonal. Since

$$x = \int g^{-1}\, \delta w = \int g^{-1} O\, db + \int \beta\, ds = \int Q\, db + \int \beta\, ds$$

for some process β, in order to find Q we need to find the differential (dx) of x.

Start by computing $d\dot{x}$,

$$\begin{aligned} d\dot{x} &= \frac{d}{dt}\, \vartheta\langle \delta v\rangle = d\vartheta\langle \dot{v}, \delta v\rangle + \delta(\vartheta\langle \dot{v}\rangle) \\ &= \Theta\langle \dot{v}, \delta v\rangle - \omega \wedge \vartheta\langle \dot{v}, \delta v\rangle + \delta(\vartheta\langle \dot{v}\rangle) \\ &= \Theta\langle \dot{v}, \delta v\rangle + \omega\langle \delta v\rangle\, \vartheta\langle \dot{v}\rangle + dh \\ &= \Theta_v\langle h, \delta x\rangle + \omega\langle \delta v\rangle\, h + dh, \qquad (10.1) \end{aligned}$$

where we have used the first structure equation ($\Theta = d\vartheta + \omega \wedge \vartheta$), $\omega\langle \dot{v}\rangle = 0$, and $\vartheta\langle \dot{v}\rangle = h$. Also compute $(d/dt)\,\omega\langle \delta v\rangle = d\omega\langle \dot{v}, \delta v\rangle + \delta(\omega\langle \dot{v}\rangle) = \Omega_v\langle h, \delta x\rangle$ so that

$$\omega\langle \delta v(t)\rangle = \int_0^t \Omega_{v(t)}\langle h, \delta x(t)\rangle. \tag{10.2}$$

Writing $x = \int Q\, db + \int \beta\, ds$, and substituting this expression for x into (10.1) using (10.2) one finds (by comparing the coefficients in front of db) that Q satisfies

$$\dot{Q}(t) = \Theta_{v(t)}\langle h, Q(t)\cdot\rangle + \int_0^t \Omega_{v(\tau)}\langle h, Q(\tau)\cdot\rangle\, h\, d\tau. \tag{10.3}$$

Because $x(0) = b$, the initial condition for (10.3) is $Q(0) = Id$.

Remark 10.2. If the curvature of ∇ is zero then $v(t) = u(t)$, which follows from the stochastic version of Eq. (2.5). In this case (10.3) is the same as (6.5), and $Q(t)$ will be orthogonal if ∇ is TSS.

For simplicity assume that ∇ is the Levi–Civita covariant derivative on M so that $\Theta \equiv 0$. If $Q(t)$ were orthogonal for all t, then $\Omega_{v(0)}\langle h, \cdot\rangle\, h$ would necessarily have to be $so(n)$ valued, since

$$\Omega_{v(0)}\langle h, \cdot\rangle\, h = \frac{d}{dt}\bigg|_0 [Q(t)^{-1}\,\dot{Q}(t)]. \tag{10.4}$$

But it is not generally true that $\Omega_{v(0)}\langle h, \cdot\rangle\, h$ is $so(n)$-valued, as can be seen by taking M to be the standard n-sphere (S^n) with the Levi–Civita connection. For $M = S^n$, $\Omega_u\langle a, b\rangle\, c = (a, c)b - (b, c)a$ for all $a, b, c \in \mathbb{R}^n$ and $u \in O(M)$, and hence $\Omega_{v(0)}\langle h, \cdot\rangle\, h$ is the non-skew symmetric linear transformation on $\mathbb{R}^n$

$$c \to (h, h)c - (h, c)h.$$

Now let us consider the second alternative for $\sigma(t)$ introduced after Example 5.1. For this example let $X\colon [0, 1] \times M \to TM$ be a smooth s-dependent vector field, such that $X(0)(o) = 0_m$—the zero vector in T_mM. Define $\sigma(t)$ using the flow of the vector field $X(s)$ by $\sigma(t)(s) = e^{tX(s)}(\sigma_o(s))$. Since, $\sigma(t)(s)$ is a smooth function of the semimartingale $s \to (s, \sigma_o(s))$, $\sigma(t)$ is still a semimartingale. Similarly, if σ_o is a Brownian semimartingale then so is $\sigma(t)$. Let T_t be defined by $T_t(\sigma_o)(s) \equiv \sigma(t)(s) \equiv e^{tX(s)}(\sigma_o(s))$, then T_t is clearly a flow on the space of semimartingales. However, we shall indicate that T_t has the quasi-invariance property iff each of the vector fields $X(s)$

(for $s \in [0, 1]$) is a Killing vector field. In other words, for each $s \in [0, 1]$ the flow $e^{tX(s)}$ should be a one parameter (t) family of isometries on M.

To investigate the quasi-invariance, again set $w(t) \equiv I^{-1} \circ H(\sigma(t))$ and $u(t) \equiv H(\sigma(t))$. Assume that σ_o is a Brownian motion on M so that $b \equiv w(0)$ is a Brownian motion on M, then $w(t)$ is a Brownian semimartingale. Again let us compute $d\dot{w}(t)$,

$$\begin{aligned} d\dot{w} = \frac{d}{dt}(\vartheta\langle\delta u\rangle) &= d\vartheta\langle\dot{u}, \delta u\rangle + \delta(\vartheta\langle\dot{u}\rangle) \\ &= \Theta\langle\dot{u}, \delta u\rangle - \omega \wedge \vartheta\langle\dot{u}, \delta u\rangle + \delta(\vartheta\langle\dot{u}\rangle) \\ &= \Theta_u\langle\xi, \delta w\rangle - \omega\langle\dot{u}\rangle\, \delta w + \delta\xi, \end{aligned} \tag{10.5}$$

where $\xi(t, s) \equiv \vartheta\langle\dot{u}(t)(s)\rangle = u^{-1}(t)(s)\, X(s)(\sigma(t, s))$. Since $s \to u(t)(s)$ is horizontal one finds that

$$\begin{aligned} \delta\xi(t, s) &= u(t, s)^{-1}\, [X'(s)(\sigma(t, s)) + \nabla_{\delta\sigma(t, s)} X(s)] \\ &= u(t, s)^{-1}\, [X'(s)(\sigma(t, s)) + \nabla_{u(t, s)\,\delta w(t)(s)} X(s)]. \end{aligned} \tag{10.6}$$

The interpretation of the last term in (10.6) requires some explanation. What is needed is a definition for the stochastic covariant differential. Let $Y \in \Gamma(TM)$ be a vector field. Define $\bar{Y}: O(M) \to \mathbb{R}^n$ by $\bar{Y}(u) = u^{-1} Y(\pi(u))$, recall the correspondence $Y \to \bar{Y}$, from $\Gamma(YM)$ to the smooth functions $\bar{Y}: O(M) \to \mathbb{R}^n$ such that $\bar{Y}(ug) = g^{-1}\bar{Y}(u)$ for $u \in O(M)$ and $g \in O(n)$, is a 1–1 correspondence. Then given an M-valued semimartingale $\sigma(s)$ define $u^{-1}\nabla_{\delta\sigma} Y \equiv d\bar{Y}\langle\delta u\rangle$, where u is a horizontal lift of σ. In the case of interest, $\delta\sigma = u\, \delta w$ or equivalently $\delta u = B\langle\delta w\rangle(u)$. So set $u^{-1}\nabla_{u\,\delta w} Y \equiv d\bar{Y}\langle B\langle\delta w\rangle(u)\rangle$. I will leave it to the interested reader to verify using these definitions the validity of (10.6).

Now insert (10.6) into (10.5) to find

$$\begin{aligned} d\dot{w}(t) = \Theta_{u(t)}&\langle\xi(t, \cdot), \delta w(t)\rangle - \omega\langle\dot{u}(t)\rangle\, \delta w(t) \\ &+ u(t)^{-1}\, [X'(\cdot)(\sigma(t)) + \nabla_{u(t)\,\delta w(t)} X(\cdot)], \end{aligned} \tag{10.7}$$

where $X'(s)(m) = (d/ds)\, X(s)(m)$.

Defining (O, α) by $w(t) = \int O(t)\, db + \int \alpha(t)\, ds$, one finds from (10.7) by considering the coefficients in front of the db terms that O satisfies

$$\dot{O} = \Theta_u\langle\xi, O\cdot\rangle + AO + u^{-1}\nabla_{uO} X, \tag{10.8}$$

where $A \equiv -\omega\langle\dot{u}\rangle$. Since $\Theta_u\langle\xi, \cdot\rangle$ and A are $so(n)$-valued processes, in order for $O(t)$ to be orthogonal for all t we must require that the linear transformation on $\mathbb{R}^n$ given by

$$a \to u_o(s)^{-1}\, \nabla_{u_o(s)a} X(s)$$

is P-a.s. skew symmetric with the null set independent of s. Here $u_0(s) \equiv u(0)(s)$. The only likely way to satisfy this condition is to require for each $s \in [0, 1]$ and $u \in O(M)$ that the linear transformation

$$a \to u^{-1} \nabla_{ua} X(s) \tag{10.9}$$

be skew symmetric. This last condition is equivalent to requiring for each $m \in M$ and $s \in [0, 1]$ that the map $(v \to \nabla_v X(s))$: $T_m M \to T_m M$ be skew symmetric with respect to the metric g, i.e., $g\langle \nabla_v X(s), v \rangle \equiv 0$ for all $s \in [0, 1]$ and $v \in TM$. If ∇ is the Levi–Civita connection, then it is well known that this condition is equivalent to X being a Killing vector field. The next lemma asserts this is still true provided that ∇ is TSS.

LEMMA 10.1. *Suppose that (M, g) is a Riemannian manifold with metric g, and that ∇ is a TSS g-compatible covariant derivative on TM. Let X be a vector field on M, then the condition that $g\langle \nabla_v X, v \rangle \equiv 0$ for all $v \in TM$ is equivalent to X being a Killing vector. Recall that X is a Killing vector field iff $L_X g = 0$, where $L_X g$ denotes the Lie derivative of g with respect to X.*

Proof. Let Y be an arbitrary vector field on M and compute

$$\begin{aligned}(L_X g)\langle Y, Y \rangle &= X(g\langle Y, Y \rangle) - 2g\langle [X, Y], Y \rangle = 2g\langle \nabla_X Y - [X, Y], Y \rangle \\ &= 2g\langle \nabla_Y X + T\langle X, Y \rangle, Y \rangle = 2g\langle \nabla_Y X, Y \rangle,\end{aligned}$$

where the last equality used torsion skew symmetric assumption on ∇. Because g is symmetric and hence so is $L_X g$, this last equation shows $L_X g \equiv 0$ iff $g\langle \nabla_Y X, Y \rangle \equiv 0$. Q.E.D.

The condition that $X(s)$ be a Killing vector field is very strong, and in fact can imply that $X(s) \equiv 0$. For example, by Bochner's Theorem (see [Bo, Theorem 1] or [W, Theorem 1]), if the Ricci curvature with respect to the Levi–Civita connection is negative definite (or at least "quasi-negative"), then g does not admit any non-trivial Killing vector fields. The absence of non-trivial Killing vector fields is what one would expect for "generic" metrics (g). The implication of these remarks is that in general the flow $T_t(\sigma_o(s)) = e^{tX(s)}(\sigma_o(s))$ can not be expected to have the quasi-invariance property for any choice of an s-dependent vector field X.

It should be noted that Lie groups and more generally homogeneous spaces do have metrics which admit non-trivial Killing vector fields. For example, if $M = G$ is a Lie group with the metric and connection given as in Example 8.1(b), then the s-dependent vector fields $X(s)(g) \equiv L_{g*} h(s)$, where h: $[0, 1] \to \mathfrak{g}$ is any function, are all Killing vector fields. (Recall from Example 5.1 that in the case of a Lie group with the left (flat) covariant derivative, all three possible shifting methods agree.) More generally, suppose that $M = G/H$ is a homogeneous space with the metric and connection defined as in Example 8.1(d). It has already been shown in Example 8.1(d) that this connection is torsion skew symmetric (TSS). Given a function h: $[0, 1] \to \mathfrak{g}$, define the s-dependent vector field (X) by $X(s)(m) = (d/dt)|_0 \, p(e^{th(s)}g)$, where $m \in M$ and g is any element in $p^{-1}(\{m\})$ and p: $G \to M$ is the canonical projection. So $X(s)$ is the generator of the 1-parameter flow given by the left action of $e^{th(s)}$ on M. Now by definition of the metric on M, each element of G acts isometrically on M, from which it follows that $X(s)$ is a Killing vector field. Therefore, in this case the shift $e^{tX(s)}(\sigma_o(s)) = e^{th(s)} \cdot \sigma_o(s)$ does have the quasi-invariance property provided that $h(0) = 0$ and h' is L^2 integrable.

The two examples in the last paragraph were studied in [MM1, Sh1, Sh2]. In [MM1] it is also shown that the above flows have the quasi-invariance property with respect to any Brownian bridge measure on $W(M)$ provided, of course, that h also satisfies $h(1) = 0$, The reader should also see [AH, Fr, Gr4] on the question of quasi-invariance in the case of compact Lie groups.

Acknowledgments

It is a pleasure to acknowledge many useful conversations with T. Frankel, L. Gross, and especially P. Fitzsimmons and Ruth Williams. I also thank T. Lohrenz for bringing to my attention a number of P. Malliavin's papers.

References

[AM1] H. Airault and P. Malliavin, Intégration géometrique sur l'espace de Wiener, *Bull. Sci. Math. (2)* **112** (1988), 3–52.

[AM2] H. Airault and P. Malliavin, Integration on loop group III, heat equation for the Wiener measure, *J. Funct. Anal.*, in press.

[AV1] H. Airault and J. Van Biesen, Géometrie riemannienne en codimension finie sur l'espace de Wiener, *C. R. Acad. Sci. Paris* **311** (1990), 125–130.

[AV2] H. Airault and J. Van Biesen, Le processus d'Ornstein–Uhlenbeck sur une sous-variété de l'espace de Wiener, *Bull. Sci. Math. (2)* **115** (1991), 185–210.

[AH] S. Albeverio and R. Hoegh-Krohn, The energy representation of Sobolev Lie groups, *Compositio Math.* **36** (1978), 37–52.

[AM] L. Auslander and R. MacKenzie, "Introduction to Differentiable Manifolds," Dover, New York, 1977.

[Bi1] J.-M. Bismut, Mécanique Aléatoire, *in* "Lecture Notes in Mathematics," Vol. 866 (A. Dold and B. Eckmann, Eds.), Springer, Berlin/Heidelberg/New York, 1981.

[Bi2] J.-M. Bismut, "Large Deviations and the Malliavin Calculus," Birkhäuser, Boston/Basel/Stuttgart, 1984.

[Bo] S. Bochner, Vector fields and Ricci curvature, *Bull. Amer. Math. Soc.* **52** (1946), 776–797.

[Ca] R. H. Cameron, The first variation of an indefinite Wiener integral, *Proc. Amer. Math. Soc.* **2** (1951), 914–924.

[CM1] R. H. Cameron and W. T. Martin, Transformations of Wiener integrals under translations, *Ann. of Math.* **45**, No. 2 (1944), 386–386.

[CM2] R. H. Cameron and W. T. Martin, Transformations of Wiener integrals under a general class of linear transformations, *Trans. Amer. Math. Soc.* **58** (1945), 184–219.

[CM3] R. H. Cameron and W. T. Martin, The transformation of Wiener integrals by non-linear transformations, *Trans. Amer. Math. Soc.* **66** (1949), 253–283.

[Cr] A. B. Cruzeiro, Equations différentielles sur l'espace de Wiener et formules de Cameron–Martin non linéaires, *J. Funct. Anal.* **54** (1983), 206–227.

[D1] B. Driver, YM_2: Continuum expectations, lattice convergence, and lassos, *Comm. Math. Phys.* **123** (1989), 575–616.

[D2] B. Driver, Classifications of bundle connection pairs by parallel translation and lassos, *J. Funct. Anal.* **83** (1989), 185–231.

[D3] B. Driver, A Cameron–Martin type quasi-invariance theorem for pinned Brownian motion on a compact Riemannian manifold, March 1992 preprint, to appear in *Trans. Amer. Math. Soc.*

[EE] Eells and Elworthy, Wiener integration on certain manifolds, *in* "Problems in Non-Linear Analysis" (G. Prodi, Ed.), pp. 67–94, Internazionale Matematico Estivo, IV Ciclo, Tome, Edizioni Cremonese, 1971.

[El1] Ellworthy, Stochastic differential equations on manifolds, *in* "London Mathematical Society Lecture Note Series," Vol. 70, Cambridge Univ. Press, London, 1982.

[EM] M. Emery, "Stochastic Calculus in Manifolds," Springer, Berlin/Heidelberg/New York, 1989.

[EL1] Epperson and Lohrenz, Diffusions of finite-energy loop spaces, preprint, November 1990.

[EL2] Epperson and Lohrenz, Brownian motion on the path space of a compact Lie group, preprint, Rice University, June 1991.

[Fr] I. B. Frenkel, Orbital theory for affine Lie algebras, *Invent. Math.* **77** (1984), 301–352.

[Ge] E. Getzler, Dirichlet forms on loop space, *Bull. Sci. Math. (2)* **113** (1989), 151–174.

[Gi] I. V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, *Theory Probab. Appl.* **5** (1960), 285–301.

[Gr1] L. Gross, Integration and nonlinear transformations in Hilbert space, *Trans. Amer. Math. Soc.* **94** (1960), 404–440.

[Gr2] L. Gross, Abstract Wiener spaces, *in* "Proceedings, 5th Berkeley Symposium Math. Stat. Prob. 2, 1965," pp. 31–42.

[Gr3] L. Gross, Logarithmic Sobolev inequalities for the heat kernel of a Lie group, *in* "White Noise Analysis: Mathematics and Applications" (T. Hida *et al.*, Eds.), pp. 108–130, World Scientific, Singapore/Teaneck/New Jersey, 1990.

[Gr4] L. Gross, Logarithmic Sobolev inequalities on loop groups, *J. Funct. Anal.* **102** (1991), 268–313.

[IW] N. Ikeda and S. Watanabe, "Stochastic Differential Equations and Diffusion Processes," North-Holland, Amsterdam/Oxford/New York, 1981.

[JL] J. D. S. Jones and R. Leandre, L^p-Chen forms on loop spaces, *in* "Stochastic Analysis" (Barlow and Bingham, Eds.), pp. 103–162, Cambridge Univ. Press, London/New York, 1991.

[Kl1] W. Klingenberg, "Lecture on Closed Geodesics," Springer, Berlin/Heidelberg/New York, 1978.

[Kl2] W. Klingenberg, Closed geodesics on Riemannian manifolds, *in* "Regional Conference Series in Mathematics," Vol. 53, Amer. Math. Soc., Providence, RI, 1982.

[Kl3] W. Klingenberg, "Riemannian Geometry," de Gruyter, Berlin/New York, 1982.

[KN] KOBAYASHI AND NOMIZU, "Foundations of Differential Geometry," Vol. 1, Wiley–Interscience, New York/London, 1963.

[K1] H.-H. KUO, Integration theory on infinite-dimensional manifolds, *Trans. Amer. Math. Soc.* **159** (1971), 57–78.

[K2] H.-H. KUO, Diffusion and Brownian motion on infinite dimensional manifolds, *Trans. Amer. Math. Soc.* **169** (1972), 439–451.

[K3] H.-H. KUO, Gaussian measures in Banach spaces, *in* "Lecture Notes in Mathematics," Vol. 463, Springer-Verlag, New York/Berlin, 1975.

[Ku1] S. KUSUOKA, The nonlinear transformation of Gaussian measure on Banach space and its absolute continuity, I, *J. Fac. Sci. Univ. Tokyo, Sect. 1A Math.* **29** (1982), 567–597.

[Ku2] S. KUSUOKA, On the foundations of Wiener–Riemannian manifolds, *in* "Stochastic Analysis, Path Integration and Dynamics: Emanations from Summer Stochastics, Warwick, 1987" (K. D. Elworthy and J.-C. Zambrini, Eds.), Wiley, pp. 130–165, New York, 1989.

[Ku3] S. KUSUOKA, "Analysis on Wiener Spaces. I. Nonlinear Maps," R.I.M.S. Rep. No. 670, October 1989.

[Ku4] S. KUSUOKA, "Analysis on Wiener Spaces. II. Differential Forms," R.I.M.S. Rep. No. 705, July 1990.

[L] S. LANG, "Differential Manifolds," Addison–Wesley, Reading, MA, 1972.

[Le] R. LEANDRE, Integration by parts formulas and rotationally invariant Sobolev calculus on free loop spaces, Strasbourg Preprint May 1992.

[M1] P. MALLIAVIN, Naturality of quasi-invariance of some measures, *in* "Proceedings, Lisbonne Conference" (A. B. Cruzerio, Ed.) Birkhäuser, Basel, 1991.

[M2] P. MALLIAVIN, Hypoellipticity in infinite dimension, *in* "Diffusion Process and Related Problems in Analysis" (M. A. Pinsky, Ed.), Vol. I, Birkhäuser, Basel, 1991.

[M3] P. MALLIAVIN, Diffusion on the loops, *in* "Proceedings, Conference in the Honour of Antoni Zygmund, Chicago, 1981–1983" (W. Beckner and A. Calderon, Eds.).

[MM1] M.-P. MALLIAVIN AND M. MALLIAVIN, Integration on loop groups. I. Quasi invariant measures, quasi invariant integration on loop groups, *J. Funct. Anal.* **93** (1990), 207–237.

[MM2] M.-P. MALLIAVIN AND M. MALLIAVIN, Integration on loop group II, asymptotic Peter–Weyl orthogonality, preprint, 1991.

[MM3] M.-P. MALLIAVIN AND M. MALLIAVIN, La représentation régulière des groupes de lacets, *in* "Proceedings, la conférence du Trientenaire, 1990," Hamburg Mathematic Seminar, Third centenary anniversary of Hamburg University, in press.

[MM4] M.-P. MALLIAVIN AND M. MALLIAVIN, An infinitesimaly quasi invariant measure on the group of diffeomorphisms of the circle, *in* "Proceedings, Hashibara Forum, Lect. Notes, August 1991" (Kashiwara and Miwa, Eds.).

[Mar] G. MARUYAMA, Notes on Wiener integrals, *Kodai Math. Seminar Rep.* **3** (1950), 41–44.

[Me] P. A. MEYER, A differential geometric formalism for the Itô calculus, *in* "Stochastic Integrals, Proceedings of the LMS Durham Symposium" (D. Williams, Ed.), pp. 256–269, Lectures Notes in Math., Vol. 851, Springer, Berlin/Heidelberg/New York, 1980.

[MS] J. W. MILNOR AND J. D. STASHEFF, "Characteristic Classes," Princeton Univ. Press, Princeton, NJ, 1974.

[No] J. R. NORIS, A complete differential formalism for stochastic calculus in manifolds, preprint, University of Cambridge, 1991.

[PS] A. PRESSLEY AND G. SEGAL, "Loop Groups," Oxford Univ. Press, Oxford/New York/Toronto, 1986.

[Pr] P. PROTTER, "Stochastic Integration and Differential Equations; A New Approach," Springer, Berlin/Heidelberg/New York, 1990.

[Ra] R. RAMER, On nonlinear transformations of Gaussian measures, *J. Funct. Anal.* **15** (1974), 166–187.

[RY] REVUZ AND YOR, "Continuous Martingales and Brownian Motion," Springer, Berlin/Heidelberg/New York, 1991.

[RW] L. C. G. ROGERS AND D. WILLIAMS, "Diffusions, Markov processes, and Martingales. Vol. 2. Itô Calculus," Wiley, Chichester/New York, 1987.

[Ru] W. RUDIN, "Real and Complex Analysis," 2nd ed., McGraw–Hill, New York, 1976.

[Sc1] L. SCHWARTZ, Semi-martingales sur des variétés et martingales conformes sur des variétés analytiques complexes, *in* "Lecture Notes in Mathematics," Vol. 780, Springer, New York/Berlin, 1980.

[Sc2] L. SCHWARTZ, Géométrie différentielle du 2^e order, semimartingales et équations différentielles stochastiques sur une variété différentielle, *in* "Seminaire de Probabilities XVI," Lecture Notes in Mathematics, Vol. 921, Springer, New York/Berlin, 1982.

[Sc3] L. SCHWARTZ, "Semimartingales and Their Stochastic Calculus on Manifolds," Presses de l'Université de Montréal, 1984.

[Sh1] ICHIRO SHIGEKAWA, Transformations of the Brownian motion on the Lie group, *in* "Stochastic Analysis: Proceedings of the Taniguchi International Symposium on Stochastic Analysis" (Katata and Kyoto, 1982, K. Itô, Ed., Japan), North–Holland, pp. 409–422, Amsterdam/New York, 1984.

[Sh2] ICHIRO SHIGEKAWA, Transformations of the Brownian motion on the Riemannian symmetric space, *Z. Wahr. verw. Geb.* **65** (1984), 493–522.

[Sp] M. SPIVAK, "A Comprehensive Introduction to Differential Geometry," 2nd ed., Vol. 1, Publish of Perish, Wilmington, DE, 1979.

[St1] D. W. STROOCK, "Lectures on Topics in Stochastic Differential Equations" (lectures given at Tata Institute of Fundamental Research, Bombay, 1982), Springer, Berlin/Heidelberg/New York, 1982.

[St2] D. W. STROOCK, "Lectures on Stochastic Analysis: Diffusion Theory," Cambridge Univ. Press, Cambridge/London/New York, 1987.

[W] H.-H. WU, The Bochner technique in differential geometry, *in* "Mathematical Reports" (J. Dieudonne, Ed.), Harwood Academic Publishers, Chur/London/Paris/New York/Melbourne, 1988.