INTEGRATION BY PARTS FOR HEAT KERNEL MEASURES REVISITED

By Bruce K. DRIVER (*)

ABSTRACT. - Stochastic calculus proofs of the integration by parts formula for cylinder functions of parallel translation on the Wiener space of a compact Riemannian manifold (M) are given. These formulas are used to prove a new probabilistic formula for the logarithmic derivative of the heat kernel on M. This new formula is well suited for generalizations to infinite dimensional manifolds.

Contents

1. Introduction .................................. 704
2. The Euclidean Prototypes ........................ 705
   2.1. Cameron’s Integration by Parts Formula .......... 706
   2.2. Notation .................................. 706
   2.3. Heat Kernel Integration by Parts .................. 709
3. Integration by Parts for “Curved” Wiener Space .......... 710
   3.1. Differential Geometric Preliminaries ............... 710
   3.2. The Basic Processes .......................... 712
   3.3. Another Integration by Parts Argument ............ 713
   4.1. Backwards Integrals .......................... 725
5. Bismut’s formula .............................. 726
6. Applications to Compact Lie Groups with Left Invariant Metrics .......... 728
7. Appendix: Geometric Identities ........................ 733

(*) This research was partially supported by NSF Grant nos. DMS 92 - 23177 and DMS 96-12651. AMS 1991 Mathematics Subject Classification. Primary 60 H 07, 60 H 30; Secondary 58 D 20.

JOURNAL DE MATHEMATIQUES PURES ET APPLIQUEES. – 0021-7824/1997/08$ 7.00
© Gauthier-Villars
1. Introduction

Let $(M^d, \langle \cdot, \cdot \rangle, \nabla, o)$ be given, where $M$ is either $\mathbb{R}^d$ or a compact connected manifold (without boundary) of dimension $d$, $\langle \cdot, \cdot \rangle$ is a Riemannian metric on $M$, $\nabla$ is a metric-compatible covariant derivative, and $o$ is a fixed base point in $M$. Let $T = T^\nabla$ and $R = R^\nabla$ denote the torsion and curvature of $\nabla$ respectively. If $M = \mathbb{R}^d$, we take $\langle \cdot, \cdot \rangle$ to be the usual dot product on $\mathbb{R}^d$ and $\nabla$ to be the Levi-Civita covariant derivative. In all cases we will assume that $\nabla$ is Torsion Skew Symmetric or TSS for short, i.e. that $\langle T(\cdot, X), Y \rangle$ is skew symmetric for all vector fields $X$ and $Y$ on $M$. Let $\{\Sigma_t\}_{t \geq 0}$, $\{\mathbb{I}_t\}_{t \geq 0}$, and $\{b(t)\}_{t \geq 0}$ be three adapted continuous processes on a filtered probability space such that $\Sigma$ is an $M$-valued Brownian motion, $\mathbb{I}$ is stochastic parallel translation along $\Sigma$, and $b(t) = \int_0^t \mathbb{I}_s^{-1} \delta \Sigma_s - \int_0^t \delta \Sigma_s$ a $T_0 M$-valued Brownian motion. See sections 3.1 and 3.2 for more details on this notation.

For any finite dimensional inner product space $V$, let $H(V)$ denote the Cameron-Martin Hilbert space of paths $h : [0, \infty) \to V$ which are absolutely continuous and satisfy

$$\langle h, h \rangle = \int_0^\infty |\dot{h}(t)|^2 dt < \infty.$$ 

Given $h \in H(T_0 M)$ let $X^h$ denote the Cameron-Martin vector field on $W(M)$ given by $X^h_t \equiv \mathbb{I}_t h(t)$. It was shown in [7] that $X^h$ may indeed be considered as a vector field on $W(M)$ in the sense that $X^h$ generates a quasi-invariant flow, at least when $h$ is $C^1$. This theorem was extended by Hsu [21, 22] to include all $h \in H(T_0 M)$. Also see Norris [33], Enchev and Stroock [17], and Lyons and Qian [29] for other approaches.

It was also shown in [7] (see Theorem 9.1 on p. 363 where $X^h$ was written as $\partial_h$) that $X^h$ may be viewed as a densely defined closed operator on the path space of $M$. This last result relies on an integration by parts formula which, in the special case of $X^h$ acting on functions of the form $f(\sigma) = F(\sigma(t))$, is due to Bismut [4]. Moreover, by Proposition 4.10 in Driver [8] (also see Enchev and Stroock [17]), it was shown that these integration by parts formula extend to cylinder functions of the parallel translation process, $\mathbb{I}$. In this paper we will give another elementary proof of the integration by parts formula for cylinder functions of $\mathbb{I}$. As a corollary (see Theorem 4.1 and Corollary 4.3 below) we find the following integration by parts formula. Let $Y$ be a smooth vector field on $M$ and $l : [0, T] \to \mathbb{R}$ be an absolutely continuous function such that $l(0) = 0$, $l(T) = 1$, and $\int_0^T l^2(t) dt < \infty$, then

$$E[(\nabla \cdot Y)(\Sigma_T)] = E\left[\left(\mathbb{I}_T^{-1} Y(\Sigma_T), \int_0^T \left(l(t) - \frac{1}{2} l(t)\text{Ric}_{\mathbb{I}}\right) \mathbb{I} \dot{b}(t)\right)\right].$$

In this formula $\nabla \cdot Y$ is the divergence of $Y$, $\text{Ric}_{\mathbb{I}} = \mathbb{I}^{-1} \text{Ric}_{\mathbb{I}}$, $\text{Ric}$ is the Ricci tensor and $\mathbb{I} \dot{b}$ denotes the backwards Itô differential.

There have been numerous proofs and extensions of integration by parts formulas on $W(M)$. See, for example, [1, 2, 15, 18, 19, 20, 27, 28, 30, 31, 33] and the references therein for some of the more recent articles. Moreover, there are many results in the literature closely related to Eq. (1.1), see for example [3, 4, 12, 14, 15, 20, 23, 32, 34, 35, 36, 37] and the references therein.
In order to explain the relationship of Eq. (1.1) to the current literature it is useful to rewrite the left hand side of this equation. Let \( o = \Sigma_0 \) and \( p_t(x, y) \) denote the heat kernel on \( M \), then:

\[
E[(\nabla \cdot Y)(\Sigma_T)] = \int_M p_T(o, x) \nabla \cdot Y(x) \, dx
\]

\[
- \int_M \langle \nabla_x p_T(o, x), Y(x) \rangle \, dx
\]

\[
= - \int_M \langle \nabla_x \ln p_T(o, x), Y(x) \rangle \, p_T(o, x) \, dx
\]

\[
= -E[(\langle \nabla \ln p_T(o, \cdot)(\Sigma_T), Y(\Sigma_T) \rangle].
\]

where \( \nabla f \) is used to denote the gradient of \( f \). Therefore, if we were to “condition” Eq. (1.1) on the set where \( \Sigma_T = x \) (\( x \in M \) is a fixed point), we would learn that

\[
(1.2) \quad \langle \nabla_x \ln p_T(o, x), v \rangle = -E\left[ \left\langle \frac{1}{2} \nabla_t v, \int_0^T \left( \frac{1}{2} l(t) - \frac{1}{2} l(t) \text{Ric} \right) \delta b(t) \right\rangle \Sigma_T = x \right].
\]

where \( v = Y(\Sigma_T) \in T_x M \). Hence, we see that Eq. (1.1) gives a probabilistic representation for the logarithmic derivative, in the second variable, of the heat kernel \( p_T(\cdot, \cdot) \). Whereas Bismut’s [4] formula (see Theorem 5.1 below and literature cited above) gives a similar representation for the logarithmic derivative, in the first variable, of the heat kernel. Of course, since \( p_T(o, x) \) is symmetric in \( o \) and \( x \), one may obviously obtain a probabilistic formula for \( \nabla_x \ln p_T(o, x) \) from Bismut’s formula.

As the above discussion indicates, when \( M \) is a finite dimensional Riemannian manifold there is no advantage of Eq. (1.2) over Bismut’s formula, Eq. (4.17) below. However, Eq. (1.1) is better suited for the purpose of generalization to the case where the finite dimensional manifold \( M \) is to be replaced by an infinite dimensional manifold. Indeed, if \( M \) is infinite dimensional one can no longer hope to write the law of \( \Sigma_T \) as a density times “Lebesgue” measure because Lebesgue measure will not exist. Thus for each \( o \in M \), we must view \( p_T(o, \cdot) \) as a measure and here we will lose the symmetry between the two arguments in \( p_T \).

In fact, the results in this paper were motivated by the desire to find integration by parts formulas for the heat kernel measures on (infinite dimensional) spaces of loops into a compact Lie group. For applications of the results in this paper to loop groups the reader is referred to Driver [10].

### 2. The Euclidean Prototypes

As a warm up for the next two sections, we will recall Cameron’s integration by parts formula on classical Wiener space along with an “elementary” stochastic calculus proof. The method of proof used here and the next section has already been described by Elton Hsu [24].
2.1. Cameron’s Integration by Parts Formula

2.2. Notation

Let \( \{b(t)\}_{t \geq 0} \) be a \( \mathbb{R}^d \)-valued standard Brownian motion on a filtered probability space \( (\mathcal{W}, \{\mathcal{F}_t\}_{t \geq 0}, \mathcal{F}, P) \). Assume that \( (\mathcal{W}, \{\mathcal{F}_t\}_{t \geq 0}, \mathcal{F}, P) \) satisfies the usual hypothesis, i.e. \( \mathcal{F} \) is \( P \)-complete, \( \mathcal{F}_t \) contains all of the null sets of \( \mathcal{F} \), and the filtration \( \{\mathcal{F}_t\}_{t \geq 0} \) is right continuous. Let \( H(\mathbb{R}^d) \) be the Cameron-Martin Hilbert space of the absolutely continuous functions \( k : [0, \infty) \to \mathbb{R}^d \) such that \( k(0) = 0 \) and

\[
(k, k) = \int_0^\infty |\dot{k}(t)|^2 \, dt < \infty.
\]

Given a partition \( \mathcal{P} = \{0 < t_1 < t_2 < \cdots < t_n = T\} \) of \([0, T]\) and \( f \in C^\infty_c((\mathbb{R}^d)^n) \) let \( f_\mathcal{P} : \mathcal{W} \to \mathbb{R} \) be defined by

\[
(f_\mathcal{P})(b(t_1), b(t_2), \ldots, b(t_n))
\]

We will call such a function \( f_\mathcal{P} \) on \( \mathcal{W} \) a cylinder function of \( b \).

**NOTATION 2.1.** Given a differentiable operator \( A \) acting on \( C^\infty_c((\mathbb{R}^d)^n) \), for \( i \in \{1, 2, \ldots, n\} \) and \( f \in C^\infty_c((\mathbb{R}^d)^n) \) let

\[
(A_i f)(x_1, x_2, \ldots, x_n) \equiv (Af(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n))(x_i),
\]

where \( x_i \in \mathbb{R}^d \) for each \( i \). That is \( A_i f \) denotes \( A \) acting on the \( i \)th variable of \( f \). If \( f_\mathcal{P} \) is a cylinder function as in Eq. (2.1), we will abuse notation and write \( A_i f_\mathcal{P} \) for \( (A_i f)(b(t_1), b(t_2), \ldots, b(t_n)) \). In particular if \( n = 1 \), then we will write \( A f_\mathcal{P} \) for \( (A f)(b(T)) \).

For \( f \in C^\infty_c(\mathbb{R}^d) \), let \( Df \in C^\infty_c((\mathbb{R}^d)^*) \) denote the differential of \( f \), i.e. \( Df(x) a \equiv \frac{d}{ds} |_{s=0} f(x + sa) \). Notice that \( D : C^\infty_c(\mathbb{R}^d) \to C^\infty_c((\mathbb{R}^d)^*) \) is a differential operator and hence

\[
(D f)(x_1, x_2, \ldots, x_n) a \equiv \frac{d}{ds} |_{s=0} f(x_1, \ldots, x_{i-1}, x_i + sa, x_{i+1}, \ldots, x_n).
\]

For \( k \in H(\mathbb{R}^d) \) and \( f_\mathcal{P} \) as above, set

\[
X^k f_\mathcal{P} = \sum_{i=1}^n (D_i f)(b(t_1), b(t_2), \ldots, b(t_n)) k(t_i).
\]

The following lemma is a simple application of Itô’s Lemma and will be left to the reader.

**LEMMA 2.2.** Let \( T > 0 \), \( \Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2} \) be the Laplacian on \( \mathbb{R}^d \) and \( f \) be a smooth vector valued function on \( \mathbb{R}^d \) with compact support. Define

\[
(e^{t \Delta/2} f)(x) = \begin{cases} 
\int_{\mathbb{R}^d} p_t(x - y) f(y) \, dy & \text{if } t > 0, \\
f(x) & \text{if } t = 0,
\end{cases}
\]
and for all \( t \in [0,T] \), \( N_t \equiv (e^{(T-t)\Delta/2}f)(b(t)) \) and
\[
W_t \equiv (De^{(T-t)\Delta/2}f)(b(t)) = (e^{(T-t)\Delta/2}Df)(b(t)).
\]

(Here \( p_t(x) \equiv (2\pi t)^{-d/2} \exp(-|x|^2/2t) \) is the convolution heat kernel on \( \mathbb{R}^d \).) Then \( N_t \) and \( W_t \) are \( L^2 \)-martingales for \( t \in [0,T] \) and \( dN_t = W_t db(t) \).

The following theorem is well known, see Cameron [5]. In order to illustrate the methods to be used in the next section we will not give Cameron's original proof which is based on the Cameron-Martin theorem but instead give a (harder) "stochastic calculus" proof.

**Theorem 2.3.** Let \( k \in H(\mathbb{R}^d) \) and \( f_p b \) be a cylinder function as in Eq. (2.1), then
\[
E[X^k f_p] = E\left[ f_p \int_0^T \langle \dot{k}, db \rangle \right].
\]

Moreover, if \( f_p \) and \( g_p \) are two cylinder functions,
\[
E[(X^k f_p)g_p] = E\left[ f_p \left( -X^k g_p + \int_0^T \langle \dot{k}, db \rangle g_p \right) \right].
\]

**Proof.** Let us first notice that Eq. (2.4) is a consequence of Eq. (2.3) with \( f_p \) replaced by \( f_p g_p = (fg)_p \) and the product rule, \( X^k(f_p g_p) = (X^k f_p)g_p + f_p(X^k g_p) \). Therefore we need only consider the proof of Eq. (2.3). Since the proof here is to illustrate the ideas to come, let us only prove Eq. (2.3) for partitions of the form \( P = \{0 < T\} \) and \( P = \{0 < u < T\} \).

**Case 1.** Suppose \( f \in C^\infty_c(\mathbb{R}^d) \), \( f_p = f(b(T)) \), and \( N_t \) and \( W_t \) are defined as in Lemma 2.2. Then \( X^k f_p = Df(b(T))k(T) = W_T k(T) \), so that
\[
E[X^k f_p] = E[W_T k(T)] = E\left[ f_p \int_0^T \langle \dot{k}, db \rangle \right].
\]

wherein the last equality we used the fact that \( W \) is an \( L^2 \)-martingale and \( k \) is bounded on \([0,T]\). Using the \( L^2 \)-isometry property of the Itô integral and Lemma 2.2,
\[
E\left[ \int_0^T W_t \dot{k}(t) dt \right] = E\left[ \int_0^T W db \int_0^T \langle \dot{k}, db \rangle \right] = E\left[ \int_0^T \langle N_T - N_0, \dot{k}, db \rangle \right] = E\left[ \int_0^T \langle k, db \rangle \right],
\]

since \( N_0 = Ef_p \) is a constant and \( N_T - f_p \). Combining Eq. (2.5) and (2.6) proves Eq. (2.3) in the case that \( P = \{0 < T\} \).

**Case 2.** Suppose \( P = \{0 < u < T\} \), \( f \in C_c((\mathbb{R}^d)^2) \), and \( f_p = f(b(u), b(T)) \). Now let
\[
f_T(t, x, y) = e^{(T-t)\Delta x} f(x, y) \equiv \int_{\mathbb{R}^d} p_{(T-t)}(y-z)f(x, z)dz,
\]
\[
N_t \equiv f_T(t, b(u), b(t)) \text{ for all } t \in [u, T]
\]
and $W_t \equiv D_2 f_T(t, b(u), b(t))$, where $D_2$ denote the differential of $f_T(t, x, y)$ in the $y$-variable. Again a short computation with Itô's lemma shows that $W_t$ and $N_t$ are $L^2$-martingales for $t \in [u, T]$ and that $dN_t = W_t \, db(t)$.

Now

$$X^k f_T = D_1 f(b(u), b(T)) k(u) + D_2 f(b(u), b(T)) k(T)$$

$$= D_1 f(b(u), b(T)) k(u) + W_T k(T),$$

so that

$$E[X^k f_T] = E[D_1 f(b(u), b(T)) k(u) + W_T k(T)]$$

$$= E[D_1 f(b(u), b(T)) k(u)] + E \left[ W_u k(u) + \int_u^T \{ dW + W \, dk \} \right]$$

$$= E[D_1 f_T(u, b(u), b(u)) k(u)] + E \left[ W_u k(u) + \int_u^T \int_u^T W_s (t) \, dt \right].$$

wherein we have used the fact the $W_t$ is $L^2$-martingale and the Markov property to show $E[D_1 f(b(u), b(T)) k(u)] = E[D_1 f_T(u, b(u), b(u)) k(u)]$. As in Eq. (2.6),

$$E \left[ \int_u^T W_s (t) \, dt \right] = E \left[ (N_T - N_u) \int_u^T \langle k, db \rangle \right] = E \left[ f_T \int_u^T \langle k, db \rangle \right].$$

Since $W_u k(u) = D_2 f_T(u, b(u), b(u)) k(u)$, combining the two above displayed equations gives:

$$E[X^k f_T] = E[D_1 f_T(u, b(u), b(u)) k(u)] + E \left[ f_T \int_u^T \langle k, db \rangle \right].$$

Letting $\tilde{f} \equiv f_T(u, b(u), b(u))$ and noting that

$$X^k \tilde{f} = D_1 f_T(u, b(u), b(u)) k(u) + D_2 f_T(u, b(u), b(u)) k(u),$$

Eq. (2.7) may be written as

$$E[X^k f_T] = E[X^k \tilde{f}] + E \left[ f_T \int_u^T \langle k, db \rangle \right].$$

By case 1) already proved,

$$E[X^k \tilde{f}] = E \left[ \tilde{f} \int_0^u \langle k, db \rangle \right] - E \left[ f_T(u, b(u), b(u)) \int_0^u \langle k, db \rangle \right]$$

$$= E \left[ f(b(u), b(T)) \int_0^u \langle k, db \rangle \right] - E \left[ f_T \int_0^T \langle k, db \rangle \right],$$

wherein the third inequality we have used the Markov property of $b$. Combining equations (2.8) and (2.9) gives (2.3). The general case can be proved similarly using induction on the number of partition points in $P$. Q.E.D.
2.3. Heat Kernel Integration by Parts

Let \( Y : \mathbb{R}^d \to \mathbb{R}^d \) be a smooth vector field on \( \mathbb{R}^d \) and \( \nabla \cdot Y \) be its divergence. Assume for simplicity that \( Y \) has compact support. A simple integration by parts shows that

\[
\int_{\mathbb{R}^d} (\nabla \cdot Y)(x)p_T(x) \, dx = - \int_{\mathbb{R}^d} Dp_T(x)Y(x) \, dx = \frac{1}{T} \int_{\mathbb{R}^d} p_T(x)\langle Y(x), x \rangle \, dx
\]

or equivalently that,

\[
E[(\nabla \cdot Y)(b(T))] = \frac{1}{T} E[(Y(b(T)), b(T))].
\]  

(2.10)

In Section 4 we will derive an analogous formula in the case where \( \mathbb{R}^d \) is replaced by a compact manifold \( M \). In order to illustrate the proofs that will be given in section 4, I will give two alternate (harder) proofs of Eq. (2.10).

**First Alternative Proof of Equation (2.10).** Choose an orthonormal basis \( S \) for \( \mathbb{R}^d \) and for \( c \in S \) set \( h_c(t) \equiv \frac{1}{T}c, \ X_t^c \equiv X_t^c \equiv h_c(t), \) and

\[
z_t^c \equiv \int_0^t \langle h_c, db \rangle = \frac{1}{T} \langle c, b(t) \rangle, \quad \forall t \in [0, \infty).
\]

Let \( 1 \) denote the constant cylinder function “one” of the Brownian motion \( b \). Since \( X^c 1 = 0, \)

\[
0 = \sum_{c \in S} E[(X^c 1)(X_T^c, Y(b(T)))] = \sum_{c \in S} E[(X^c)^*(X_T^c, Y(b(T)))]
\]

\[
= \sum_{c \in S} E[(-X^c + z_T^c)(c, Y(b(T)))]
\]

\[
= \sum_{c \in S} E[-X^c(c, Y(b(T)))] + \frac{1}{T} E(Y(b(T), b(T)),
\]

wherein the second equality we have used Eq. (2.4). This proves Eq. (2.10), since after unwinding the definitions we have that

\[
\sum_{c \in S} X^c(c, Y(b(T))) = (\nabla \cdot Y)(b(T)).
\]

Q.E.D.

**Second Alternative Proof of Equation (2.10).** For \( t \in [0, T], \) set

\[
Y_t(x) \equiv (e^{\frac{t}{T}}Y)(x) \equiv \int_{\mathbb{R}^d} p(T-t)(x-y)Y(y) \, dy
\]
and

\begin{equation}
Q(t) = t(Y_t(b(t)) - (b(t), Y_t(b(t))).
\end{equation}

Using \( \frac{\partial Y_t}{\partial t} = -\frac{1}{2} \Delta Y_t \), \( \frac{\partial (\nabla \cdot Y_t)}{\partial t} = -\frac{1}{2} \Delta (\nabla \cdot Y_t) \) \((^1)\), and Itô's lemma, we find

\[
dQ_t = (\nabla \cdot Y_t)(b(t))dt + (\nabla \cdot Y_t)(b(t)) - (d(b(t), Y_t(b(t)))) - (d(db(t), Y_t(b(t)))) - (db(t), (\nabla Y_t)(b(t))).
\]

Therefore, \( Q_t \) is an \( L^2 \)-martingale and in particular \( EQ_T = EQ_0 \). This proves Eq. (2.10), since \( Q_T = T(Y_t(b(T)) - b(T), Y(b(T))) \) while \( Q_0 = 0 \). Q.E.D.

3. Integration by Parts for "Curved" Wiener Space

3.1. Differential Geometric Preliminaries

Let \((M^d, \langle \cdot, \cdot \rangle, \nabla, o)\) be given, where \( M \) is a compact connected manifold (without boundary) of dimension \( d \), \( \langle \cdot, \cdot \rangle \) is a Riemannian metric on \( M \), \( \nabla \) is a \( \langle \cdot, \cdot \rangle \)-compatible covariant derivative, and \( o \) is a fixed base point in \( M \). Let \( T = T^\nabla \) and \( R = R^\nabla \), denote the torsion and curvature of \( \nabla \) respectively. So

\[
R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z
\]

and

\[
T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]
\]

for all smooth vector fields \( X \) and \( Y \) on \( M \).

**Standing Assumption:** The covariant derivative (\( \nabla \)) is assumed to be Torsion Skew Symmetric or TSS for short. That is to say \( \langle T(X, Y), Y \rangle \equiv 0 \) for all vector fields \( X \) and \( Y \) on \( M \).

The Laplacian (\( \Delta \)) with respect to (\( \nabla \)) is the second order differential operator acting on the smooth functions \( f \in C^\infty(M) \) given by \( \Delta f = tr\nabla df = \sum_{j=1}^n \langle E_j E_i f - df \langle \nabla E_j, E_i \rangle \rangle \}

where \( \{E_i\}_{i=1}^n \) is a local orthonormal frame. We recall from Driver [6] that this Laplacian is the same as the Levi-Civita Laplacian due to the TSS condition on (\( \nabla \)).

For \( m \in M \), let \( O_m(M) \) denote the collection of linear isometries of \( T_o M \) to \( T_m M \), \( O(M) \equiv \bigcup_{m \in M} M \), and \( \pi : O(M) \rightarrow M \) denote the projection map, \( \pi(O_m(M)) = \{m\} \) for all \( m \in M \). The principal bundle \( \pi : O(M) \rightarrow M \) is a convenient representation for the orthogonal frame bundle over \( M \). The structure group of this bundle is the group \( O(T_o M) \) of isometries of \( T_o M \). Let \( so(T_o M) \) be the Lie algebra of \( O(T_o M) \) consisting of skew-symmetric linear transformations on \( T_o M \).

\(^1\) We have used the fact that the Laplacian and the divergence operator commute on \( \mathbb{R}^n \). For general Riemannian manifolds this is not the case. It is at this point that the geometry enters into the argument.
Given smooth paths $u$ in $O(M)$ and $\sigma$ in $M$ such that $u(s) \in O_{\sigma(s)}(M)$, let $\nabla u(s)/ds : T_0 M \to T_{\sigma(s)} M$ denote the linear operator defined by $(\nabla u(s)/ds)a = \nabla(u(s)a)/ds$ for all $a \in T_0 M$. Notice that $V(s) = u(s)a$ is a vector field along $\sigma$ so that $\nabla V(s)/ds = \nabla(u(s)a)/ds$ makes sense.

**Definition 3.1 (Canonical 1-form).** Let $\theta$ be the $T_0 M$-valued 1-form on $O(M)$ given by $\theta(\xi) \equiv (\pi(\xi))^{-1}\pi_*\xi$. In particular, if $s \to u(s) \in O(M)$ is a smooth path and $\sigma(s) = \pi(u(s))$ then $\theta(u'(s)) = u(s)^{-1}\sigma'(s)$.

**Definition 3.2 (Connection 1-form).** Let $\omega$ be the $so(T_0 M)$-valued connection 1-form on $O(M)$ given by $\omega(u'(0)) \equiv u(0)^{-1}\nabla u(s)/ds|_{s=0}$, where $s \to u(s)$ is any smooth curve in $O(M)$. Notice that a path $u$ is parallel or horizontal in $O(M)$ iff $\omega(u') \equiv 0$.

**Definition 3.3 (Horizontal Vector Fields).** For $a \in T_0 M$ and $u \in O(M)$ let $B_a(u) \in T_u O(M)$ be defined by $\omega(B_a(u)) = 0$ (i.e. $B_a(u)$ is horizontal) and $\pi_* B_a(u) = ua$.

Let $S$ be an orthonormal basis for $T_0 M$. Let $L$ denote the flat or Bochner Laplacian on $O(M)$, $L \equiv \sum_{c \in S} B_c^2$. A key fact about the Bochner Laplacian is that for all $f \in C^\infty(M)$, $L(f \circ \pi) = (\Delta f) \circ \pi$.

Given a complete vector field $Z$ on a manifold $Q$, let $e^{tZ}$ denote the flow of $Z$. So for each $t \in \mathbb{R}$, $e^{tZ} : Q \to Q$ is diffeomorphism and if $\sigma(t) = e^{tZ}(q)$ for some $q \in Q$, then $d\sigma(t)/dt = Z(\sigma(t))$ with $\sigma(0) = q$.

**Notation 3.4.** For $a, b \in T_0 M$ and an isometry $u : T_0 M \to T_m M$ (i.e. $u \in O_m(M)$), define

$$\Omega_u(a, b) = u^{-1}R^u(a, ub),$$

$$\text{Ric}_u a = \sum_{c \in S} \Omega_u(a, c)c,$$

$$\Theta_u(a, b) = u^{-1}T^u(a, ub) \in T_{\sigma(s)} M$$

and

$$\tilde{\Theta}_u a = \sum_{c \in S} (B_c \Theta)_u(a, c),$$

where

$$(B_c \Theta)_u(a, c) \equiv \left. \frac{d}{ds} \right|_0 \Theta_{e^{tB_c}(u)}(a, c).$$

So $\Omega$, Ric , and $\Theta$ are the equi-invariant forms of the curvature tensor, the Ricci tensor, and the torsion tensor respectively. Similarly, $\tilde{\Theta}_u$ is the equi-invariant form of a contraction of $\nabla^u T^u$. It is well known that the Ricci tensor, Ric , is symmetric if $\nabla$ is the Levi-Civita covariant derivative on $M$, i.e. $T^\nabla \equiv 0$. More generally we have the following lemma.

**Lemma 3.5.** Suppose that $\nabla$ is torsion skew symmetric, then

$$\text{Ric}^*_u = \text{Ric}_u + \tilde{\Theta}_u.$$

**Proof.** By Theorem 9.4 of [8] for any $a, b, c \in T_0 M$,

$$\Omega(b, a)c, b) - \Omega(b, c)a, b) - (B_b \Theta(a, c), b) = 0.$$
Using $\Omega(a, b) = -\Omega(b, a)$, $(\Theta(a, c), b) = -\langle \Theta(a, c), b \rangle$ (since $\nabla$ is TSS), and the fact that $\Omega_{\alpha}(a, b) \in \mathfrak{so}(T_{\alpha}M)$, Eq. (3.2) may be written as,

$$\langle \Omega(a, b), c \rangle = \langle \Omega(a, c), b \rangle + \langle B_{\alpha} \Theta(a, b), c \rangle = 0.$$  

Summing the above equation over $b \in S$ yields,

$$\langle \text{Ric}, a, c \rangle = \langle \text{Ric}, c, a \rangle + \langle \Theta, a, c \rangle = 0$$

from which Eq. (3.1) follows. Q.E.D.

3.2. The Basic Processes

**NOTATION**

3.6. Given a (vector valued) semi-martingale $X$, $\delta X$ will denote the Stratonovich differential of $X$ while $dX$ will denote the Itô differential of $X$.

Let $(W, \{\mathcal{F}_{t}\}_{t \geq 0}, \mathcal{F}, \mathcal{P})$ be a filtered probability space satisfying the usual hypothesis, i.e., $\mathcal{F}$ is $\mathcal{P}$-complete, $\mathcal{F}_{t}$ contains all of the null sets of $\mathcal{F}$, and the filtration $\{\mathcal{F}_{t}\}_{t \geq 0}$ is right continuous. We also assume that there are three adapted continuous processes $\{\Sigma_{t}\}_{t \geq 0}$, $\{/t\}_{t \geq 0}$, and $\{b(t)\}_{t \geq 0}$ on $(W, \{\mathcal{F}_{t}\}_{t \geq 0}, \mathcal{F}, \mathcal{P})$ with values in $M$, $O(M)$, and $T_{\alpha}M$ respectively such that $\Sigma$ is a $M$-valued Brownian motion, $/t$ is parallel translation along $\Sigma$, and $b(t)$ is the "undevelopment" of $\Sigma$. To be more precise, we are assuming:

1. $\Sigma_{0} = o$ and $\Sigma$ is a diffusion process on $M$ with generator $\frac{1}{2}\Delta$, i.e. for all $f \in C^\infty(M)$,

$$f(\Sigma_{t}) - f(\Sigma_{0}) - \int_{0}^{t} (\Delta f)(\Sigma_{\tau}) d\tau$$

is a martingale.

2. $/t = I_{t, M}, \pi \circ /t = \Sigma_{t}$ for all $t \geq 0$, and $\int_{0}^{t} \omega(\delta /t) = 0$ is the zero process.

3. $b$ is an $T_{\alpha}M$-valued Brownian motion,

$$b(t) = \int_{0}^{t} \Theta(b /t),$$

and $/t$ solves the Stratonovich stochastic differential equation,

$$\delta /t = B_{bb(t)}(/t) = \sum_{c \in S} B_{c}(/t) \delta b_{c}(t),$$

where $b_{c}(t) \equiv \langle b(t), c \rangle$.

The fact that such processes exist is well known by the works of Eells and Elworthy and Malliavin, see for example Elworthy [13], Kunita [Ku], Malliavin [M0, M1], Emery [16], Theorem 3.3, p. 297 in [7], and also [9]. For later purposes, recall that a process satisfying item 2 above is called **stochastic parallel translation** along $\Sigma$ and that stochastic parallel translation exists uniquely along any $M$-valued semi-martingale $\Sigma$.

Now let $W(M)$ denote the space of continuous paths $\sigma : [0, \infty) \to M$ such that $\sigma(0) = o$ and $H(T_{\alpha}M)$ be the Cameron-Martin space:

$$H(T_{\alpha}M) \equiv \left\{ k : [0, \infty) \to T_{\alpha}M \mid k \text{ is absolutely continuous and } \int_{0}^{\infty} |k(t)|^{2} dt < \infty \right\}.$$
Given $k$ in the Cameron-Martin space $H(T_o M)$, let $X^k$ denote the Cameron-Martin vector field given by $X^k_t \equiv /t k(t)$. (Notice that $\{X^k_t\}_{t \geq 0}$ is a $TM$-valued process on $W$.) It was shown in [D5]: if $W = W(M)$, $\Sigma$ is the canonical process (i.e. $\Sigma_t(\sigma) = \sigma(t)$ for $t \in [0, \infty)$ and $\sigma \in W(M)$), $P$ is Wiener measure on $W(M)$ and $k$ is $C^1$, then $X^k$ may be considered as a vector field on $W(M)$ which generates a quasi-invariant flow. This theorem was extended by Hsu [21, 22] to include all $k \in H(T_o M)$. Also see Norris [33], Enchev and Stroock [17], and Lyons and Qian [29] for other approaches.

It was also shown in [7] (where $X^k$ was written as $\partial_k$), Theorem 9.1, p. 363 , that $X^k$ may be viewed as a densely defined closed operator on $L^2(W(M), P)$. This last result relies on an integration by parts formula which in the special case of $X^k$ acting on functions of the form $f(\sigma) = F(\sigma(t))$ is due to Bismut [4]. There have been numerous proofs and extensions of integration by parts formulas on $W(M)$, see for example [1, 2, 8, 15, 18, 19, 20, 27, 28, 30, 31, 33] and the references therein for some of the more recent articles. In the next subsection, we will give an alternate proof of this integration by parts formula. The method to be used here has also been discussed by Elton Hsu [24]. My motivation for developing this method stems from its application to the situation where $M$ is replaced by an (infinite dimensional) loop group, see Driver [10].

### 3.3. Another Integration by Parts Argument

**NOTATION 3.7.** Suppose that $k \in H(T_o M)$ is given, let

$$
A^k_t \equiv \int_0^t \Omega_{i/j} \langle k(\tau), dB(\tau) \rangle = \int_0^t \Omega_{i/j} \langle k(\tau), db(\tau) \rangle + \frac{1}{2} \sum_{a \in S} \int_0^t (B_a \Omega_{i/j}) \langle k(\tau), a \rangle d\tau.
$$

where

$$
(B_a \Omega_u)(b,c) \equiv \left. \frac{d}{ds} \right|_0 \Omega_{e^{sA}(u)}(b,c).
$$

**NOTATION 3.8.** To each $A \in C^\infty(O(M) \rightarrow so(T_o M))$, let $\hat{A}$ denote the vertical vector field on $O(M)$ defined by

$$
(\hat{A} F)(u) \equiv \left. \frac{d}{ds} \right|_0 F(ue^{sA(u)})
$$

for all $F \in C^\infty(O(M))$. An important special case is when $A \in so(T_o M)$ is a constant function.

**DEFINITION 3.9 (Cylinder Function).** Given $T > 0$, a partition $P \equiv \{0 < t_1 < t_2 < \cdots < t_n = T\}$ of $[0, T]$, and $F \in C^\infty(O(M)^n)$ let $F_P$ denote the function on $W$ defined by

$$
F_P = F(//t_1, //t_2, \cdots, //t_n).
$$

Any function $F_P$ of the form given in Eq. (3.8) will be called a cylinder function of parallel translation or more simply a cylinder function of //$. When given a cylinder function $f$ of $//$ which may be written in the form $F_P$, we will say that $f$ is based on $P$ and the degree of $f$ is less than or equal to $n$. 


NOTATION 3.10. - Given a differential operator $A$ acting on $C^\infty(O(M))$, for $i \in \{1, 2, \ldots, n\}$ and $F \in C^\infty(O(M)^n)$ let

$$(A_i F)(u_1, u_2, \ldots, u_n) \equiv (AF(u_1, \ldots, u_{i-1},', u_{i+1}, \ldots, u_n))(u_i),$$

where $u_i \in O(M)$ for each $i$. That is $A_i F$ denotes $A$ acting on the $i$-th variable of $F$. If $F_P$ is a cylinder function as in Eq. (3.8), we will abuse notation and write $A_i F_P$ for $(A_i F)(/t_1, /t_2, \ldots, /t_n)$. In particular if $n = 1$, then $A F_P = (AF)(/t)$.

Let $D : C^\infty(O(M)) \to C^\infty(O(M)) \to (T_o M \oplus so(T_o M))^*$ be the first order differential operator given by

$$DF(u)(a | A) = ((B_a | \hat{A}) F)(u).$$

Hence by the above notation, $D_i F_P : W \to (T_o M \oplus so(T_o M))^*$ is defined by:

$$D_i F_P a = \left. \frac{d}{ds} \right|_0 F(/t_1, \ldots, /t_{i-1}, e^{sA}(/t_i), /t_{i+1}, \ldots, /t_n)$$

and

$$D_i F_P A = \left. \frac{d}{ds} \right|_0 F(/t_1, \ldots, /t_{i-1}, /t_i e^{sA}, /t_{i+1}, \ldots, /t_n),$$

where $a \in T_o M$ and $A \in so(T_o M)$.

DEFINITION 3.11. - Given $k$ in the Cameron-Martin space $H(T_o M)$, let $X^k$ denote the Cameron-Martin vector field on $W(M)$ given by $X^k_t \equiv /t_k(t)$. We let $X^k$ act on the cylinder function $F_P$ given in (3.8) via

$$(3.11) \quad X^k F_P = \sum_{i=1}^n D_i F_P (k(t_i) - A^k_{t_i}).$$

To motivate this definition, suppose for the moment that $W = W(M)$ and $P$ is Wiener measure on $W$. Recall that $X^k$ should be thought of as a vector field on $W(M)$. Noting that parallel translation $/t$ is an almost everywhere defined function from $W$ to $O(M)$, it makes sense to try to compute the differential $((/t)_* X^k)$ of $/t$ in the direction determined by the vector field $X^k$. The result is

$$(3.12) \quad ((/t)_* X^k) = B_{k(t)}(/t) - \hat{A}^k_t(/t),$$

see Theorem 2.2 on p. 282 of [7] for a proof of (3.12) for smooth paths and the proof of Theorem 5.1 on p. 320 of [7] for the stochastic version. So if $F \in C^\infty(O(M))$,

$$(3.13) \quad X^k F((/T)) = ((/T)_* X^k) F = (DF)(/T)(k(T) - A^k_T).$$

Clearly, Eq. (3.11) is the natural generalization of Eq. (3.13) to cylinder functions of $/t$ of degree greater than one.

The main result of this section is the following integration by parts theorem (see Theorem 9.1 in [7]) which will be proved at the end of this section.
THEOREM 3.12. Let $k \in H(T_0 M)$ and $F_\tau$ be a cylinder function of $//$, as in Eq. (3.8), then

\begin{equation}
E[X^k F_\tau] = E[z_T F_\tau],
\end{equation}

where

\begin{equation}
z_t = \int_0^t \left\langle \left\{ \dot{k}(\tau) + \frac{1}{2} \text{Ric}_{//}^* k(\tau) \right\}, db(\tau) \right\rangle.
\end{equation}

Let $G_T$ be the $\sigma$-subalgebra of $\mathcal{F}_T$ which is the completion of the $\sigma$-algebra generated by $\{///\}_{\tau \in [0, T]}$. Note that one would get the same $\sigma$-algebras if $//$, was replaced by $b(t)$ or $\Sigma_t$ above.

COROLLARY 3.13. For each $T > 0$ and $k \in H(T_0 M)$, $X^k$ may be considered to be a densely defined closable operator on $L^2(W, G_T, P)$ with domain $\mathcal{D}(X^k)$ consisting of the cylinder functions based on $[0, T]$. Moreover, $-X^k + z_T \in (X^k)^*$, where $(X^k)^*$ denotes the $L^2(W, G_T, P)$-adjoint of $X^k$.

Proof. Let $F_\tau$ and $G_\tau$ be two cylinder functions of $//$. Upon noting that $X^k(F_\tau G_\tau) = (X^k F_\tau)G_\tau + F_\tau X^k G_\tau$, Theorem 3.12 with $F_\tau$ replaced by $(F_\tau G_\tau)$ implies

\begin{equation}
E[(X^k F_\tau)G_\tau] = E[F_\tau (-X^k + z_T) G_\tau].
\end{equation}

Suppose that $F_\tau$ and $H_\tau$ are two cylinder functions based on $[0, T]$ ($\mathcal{P}'$ is possibly another partition of $[0, T]$) such that $F_\tau = H_\tau$ a.e. Let $\mathcal{P}$ be any partition of $[0, T]$ such that $\mathcal{P} \cup \mathcal{P}' \subset \mathcal{P}$ and $G_{\mathcal{P}}$ be any cylinder function of $//$ based on $\mathcal{P}$. Then by Eq. (3.16) with $F_\tau$ replaced by $F_\tau - H_\tau$, and $G_{\mathcal{P}}$ replaced by $G_{\mathcal{P}}$, $G_{\mathcal{P}}$ which is $G_{\mathcal{P}}$ as above are dense in $L^2(W, G_{\mathcal{P}}, P)$, it follows that $X^k F_\tau - X^k H_\tau = 0$ a.e. This shows that $X^k$ is well defined as an operator on $L^2(W, G_{\mathcal{P}}, P)$ with domain $\mathcal{D}(X^k)$. Given this, the assertion that $-X^k + z_T \in (X^k)^*$ follows directly from Eq. (3.16). This in turn shows that $(X^k)^*$ is densely defined which implies that $X^k$ is closable. Q.E.D.

Before beginning the proof of Theorem 3.12, let us recall some facts about the degenerate parabolic partial differential equation,

\begin{equation}
\frac{\partial G(t, u)}{\partial t} = \frac{1}{2} \mathcal{L} G(t, u) \quad \text{with} \quad G(0, \cdot) = F(\cdot),
\end{equation}

where $F \in C^\infty(O(M))$. The following theorem is a restatement of Theorem 3.1, p. 259 in Ikeda and Watanabe [25] specialized to the setting of this paper.

THEOREM 3.14. To each $F \in C^\infty(O(M))$, there is a unique function $G \in C([0, \infty) \times O(M))$, such that $G \in C^\infty([0, \infty) \times O(M))$ and $G$ solves Eq. (3.18).
Moreover, if \( \gamma^u_t \) denotes the solution to the stochastic differential equation (3.4) with \( \gamma_0 = u \in O(M) \), then

\[
G(t, u) = E[ F(\gamma^u_t) ].
\]

(In the sequel we will often write \( E[ F(\gamma^u_t) ] \) as \( (e^{t\mathcal{L}/2} F)(u) \).)

The following lemma is a slight extension of the above Theorem 3.14. The proof is a very minor modification of the proof given for theorem 3.1 on p. 259 of Ikeda and Watanabe [25] and hence will be omitted.

**Lemma 3.15.** Suppose that \( Q \) is another compact manifold and \( F \in C^\infty(Q \times O(M)) \). Let

\[
e^{t\mathcal{L}} F(q, u) \equiv E[ F(q, /r^u) ],
\]

then the function \( (t, q, u) \rightarrow e^{t\mathcal{L}} F(q, u) \) is continuous on \([0, \infty) \times Q \times O(M)\) and infinitely differentiable on \((0, \infty) \times O(M) \times Q\).

**Remark 3.16.** If \( F = f \circ \pi \), where \( f \in C^\infty(M) \), then \( \mathcal{L} F = (\Delta f) \circ \pi \) and \( e^{t\mathcal{L}} F = (e^{t\mathcal{L}} f) \circ \pi \).

**Lemma 3.17.** Fix \( T > 0 \), \( \mathcal{P} = \{0 < t_1 < t_2 < \ldots < t_n = T\} \) and \( F \in C^\infty(O(M)^n) \). Define \( \bar{F} \equiv (\gamma_1, \gamma_2, \ldots, \gamma_{n-1}) \) (an \( O(M)^{(n-1)} \)-valued process), \( F_T(t, u) \equiv ((e^{(T-t)\mathcal{L}/2} F(\gamma_{t+1}, \cdot))(u) \) (a random function of \((t, u) \in [t_{n-1}, T] \times O(M)\)), and

\[
N_t \equiv F_T(t, /r) \text{ for } t \in [t_{n-1}, T],
\]

then \( \{N_t : t \in [t_{n-1}, T]\} \) is a martingale and \( dN_t = W_t dB_t \), where \( W_t \) is the \((T_n M \oplus \text{so}(T_n M))^*\)-valued process defined by

\[
W_t(a + A) \equiv ((B_a + \dot{A}) F_T)(t, /r) \text{ for all } t \in [t_{n-1}, T],
\]

where \((a + A) \in T_n M \oplus \text{so}(T_n M)\). Moreover, the differential of \( W_t \) is given by

\[
dW_t(a + A) = \left( B_{\delta_b(t)} (B_a + \dot{A}) F_T \right)(t, /r)
+ \frac{1}{2} \left( W_t(Ric_{/r})a + \sum_{c \in S} (B_c \Omega_{/r} c, a) \right) dt
- \sum_{c \in S} (B_c \dot{\Omega}(c, a) F_T)(t, /r) \, dt.
\]

**Proof.** By definition of solving the stochastic differential equation in (3.4) and Itô’s Lemma, for \( t \in [t_{n-1}, T] \),

\[
dN_t = \sum_{a \in S} (B_a F_T)(t, /r) \, dB_a(t) - \frac{1}{2} (\mathcal{L} F_T)(t, /r) \, dt
= \sum_{a \in S} (B_a F_T)(t, /r) \, db_a(t) + \frac{1}{2} \sum_{a \in S} (B^2_a F_T)(t, /r) \, dt - \frac{1}{2} (\mathcal{L} F_T)(t, /r) \, dt
= W_t dB_t.
\]
A similar computation shows that

\begin{equation}
(3.22) \quad dW_t(a + A) = \sum_{c \in S} (B_c(B_a + \Delta)F_T)(t, /c)db_c + \frac{1}{2}([\mathcal{L}, (B_a + \Delta)]F_T)(t, /c)dt.
\end{equation}

Using the commutator formulas in Lemma 7.3 of the Appendix, Eq. (3.22) may be expressed as

\begin{equation}
dW_t(a + A) = \sum_{c \in S} (B_c(B_a + \Delta)F_T)(t, /c)db_c + \frac{1}{2} \sum_{c \in S} \left( -2B_c\hat{\Omega}(c, a) + H_{\text{Ric}}c_{a} + (B_c\Omega(c, a)) \right)F_T(/c)dt,
\end{equation}

from which Eq. (3.21) follows. Q.E.D.

The following corollary is a special case of the above lemma.

**Corollary 3.18.** - Let \( P = \{0 < T\} \), \( F_p = f(\Sigma_T) = F(\Sigma_T) \), where \( f \in \mathcal{C}^\infty(M) \) and \( F \equiv f \circ \pi \). Define

\[ N_t \equiv (e^{(T-t)\Delta/2}f)(\Sigma_t) \text{ for all } t \in [0, T]. \]

Then \( N_t \) is a martingale and \( dN_t = W_tdb(t) \), where \( W_t \) is the \((T_0, M)^*\)-valued process defined by

\begin{equation}
(3.23) \quad W_t a \equiv (\nabla^2_{/\partial_\alpha}e^{(T-t)\Delta/2}f)(\Sigma_t) \text{ for all } t \in [0, T].
\end{equation}

Moreover, the differential of \( W_t \) is given by

\begin{equation}
(3.24) \quad dW_t a = (\nabla^2_{/\partial_\alpha}e^{(T-t)\Delta/2}f)(\Sigma_t) + \frac{1}{2} W_t Ric^{*}_{/\partial_\alpha}a dt,
\end{equation}

where for \( f \in \mathcal{C}^\infty(M) \), \( \nabla^2 f \) is the tensor field defined by

\[ \nabla^2_{X \otimes Y} f \equiv X(Y f) - (\nabla_X Y) f. \]

Here \( X \) and \( Y \) are arbitrary vector fields on \( M \).

**Proof.** - This is a straightforward application of Lemma 3.17 using the following remarks. For \( F = f \circ \pi, (\hat{A}F)(u) = 0, (B_a F)(u) = \nabla_{ua} f, (D_a D_c F)(u) - \nabla^2_{ua\otimes ub} f, \) and \( e^{t\Delta} F = (e^{t\Delta} f) \circ \pi \).

Q.E.D.

The proof of Theorem 3.12 is slightly cumbersome. So before proving it, let us practice with the following special case which already contains the main ideas.

**Theorem 3.19.** - Let \( k \in H(T_0, M) \), \( f \in \mathcal{C}^\infty(M) \), and \( F_p = f(\Sigma_T) \), then

\begin{equation}
(3.25) \quad E[X^k F_p] = E[z T F_p],
\end{equation}

where \( z \) is defined in (3.15).
Proof. – Let
\begin{equation}
K(t) = k(t) + \frac{1}{2} \int_0^t \text{Ric}^{\ast} \gamma, k(\tau) d\tau,
\end{equation}
so that
\begin{equation}
\dot{K}(t) = \dot{k}(t) + \frac{1}{2} \text{Ric}^{\ast} \gamma, k(t) \text{ with } K(0) = 0.
\end{equation}

Let $W_t$ be defined as in Eq. (3.23) and notice that
\begin{equation}
X_k F_p = (B_k(T) \gamma) (T) = W_T k(T)
\end{equation}
\begin{equation}
= W_0 k(0) + \int_0^T \{ W dk + \frac{1}{2} \text{Ric} \gamma, k dt \} +
\end{equation}
\begin{equation}
+ \int_0^T \left( \nabla^2_{\gamma} \gamma, \delta \gamma, dW(t) \right) (\Sigma_t).
\end{equation}

wherein the last equality we have used Eq. (3.24) of Corollary 3.18. Taking expectations of this identity gives:
\begin{equation}
E[X_k F_p] = E \int_0^T \left\{ W dk + \frac{1}{2} \text{Ric} \gamma, k dt \right\} = E \int_0^T W_t \dot{K}(t) dt.
\end{equation}

Since $N_T = N_0 + \int_0^T W_t db(t)$, Eq. (3.27) and the $L^2$-isometry properties of the Itô’s integral implies
\begin{equation}
E \int_0^T W_t \dot{K}(t) dt = E \left( \int_0^T W_t db(t) \cdot \int_0^T \langle \dot{K}, db \rangle \right)
\end{equation}
\begin{equation}
= E \left( (N_T - N_0) \cdot \int_0^T \langle \dot{K}, db \rangle \right)
\end{equation}
\begin{equation}
= E \left\{ N_T \int_0^T \langle \dot{K}, db \rangle \right\} = E \left\{ F_p \int_0^T \langle \dot{K}, db \rangle \right\}.
\end{equation}

Q.E.D.

Proof of Theorem 3.12. – Let $F_p$ be a cylinder function of $\gamma$ as in Eq. (3.8), $K(t)$ be as in Eq. (3.26) $W_t$ as in (3.23), and $\gamma(t) = (\gamma,\gamma,\ldots,\gamma)$, $W_T(a + A) = (B_a F + \dot{A})(T)$.
\begin{equation}
EX_k F_p = \sum_{i=1}^n E[D_i F(\gamma, \gamma, \gamma) (k(t_i) - A_{t_i})]
\end{equation}
\begin{equation}
= \sum_{i=1}^{n-1} E[D_i F(\gamma, \gamma, \gamma) (k(t_i) - A_{t_i})] + E[W_T (k(T) - A_T)],
\end{equation}
where $A_t = A^k_t$ as in Notation 3.7.
Now by Itô’s Lemma, Eq. (3.26), Lemma 3.17 and Eq. (3.5),

\[ E[W_T(k(T) - A_T)] = E[W_{t_{n-1}}(k(t_{n-1}) - A_{t_{n-1}})] + T_1 + T_2 + T_3, \]

where

\[ T_1 = E \int_{t_{n-1}}^T W d(k - A) = E \int_{t_{n-1}}^T \left\{ W \left( dk - \frac{1}{2} \sum_{c \in S}(B_c \Omega)_{/i,}(k(t), c) dt \right) \right\}, \]

\[ T_2 = E \int_{t_{n-1}}^T dW(k - A) \]

\[ = -E \int_{t_{n-1}}^T \sum_{c \in S}(B_c \Omega(c, k(t))) F_T(t, /i) dt + \frac{1}{2} E \int_{t_{n-1}}^T W_t (\text{Ric}^+ k(t) + \sum_{c \in S}(B_c \Omega)_{/i,}(c, k(t))) dt \]

and

\[ T_3 = -E \int_{t_{n-1}}^T dW dA \]

\[ = -E \int_{t_{n-1}}^T \sum_{c \in S}(B_c \Omega(c, k(t))) F_T(t, /i) \big|_{A=\Omega_{/i,}}(k(t), c) dt \]

\[ = -E \int_{t_{n-1}}^T \sum_{c \in S}(B_c \Omega(k(t), c)) F_T(t, /i) dt + E \int_{t_{n-1}}^T \sum_{c \in S}((B_c \Omega)_{/i,}(k(t), c)) F_T(t, /i) dt \]

\[ = -E \int_{t_{n-1}}^T \sum_{c \in S}((B_c \Omega(k(t), c))) F_T(t, /i) dt + E \int_{t_{n-1}}^T \sum_{c \in S}(W_t ((B_c \Omega)_{/i,}(k(t), c)) dt. \]

Therefore, we obtain

\[ T_1 + T_2 = E \int_{t_{n-1}}^T \{ W(dK - (B_c \Omega)_{/i,}(k(t), c) dt) \} \]

\[ - E \int_{t_{n-1}}^T \sum_{c \in S}(B_c \Omega(c, k(t)) F_T(t, /i) dt, \]
and

\[ T_1 + T_2 + T_3 = E \int_{t_{n-1}}^{T} WdK \]
\[ = E \left( \int_{t_{n-1}}^{T} Wdb \cdot \int_{t_{n-1}}^{T} \langle \dot{K}, db \rangle \right) \]
\[ = E \left( (N_T - N_{t_{n-1}}) \cdot \int_{t_{n-1}}^{T} \langle \dot{K}, db \rangle \right) \]
\[ = E \left( F_{\nu} \int_{t_{n-1}}^{T} \langle \dot{K}, db \rangle \right), \]

where we have used \( dM = Wdb \) in the third equality and

\[ E \left( N_{t_{n-1}} \int_{t_{n-1}}^{T} \langle \dot{K}, db \rangle \right) = 0 \]

and \( N_T = F_{\nu} \) in the last. Hence

\[ E[W_T(k(T) - A_T)] = E[W_{t_{n-1}}(k(t_{n-1}) - A_{t_{n-1}})] + E \left( F_{\nu} \int_{t_{n-1}}^{T} \langle \dot{K}, db \rangle \right) \]

which combined with Eq. (3.28) gives

\[ (3.29) \quad E^k F_{\nu} = \sum_{i=1}^{n-1} E[D_i F(\langle \mathcal{T}, \mathcal{T} \rangle)(k(t_i) - A_{t_i})] + E \left( F_{\nu} \int_{t_{n-1}}^{T} \langle \dot{K}, db \rangle \right). \]

The rest of the proof will now proceed by induction on \( n \). The case \( n = 1 \) follows directly by taking \( t_{n-1} = 0 \) in Eq. (3.29).

Now suppose that Eq. (3.14) holds for all cylinder functions \( F_{\nu} \) of degree \( n - 1 \) or less. We wish to show that it holds for a cylinder function of degree \( n \). To simplify notation, let

\[ \bar{F}(u_1, \ldots, u_{n-1}) \equiv e^{(T-t_{n-1})\mathcal{L}n/2 F}(u_1, \ldots, u_{n-1}, u_{n-1}) \]
\[ = E[F(u_1, u_2, \ldots, u_{n-1}, \langle \mathcal{T}, t_{n-1} \rangle)] \]

(see Theorem 3.14 and Lemma 3.15) and \( \bar{F}_{\nu} \) denote the cylinder function of \( \langle \rangle \) given by

\[ \bar{F}_{\nu} = \bar{F}(\langle \rangle) = e^{(T-t_{n-1})\mathcal{L}n/2 F}(\langle \rangle, \langle \mathcal{T}, t_{n-1} \rangle). \]

Notice that

\[ (3.30) \quad X^k \bar{F}_{\nu} = \sum_{i=1}^{n-2} (D_i e^{(T-t_{n-1})\mathcal{L}n/2 F}(\langle \mathcal{T}, t_{n-1} \rangle)(k(t_i) - A_{t_i}) \]
\[ + (D_{n-1} e^{(T-t_{n-1})\mathcal{L}n/2 F}(\langle \mathcal{T}, t_{n-1} \rangle)(k(t_{n-1}) - A_{t_{n-1}}) \]
\[ + (D_n e^{(T-t_{n-1})\mathcal{L}n/2 F}(\langle \rangle, \langle \mathcal{T}, t_{n-1} \rangle)(k(t_{n-1}) - A_{t_{n-1}}). \]
By the Markov property or by Itô’s Lemma, for any $i < n$,

$$E[D_i F(\langle \langle \cdot \rangle \rangle, \langle \langle \tau \rangle \rangle)(k(t_i) - A(t_i))] = E((e^{(T-t_{n-1})L_n/2} D_i F(\langle \langle \cdot \rangle \rangle, \langle \langle \tau \rangle \rangle)(k(t_i) - A(t_i))).$$

Therefore

$$\sum_{i=1}^{n-1} E[D_i F(\langle \langle \cdot \rangle \rangle, \langle \langle \tau \rangle \rangle)(k(t_i) - A(t_i))] + E[W_{t_{n-1}}(k(t_{n-1}) - A(t_{n-1}))]
+ E((e^{(T-t_{n-1})L_n/2} D_n F(\langle \langle \cdot \rangle \rangle, \langle \langle \tau \rangle \rangle)(k(t_{n-1}) - A(t_{n-1}))
+ E(D_n(e^{(T-t_{n-1})L_n/2} D_n F)(\langle \langle \cdot \rangle \rangle, \langle \langle \tau \rangle \rangle)(k(t_{n-1}) - A(t_{n-1})).$$

Using the fact that $D_t$ commutes with $L_n$ for $i < n$, this last expression combined with Eq. (3.30) and the definition of $W_t$ in Eq. (3.20) shows:

$$\sum_{i=1}^{n-1} E[D_i F(\langle \langle \cdot \rangle \rangle, \langle \langle \tau \rangle \rangle)(k(t_i) - A(t_i))] + E[W_{t_{n-1}}(k(t_{n-1}) - A(t_{n-1}))] = E[X^k \tilde{F}_P].$$

So by Eq. (3.32) and Eq. (3.29),

$$EX^k F_P = E[X^k \tilde{F}_P] + E \left( F_P \int_{t_{n-1}}^{T} \langle \hat{K}, db \rangle \right).$$

Applying the induction hypothesis, this gives

$$EX^k F_P = E \left( F_P \int_{0}^{t_{n-1}} \langle \hat{K}, db \rangle \right) + E \left( F_P \int_{t_{n-1}}^{T} \langle \hat{K}, db \rangle \right)
= E \left( F_P \int_{0}^{t_{n-1}} \langle \hat{K}, db \rangle \right) + E \left( F_P \int_{t_{n-1}}^{T} \langle \hat{K}, db \rangle \right)
= E \left( F_P \int_{0}^{T} \langle \hat{K}, db \rangle \right).$$

wherein the second equality we have again used the Markov property to replace $\tilde{F}_P$ by $F_P$.

**Q.E.D.**


The notation of Section 3 will be use in this section as well. Also if $Y$ is a smooth vector field on $M$, let $\nabla \cdot Y$ denote the divergence of $Y$ i.e.

$$\nabla \cdot Y)(m) \equiv \text{tr} \nabla Y = \sum_{i=1}^{d} \langle \nabla_{E_i} Y, E_i \rangle,$$

where $\{E_i\}_{i=1}^{d}$ is any orthonormal basis for $T_m M$.
THEOREM 4.1. Let $T > 0$ and $l \in H(R)$ such that $l(t) = 1$. Then for all smooth vector fields $Y$ on $M$, 

\begin{align*}
E[(\nabla \cdot Y)(\Sigma_T)] &= E \left[ \langle f_{-T}^{-1} Y(\Sigma_T), \int_0^T \left( i(t) - \frac{1}{2} l(t) \text{Ric} _{f_{-t}} \right) dt \rangle \right].
\end{align*}

where for $u \in O(M)$,

\begin{align*}
J_u &= \sum_{c \in S} (B_c \text{Ric} (\cdot)c)u = \frac{d}{dt} \left| \sum_{c \in S} \text{Ric} _{\epsilon B_c(u)c} \right|
= u^{-1} \sum_{c \in S} (\nabla u \text{Ric} )uc.
\end{align*}

COROLLARY 4.2. If $Y$ is a smooth vector field on $M$ and $f \in C^\infty(M)$, then

\begin{align*}
E[(Yf)(\Sigma_T)] &= E \left[ f(\Sigma_T) \langle f_{-T}^{-1} Y(\Sigma_T), \int_0^T \left( i(t) - \frac{1}{2} l(t) \text{Ric} _{f_{-t}} \right) dt \rangle \right] \\
&- \frac{1}{2} \int_0^T l(t) J_{f_{-t}} dt \right] - E[f(\Sigma_T)(\nabla \cdot Y)(\Sigma_T)]
\end{align*}

and

\begin{align*}
E[\nabla \cdot Y(\Sigma_T)] &= E \left[ \langle f_{-T}^{-1} Y(\Sigma_T), \frac{b(T)}{T} - \frac{1}{2} \int_0^T t \{ \text{Ric} _{f_{-t}} dt \} \rangle \right].
\end{align*}

Proof. Because $\nabla \cdot (fY) = Yf + f\nabla \cdot Y$, Eq. (4.4) follows by replacing $Y$ by $fY$ in Eq. (4.2). Taking $l(t) = t/T$ in Eq. (4.2) proves Eq. (4.5). Q.E.D.

First proof of Theorem 4.1. This proof is modeled on the first proof of Theorem 2.3. Choose an orthonormal basis $S$ for $T_0M$ and for $c \in S$ set $h_c(t) = l(t)c$ and $X_t^c \equiv X_t^c \equiv \langle h_c(t) \rangle_f(t)$ for $t \in [0, \infty)$. Then using the integration by parts Theorem 3.19,

\begin{align*}
\sum_{c \in S} E(X^c_{-T}|X^c_T, Y(\Sigma_T)) &= \sum_{c \in S} E[(X^c_{-T})^* < X^c_T, Y(\Sigma_T)] \\
&= \sum_{c \in S} E[(-X^c + z^c_T)|X^c_T, Y(\Sigma_T)] =: E(-I + II),
\end{align*}

where

$I \equiv \sum_{c \in S} X^c < X^c_T, Y(\Sigma_T) >$ and $II \equiv \sum_{c \in S} z^c_T < X^c_T, Y(\Sigma_T) >.$
and
\[
\zeta_T^c \equiv \int_0^T \left\langle \frac{1}{2} \text{Ric}_{\mathcal{T}}^*, h_c(t) + \dot{h}_c(t), db(t) \right\rangle \\
= \int_0^T \left\langle \frac{1}{2} \text{Ric}_{\mathcal{T}}^* - i(t), c, db(t) \right\rangle \\
= \int_0^T \left\langle \frac{1}{2} \text{Ric}_{\mathcal{T}}^* + i(t), db(t), c \right\rangle.
\]

Notice that \( g = \langle X_T^c, Y(\Sigma_T) \rangle = F(\mathcal{T}) \), where \( F(u) = \langle c, u^{-1}Y(\pi(u)) \rangle \). Hence \( g = \langle X_T^c, Y(\Sigma_T) \rangle \) is a cylinder function of \( \mathcal{T} \) as defined in Section 3 and moreover by Eq. (3.11),
\[
\langle X^c g \rangle = DF(\mathcal{T})(-A_T^c + h_c(T)).
\]

Note
\[
\langle 4.6 \rangle \quad DF(u)A = \frac{d}{du} \left\langle c, e^{-A}u^{-1}Y(\pi(u)) \right\rangle = -\langle c, Au^{-1}Y(\pi(u)) \rangle, \quad \forall A \in so(T_sM)
\]
and
\[
\langle 4.7 \rangle \quad DF(u)a = \langle c, w^{-1}\nabla_{wa}Y \rangle - \langle uc, \nabla_{ua}Y \rangle \quad \forall A \in T_sM.
\]

Using equations (4.6-4.8) and Eq. (3.5),
\[
I = \sum_{c \in S} \left\langle \langle /\mathcal{T}c, \nabla /\mathcal{T}c, Y \rangle + \langle c, A_T^c /\mathcal{T}^{-1}Y(\Sigma_T) \rangle \right\rangle \\
= (\nabla \cdot Y)(\Sigma_T) + \sum_{c \in S} \left\langle c, \left( \int_0^T l(t)\Omega_{/\mathcal{T}}, \langle c, \delta b(t) \rangle /\mathcal{T}^{-1}Y(\Sigma_T) \right) \right\rangle \\
= (\nabla \cdot Y)(\Sigma_T) + \sum_{c \in S} \left\langle \left( \int_0^T l(t)\Omega_{/\mathcal{T}}, \langle \delta b(t), c \rangle /\mathcal{T}^{-1}Y(\Sigma_T) \right) \right\rangle \\
= (\nabla \cdot Y)(\Sigma_T) + \left\langle \left( \int_0^T l(t)\text{Ric}_{/\mathcal{T}}, \delta b(t), /\mathcal{T}^{-1}Y(\Sigma_T) \right) \right\rangle \\
= (\nabla \cdot Y)(\Sigma_T) + \left\langle \left( \int_0^T l(t)\text{Ric}_{/\mathcal{T}}, db(t), + \frac{1}{2} \int_0^T l(t)J_{/\mathcal{T}}, dt, /\mathcal{T}^{-1}Y(\Sigma_T) \right) \right\rangle,
\]
where we have used
\[
\langle 4.9 \rangle \quad \text{Ric}_{/\mathcal{T}}, \delta b(t) = \text{Ric}_{/\mathcal{T}}, db(t) + \frac{1}{2} \sum_{c \in S} (B, \text{Ric c})_{/\mathcal{T}}, dt \\
= \text{Ric}_{/\mathcal{T}}, db(t) + \frac{1}{2} J_{/\mathcal{T}}, dt.
\]

Similarly, we have
\[
II = \sum_{c \in S} z_T^c (\langle /\mathcal{T}c, Y(\Sigma_T) \rangle) \\
= \sum_{c \in S} \left\langle c, /\mathcal{T}^{-1}Y(\Sigma_T) \right\rangle \left\langle \int_0^T \left[ \frac{l(t)}{2} \text{Ric}_{/\mathcal{T}}, + \dot{i}(t) \right] db(t), c \right\rangle \\
= \left\langle \int_0^T \left[ \frac{l(t)}{2} \text{Ric}_{/\mathcal{T}}, + \dot{i}(t) \right] db(t), /\mathcal{T}^{-1}Y(\Sigma_T) \right\rangle.
\]
Combining the above expressions shows
\[
0 = E[-I + II] = E \left[ - (\nabla \cdot Y)(\Sigma_T) + \left\langle \int_0^T \dot{t}(t) db(t), \int_T^{-1} Y(\Sigma_T) \right\rangle \right] - \frac{1}{2} E \left\langle \int_0^T l(t) \text{Ric}_{/\epsilon} db(t) + \int_0^T l(t) J_{/\epsilon} dt, \int_T^{-1} Y(\Sigma_T) \right\rangle.
\]
This finishes the first proof of the Theorem. Q.E.D.

Second proof of Theorem 4.1. - This proof will follow the strategy of the second proof of Eq. (2.10) of Section 2. To simplify notation, we will write \(dQ_t \sim dV_t\) if \(Q\) and \(V\) are semi-martingales such that \(Q_t - V_t\) is a martingale.

Let \(\tilde{Y}(u) \equiv u^{-1} Y(\pi(u))\) and set \(Y_t(u) \equiv (e^{(T-t)\Delta/2} Y)(u) = (e^{(T-t)\zeta/2} \tilde{Y})(u)\). Define
\[
B \cdot Y_t \equiv \sum_{c \in S} \langle B_c Y_t, c \rangle = \sum_{c \in S} B_c Y_t(c)
\]
and set
\[
Q_t \equiv t(\langle B \cdot Y_t \rangle) - \left\langle Y_t, \int_0^t \dot{t} db \right\rangle.
\]
By Itô’s Lemma, as in the proof of Lemma 3.17,
\[
d[Y_t(\langle /\epsilon \rangle)] = \sum_{c \in S} (B_c Y_t(\langle /\epsilon \rangle)) db_c(t) = (B_{\epsilon(t)} Y_t(\langle /\epsilon \rangle)
\]
and
\[
d[B \cdot Y_t(\langle /\epsilon \rangle)] = \sum_{c \in S} B_c B \cdot Y_t(\langle /\epsilon \rangle) db_c(t) - B \cdot \left( \frac{L}{2} Y_t(\langle /\epsilon \rangle) \right) dt
\]
\[
= \sum_{c \in S} B_c B \cdot Y_t(\langle /\epsilon \rangle) db_c(t) + \frac{1}{2} \sum_{c \in S} B_c^2 B \cdot Y_t(\langle /\epsilon \rangle) dt - B \cdot \left( \frac{L}{2} Y_t(\langle /\epsilon \rangle) \right) dt
\]
\[
= \sum_{c \in S} B_c B \cdot Y_t(\langle /\epsilon \rangle) db_c(t) + \frac{1}{2} \left[ \mathcal{L} \cdot B \cdot Y_t(\langle /\epsilon \rangle) \right] dt.
\]
Therefore, we obtain
\[
dQ_t \equiv \dot{t}(B \cdot Y_t(\langle /\epsilon \rangle)) dt + \frac{d(t)}{2} \left[ (\mathcal{L}, B \cdot Y_t) \langle /\epsilon \rangle \right] dt - \left\langle (B_{\epsilon(t)} Y_t(\langle /\epsilon \rangle), \dot{t}(t) db(t)) \right\rangle
\]
\[
= \left\{ \dot{t}(B \cdot Y_t(\langle /\epsilon \rangle)) + \frac{d(t)}{2} \left[ (\mathcal{L}, B \cdot Y_t) \langle /\epsilon \rangle - \sum_{c \in S} \langle (B_c Y_t(\langle /\epsilon \rangle), c) \dot{t}(t) \rangle \right] \right\} dt
\]
\[
= \frac{d(t)}{2} \left[ (\mathcal{L}, B \cdot Y_t) \langle /\epsilon \rangle \right] dt
\]
\[
- \frac{d(t)}{2} \left\{ \langle Y_t(\langle /\epsilon \rangle), J_{/\epsilon} \rangle + \sum_{c \in S} \langle (B_c Y_t(\langle /\epsilon \rangle), \text{Ric}_{/\epsilon}, c) \rangle \right\} dt,
\]
wherein the last equality we have used the commutator formula in Eq. (7.6) of Corollary 7.4 from the Appendix below.

Unlike the proof of Eq. (2.10), \( Q_t \) is not a martingale and hence we will have to modify \( Q \). Set \( N_t = Q_t + V_t \) where

\[
V_t = \frac{1}{2} \left( \int_0^t l(\tau) \{ J_{//}, d\tau + \text{Ric}_{//}, db(\tau) \}, Y_t(/\!/t) \right).
\]

Notice that

\[
dV_t \equiv \frac{l(t)}{2} \{ J_{//}, Y_t(/\!/t) \} dt + \frac{l(t)}{2} \{ \text{Ric}_{//}, db(t), (B_{db(t)} Y_t)(/\!/t) \} dt
\]

\[
= \frac{l(t)}{2} \{ J_{//}, Y_t(/\!/t) \} dt + \frac{l(t)}{2} \sum_{c \in S} \{ \text{Ric}_{//}, c, (B_c Y_t)(/\!/t) \} dt
\]

\[
\approx -dQ_t.
\]

Hence we have shown that \( dN_t = dQ_t + dV_t \approx 0 \), that is \( N \) is a martingale. Therefore we conclude that \( EN_T = EN_0 = 0 \), since \( N_0 = Q_0 + V_0 = 0 \). Since \( B \cdot \tilde{Y} = (\nabla \cdot Y) \circ \pi \), and \( l(T)(B \cdot Y_T)(/\!/T) = (B \cdot Y)(/\!/T) = (\nabla \cdot Y)(\Sigma_T) \),

\[
N_T = Q_T + V_T
\]

\[
= l(T)(B \cdot Y_T)(/\!/T) - \left\langle Y_T(/\!/T), \int_0^T db \right\rangle
\]

\[
+ \frac{1}{2} \left\langle \int_0^T l(t) \{ J_{//}, d\tau + \text{Ric}_{//}, db(\tau) \}, Y_T(/\!/T) \right\rangle
\]

\[
= (\nabla \cdot Y)(\Sigma_T) - \left\langle /\!/T^{-1} Y(\Sigma_T), \int_0^T l(t)db(t) \right\rangle
\]

\[
+ \frac{1}{2} \left\langle \int_0^T l(t) \{ J_{//}, dt + \text{Ric}_{//}, db(t) \}, /\!/T^{-1} Y_T(\Sigma_T) \right\rangle.
\]

It is now clear the statement \( EN_T = 0 \) is equivalent to Eq. (4.2). Q.E.D.

### 4.1. Backwards Integrals

The “\( J \)” term in Eq. (4.2) is somewhat undesirable, since it involves derivatives of the Ricci tensor. This is in contrast with Bismut’s formula which is reviewed in the next section. Before ending this section I would like to point out that by using a “backwards” Itô’s integral, we may eliminate the “\( J \)” term.

Let \( \pi = \{0 = t_0 < t_1 < t_2 < \cdots < t_n \to \infty\} \) denote a partition of \([0, \infty)\), \(|\pi| \equiv \max_i |t_{i+1} - t_i|\). For \( \tau = t_i \in \pi \), let \( \tau^+ \equiv t_{i+1} \) be the successor to \( \tau \) in \( \pi \). Suppose that \( V \) is a finite dimensional vector space, \( X \) is a \( V \)-valued continuous semi-martingale and \( A \) is an \( \text{End}(V) \)-valued continuous semi-martingale. Then the backwards stochastic integral of \( A \) relative to \( X \) is

\[
\int_0^T AdX = \lim_{|\pi| \to 0} \sum_{\tau \in \pi} A(t \wedge \tau^+)(X(t \wedge (\tau^+)) - X(\tau \wedge t)),
\]
where the limit exists in probability uniformly for \( t \) in compact subset of \([0, \infty)\). Recall that the forward and Stratonovich integrals may be defined similarly as

\[
\int_0^t A \delta X \equiv \lim_{|\tau| \to 0} \sum_{\tau \in \pi} \frac{1}{2} (A(t \wedge \tau +) + A(t)) (X(t \wedge (\tau +)) - X(t \wedge t))
\]

and

\[
\int_0^t A \delta X \equiv \lim_{|\tau| \to 0} \sum_{\tau \in \pi} \frac{1}{2} (A(t \wedge \tau +) + A(\tau)) (X(t \wedge (\tau +)) - X(t \wedge t))
\]

respectively. We also know that the joint quadratic variation “\( \int d A \delta X \)” is given by

\[
\int d A \delta X = \lim_{|\tau| \to 0} \sum_{\tau \in \pi} (A(t \wedge \tau +) - A(t)) (X(t \wedge (\tau +)) - X(t \wedge t)).
\]

It is a trivial exercise to prove that

\[
\int_0^t A \partial X = \int_0^t A \delta X + \int_0^t d A \delta X.
\]

From the computation given for Eq. (4.9) we know that

\[
d(Ric_{/\!\!/}, \partial(t)) = J_{/\!\!/}, d t.
\]

Using this equation and Eq. (4.14) we find that Theorem 4.1 may be written as:

**Corollary 4.3.** — Let \( T > 0 \) and \( l \in H(\mathbb{R}) \) such that \( l(T) = 1 \). Then for all smooth vector fields \( Y \) on \( M \),

\[
E[(\nabla \cdot Y)(\Sigma_T)] = E \left[ \left( l(t)^{-1} Y(\Sigma_T), \int_0^T \left( \dot{l}(t) - \frac{1}{2} l(t)Ric_{/\!\!/}, \dot{\partial} \right) db(t) \right) \right].
\]

**5. Bismut’s formula**

For the sake of comparison and completeness, let us include Bismut’s formula, see Eq. (2.77) in [4]. (Note: Bismut uses \( \delta \) for the Itô’s differential and \( d \) for the Stratonovich differential.)

**Theorem 5.1 (Bismut).** — Let \( f : M \rightarrow \mathbb{R} \) be a smooth function, then for any \( 0 < t \leq T \),

\[
\tilde{\nabla} (e^{T \Delta t / 2} f)(\omega) = t^{-1} E \left[ \left( \int_0^t Q_r db_r \right) (e^{(T-t) \Delta t / 2} f)(\Sigma_r) \right]
\]

\[
= t^{-1} E \left[ \left( \int_0^t Q_r db_r \right) f(\Sigma_T) \right].
\]
where $Q_t$ is the unique solution to the differential equation:

\[
\frac{dQ_t}{dt} = -\frac{1}{2} Q_t \text{Ric}' t, \quad \text{with } Q_0 = I.
\]

**Proof.** The proof given here is modeled on Remark 6 on p. 84 in Bismut [4] and the proof of Theorem 2.1 in Elworthy and Li [14]. Also see Norris [33].

For $(t, m) \in [0, T] \times M$ let

\[
F(t, m) \equiv (e^{(T-t)\Delta/2} f)(m).
\]

Also let $Q_t \in \text{End}(T_o M)$ be an adapted continuous process to be chosen later. Consider $z_t \equiv \left( \int_0^t Q_r db(r) \right) F(t, \Sigma_t)$. We wish to compute the differential of $z_t$. First notice that

\[
dF(t, \Sigma_t) = (\nabla_{/r, \text{mart}} F(t, \Sigma_t))(\Sigma_t)
\]

\[
= \sum_{c \in S} (Q_r, (\nabla_{/r, c} F(t, \Sigma_t))(\Sigma_t)) dt.
\]

From this we conclude that:

\[
\begin{align*}
E z_t &= E z_0 + E \int_0^t \sum_{c \in S} Q_r c(\nabla_{/r, c} F(r, \Sigma_r))(\Sigma_r) dr \\
&= 0 + E \int_0^t \sum_{c \in S} Q_r c(\nabla_{/r, c} F(r, \Sigma_r)) dr \\
&= E \int_0^t \sum_{c \in S} Q_r c(\nabla_{/r, c}^{-1} F(r, \Sigma_r)) dr \\
&= E \int_0^t Q_r \nabla_{/r, r}^{-1} F(r, \Sigma_r) dr.
\end{align*}
\]

Suppose that we can choose $Q_r$ such that $Q_r \nabla_{/r, r}^{-1} F(r, \Sigma_r)$ is a martingale and $Q_0 = I$. For such a $Q$,

\[
E z_t = E \int_0^t Q_r \nabla_{/r, r}^{-1} F(r, \Sigma_r) dr - t \nabla (e^{T\Delta/2} f)(o).
\]

That is to say

\[
\nabla (e^{T\Delta/2} f)(o) = t^{-1} E \left[ \left( \int_0^t Q_r db(r) \right) (e^{(T-t)\Delta/2} f)(\Sigma_t) \right]
\]

\[
= t^{-1} E \left[ \left( \int_0^t Q_r db(r) \right) f(\Sigma_T) \right],
\]

wherein the second inequality we have used the Markov property of $\Sigma_t$. 

JOURNAL DE MATHÉMATIQUES PURES ET APPLIQUÉES
We will now show how to choose $Q_t$. To this end let $W_t \equiv \int_t^1 \nabla F(t, \Sigma_t)$. Then by Corollary 3.18 we know that
\[
dW_t = d(\text{martingale}) + \frac{1}{2} \text{Ric}_W(t) dt.
\]
So if $Q$ is the solution to Eq. (5.1), then we will have that $Q_t W_t$ is a martingale by Itô's Lemma. Q.E.D.

Let us write Theorem 5.1 out in the special case where $M = G$ is a compact Lie group. Suppose $(\cdot, \cdot)$ is an $Ad_G$-invariant inner product on $g = T_e G$ — the Lie algebra of $G$. We extend $(\cdot, \cdot)$ to a metric on $G$ by requiring left translations to be isometries. Since $(\cdot, \cdot)$ is $Ad_G$-invariant, it follows that $(\cdot, \cdot)$ is a bi-invariant metric on $G$. Let $\nabla$ denote the left covariant derivative, i.e. $\nabla_A B = 0$ for all $A, B \in g$ where $A$ denotes the unique left invariant vector field on $G$ such that $A(e) = A$. Also let $\{b(t)\}_{t \geq 0}$ be a $g$-valued Brownian motion and $\Sigma_t$ denote the solution to the stochastic differential equation:

\[
\delta \Sigma_t = \Sigma_t db(t) \text{ with } \Sigma_0 = e \in G.
\]

More precisely, Eq. (5.2) is short hand for the stochastic differential equation

\[
\delta \Sigma_t = \sum_{A \in S} \nabla_A \Sigma_t b^A(t) \text{ with } \Sigma_0 = e \in G.
\]

where $b^A(t) \equiv \langle b(t), A \rangle$. Now $\nabla$ is a flat connection and parallel translation (relative to $\nabla$) along $\Sigma_t$ is $\int_t^1 \nabla \cdot$. In particular, the triple of processes $(\Sigma, \int, b)$ satisfy the assumptions described in Section 3.2.

**Proposition 5.2.** Suppose $(G, (\cdot, \cdot), \nabla, \Sigma, \int, b)$ are as above. Then

\[
(\nabla_t e^{T/2} f)(e) = t^{-1} E\cdot [b(t) f(\Sigma_T)]
\]

Proof. The covariant derivative is flat, and hence the Ricci tensor is zero. Thus $Q_t \equiv I$ in this case. Therefore Eq. (5.4) is a direct consequence of Theorem 5.1. Q.E.D.

6. Applications to Compact Lie Groups with Left Invariant Metrics

For this section we will suppose, to avoid technical complications, that $G$ is a compact Lie group. (Actually the results of this section are valid for general unimodular Lie groups.) Let $o = e \in G$, $g = T_e G$ be the Lie algebra of $G$ ($e \in G$ is the identity element). $(\cdot, \cdot)$ is an inner product on $g$ which we also view as a left invariant Riemannian metric on $G$, and $\nabla$ is the Levi-Civita covariant derivative associated to $(\cdot, \cdot)$. Notice that we do not assume here that $(\cdot, \cdot)$ on $g$ is $Ad$-invariant or equivalently that the Riemannian metric made from $(\cdot, \cdot)$ is bi-invariant.

**Notation 6.1.** If $h \in g$, let $\hat{h}$ denote the unique left invariant vector field on $G$ such that $\hat{h}(e) = h$. Also let $D$ denote the Lie algebra version of $\nabla$, i.e. for each $h \in g$, $D_h$ is the linear operator on $g$ defined by $D_h k \equiv (\nabla_{\hat{h}} k)(e)$. 
One of the major simplifications when working with Lie groups is that computations involving $G$ may often be reduced to computations essentially only involving the Lie algebra $\mathfrak{g}$. For example, since $(\cdot, \cdot)$ is a left-invariant metric on $G$, one has $\nabla_h k = (D_h k)$.

Now suppose that $\sigma$ and $\nu$ are $C^1$-curves in $G$ and $\mathcal{T}G$ respectively and that $\beta(t) = \int_0^t L_{\sigma(r)^{-1} \sigma'(r)} \nu(t) \tau(\sigma(t))$ and $a(t) \equiv L_{\sigma(t)^{-1} \sigma'(t)} \nu(t)$. If $\nabla \nu(t)/dt$ denotes the covariant differential of $\nu$ along $\sigma$, then:

\[
(6.1) \quad \frac{\nabla \nu(t)}{dt} = L_{\sigma(t)^*} \left( \frac{da(t)}{dt} + D_{\beta(t)} a(t) \right).
\]

To prove this equation, notice that $\nu$ may be written as

\[
(6.2) \quad \nu(t) = L_{\sigma(t)^*} a(t) = \langle a(t) \rangle \langle \sigma(t) \rangle = \sum_{h \in \mathfrak{g}} \langle a(t), h \rangle \hat{h}(\sigma(t)),
\]

where, as above, $S$ is an orthonormal basis for $\mathfrak{g}$. Hence,

\[
\sum_{h \in \mathfrak{g}} \left\{ \left( \frac{da(t)}{dt}, h \right) \hat{h}(\sigma(t)) + \langle a(t), h \rangle \nabla_{a(t)} h \right\}
= \left. L_{\sigma(t)^*} \frac{da(t)}{dt} + \sum_{h \in \mathfrak{g}} \langle a(t), h \rangle (D_{L_{\sigma(t)^{-1} \sigma'(t)}} \hat{h})(\sigma(t)) \right|_{\sigma(t)},
\]

Now let $\{\beta(t)\}_{t \geq 0}$ be a $\mathfrak{g}$-valued Brownian motion on the filtered probability space $(\mathcal{W}, \sigma_1, \mathcal{F}, \mathcal{P})$ with covariance $E(\langle \beta(t), h \rangle \cdot \langle \beta(t), k \rangle) = t \wedge \tau(h, k)$ for all $h, k \in \mathfrak{g}$. Also let $\{\Sigma_t\}_{t \geq 0}$ be the solution to the stochastic differential equation,

\[
(6.3) \quad \delta \Sigma_t = L_{\Sigma_t} \delta \beta(t) \text{ with } \Sigma_0 = e.
\]

This last equation may be written more explicitly as

\[
(6.4) \quad \delta \Sigma_t = \sum_{h \in \mathfrak{g}} \hat{h}(\Sigma_t) \delta \beta^h(t) \text{ with } \Sigma_0 = e,
\]

where $S \subset \mathfrak{g}$ is an orthonormal basis, $\beta^h(t) \equiv \langle \beta(t), h \rangle$ (a real Brownian motion for $h \in S$) and $\hat{h}$ denotes the left invariant vector field on $G$ such that $\hat{h}(e) = h$.

**Lemma 6.2.** The process $\{\Sigma_t\}_{t \geq 0}$ is a Brownian motion on the Riemannian manifold $(G, (\cdot, \cdot))$.

**Proof.** Since $G$ is compact it is also unimodular. For unimodular Lie groups it is well known that the Laplace Beltrami operator $\Delta$ is given by $\Delta = \sum_{h \in S} \hat{h}^2$, see for example Remark 2.2 in [11]. Therefore for $f \in C^\infty(G),
\[
d(f(\Sigma_t)) = \sum_{h \in S} (\hat{h} f)(\Sigma_t) \delta \beta^h(t)
= \sum_{h \in S} (\hat{h} f)(\Sigma_t) d\beta^h(t) + \frac{1}{2} \sum_{h \in S} (\hat{h}^2 f)(\Sigma_t) dt.
\]

\[
JOURNAL DE MATHÉMATIQUES PURES ET APPLIQUÉES
\]
which shows that $\Sigma$ satisfies the martingale characterization of a Brownian motion, see Section 3.2.

Continuing the notation of Section 3.2, let $//\ $ denote stochastic parallel translation (for the Levi-Civita covariant derivative $\nabla$) along $\Sigma$ and $b$ be the $g$-valued Brownian motion defined in Eq. (3.3). In the next theorem we will give a more concrete description of the processes $//\ $ and $b$.

**Theorem 6.3.** Let $\beta, \Sigma, //\ $ and $b$ be the processes described above and set

\begin{equation}
U(t) \equiv L_{\Sigma(t)^{-1}} //_t.
\end{equation}

Then $U(t)$ is the $O(g)$-valued adapted and continuous process satisfying the stochastic differential equation

\begin{equation}
dU(t) + D_{\beta(t)} U(t) = 0 \text{ with } U(0) = I_g.
\end{equation}

\begin{equation}
b(t) = \int_0^t U(\tau)^{-1} \delta \beta(\tau),
\end{equation}

and also

\begin{equation}
b(t) = \int_0^t U(\tau)^{-1} d\beta(\tau).
\end{equation}

**Proof.** The proof of Eq. (6.7) is easy:

\[ \delta b(t) = //^{-1} \delta \Sigma_{//} = //^{-1} L_{\Sigma(t)^{-1}} \delta \beta(t) = U(t)^{-1} \delta \beta(t). \]

We now will show that Eq. (6.6) holds. Define $F : O(G) \to O(g)$ by $F(p) = L_{\pi(p)^{-1}} p$, where $\pi : O(G) \to G$ is the fiber projection map. (Recall that $O(G) = \cup_{g \in G} O_g(G)$ and $O_g(G)$ is the set of isometries from $g = T_e G$ to $T_g G$.) Since $U(t) = F(//_t)$, in order to find the equation satisfied by $U(t)$ we will need the horizontal derivative of $F$.

**Claim.** For $h \in g$, let $B_h$ be the associated horizontal vector field on $O(G)$ defined in Definition 3.3. Then $(B_h F)(p) = -D_{L_{\pi(p)^{-1}} p} F(p)$.

To prove this claim let $g(t)$ be a curve in $G$ such that $g(0) = \pi(p)$ and $\dot{g}(0) = ph$. Define $O(t) \in O(g)$ to be the solution the ordinary differential equation

\begin{equation}
dO(t)/dt + D_{\alpha(t)} O(t) = 0 \text{ with } O(0) = L_{g(0)^{-1}} p,
\end{equation}

where $\alpha(t) \equiv L_{g(t)^{-1}} \dot{g}(t) \in g$. (Notice that $\alpha(0) = L_{\pi(p)^{-1}} ph$.) Let $p(\cdot)$ be the curve in $O(G)$ defined by $p(t) \equiv L_{g(t)^{-1}} O(t) \in O(G)$. By Eq. (6.1), $\nabla p(t)/dt = 0$ so $p$ is horizontal. Also we have $p(0) = p$, and $\pi \circ \dot{p}(0) = \dot{g}(0) = ph$ and hence $\dot{p}(0) = B_h (p)$. Therefore:

\[
(B_h F)(p) = \left. \frac{d}{dt} \right|_{t=0} F(p(t)) = \left. \frac{d}{dt} \right|_{t=0} L_{g(t)^{-1}} p(t)
\]

\[
= \left. \frac{d}{dt} \right|_{t=0} L_{g(t)^{-1}} L_{g(t^*)} O(t)
\]

\[
= \left. \frac{d}{dt} \right|_{t=0} O(t) = -D_{\alpha(0)} O(0)
\]

\[
= -D_{L_{\pi(p)^{-1}} p} F(p).
\]
This proves the claim.
Because of the claim, we now know that
\[ dU(t) = dF(t) = (B_{bb(t)}F)(t) \]
\[ = -D_{(R_{b(t)}^{-1})t}F(t) \]
\[ = -D_{(R_{t}^{-1})t}F(t) \]
\[ = -D_{(U(t)bb(t))}U(t) \]
which in view of Eq. (6.7) proves Eq. (6.6).
To complete the proof we need only verify Eq. (6.8).
Now, (6.10)
\[ b(t) = \int_0^t U(t)^{-1} \delta \beta(t) = \int_0^t U(t)^{-1} d\beta(t) + V_t. \]
where \( V \) is the process of bounded variation given by
\[ V(t) = \frac{1}{2} \int_0^t d(U(\tau)^{-1})d\beta(\tau) = \]
\[ -\frac{1}{2} \int_0^t U(\tau)^{-1}D_{d\beta(\tau)}d\beta(\tau) - \frac{1}{2} \int_0^t U(\tau)^{-1} \sum_{h \in S} D_h h d\tau. \]
Using Lévy's criteria, it is easily checked that \( \int_0^t U(t)^{-1} d\beta(t) \) is a Brownian motion. But we know \textit{a priori} that \( b(t) \) is also a Brownian motion on \( g \). In order for these statements to be consistent with Eq. (6.10) it is necessary for \( V \equiv 0 \), \textit{i.e.} Eq. (6.8) is valid. \textit{Q.E.D.}

As a bi-product of the proof we have shown that
\[ \sum_{h \in S} D_h h = 0. \]
We could also verify this equation more directly as follows. For \( k \in g \) let
\[ e^{ik}(g) = R_{e^{ik}}(g) \equiv ge^{ik}, \]
then \( e^{ik} \) is the flow of \( \tilde{k} \). Because \( G \) is compact and hence unimodular, the Riemannian volume form on \( G \) which is a left invariant volume form is also a right invariant volume form. Therefore right translation preserve the Riemannian volume form and hence the flow \( e^{ik} \) preserves the Riemannian volume form. Consequently, the divergence of \( \tilde{k} \) is zero. On the other hand, we may also compute the divergence of \( \tilde{k} \) using the Levi-Civita covariant derivative \( \nabla \) via,
\[ \nabla \cdot \tilde{k} = \text{tr} \nabla \tilde{k} = \sum_{h \in S} \langle \nabla_h \tilde{k}, h \rangle = \sum_{h \in S} \langle D_h k, h \rangle = - \sum_{h \in S} \langle k, D_h h \rangle, \]
wherein the last equality we have used the metric compatibility of \( \nabla \) to conclude that \( D_h \) is skew adjoint on \( g \). Since, as already noted, \( \nabla \cdot \tilde{k} = 0 \) for all \( k \in g \), Eq. (6.11) follows from Eq. (6.12).
We now wish to write down Theorem 4.1 in the context of this section.

**Corollary 6.4.** Let \((G, \langle \cdot, \cdot \rangle, \nabla)\) be as above. Also let \(f \in C^\infty(G), h \in \mathfrak{g}, T > 0, \) and \(l \in H(\mathbb{R})\) such that \(l(T) = 1,\) then

\[
(6.13) \quad E[(\hat{h}f)(\Sigma_T)] = E \left[ f(\Sigma_T) \left( U(T)^{-1}h, \int_0^T U^{-1}(t) \left( \dot{l}(t) - \frac{1}{2} l(t) Ric_\epsilon \right) dt \right) \right] 
- \frac{1}{2} E \left[ f(\Sigma_T) \left( U(T)^{-1}h, \int_0^T l(t)U(t)^{-1}J_\epsilon dt \right) \right],
\]

where \(Ric_\epsilon\) is the Ricci tensor restricted to \(\mathfrak{g} = T_e G\) and \(J\) is defined in Eq. (4.3).

**Proof.** Applying Theorem 4.1 with \(Y = h,\) using the fact that \(\nabla \cdot (f \dot{h}) = \dot{h}f + f \nabla \cdot \dot{h} = \dot{h}f,\) shows that

\[
(6.14) \quad E[(\hat{h}f)(\Sigma_T)] = E \left[ \left( f(\Sigma_T)/f(T)^{-1}\hat{h}(\Sigma_T), \int_0^T \left( \dot{l}(t) - \frac{1}{2} l(t) Ric_\epsilon \right) dt \right) \right].
\]

Since \(\langle \cdot, \cdot \rangle\) is left invariant, it follows that \(Ric\) and \(\nabla Ric\) are also left invariant, i.e. \(L_{g*}^{-1}Ric = Ric\) and \(L_{g*}^{-1}(\nabla_{t_e g} Ric) = (\nabla_{t_e g} Ric)k\) for all \(g \in G,\) therefore.

\[
Ric_{\parallel/t} = Ric_{L_{\Sigma_\epsilon} U(t)} = U^{-1}(t)Ric_{\Sigma_\epsilon} U(t) = U^{-1}(t)Ric U(t),
\]

and

\[
J_{\parallel/t} = J_{L_{\Sigma_\epsilon} U(t)} = \sum_{k \in S} (L_{\Sigma_\epsilon} U(t))^{-1} (\nabla_{t_{\Sigma_\epsilon} U(t)} Ric) L_{\Sigma_\epsilon} U(t) h
\]

\[
= \sum_{k \in S} U^{-1}(t)(\nabla_{U(t)} Ric) U(t) h = U^{-1}(t)J_\epsilon.
\]

Also notice that

\[
/_{T}^{-1}\hat{h}(\Sigma_T) = U^{-1}(T)L_{\Sigma_\epsilon}^{-1}\hat{h}(\Sigma_T) = U^{-1}(T)h.
\]

Using these last three equalities and Eq. (6.8) in Eq. (6.14), one finds that:

\[
E[(\hat{h}f)(\Sigma_T)] = E \left[ f(\Sigma_T) \left( U^{-1}(T)h, \int_0^T \left( \dot{l}(t) - \frac{1}{2} l(t)U^{-1}(t)Ric_\epsilon U(t) \right) dt \right) \right] 
- \frac{1}{2} E \left[ U^{-1}(T)h, \int_0^T l(t)U(t)^{-1}J_\epsilon dt \right],
\]

from which Eq. (6.13) clearly follows. Q.E.D.

**Corollary 6.5.** Let \((G, \langle \cdot, \cdot \rangle, \nabla)\), \(f \in C^\infty(G), h \in \mathfrak{g}, T > 0,\) and \(l \in H(\mathbb{R})\) such that \(l(T) = 1\) be as in Corollary 6.4. Then

\[
(6.15) \quad E[(\hat{h}f)(\Sigma_T)] = E \left[ f(\Sigma_T) \left( U(T)^{-1}h, \int_0^T U^{-1}(t) \left( \dot{l}(t) - \frac{1}{2} l(t)Ric_\epsilon \right) dt \right) \right]
\]

\[
(6.16) \quad = E \left[ f(\Sigma_T) \int_0^T \left( U(t)^{-1}h, \left( \dot{l}(t) - \frac{1}{2} l(t)Ric_\epsilon \right) dt \right) \right].
\]
where $\overrightarrow{d\beta}$ denotes the backwards Itô’s differential and $U(t,T) = U(t)U(T)^{-1}$ for $0 \leq t \leq T$.

Remark 6.6. – Notice that the process $U(t,T)h$ is not adapted to the forward filtration. The stochastic integral in Eq. (6.16) is to be interpreted as a limit in probability of Riemann sums of the form in Eq. (4.10) with $A(t) = (\dot{l}(\tau) - \frac{1}{2}l(\tau)\text{Ric}^\tau)U(t,T)h$ and $X(t) = \beta(t)$. The convergence (in probability uniformly on compact subsets of $[0, \infty)$) of these sums follows from the corresponding convergence of the Riemann sums defining the stochastic integral in Eq. (6.15). In fact $U(t,T)$ solves the (non-adapted to the forward filtration) Stratonovich differential equation,

$$dU(t,T) + D_{\beta(t)}U(t,T) = 0 \text{ with } U(T,T) = 1.$$  

From this it follows that $U(t,T)$ may be chosen to be $\sigma(\beta(\tau) - \beta(T) : t < \tau < T)$-adapted. Hence the backward Itô’s integral is an adapted integral when “run” in reverse time. This fact will be exploited in Section 6 in Driver [10] where the reader may find more details on this remark.

Proof of Corollary 6.5. – Equation (6.16) follows directly from (6.15) and the proof of Eq. (6.15) is basically the same as the proof of Corollary 4.3. Just apply Eq. (4.14) to Eq. (6.13) using

$$d(U^{-1}(t)\text{Ric}_\tau)\overrightarrow{d\beta}(t) = (U^{-1}(t)D_{\beta(t)}\text{Ric}_\tau)\overrightarrow{d\beta}(t)$$

and

$$d(U^{-1}(t))\overrightarrow{d\beta}(t) = U^{-1}(t)D_{\beta(t)}\overrightarrow{d\beta}(t)$$

wherein the last equality we used Eq. (6.11). Q.E.D.

7. Appendix: Geometric Identities

The purpose of this Appendix is to recall some basic commutator formulas for the operators used in the body of the paper. First recall that if $A : O(M) \to \text{so}(T_oM)$ is a smooth function let $\hat{A}$ denote the vertical vector field on $O(M)$ defined as: $(\hat{A}f)(u) \equiv \frac{d}{dt}|_0 f(ue^{tA(u)})$. Let $\Omega(a,b)$ denote the vertical vector field $\hat{A}$, where $A(u) \equiv \Omega_u(a,b)$. Similarly, let $B_{\circ(a,b)}$ denote the horizontal vector field on $O(M)$ defined by $u \in O(M) \to B_{\circ(a,b)}(u) \in TO(M)$. The following lemma is standard, for a proof see Kobayshi and Nomizu [26] or Lemma 9.2 in [8] for example.

**Lemma 7.1.** – Let $a, b \in T_oM$ and $A, C \in \text{so}(T_oM)$, then

1. $[\hat{A}, \hat{C}] = [A, C]$, and
2. $[B_{\circ(a,b)}, B_{\circ(a,b)}] = -\Omega(a,b) - B_{\circ(a,b)}$,
3. $[\hat{A}, B_a] = B_{Aa}$.
Remark 7.2. The last commutator formula easily generalizes to

(7.1) \[ [A, B] = B_{Aa} - B_{aA} \]

when \( A : O(M) \to \text{so}(T_o M) \) is a non-constant smooth function.

Lemma 7.3. Let \( a \in T_o M \) and \( A \in \text{so}(T_o M) \), then \( [\mathcal{L}, \hat{A}] = 0 \) and

(7.2) \[ [\mathcal{L}, B_a] = \sum_{c \in S} \{ (B_a, \Omega(c, a)) - 2B_a, \hat{\Omega}(c, a) \} + B_{\text{Ric}, a} \]

Proof. We compute,

(7.3) \[ [\mathcal{L}, \hat{A}] = \sum_{a \in S} B_a [B_a, \hat{A}] + [B_a, \hat{A}] B_a = - \sum_{a \in S} \{ B_a B_{Aa} + B_{Aa} B_a \} \]

\[ = - \sum_{a, c \in S} \{ B_a B_c + B_b B_a \} (Aa, c) \]

which is zero, since \( (Aa, c) \) is skew symmetric in \( a \) and \( c \) while \( \{ B_a B_c + B_c B_a \} \) is symmetric. Similarly,

\[ [\mathcal{L}, B_a] = \sum_{c \in S} [B^2_a B_c] = \sum_{c \in S} \{ B_c [B_a, B_c] + [B_c, B_a] B_c \} \]

\[ = - \sum_{c \in S} \{ B_c \hat{\Omega}(c, a) + \hat{\Omega}(c, a) B_c \} - \sum_{c \in S} \{ B_a B_{\Omega(c, a)} + B_{\Omega(c, a)} B_c \} \]

\[ = - \sum_{c \in S} \{ 2B_c \hat{\Omega}(c, a) + [\hat{\Omega}(c, a), B_c] \} \]

\[ = - \sum_{c \in S} \{ \sum_{b \in S} (B_b B_c + B_c B_b) (\Omega(c, a), b) + B_{B_c \Omega(c, a)} \} \]

Now

\[ \sum_{b, c \in S} (B_b B_c + B_c B_b) (\Omega(c, a), b) = 0. \]

since \( (B_b B_c + B_c B_b) \) is symmetric in \( b \) and \( c \) while \( (\Omega(c, a), b) \) is anti-symmetric in \( b \) and \( c \) because of the TSS assumption. Also

\[ \sum_{c \in S} \{ [\hat{\Omega}(c, a), B_c] + B_{B_c \Omega(c, a)} \} = \sum_{c \in S} \{ B_{\Omega(c, a) c} - (B_c \Omega(c, a)) \} - B_{\hat{\Omega} a} \]

\[ = -B_{\text{Ric}, a} \sum_{c \in S} (B_c \Omega(c, a)) - B_{\hat{\Omega} a} \]

\[ = -B_{\text{Ric}, a + \hat{\Omega} a} - \sum_{c \in S} (B_c \Omega(c, a)) \]

\[ = -B_{\text{Ric}, a} - \sum_{c \in S} (B_c \Omega(c, a)) \]

Assembling the last three displayed equations proves Eq. (7.2). Q.E.D.
We now apply the formula for $[\mathcal{L}, B_a]$ to functions on $O(M)$ coming from functions and vector fields on $M$. To state the next result, if $Z : O(M) \to T_o M$ is a smooth function, let

$$B \cdot Z \equiv \sum_{a \in S} \langle B_a Z, a \rangle = \sum_{a \in S} B_a \langle Z, a \rangle.$$  

**Corollary 7.4.** Suppose that $F \in C^\infty(O(M))$ and $Z \in C^\infty(O(M) \to T_o M)$ such that $F(uq) = F(u)$ and $Z(uq) = q^{-1}Z(u)$ for all $q \in O(T_o M)$, i.e. $F = f \circ \pi$ and $Z(u) = u^{-1}Y(\pi(u))$, where $f(Y)$ is a smooth function (vector field) on $M$, then,

$$[\mathcal{L}, B_a] F = B \text{Ric}_{\pi*} F$$

and

$$[\mathcal{L}, B_a] Z \equiv \mathcal{L}(B \cdot Z) - B \cdot LZ = -\langle Z, J \rangle - \sum_{a \in S} (B_a Z, \text{Ric} a),$$

where

$$J \equiv \sum_{a \in S} B_a \text{Ric} a.$$

**Remark 7.5.** The commutator formula in Eq. (7.6) may be written directly for the vector field $Y$ as,

$$[\Delta, \nabla :] Y \equiv \Delta(\nabla \cdot Y) - \nabla \cdot \Delta Y = -\langle Y, j \rangle - \langle \nabla Y, \text{Ric} \rangle,$$

where $j$ is the vector field on $M$ such that $j(m) = \sum_{c \in S} \langle \nabla c \text{Ric} \rangle e$ for all $m \in M$ and any orthonormal basis $S$ of $T_m M$.

**Proof.** Suppose that $\rho : G = SO(n) \to \text{Aut}(V)$ is a representation, $A : O(M) \to \text{so}(T_o M)$, and $W : O(M) \to V$ are smooth functions and that $W(uq) = \rho(g^{-1})W(u)$ for all $u \in O(M)$ and $g \in G$, then

$$\left(\dot{A}W\right)(u) = \frac{d}{dt} \bigg|_0 W(ue^{tA(u)}) = \frac{d}{dt} \bigg|_0 \rho(e^{-tA(u)})W(u) = -\rho_*(A(u))W(u),$$

where $\rho_*(A) \equiv \frac{d}{dt} |_0 \rho_*(e^{tA})$. In particular,

$$\left(\dot{A}Z\right)(u) = \frac{d}{dt} \bigg|_0 Z(ue^{tA(u)}) = \frac{d}{dt} \bigg|_0 e^{-tA(u)}u^{-1}Y(\pi(u)) = -A(u)Z(u),$$

which we abbreviate as: $\dot{A} Z = -AZ$.

We then have the following important commutator formula:

$$[\mathcal{L}, B_a] W = \sum_{c \in S} \left\{(B_c \Omega(c, a) - 2B_c \tilde{\Omega}(c, a)) W + B_{\text{Ric}, c} W\right\} + B_{\text{Ric}, c} W$$

$$= \sum_{c \in S} \left\{-\rho_*(B_c \Omega(c, a)) W + 2B_c(\rho_*(\Omega(c, a)) W)\right\} + B_{\text{Ric}, c} W$$

$$- \sum_{c \in S} \left\{\rho_*(B_c \Omega(c, a)) W + 2\rho_*(\Omega(c, a)) B_c W\right\} + B_{\text{Ric}, c} W.$$
Eq. (7.5) follows by taking \( W = F \), in which case \( \rho_* = 0 \). If \( W = Z \) in the above equation (in which case \( \rho_* = Id \)), then
\[
[\mathcal{L}, B_\alpha] Z = \sum_{c \in S} \{(B,\Omega\langle c, a \rangle)Z + 2\Omega\langle c, a \rangle B_\alpha Z\} + B \text{Ric} \cdot a \cdot Z.
\]
Hence, we have
\[
\sum_{a \in S} [\mathcal{L}, B_\alpha] Z, a = \sum_{a \in S} \{(B,\Omega\langle c, a \rangle)Z, a + 2\Omega\langle c, a \rangle B_\alpha Z, a\} + \sum_{a \in S} B \text{Ric} \cdot a \langle Z, a \rangle
\]
\[
= \sum_{a \in S} \{-\langle Z, (B_\alpha \Omega\langle c, a \rangle) a \rangle - 2\langle B_\alpha Z, \Omega\langle c, a \rangle a \rangle + \langle B_\alpha Z, a \rangle \langle \text{Ric}, a, c \rangle\}
\]
\[
= -\sum_{c \in S} \{(Z, B_\alpha \text{Ric} c) + 2\langle B_\alpha Z, \text{Ric} c \rangle - \langle B_\alpha Z, \text{Ric} c \rangle\}
\]
\[
= -(Z, J) - \sum_{c \in S} \langle B_\alpha Z, \text{Ric} c \rangle.
\]
which is Eq. (7.6).

Acknowledgments

I would like to thank David Elworthy for illuminating discussions on topics related to this work.

REFERENCES


[9] B. K. Driver, A primer on Riemannian geometry and stochastic analysis on path spaces, ETH (Zürich, Switzerland) preprint series.


INTEGRATION BY PARTS FOR HEAT KERNEL MEASURES

[24] E. P. Hsu, Lectures given at the IAS/Park City Mathematics Institute, Summer Session, held June 23-July 13, 1996 at the Institute for Advanced Study.

(Manuscript received December 30, 1996.)

B. K. Driver
Department of Mathematics, 0112
University of California, San Diego
La Jolla, CA 92093-0112
Email: driver@euclid.ucsd.edu

JOURNAL DE MATHÉMATIQUES PURES ET APPLIQUÉES