J. Math. Pures Appl.,
76, 1997, *p.* 703-737

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

(*) 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 MATHÉMATIQUES PURES ET APPLIQUÉES. – 0021-7824/1997/08/$ 7.00

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, ∇ 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 ∇ respectively. If $M = \mathbb{R}^d$, we take $\langle\cdot,\cdot\rangle$ to be the usual dot product on $\mathbb{R}^d$ and ∇ to be the Levi Civita covariant derivative. In all cases we will assume that ∇ is *Torsion Skew Symmetric* or *TSS* for short, *i.e.* that $\langle T\langle X, Y\rangle, Y\rangle \equiv 0$ for all vector fields X and Y on M. Let $\{\Sigma_t\}_{t\geq 0}$, $\{//_t\}_{t\geq 0}$, and $\{b(t)\}_{t\geq 0}$ be three adapted continuous processes on a filtered probability space such that Σ is an M-valued Brownian motion, $//$ is stochastic parallel translation along Σ, and $b(t) \equiv \int_0^t //_\tau^{-1}\delta\Sigma_\tau$ – a T_oM–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

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

Given $h \in H(T_oM)$ let X^h denote the *Cameron-Martin vector field* on $W(M)$ given by $X_t^h \equiv //_t h(t)$. 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 $h \in H(T_oM)$. 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 ∂_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, $//$. In this paper we will give another elementary proof of the integration by parts formula for cylinder functions of $//$. 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 $\int_0^T \dot{l}^2(t)\, dt < \infty$, then

$$E[(\nabla\cdot Y)(\Sigma_T)] = E\left[\left\langle //_T^{-1}Y(\Sigma_T), \int_0^T \left(\dot{l}(t) - \frac{1}{2}l(t)\mathrm{Ric}_{//_t}\right)\overleftarrow{db}(t)\right\rangle\right]. \tag{1.1}$$

In this formula $\nabla\cdot Y$ is the divergence of Y, $\mathrm{Ric}_{//_t} = //_t^{-1}\mathrm{Ric}\, //_t$, Ric is the Ricci tensor and $\overleftarrow{db}$ 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:

$$\begin{aligned} E[(\nabla \cdot Y)(\Sigma_T)] &= \int_M p_T(o,x) \nabla \cdot Y(x)\, dx \\ &= -\int_M \langle \vec{\nabla}_x p_T(o,x), Y(x)\rangle\, dx \\ &= -\int_M \langle \vec{\nabla}_x \ln p_T(o,x), Y(x)\rangle\, p_T(o,x) dx \\ &= -E[\langle (\vec{\nabla} \ln p_T(o,\cdot))(\Sigma_T), Y(\Sigma_T)\rangle], \end{aligned}$$

where $\vec{\nabla} f$ 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

$$\text{(1.2)} \quad \langle \vec{\nabla}_x \ln p_T(o,x), v\rangle = -E\left[\left\langle //_T^{-1} v, \int_0^T \left(\dot{l}(t) - \frac{1}{2} l(t) \mathrm{Ric}_{//_t}\right) \overleftarrow{db}(t)\right\rangle \middle| \Sigma_T = x\right],$$

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 $\vec{\nabla}_x \ln p_T(o, x)$ 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 Σ_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

$$(k,k) \equiv \int_0^\infty |\dot{k}(t)|^2 dt < \infty.$$

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

$$f_{\mathcal{P}} \equiv f(b(t_1), b(t_2)\ldots, b(t_n)) \tag{2.1}$$

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

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

$$(\mathcal{A}_i f)(x_1, x_2, \ldots, x_n) \equiv (\mathcal{A} f(x_1, \ldots x_{i-1}, \cdot, x_{i+1}, \ldots, x_n))(x_i),$$

where $x_i \in \mathbb{R}^d$ for each i. That is $\mathcal{A}_i f$ denotes $\mathcal{A}$ 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 $\mathcal{A}_i f_{\mathcal{P}}$ for $(\mathcal{A}_i f)(b(t_1), b(t_2)\ldots, b(t_n))$. In particular if $n = 1$, then we will write $\mathcal{A} f_{\mathcal{P}}$ for $(\mathcal{A} f)(b(T))$.*

For $f \in C^\infty(\mathbb{R}^d)$, let $Df \in C^\infty(\mathbb{R}^d, (\mathbb{R}^d)^*)$ denote the differential of f, *i.e.* $Df(x)a \equiv \frac{d}{ds}|_0 f(x + sa)$. Notice that $D : C^\infty(\mathbb{R}^d) \to C^\infty(\mathbb{R}^d, (\mathbb{R}^d)^*)$ is a differential operator and hence

$$(D_i f)(x_1, x_2, \ldots, x_n)a \equiv \frac{d}{ds}|_0 f(x_1, \ldots x_{i-1}, x_i + sa\cdot, x_{i+1}, \ldots, x_n).$$

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

$$X^k f_{\mathcal{P}} = \sum_{i=1}^n (D_i f)(b(t_1), b(t_2)\ldots, b(t_n))k(t_i). \tag{2.2}$$

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

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

$$(e^{t\Delta/2} f)(x) \equiv \begin{cases} \int_{\mathbb{R}^d} p_t(x-y) f(y)\, dy & \text{if } t > 0, \\ f(x) & \text{if } t = 0, \end{cases}$$

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* $f_{\mathcal{P}}$ *be a cylinder function as in Eq. (2.1), then*

$$E[X^k f_{\mathcal{P}}] = E\left[f_{\mathcal{P}}\int_0^T \langle \dot{k}, db\rangle\right]. \tag{2.3}$$

Moreover, if $f_{\mathcal{P}}$ *and* $g_{\mathcal{P}}$ *are two cylinder functions,*

$$E[(X^k f_{\mathcal{P}})g_{\mathcal{P}}] = E\left[f_{\mathcal{P}}\left(-X^k g_{\mathcal{P}} + \int_0^T \langle \dot{k}, db\rangle g_{\mathcal{P}}\right)\right]. \tag{2.4}$$

Proof. – Let us first notice that Eq. (2.4) is a consequence of Eq. (2.3) with $f_{\mathcal{P}}$ replaced by $f_{\mathcal{P}}g_{\mathcal{P}} = (fg)_{\mathcal{P}}$ and the product rule, $X^k(f_{\mathcal{P}}g_{\mathcal{P}}) = (X^k f_{\mathcal{P}})g_{\mathcal{P}} + f_{\mathcal{P}}(X^k g_{\mathcal{P}})$. 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 $\mathcal{P} = \{0 < T\}$ and $\mathcal{P} = \{0 < u < T\}$.

Case 1. – Suppose $f \in C_c^\infty(\mathbb{R}^d)$, $f_{\mathcal{P}} = f(b(T))$, and N_t and W_t are defined as in Lemma 2.2. Then $X^k f_{\mathcal{P}} = Df(b(T))k(T) = W_T k(T)$, so that

$$E[X^k f_{\mathcal{P}}] = E[W_T k(T)] = E\left[\int_0^T \{dWk + Wdk\}\right] = E\left[\int_0^T W_t \dot{k}(t)\,dt\right], \tag{2.5}$$

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,

$$\begin{aligned} E\left[\int_0^T W_t \dot{k}(t)\,dt\right] &= E\left[\int_0^T W db \int_0^T \langle \dot{k}, db\rangle\right] \\ &= E\left[(N_T - N_0)\int_0^T \langle \dot{k}, db\rangle\right] = E\left[f_{\mathcal{P}}\int_0^T \langle \dot{k}, db\rangle\right], \end{aligned} \tag{2.6}$$

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

Case 2. – Suppose $\mathcal{P} = \{0 < u < T\}$, $f \in C_c((\mathbb{R}^d)^2)$, and $f_{\mathcal{P}} = f(b(u), b(T))$. Now let

$$f_T(t,x,y) \equiv e^{\frac{(T-t)\Delta_2}{2}} f(x,y) \equiv \int_{\mathbb{R}^d} p_{(T-t)}(y-z) f(x,z)dz,$$

$$N_t \equiv f_T(t, b(u), b(t)) \text{ for all } t \in [u,T]$$

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

$$\begin{aligned} X^k f_{\mathcal{P}} &= D_1 f(b(u), b(T)) k(u) + D_2 f(b(u), b(T)) k(T) \\ &= D_1 f(b(u), b(T)) k(u) + W_T k(T), \end{aligned}$$

so that

$$\begin{aligned} E[X^k f_{\mathcal{P}}] &= E[D_1 f(b(u), b(T)) k(u) + W_T k(T)] \\ &= E[D_1 f(b(u), b(T)) k(u)] + E\left[W_u k(u) + \int_u^T \{dWk + Wdk\}\right] \\ &= E[D_1 f_T(u, b(u), b(u)) k(u)] + E\left[W_u k(u) + \int_u^T W_t \dot{k}(t)\, dt\right], \end{aligned}$$

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),

$$E\left[\int_u^T W_t \dot{k}(t)\, dt\right] = E\left[(N_T - N_u) \int_u^T \langle \dot{k}, db \rangle\right] = E\left[f_{\mathcal{P}} \int_u^T \langle \dot{k}, db \rangle\right].$$

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

(2.7)

$$E[X^k f_{\mathcal{P}}] = E[D_1 f_T(u, b(u), b(u)) k(u) + D_2 f_T(u, b(u), b(u)) k(u)] + E\left[f_{\mathcal{P}} \int_u^T \langle \dot{k}, db \rangle\right].$$

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

$$E[X^k f_{\mathcal{P}}] = E[X^k \tilde{f}] + E\left[f_{\mathcal{P}} \int_u^T \langle \dot{k}, db \rangle\right]. \tag{2.8}$$

By case 1) already proved,

$$\begin{aligned} E[X^k \tilde{f}] &= E\left[\tilde{f} \int_0^u \langle \dot{k}, db \rangle\right] = E\left[f_T(u, b(u), b(u)) \int_0^u \langle \dot{k}, db \rangle\right] \\ &= E\left[f(b(u), b(T)) \int_0^u \langle \dot{k}, db \rangle\right] = E\left[f_{\mathcal{P}} \int_0^u \langle \dot{k}, db \rangle\right], \end{aligned} \tag{2.9}$$

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 $\mathcal{P}$. 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,

$$E[(\nabla \cdot Y)(b(T)] = \frac{1}{T} E[\langle Y(b(T)), b(T)\rangle]. \tag{2.10}$$

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 $h_c(t) \equiv \frac{t}{T} c$, $X_t^c \equiv X_t^{h_c} \equiv h_c(t)$, and

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

Let $\mathbf{1}$ denote the constant cylinder function "one" of the Brownian motion b. Since $X^c\mathbf{1} = 0$,

$$\begin{aligned} 0 &= \sum_{c\in S} E[(X^c\mathbf{1})\langle X_T^c, Y(b(T))\rangle] = \sum_{c\in S} E[(X^c)^* \langle X_T^c, Y(b(T))\rangle] \\ &= \sum_{c\in S} E[(-X^c + z_T^c)\langle c, Y(b(T))\rangle] \\ &= \sum_{c\in S} E[-X^c\langle c, Y(b(T))\rangle] + \frac{1}{T} E\langle Y(b(T), b(T)\rangle, \end{aligned}$$

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

$$\sum_{c\in S} X^c\langle c, Y(b(T))\rangle = (\nabla \cdot Y)(b(T)).$$

Q.E.D.

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

$$Y_t(x) \equiv (e^{\frac{T-t}{2}\Delta} Y)(x) \equiv \int_{\mathbb{R}^d} p_{(T-t)}(x-y) Y(y) dy$$

and

$$Q_t \equiv t(\nabla \cdot Y_t)(b(t)) - \langle b(t), Y_t(b(t))\rangle. \tag{2.11}$$

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

$$\begin{aligned} dQ_t &= (\nabla \cdot Y_t)(b(t))dt + t(\partial_{db(t)}\nabla \cdot Y_t)(b(t)) \\ &\quad - \langle db(t), Y_t(b(t))\rangle - \langle(\partial_{db(t)}Y_t)(b(t)), b(t)\rangle - \langle db(t), (\partial_{db(t)}Y_t)(b(t))\rangle \\ &= t(\partial_{db(t)}\nabla \cdot Y_t)(b(t)) - \langle db(t), Y_t(b(t))\rangle - \langle(\partial_{db(t)}Y_t)(b(t)), b(t)\rangle. \end{aligned}$$

Therefore, Q_t is an L^2-martingale and in particular $EQ_T = EQ_0$. This proves Eq. (2.10), since $Q_T = T(\nabla \cdot Y)(b(T)) - b(T) \cdot Y(b(T))$ 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, ∇ 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 ∇ respectively. So

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

and

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

for all smooth vector fields X, Y, and Z on M.

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

The Laplacian (Δ) with respect to (∇) is the second order differential operator acting on the smooth functions $f \in C^\infty(M)$ given by $\Delta f \equiv \mathrm{tr}\nabla df \equiv \sum_{i=1}^n \{E_iE_i f - df\langle\nabla_{E_i}E_i\rangle\}$, 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 (∇).

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

(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 σ in M such that $u(s) \in O_{\sigma(s)}(M)$, let $\nabla u(s)/ds : T_oM \to T_{\sigma(s)}M$ denote the linear operator defined by $(\nabla u(s)/ds)a = \nabla(u(s)a)/ds$ for all $a \in T_oM$. Notice that $V(s) \equiv u(s)a$ is a vector field along σ so that $\nabla V(s)/ds = \nabla(u(s)a)/ds$ makes sense.

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

DEFINITION 3.2 (Connection 1-form). – *Let ω be the $so(T_oM)$-valued connection 1-form on $O(M)$ given by $\omega\langle u'(0)\rangle \equiv u(0)^{-1}\nabla u(s)/ds|_{s=0}$, 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 $\omega\langle u'\rangle \equiv 0$.*

DEFINITION 3.3 (Horizontal Vector Fields). – *For $a \in T_oM$ and $u \in O(M)$ let $B_a(u) \in T_uO(M)$ be defined by $\omega\langle B_a(u)\rangle = 0$ (i.e. $B_a(u)$ is horizontal) and $\pi_*B_a(u) = ua$.*

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

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 $a, b \in T_oM$ and an isometry $u : T_oM \to T_mM$ (i.e. $u \in O_m(M)$), define*

$$\Omega_u\langle a, b\rangle \equiv u^{-1}R^\nabla\langle ua, ub\rangle u \in so(T_oM),$$

$$Ric_u a \equiv \sum_{c\in S} \Omega_u\langle a, c\rangle c$$

$$\Theta_u\langle a, b\rangle \equiv u^{-1}T^\nabla\langle ua, ub\rangle \in T_oM$$

and

$$\bar{\Theta}_u a \equiv \sum_{c\in S} (B_c\Theta)_u\langle a, c\rangle),$$

where

$$(B_c\Theta)_u\langle a, c\rangle \equiv \frac{d}{ds}\bigg|_0 \Theta_{e^{sB_c}(u)}\langle a, c\rangle.$$

So Ω, Ric, and Θ are the equi-invariant forms of the curvature tensor, the Ricci tensor, and the torsion tensor respectively. Similarly, $\bar{\Theta}_u$ is the equi-invariant form of a contraction of ∇T^∇. It is well known that the Ricci tensor, Ric, is symmetric if ∇ 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 ∇ is torsion skew symmetric, then*

$$Ric_u^* = Ric_u + \bar{\Theta}_u. \tag{3.1}$$

Proof. – By Theorem 9.4 of [8] for any $a, b, c \in T_oM$,

$$\langle\Omega\langle b, a\rangle c, b\rangle - \langle\Omega\langle b, c\rangle a, b\rangle - \langle B_b\Theta\langle a, c\rangle, b\rangle = 0. \tag{3.2}$$

Using $\Omega\langle b, a\rangle = -\Omega\langle a, b\rangle$, $\langle\Theta\langle a, c\rangle, b\rangle = -\langle\Theta\langle a, c\rangle, b\rangle)$ (since ∇ is TSS), and the fact that $\Omega_u\langle a, b\rangle \in so(T_oM)$, Eq. (3.2) may be written as,

$$\langle\Omega\langle a, b\rangle b, c\rangle - \langle\Omega\langle c, b\rangle b, a\rangle + \langle B_b\Theta\langle a, b\rangle, c\rangle = 0.$$

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

$$\langle \mathrm{Ric}\, a, c\rangle - \langle \mathrm{Ric}\, c, a\rangle + \langle\tilde{\Theta}a, c\rangle = 0$$

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

3.2. The Basic Processes

NOTATION 3.6. – *Given a (vector valued) semi-martingale X, δX will denote the Stratonovich differential of X while dX will denote the Itô differential of X.*

Let $(\mathcal{W}, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{F}, P)$ be a filtered probability space satisfying 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. 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 $(\mathcal{W}, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{F}, P)$ with values in M, $O(M)$, and T_oM respectively such that Σ is a M-valued Brownian motion, $//$ is parallel translation along Σ, and $b(t)$ is the "undevelopment" of Σ. To be more precise, we are assuming:

1. $\Sigma_0 = o$ and Σ is a diffusion process on M with generator $\frac{1}{2}\Delta$, *i.e.* for all $f \in C^\infty(M)$,

$$f(\Sigma_t) - f(\Sigma_0) - \frac{1}{2}\int_0^t (\Delta f)(\Sigma_\tau)d\tau$$

is a martingale.

2. $//_0 = I_{T_oM}$, $\pi \circ //_t = \Sigma_t$ for all $t \geq 0$, and $\int_0^t \omega\langle\delta //_\tau\rangle$ is the zero process.
3. b is an T_oM-valued Brownian motion,

$$b(t) = \int_0^t \theta\langle\delta //_\tau\rangle, \tag{3.3}$$

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

$$\delta //_t = B_{\delta b(t)}(//_t) \equiv \sum_{c\in S} B_c(//_t)\delta b_c(t), \tag{3.4}$$

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 Σ and that stochastic parallel translation exists uniquely along any M-valued semi-martingale Σ.

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

$$H(T_oM) \equiv \left\{k : [0, \infty) \to T_oM \,\middle|\, k \text{ is absolutely continuous and } \int_0^\infty |\dot{k}(t)|^2 dt < \infty\right\}.$$

Given k in the Cameron-Martin space $H(T_oM)$, let X^k denote the *Cameron-Martin vector field* given by $X_t^k \equiv //_t k(t)$. (Notice that $\{X_t^k\}_{t\geq 0}$ is a TM-valued process on $\mathcal{W}$.) It was shown in [D5]: if $\mathcal{W} = W(M)$, Σ 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 $k \in H(T_oM)$. 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 ∂_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* $k \in H(T_oM)$ *is given, let*

(3.5)
$$A_t^k \equiv \int_0^t \Omega_{//_\tau}\langle k(\tau), \delta b(\tau)\rangle = \int_0^t \Omega_{//_\tau}\langle k(\tau), db(\tau)\rangle + \frac{1}{2}\sum_{a\in S}\int_0^t (B_a\Omega)_{//_\tau}\langle k(\tau), a\rangle d\tau,$$

where

$$(B_a\Omega)_u\langle b, c\rangle \equiv \frac{d}{ds}\bigg|_0 \Omega_{e^{tB_a}(u)}\langle b, c\rangle. \tag{3.6}$$

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

$$(\hat{A}F)(u) \equiv \frac{d}{ds}\bigg|_0 F(ue^{sA(u)}) \tag{3.7}$$

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

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

$$F_\mathcal{P} = F(//_{t_1}, //_{t_2} \dots, //_{t_n}). \tag{3.8}$$

Any function $F_\mathcal{P}$ *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* $F_\mathcal{P}$, *we will say that* f *is based on* $\mathcal{P}$ *and the* **degree** *of* f *is less than or equal to* n.

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

$$(\mathcal{A}_i F)(u_1, u_2, \ldots, u_n) \equiv (\mathcal{A}F(u_1, \ldots u_{i-1}, \cdot, u_{i+1}, \ldots, u_n))(u_i),$$

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

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

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

Hence by the above notation, $D_i F_\mathcal{P} : \mathcal{W} \to (T_oM \oplus so(T_oM))^*$ is defined by:

$$D_i F_\mathcal{P} a \equiv \frac{d}{ds}\bigg|_0 F(//_{t_1}, \ldots, //_{t_{i-1}}, e^{sB_a}(//_{t_i}), //_{t_{i+1}}, \ldots, //_{t_n}) \text{ and} \tag{3.9}$$

$$D_i F_\mathcal{P} A \equiv \frac{d}{ds}\bigg|_0 F(//_{t_1}, \ldots, //_{t_{i-1}}, //_{t_i} e^{sA}, //_{t_{i+1}}, \ldots, //_{t_n}), \tag{3.10}$$

where $a \in T_oM$ and $A \in so(T_oM)$.

DEFINITION 3.11. – *Given k in the Cameron-Martin space $H(T_oM)$, let X^k denote the* Cameron-Martin vector field *on $W(M)$ given by $X_t^k \equiv //_t k(t)$. We let X^k act on the cylinder function $F_\mathcal{P}$ given in (3.8) via*

$$X^k F_\mathcal{P} = \sum_{i=1}^n D_i F_\mathcal{P}(k(t_i) - A_{t_i}^k). \tag{3.11}$$

To motivate this definition, suppose for the moment that $\mathcal{W} = W(M)$ and P is Wiener measure on $\mathcal{W}$. 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 $\mathcal{W}$ 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

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

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 $F \in C^\infty(O(M))$,

$$X^k(F(//_T)) = ((//_T)_* X^k)F = (DF)(//_T)(k(T) - A_T^k). \tag{3.13}$$

Clearly, Eq. (3.11) is the natural generalization of Eq. (3.13) to cylinder functions of $//$ 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 $k \in H(T_oM)$ and $F_{\mathcal{P}}$ be a cylinder function of $//$, as in Eq.* (3.8), *then*

$$E[X^k F_{\mathcal{P}}] = E[z_T F_{\mathcal{P}}], \tag{3.14}$$

where

$$z_t \equiv \int_0^t \left\langle \left\{ \dot{k}(\tau) + \frac{1}{2} Ric^*_{//_\tau} k(\tau) \right\}, db(\tau) \right\rangle. \tag{3.15}$$

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

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

Proof. – Let $F_{\mathcal{P}}$ and $G_{\mathcal{P}}$ be two cylinder functions of $//$. Upon noting that $X^k(F_{\mathcal{P}}G_{\mathcal{P}}) = (X^kF_{\mathcal{P}})G_{\mathcal{P}} + F_{\mathcal{P}}X^kG_{\mathcal{P}}$, Theorem 3.12 with $F_{\mathcal{P}}$ replaced by $(F_{\mathcal{P}}G_{\mathcal{P}})$ implies

$$E[(X^kF_{\mathcal{P}})G_{\mathcal{P}}] = E[F_{\mathcal{P}}(-X^k + z_T)G_{\mathcal{P}}]. \tag{3.16}$$

Suppose that $F_{\mathcal{P}}$ and $H_{\mathcal{P}'}$ are two cylinder functions based on $[0, T]$ ($\mathcal{P}'$ is possibly another partition of $[0, T]$) such that $F_{\mathcal{P}} = H_{\mathcal{P}'}$ a.e. Let $\tilde{\mathcal{P}}$ be any partition of $[0, T]$ such that $\mathcal{P} \cup \mathcal{P}' \subset \tilde{\mathcal{P}}$ and $G_{\tilde{\mathcal{P}}}$ be any cylinder function of $//$ based on $\tilde{\mathcal{P}}$. Then by Eq. (3.16) with $F_{\mathcal{P}}$ replaced by $F_{\mathcal{P}} - H_{\mathcal{P}'}$ and $G_{\mathcal{P}}$ replaced by $G_{\tilde{\mathcal{P}}}$,

$$E[(X^kF_{\mathcal{P}} - X^kH_{\mathcal{P}'})G_{\tilde{\mathcal{P}}}] = E[(F_{\mathcal{P}} - H_{\mathcal{P}'})(-X^k + z_T)G_{\tilde{\mathcal{P}}}] = 0. \tag{3.17}$$

Since $(X^kF_{\mathcal{P}} - X^kH_{\mathcal{P}'})$ is $\mathcal{G}_T$ measurable and the collection of cylinder functions $G_{\tilde{\mathcal{P}}}$ as above are dense in $L^2(\mathcal{W}, \mathcal{G}_T, P)$, it follows that $X^kF_{\mathcal{P}} - X^kH_{\mathcal{P}'} = 0$ a.e. This shows that X^k is well defined as an operator on $L^2(\mathcal{W}, \mathcal{G}_T, P)$ with domain $\mathcal{D}(X^k)$. Given this, the assertion that $-X^k + z_T \subset (X^k)^*$ 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,

$$\partial G(t, u)/\partial t = \frac{1}{2}\mathcal{L}G(t, u) \text{ with } G(0, \cdot) = F(\cdot), \tag{3.18}$$

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 $//_t^u$ *denotes the solution to the stochastic differential equation (3.4) with* $//_0^u = u \in O(M)$, *then*

$$G(t, u) = E[F(//_t^u)]. \tag{3.19}$$

(In the sequel we will often write $E[F(//_t^u)]$ *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*

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

then the function $(t, q, u) \to e^{t\mathcal{L}}F(q, u)$ *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 $e^{t\mathcal{L}}F = (e^{t\Delta}f) \circ \pi$.

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

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

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

$$W_t(a + A) \equiv ((B_a + \hat{A})F_T)(t, //_t) \text{ for all } t \in [t_{n-1}, T], \tag{3.20}$$

where $(a + A) \in T_oM \oplus so(T_oM)$. *Moreover, the differential of* W_t *is given by*

$$\begin{aligned} dW_t(a + A) = {} & (B_{db(t)}(B_a + \hat{A})F_T)(t, //_t) \\ & + \frac{1}{2}W_t(Ric^*_{//_t}a + \sum_{c \in S}(B_c\Omega)_{//_t}\langle c, a\rangle)dt \\ & - \sum_{c \in S}(B_c\hat{\Omega}\langle c, a\rangle F_T)(t, //_t)\, dt. \end{aligned} \tag{3.21}$$

Here $\hat{\Omega}\langle c, a\rangle \equiv \hat{A}$ *as in Eq. (3.7) where* $A(u) \equiv \Omega_u\langle c, a\rangle$.

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

$$\begin{aligned} dN_t &= \sum_{a \in S}(B_aF_T)(t, //_t)\delta b_a(t) - \frac{1}{2}(\mathcal{L}F_T)(t, //_t)dt \\ &= \sum_{a \in S}(B_aF_T)(t, //_t)db_a(t) + \frac{1}{2}\sum_{a \in S}(B_a^2F_T)(t, //_t)dt - \frac{1}{2}(\mathcal{L}F_T)(t, //_t)dt \\ &= W_tdb(t). \end{aligned}$$

A similar computation shows that

$$dW_t(a+A) = \sum_{c\in S}(B_c(B_a+\hat{A})F_T)(t,//_t)db_c + \frac{1}{2}([\mathcal{L},(B_a+\hat{A})]F_T)(t,//_t)dt. \tag{3.22}$$

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

$$dW_t(a+A) = \sum_{c\in S}(B_c(B_a+\hat{A})F_T)(t,//_t)db_c + \frac{1}{2}\sum_{c\in S}(\{-2B_c\hat{\Omega}\langle c,a\rangle + B_{\mathrm{Ric}^*a} + (B_c\Omega\langle c,a\rangle)\hat{}\}F_T)(//_t)dt,$$

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

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

COROLLARY 3.18. – *Let $\mathcal{P} = \{0 < T\}$, $F_{\mathcal{P}} = f(\Sigma_T) = F(//_T)$, where $f \in C^\infty(M)$ 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 $dN_t = W_t db(t)$, where W_t is the $(T_oM)^*$-valued process defined by

$$W_t a \equiv (\nabla_{//_t a}e^{(T-t)\Delta/2}f)(\Sigma_t) \text{ for all } t\in[0,T]. \tag{3.23}$$

Moreover, the differential of W_t is given by

$$dW_t a = (\nabla^2_{//_t db(t)\otimes //_t a}e^{(T-t)\Delta/2}f)(\Sigma_t) + \frac{1}{2}W_t Ric^*_{//_t} a\, dt, \tag{3.24}$$

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

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

Here X and Y are arbitrary vector fields on M.

Proof. – This is a straight forward application of Lemma 3.17 using the following remarks. For $F = f\circ\pi$, $(\hat{A}F)(u) = 0$, $(B_aF)(u) = \nabla_{ua}f$, $(B_aB_cF)(u) = \nabla^2_{ua\otimes ub}f$, and $e^{t\mathcal{L}}F = (e^{t\Delta}f)\circ\pi$. 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 $k\in H(T_oM)$, $f\in C^\infty(M)$, and $F_{\mathcal{P}} = f(\Sigma_T)$, then*

$$E[X^kF_{\mathcal{P}}] = E[z_TF_{\mathcal{P}}], \tag{3.25}$$

where z is defined in (3.15).

Proof. – Let

$$K(t) = k(t) + \frac{1}{2}\int_0^t \mathrm{Ric}^*_{//_\tau} k(\tau)d\tau, \tag{3.26}$$

so that

$$\dot K(t) = \dot k(t) + \frac{1}{2}\mathrm{Ric}^*_{//_t} k(t) \text{ with } K(0) = 0.$$

Let W_t be defined as in Eq. (3.23) and notice that

$$\begin{aligned} X^k F_{\mathcal{P}} &= (B_{k(T)}F)(//_T) = W_T k(T) \\ &= W_0 k(0) + \int_0^T \{W dk + dW k\} \\ &= \int_0^T \left\{W dk + \frac{1}{2} W \mathrm{Ric}^*_{//_t} k dt\right\} + \\ &\quad + \int_0^T (\nabla^2_{//_t db(t) \otimes //_t a} e^{(T-t)\Delta/2} f)(\Sigma_t), \end{aligned}$$

wherein the last equality we have used Eq. (3.24) of Corollary 3.18. Taking expectations of this identity gives:

$$E[X^k F_{\mathcal{P}}] = E\int_0^T \left\{W dk + \frac{1}{2} W \mathrm{Ric}^*_{//_t} k dt\right\} = E\int_0^T W_t \dot K(t)\, dt. \tag{3.27}$$

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

$$\begin{aligned} E\int_0^T W_t \dot K(t)\, dt &= E\left(\int_0^T W_t db(t) \cdot \int_0^T \langle \dot K, db\rangle\right) \\ &= E\left((N_T - N_0) \cdot \int_0^T \langle \dot K, db\rangle\right) \\ &= \\ &= E\left\{N_T \int_0^T \langle \dot K, db\rangle\right\} = E\left\{F_{\mathcal{P}} \int_0^T \langle \dot K, db\rangle\right\}. \end{aligned}$$

Q.E.D.

Proof of Theorem 3.12. – Let $F_{\mathcal{P}}$ be a cylinder function of $//$ as in Eq. (3.8), $K(t)$ be as in Eq. (3.26) W_t as in (3.23), and $\overrightarrow{//} \equiv (//_{t_1}, //_{t_2}, \ldots, //_{t_{n-1}})$. Since $W_T(a + A) = (B_a F + \hat{A} F)(//_T)$,

$$\begin{aligned} EX^k F_{\mathcal{P}} &= \sum_{i=1}^{n} E[D_i F(\overrightarrow{//}, //_T)(k(t_i) - A_{t_i})] \\ &= \sum_{i=1}^{n-1} E[D_i F(\overrightarrow{//}, //_T)(k(t_i) - A_{t_i})] + E[W_T(k(T) - A_T)], \end{aligned} \tag{3.28}$$

where $A_t = A_t^k$ 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

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

$$\begin{aligned} T_2 &\equiv E\int_{t_{n-1}}^{T} dW(k - A) \\ &= -E\int_{t_{n-1}}^{T} \sum_{c\in S}(B_c\hat{\Omega}\langle c, k(t)\rangle F_T)(t, //_t)\, dt \\ &\quad + \frac{1}{2}E\int_{t_{n-1}}^{T} W_t(\mathrm{Ric}^{*}_{//_t}k(t) + \sum_{c\in S}(B_c\Omega)_{//_t}\langle c, k(t)\rangle)dt \end{aligned}$$

and

$$\begin{aligned} T_3 &\equiv -E\int_{t_{n-1}}^{T} dW dA \\ &= -E\int_{t_{n-1}}^{T} \sum_{c\in S}(B_c\hat{A}F_T)(t, //_t)|_{A=\Omega_{//_\tau}\langle k(t), c\rangle}\, dt \\ &= -E\int_{t_{n-1}}^{T} \sum_{c\in S}(B_c\hat{\Omega}\langle k(t), c\rangle F_T)(t, //_t)\, dt \\ &\quad + E\int_{t_{n-1}}^{T} \sum_{c\in S}(((B_c\Omega)_{//_t}\langle k(t), c\rangle)\hat{} F_T)(t, //_t)\, dt \\ &= -E\int_{t_{n-1}}^{T} \sum_{c\in S}(((B_c\hat{\Omega}\langle k(t), c\rangle) F_T)(t, //_t)\, dt \\ &\quad + E\int_{t_{n-1}}^{T} \sum_{c\in S}(W_t((B_c\Omega)_{//_t}\langle k(t), c\rangle)\, dt. \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} T_1 + T_2 &= E\int_{t_{n-1}}^{T} \{W(dK - (B_c\Omega)_{//_t}\langle k(t), c\rangle dt)\} \\ &\quad - E\int_{t_{n-1}}^{T} \sum_{c\in S}(B_c\hat{\Omega}\langle c, k(t)\rangle F_T)(t, //_t)dt, \end{aligned}$$

and

$$\begin{aligned}T_1+T_2+T_3 &= E\int_{t_{n-1}}^{T} W dK\\ &= E\Big(\int_{t_{n-1}}^{T} W db \cdot \int_{t_{n-1}}^{T} \langle \dot{K}, db\rangle\Big)\\ &= E\Big((N_T - N_{t_{n-1}}) \cdot \int_{t_{n-1}}^{T} \langle \dot{K}, db\rangle\Big)\\ &= E\Big(F_{\mathcal{P}} \int_{t_{n-1}}^{T} \langle \dot{K}, db\rangle\Big),\end{aligned}$$

where we have used $dM = W db$ in the third equality and

$$E\Big(N_{t_{n-1}} \int_{t_{n-1}}^{T} \langle \dot{K}, db\rangle\Big) = 0$$

and $N_T = F_{\mathcal{P}}$ in the last. Hence

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

which combined with Eq. (3.28) gives

$$\begin{aligned}(3.29)\quad EX^k F_{\mathcal{P}} = \sum_{i=1}^{n-1} E[D_i F(\overrightarrow{//}, //_T)(k(t_i) - A_{t_i})] + E[W_{t_{n-1}}(k(t_{n-1}) - A_{t_{n-1}})]\\ + E\Big(F_{\mathcal{P}} \int_{t_{n-1}}^{T} \langle \dot{K}, db\rangle\Big).\end{aligned}$$

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 $F_{\mathcal{P}}$ of degree $n - 1$ or less. We wish to show that it holds for a cylinder function of degree n. To simplify notation, let

$$\begin{aligned}\tilde{F}(u_1, \ldots, u_{n-1}) &\equiv (e^{(T-t_{n-1})\mathcal{L}_n/2} F)((u_1, \ldots, u_{n-1}, u_{n-1}))\\ &= E[F(u_1, u_2, \ldots, u_{n-1}, //^{u_n}_{(T-t_{n-1})})]\end{aligned}$$

(see Theorem 3.14 and Lemma 3.15) and $\tilde{F}_{\mathcal{P}}$ denote the cylinder function of $//$ given by

$$\tilde{F}_{\mathcal{P}} = \tilde{F}(\overrightarrow{//}) = (e^{(T-t_{n-1})\mathcal{L}_n/2} F)(\overrightarrow{//}, //_{t_{n-1}}).$$

Notice that

$$\begin{aligned}(3.30)\qquad X^k \tilde{F}_{\mathcal{P}} = \sum_{i=1}^{n-2} (D_i e^{(T-t_{n-1})\mathcal{L}_n/2} F)(\overrightarrow{//}, //_{t_{n-1}})(k(t_i) - A_{t_i})\\ + (D_{n-1} e^{(T-t_{n-1})\mathcal{L}_n/2} F)(\overrightarrow{//}, //_{t_{n-1}})(k(t_{n-1}) - A_{t_{n-1}})\\ + (D_n e^{(T-t_{n-1})\mathcal{L}_n/2} F)(\overrightarrow{//}, //_{t_{n-1}})(k(t_{n-1}) - A_{t_{n-1}}).\end{aligned}$$

By the Markov property or by Itô's Lemma, for any $i < n$,
(3.31)

$$E[D_iF(\overrightarrow{//},//_T)(k(t_i) - A_{t_i})] = E((e^{(T-t_{n-1})\mathcal{L}_n/2}D_iF)(\overrightarrow{//},//_{t_{n-1}})(k(t_i) - A_{t_i})).$$

Therefore

$$\begin{aligned}\sum_{i=1}^{n-1} ED_iF(\overrightarrow{//},//_T)(k(t_i) - A_{t_i}) &+ E[W_{t_{n-1}}(k(t_{n-1}) - A_{t_{n-1}})] \\ &= \sum_{i=1}^{n-2} E((e^{(T-t_{n-1})\mathcal{L}_n/2}D_iF)(\overrightarrow{//},//_{t_{n-1}})(k(t_i) - A_{t_i})) \\ &\quad + E((e^{(T-t_{n-1})\mathcal{L}_n/2}D_{n-1}F)(\overrightarrow{//},//_{t_{n-1}})(k(t_{n-1}) - A_{t_{n-1}})) \\ &\quad + E(D_n(e^{(T-t_{n-1})\mathcal{L}_n/2}F)(\overrightarrow{//},//_{t_{n-1}})(k(t_{n-1}) - A_{t_{n-1}})).\end{aligned}$$

Using the fact that D_i commutes with $\mathcal{L}_n$ for $i < n$, this last expression combined with Eq. (3.30) and the definition of W_t in Eq. (3.20) shows:

$$\sum_{i=1}^{n-1} E[D_iF(\overrightarrow{//},//_T)(k(t_i) - A_{t_i})] + E[W_{t_{n-1}}(k(t_{n-1}) - A_{t_{n-1}})] = E[X^k\tilde{F}_{\mathcal{P}}]. \tag{3.32}$$

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

$$EX^kF_{\mathcal{P}} = E[X^k\tilde{F}_{\mathcal{P}}] + E\left(F_{\mathcal{P}}\int_{t_{n-1}}^T \langle\dot{K}, db\rangle\right). \tag{3.33}$$

Applying the induction hypothesis, this gives

$$\begin{aligned}EX^kF_{\mathcal{P}} &= E\left(\tilde{F}_{\mathcal{P}}\int_0^{t_{n-1}}\langle\dot{K}, db\rangle\right) + E\left(F_{\mathcal{P}}\int_{t_{n-1}}^T\langle\dot{K}, db\rangle\right) \\ &= E\left(F_{\mathcal{P}}\int_0^{t_{n-1}}\langle\dot{K}, db\rangle\right) + E\left(F_{\mathcal{P}}\int_{t_{n-1}}^T\langle\dot{K}, db\rangle\right) \\ &= E\left(F_{\mathcal{P}}\int_0^T\langle\dot{K}, db\rangle\right),\end{aligned}$$

wherein the second equality we have again used the Markov property to replace $\tilde{F}_{\mathcal{P}}$ by $F_{\mathcal{P}}$. Q.E.D.

4. Consequences for Heat Kernel Measures on M.

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.*

$$(\nabla \cdot Y)(m) \equiv \mathrm{tr}\nabla Y = \sum_{i=1}^{d}\langle\nabla_{E_i}Y, E_i\rangle, \tag{4.1}$$

where $\{E_i\}_{i=1}^d$ is any orthonormal basis for T_mM.

THEOREM 4.1. – *Let $T > 0$ and $l \in H(\mathbb{R})$ such that $l(T) = 1$. Then for all smooth vector fields Y on M,*

(4.2)
$$E[(\nabla \cdot Y)(\Sigma_T)] = E\left[\left\langle //_T^{-1} Y(\Sigma_T), \int_0^T \left(\dot{l}(t) - \frac{1}{2} l(t) Ric_{//_t}\right) db(t) - \frac{1}{2}\int_0^T l(t) J_{//_t} dt \right\rangle\right].$$

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

$$(4.3) \qquad J_u = \sum_{c \in S} (B_c Ric_{(\cdot)} c)_u = \frac{d}{dt}\bigg|_0 \sum_{c \in S} Ric_{e^{tB_c}(u)} c = u^{-1} \sum_{c \in S} (\nabla_{uc} Ric) uc.$$

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

$$(4.4) \qquad E[(Yf)(\Sigma_T)] = E\left[f(\Sigma_T)\left\langle //_T^{-1} Y(\Sigma_T), \int_0^T \left(\dot{l}(t) - \frac{1}{2} l(t) Ric_{//_t}\right) db(t) - \frac{1}{2}\int_0^T l(t) J_{//_t} dt \right\rangle\right] - E[f(\Sigma_T)(\nabla \cdot Y)(\Sigma_T)]$$

and

$$(4.5) \qquad E[\nabla \cdot Y(\Sigma_T)] = E\left[\left\langle //_T^{-1} Y(\Sigma_T), \frac{b(T)}{T} - \frac{1}{2T}\int_0^T t\{Ric_{//_t} db(t) + J_{//_t} dt\}\right\rangle\right].$$

Proof. – Because $\nabla \cdot (fY) = Yf + f\nabla \cdot Y$, Eq. (4.4) follows by replacing Y by fY in Eq. (4.2). Taking $l(t) \equiv t/T$ 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 $T_o M$ and for $c \in S$ set $h_c(t) \equiv l(t)c$ and $X_t^c \equiv X_t^{h_c} \equiv //_t h_c(t)$ for $t \in [0, \infty)$. Then using the integration by parts Theorem 3.19,

$$0 = \sum_{c \in S} E(X^c 1)\langle X_T^c, Y(\Sigma_T)\rangle = \sum_{c \in S} E[(X^c)^* \langle X_T^c, Y(\Sigma_T)\rangle] = \sum_{c \in S} E[(-X^c + z_T^c)\langle X_T^c, Y(\Sigma_T)\rangle] =: E(-I + II),$$

where

$$I \equiv \sum_{c \in S} X^c \langle X_T^c, Y(\Sigma_T)\rangle \text{ and } II \equiv \sum_{c \in S} z_T^c \langle X_T^c, Y(\Sigma_T)\rangle,$$

and

$$\begin{aligned}
z_T^c &\equiv \int_0^T \left\langle \frac{1}{2}\mathrm{Ric}^*_{//_t} h_c(t) + \dot{h}_c(t), db(t) \right\rangle \\
&= \int_0^T \left\langle \left[\frac{l(t)}{2}\mathrm{Ric}^*_{//_t} + \dot{l}(t) \right] c, db(t) \right\rangle \\
&= \int_0^T \left\langle \left[\frac{l(t)}{2}\mathrm{Ric}_{//_t} + \dot{l}(t) \right] db(t), c \right\rangle.
\end{aligned}$$

Notice that $g \equiv \langle X_T^c, Y(\Sigma_T)\rangle = F(//_T)$, where $F(u) \equiv \langle c, u^{-1}Y(\pi(u))\rangle$. Hence $g = \langle X_T^c, Y(\Sigma_T)\rangle$ is a cylinder function of $//$ as defined in Section 3 and moreover by Eq. (3.11),

$$X^c g = DF(//_T)(-A_T^{h_c} + h_c(T)). \tag{4.6}$$

Note

$$DF(u)A = \frac{d}{d\epsilon}\bigg|_0 \langle c, e^{-\epsilon A}u^{-1}Y(\pi(u))\rangle = -\langle c, Au^{-1}Y(\pi(u))\rangle, \quad \forall A \in so(T_oM) \tag{4.7}$$

and

$$DF(u)a = \langle c, u^{-1}\nabla_{ua}Y\rangle = \langle uc, \nabla_{ua}Y\rangle \quad \forall A \in T_oM. \tag{4.8}$$

Using equations (4.6-4.8) and Eq. (3.5),

$$\begin{aligned}
I &= \sum_{c\in S}\{\langle //_T c, \nabla_{//_T c}Y\rangle + \langle c, A_T^{h_c} //_T^{-1} Y(\Sigma_T)\rangle\} \\
&= (\nabla\cdot Y)(\Sigma_T) + \sum_{c\in S}\left\langle c, \left(\int_0^T l(t)\Omega_{//_t}\langle c, \delta b(t)\rangle\right) //_T^{-1}Y(\Sigma_T)\right\rangle \\
&= (\nabla\cdot Y)(\Sigma_T) + \sum_{c\in S}\left\langle \left(\int_0^T l(t)\Omega_{//_t}\langle \delta b(t), c\rangle\right) c, //_T^{-1}Y(\Sigma_T)\right\rangle \\
&= (\nabla\cdot Y)(\Sigma_T) + \left\langle \int_0^T l(t)\mathrm{Ric}_{//_t}\delta b(t), //_T^{-1}Y(\Sigma_T)\right\rangle \\
&= (\nabla\cdot Y)(\Sigma_T) + \left\langle \int_0^T l(t)\mathrm{Ric}_{//_t} db(t) + \frac{1}{2}\int_0^T l(t)J_{//_t}dt, //_T^{-1}Y(\Sigma_T)\right\rangle,
\end{aligned}$$

where we have used

$$\begin{aligned}
\mathrm{Ric}_{//_t}\delta b(t) &= \mathrm{Ric}_{//_t} db(t) + \frac{1}{2}\sum_{c\in S}(B_c \mathrm{Ric}\, c)_{//_t} dt \\
&= \mathrm{Ric}_{//_t} db(t) + \frac{1}{2}J_{//_t}dt.
\end{aligned} \tag{4.9}$$

Similarly, we have

$$\begin{aligned}
II &= \sum_{c\in S} z_T^c \langle //_T c, Y(\Sigma_T)\rangle \\
&= \sum_{c\in S}\langle c, //_T^{-1}Y(\Sigma_T)\rangle \left\langle \int_0^T \left[\frac{l(t)}{2}\mathrm{Ric}_{//_t} + \dot{l}(t)\right] db(t), c\right\rangle \\
&= \left\langle \int_0^T \left[\frac{l(t)}{2}\mathrm{Ric}_{//_t} + \dot{l}(t)\right] db(t), //_T^{-1}Y(\Sigma_T)\right\rangle.
\end{aligned}$$

Combining the above expressions shows

$$\begin{aligned}
0 &= E[-I+II] \\
&= E\left[-(\nabla\cdot Y)(\Sigma_T) + \left\langle \int_0^T \dot{l}(t)db(t), //_T^{-1}Y(\Sigma_T)\right\rangle\right] \\
&\quad -\frac{1}{2}E\left\langle \int_0^T l(t)\mathrm{Ric}_{//_t}db(t) + \int_0^T l(t)J_{//_t}dt, //_T^{-1}Y(\Sigma_T)\right\rangle.
\end{aligned}$$

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 $dQ_t \cong dV_t$ 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 $Y_t(u) \equiv (e^{(T-t)\Delta/2}Y\tilde{)}(u) = (e^{(T-t)\mathcal{L}/2}\tilde{Y})(u)$. Define

$$B\cdot Y_t \equiv \sum_{c\in S}\langle B_cY_t, c\rangle = \sum_{c\in S} B_c\langle Y_t, c\rangle$$

and set

$$Q_t \equiv l(t)(B\cdot Y_t)(//_t) - \left\langle Y_t(//_t), \int_0^t \dot{l}db\right\rangle.$$

By Itô's Lemma, as in the proof of Lemma 3.17,

$$d[Y_t(//_t)] = \sum_{c\in S}(B_cY_t)(//_t)db_c(t) = (B_{db(t)}Y_t)(//_t)$$

and

$$\begin{aligned}
d[B\cdot Y_t(//_t)] &= \sum_{c\in S} B_cB\cdot Y_t(//_t)\delta b_c(t) - B\cdot\left(\frac{\mathcal{L}}{2}Y_t\right)(//_t)dt \\
&= \sum_{c\in S} B_cB\cdot Y_t(//_t)db_c(t) + \frac{1}{2}\sum_{c\in S} B_c^2B\cdot Y_t(//_t)dt - B\cdot\left(\frac{\mathcal{L}}{2}Y_t\right)(//_t)dt \\
&= \sum_{c\in S} B_cB\cdot Y_t(//_t)db_c(t) + \frac{1}{2}[\mathcal{L}, B\cdot]Y_t(//_t)dt.
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
dQ_t &\cong \dot{l}(t)(B\cdot Y_t)(//_t)\,dt + \frac{l(t)}{2}([\mathcal{L}, B\cdot]Y_t)(//_t)\,dt - \langle (B_{db(t)}Y_t)(//_t), \dot{l}(t)db(t)\rangle \\
&= \left\{\dot{l}(t)(B\cdot Y_t)(//_t) + \frac{l(t)}{2}([\mathcal{L}, B\cdot]Y_t)(//_t) - \sum_{c\in S}\langle (B_cY_t)(//_t), c\rangle\dot{l}(t)\right\}dt \\
&= \frac{l(t)}{2}([\mathcal{L}, B\cdot]Y_t)(//_t)\,dt \\
&= -\frac{l(t)}{2}\{\langle Y_t(//_t), J_{//_t}\rangle + \sum_{c\in S}\langle (B_cY_t)(//_t), \mathrm{Ric}_{//_t}c\rangle)\}\,dt,
\end{aligned}$$

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 $N_t \equiv Q_t + V_t$ where

$$V_t \equiv \frac{1}{2}\left\langle \int_0^t l(\tau)\{J_{//_\tau}\, d\tau + \mathrm{Ric}_{//_\tau}\, db(\tau)\}, Y_t(//_t)\right\rangle.$$

Notice that

$$\begin{aligned} dV_t &\cong \frac{l(t)}{2}\langle J_{//_t}, Y_t(//_t)\rangle\, dt + \frac{l(t)}{2}\langle \mathrm{Ric}_{//_t} db(t), (B_{db(t)}Y_t)(//_t)\rangle \\ &= \frac{l(t)}{2}\langle J_{//_t}, Y_t(//_t)\rangle\, dt + \frac{l(t)}{2}\sum_{c\in S}\langle \mathrm{Ric}_{//_t} c, (B_c Y_t)(//_t)\rangle\, dt \\ &\cong -dQ_t. \end{aligned}$$

Hence we have shown that $dN_t = dQ_t + dV_t \cong 0$, 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 $l(T)(B \cdot Y_T)(//_T) = (B \cdot \tilde{Y})(//_T) = (\nabla \cdot Y)(\Sigma_T)$,

$$\begin{aligned} N_T &= Q_T + V_T \\ &= l(T)(B \cdot Y_T)(//_T) - \left\langle Y_T(//_T), \int_0^T \dot{l}\, db\right\rangle \\ &\quad + \frac{1}{2}\left\langle \int_0^T l(\tau)\{J_{//_\tau}\, d\tau + \mathrm{Ric}_{//_\tau}\, db(\tau)\}, Y_T(//_T)\right\rangle \\ &= (\nabla \cdot Y)(\Sigma_T) - \left\langle //_T^{-1} Y(\Sigma_T), \int_0^T \dot{l}(t)\, db(t)\right\rangle \\ &\quad + \frac{1}{2}\left\langle \int_0^T l(t)\{J_{//_t}\, dt + \mathrm{Ric}_{//_t}\, db(t)\}, //_T^{-1} Y_T(\Sigma_T)\right\rangle. \end{aligned}$$

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 τ in π. Suppose that V is a finite dimensional vector space, X is a V-valued continuous semi-martingale and A is a $End(V)$-valued continuous semi-martingale. Then the *backwards stochastic integral* of A relative to X is

$$\int_0^t A\overleftarrow{dX} \equiv \lim_{|\pi|\to 0} \sum_{\tau\in\pi} A(t\wedge\tau+)(X(t\wedge(\tau+)) - X(\tau\wedge t)), \tag{4.10}$$

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

$$\int_0^t A\overleftarrow{dX} \equiv \lim_{|\pi|\to 0}\sum_{\tau\in\pi} A(\tau)(X(t\wedge(\tau+)) - X(\tau\wedge t)) \tag{4.11}$$

and

$$\int_0^t A\delta X \equiv \lim_{|\pi|\to 0}\sum_{\tau\in\pi} \frac{1}{2}(A(t\wedge\tau+) + A(\tau))(X(t\wedge(\tau+)) - X(\tau\wedge t)) \tag{4.12}$$

respectively. We also know that the joint quadratic variation "$\int dAdX$" is given by

$$\int dAdX = \lim_{|\pi|\to 0}\sum_{\tau\in\pi}(A(t\wedge\tau+) - A(\tau))(X(t\wedge(\tau+)) - X(\tau\wedge t)). \tag{4.13}$$

It is a trivial exercise to prove that

$$\int_0^t A\overleftarrow{dX} = \int_0^t AdX + \int_0^t dAdX. \tag{4.14}$$

and

$$\int_0^t A\delta X = \int_0^t AdX + \frac{1}{2}\int_0^t dAdX = \int_0^t A\overleftarrow{dX} - \frac{1}{2}\int_0^t dAdX. \tag{4.15}$$

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

$$d(\mathrm{Ric}_{//_t})db(t) = J_{//_t}dt. \tag{4.16}$$

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,*

$$E[(\nabla\cdot Y)(\Sigma_T)] = E\left[\left\langle //_T^{-1}Y(\Sigma_T), \int_0^T \left(\dot{l}(t) - \frac{1}{2}l(t)Ric_{//_t}\right)\overleftarrow{db}(t)\right\rangle\right]. \tag{4.17}$$

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 δ for the Itô's differential and d for the Stratonovich differential.)

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

$$\begin{aligned}\vec{\nabla}(e^{T\Delta/2}f)(o) &= t^{-1}E\left[\left(\int_0^t Q_r db(r)\right)(e^{(T-t)\Delta/2}f)(\Sigma_t)\right]\\ &= t^{-1}E\left[\left(\int_0^t Q_r db(r)\right)f(\Sigma_T)\right],\end{aligned}$$

where Q_t is the unique solution to the differential equation:

$$dQ_t/dt = -\frac{1}{2}Q_t Ric^*_{//_t} \text{ with } Q_0 = I. \tag{5.1}$$

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 $Q_t \in End(T_o M)$ 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 $dF(t,\Sigma_t) = (\nabla_{//_t db(t)} F(t,\cdot))(\Sigma_t)$ and therefore:

$$\begin{aligned} dz_t &= Q_t db(t) \cdot F(t,\Sigma_t) + \left(\int_0^t Q_r db(r)\right)(\nabla_{//_t db(t)} F(t,\cdot))(\Sigma_t) \\ &\quad + \sum_{c\in S} \langle Q_t c, (\nabla_{//_t c} F(t,\cdot))(\Sigma_t)\rangle\, dt. \end{aligned}$$

From this we conclude that:

$$\begin{aligned} Ez_t &= Ez_0 + E\int_0^t \sum_{c\in S} Q_r c (\nabla_{//_r c} F(r,\cdot))(\Sigma_r)\, dr \\ &= 0 + E\int_0^t \sum_{c\in S} Q_r c \langle //_r c, \vec{\nabla} F(r,\Sigma_r)\rangle\, dr \\ &= E\int_0^t \sum_{c\in S} Q_r c \langle c, //_r^{-1} \vec{\nabla} F(r,\Sigma_r)\rangle\, dr \\ &= E\int_0^t Q_r //_r^{-1} \vec{\nabla} F(r,\Sigma_r)\, dr. \end{aligned}$$

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

$$Ez_t = E\int_0^t Q_0 //_0^{-1} \vec{\nabla} F(0,\Sigma_0)\, dr = t\vec{\nabla}(e^{T\Delta/2} f)(o).$$

That is to say

$$\begin{aligned} \vec{\nabla}(e^{T\Delta/2} f)(o) &= t^{-1} E\left[\left(\int_0^t Q_r db(r)\right)(e^{(T-t)\Delta/2} f)(\Sigma_t)\right] \\ &= t^{-1} E\left[\left(\int_0^t Q_r db(r)\right) f(\Sigma_T)\right], \end{aligned}$$

wherein the second inequality we have used the Markov property of Σ_t.

We will now show how to choose Q_t. To this end let $W_t \equiv //_t^{-1}\vec{\nabla}F(t,\Sigma_t)$. Then by Corollary 3.18 we know that

$$dW_t = d(\text{martingale}) + \frac{1}{2}\text{Ric}^*_{//_t} W_t dt.$$

So if Q is the solution to Eq. (5.1), then we will have that Q_tW_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 $\langle\cdot,\cdot\rangle$ is an Ad_G-invariant inner product on $\mathfrak{g} = T_eG$ — the Lie algebra of G. We extend $\langle\cdot,\cdot\rangle$ to a metric on G by requiring left translations to be isometries. Since $\langle\cdot,\cdot\rangle$ is Ad_G-invariant, it follows that $\langle\cdot,\cdot\rangle$ is a bi-invariant metric on G. Let ∇ denote the left covariant derivative, *i.e.* $\nabla_{\tilde{A}}\tilde{B} = 0$ for all $A, B \in \mathfrak{g}$ where $\tilde{A}$ denotes the unique left invariant vector field on G such that $\tilde{A}(e) = A$. Also let $S \subset \mathfrak{g}$ be a fixed orthonormal basis of $\mathfrak{g}$. Let $\{b(t)\}_{t\geq 0}$ be a $\mathfrak{g}$-valued Brownian motion and Σ_t denote the solution to the stochastic differential equation:

$$\delta\Sigma_t = \Sigma_t\delta b(t) \text{ with } \Sigma_0 = e \in G. \tag{5.2}$$

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

$$\delta\Sigma_t = \sum_{A\in S}\tilde{A}(\Sigma_t)\delta b^A(t) \text{ with } \Sigma_0 = e \in G, \tag{5.3}$$

where $b^A(t) \equiv \langle b(t), A\rangle$. Now ∇ is a flat connection and parallel translation (relative to ∇) along $\Sigma_{(\cdot)}$ is $//_t \equiv L_{\Sigma_t *}$. In particular, the triple of processes $(\Sigma, //, b)$ satisfy the assumptions described in Section 3.2.

PROPOSITION 5.2. – *Suppose $(G, \langle\cdot,\cdot\rangle, \nabla, \Sigma, //, b)$ are as above. Then*

$$(\vec{\nabla}e^{T\Delta/2}f)(e) = t^{-1}E_e[b(t)f(\Sigma_T)] \tag{5.4}$$

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 $o = e \in G$, $\mathfrak{g} = T_eG$ be the Lie algebra of G ($e \in G$ is the identity element), $\langle\cdot,\cdot\rangle$ is an inner product on $\mathfrak{g}$ which we also view as a left invariant Riemannian metric on G, and ∇ is the Levi-Civita covariant derivative associated to $\langle\cdot,\cdot\rangle$. Notice that we do **not** assume here that $\langle\cdot,\cdot\rangle$ on $\mathfrak{g}$ is Ad-invariant or equivalently that the Riemannian metric made from $\langle\cdot,\cdot\rangle$ is bi-invariant.

NOTATION 6.1. – *If $h \in \mathfrak{g}$, let $\tilde{h}$ denote the unique left invariant vector field on G such that $\tilde{h}(e) = h$. Also let D denote the Lie algebra version of ∇,* i.e. *for each $h \in \mathfrak{g}$, D_h is the linear operator on $\mathfrak{g}$ defined by $D_hk \equiv (\nabla_{\tilde{h}}\tilde{k})(e)$.*

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 $\langle\cdot,\cdot\rangle$ is a left-invariant metric on G, one has $\nabla_{\tilde{h}}\tilde{k} = (D_h k)\tilde{}$.

Now suppose that σ and v are C^1-curves in G and TG respectively and that $\beta(t) = \int_0^t L_{\sigma(\tau)^{-1}*}\dot{\sigma}(\tau)d\tau$ and $a(t) \equiv L_{\sigma(t)^{-1}*}v(t)$. If $\nabla v(t)/dt$ denotes the covariant differential of v along σ, then:

$$\frac{\nabla v(t)}{dt} = L_{\sigma(t)*}\left(\frac{da(t)}{dt} + D_{\dot{\beta}(t)}a(t)\right). \tag{6.1}$$

To prove this equation, notice that v may be written as

$$v(t) = L_{\sigma(t)*}a(t) = (a(t))\tilde{}(\sigma(t)) = \sum_{h\in S}\langle a(t),h\rangle\tilde{h}(\sigma(t)), \tag{6.2}$$

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

$$\begin{aligned}\frac{\nabla v(t)}{dt} &= \sum_{h\in S}\left\{\left\langle\frac{da(t)}{dt},h\right\rangle\tilde{h}(\sigma(t)) + \langle a(t),h\rangle\nabla_{\dot{\sigma}(t)}\tilde{h}\right\}\\ &= L_{\sigma(t)*}\frac{da(t)}{dt} + \sum_{h\in S}\langle a(t),h\rangle(D_{(L_{\sigma(t)^{-1}*}\dot{\sigma}(t))}h)\tilde{}(\sigma(t))\\ &= L_{\sigma(t)*}\left(\frac{da(t)}{dt} + D_{\dot{\beta}(t)}a(t)\right).\end{aligned}$$

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

$$\delta\Sigma_t = L_{\Sigma_t*}\delta\beta(t) \text{ with } \Sigma_0 = e. \tag{6.3}$$

This last equation may be written more explicitly as

$$\delta\Sigma_t = \sum_{h\in S}\tilde{h}(\Sigma_t)\delta\beta^h(t) \text{ with } \Sigma_0 = e, \tag{6.4}$$

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 $\tilde{h}$ denotes the left invariant vector field on G such that $\tilde{h}(e) = h$.

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

Proof. – Since G is compact it is also unimodular. For unimodular Lie groups it is well known that the Laplace Beltrami operator Δ is given by $\Delta = \sum_{h\in S}\tilde{h}^2$, see for example Remark 2.2 in [11]. Therefore for $f\in C^\infty(G)$,

$$\begin{aligned}d(f(\Sigma_t)) &= \sum_{h\in S}(\tilde{h}f)(\Sigma_t)\delta\beta^h(t)\\ &= \sum_{h\in S}(\tilde{h}f)(\Sigma_t)d\beta^h(t) + \frac{1}{2}\sum_{h\in S}(\tilde{h}^2 f)(\Sigma_t)dt\\ &= \sum_{h\in S}(\tilde{h}f)(\Sigma_t)d\beta^h(t) + \frac{1}{2}(\Delta f)(\Sigma_t)dt,\end{aligned}$$

which shows that Σ satisfies the martingale characterization of a Brownian motion, see Section 3.2. Q.E.D.

Continuing the notation of Section 3.2, let $//$ denote stochastic parallel translation (for the Levi-Civita covariant derivative ∇) along Σ and b be the $\mathfrak{g}$-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 β, Σ, $//$, and b be the processes described above and set*

$$U(t) \equiv L_{\Sigma(t)^{-1}*}//_t. \tag{6.5}$$

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

$$dU(t) + D_{\delta\beta(t)}U(t) = 0 \text{ with } U(0) = I_{\mathfrak{g}}, \tag{6.6}$$

$$b(t) \equiv \int_0^t U(\tau)^{-1}\delta\beta(\tau), \tag{6.7}$$

and also

$$b(t) = \int_0^t U(\tau)^{-1}d\beta(\tau). \tag{6.8}$$

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

$$\delta b(t) = //_t^{-1}\delta\Sigma_t = //_t^{-1}L_{\Sigma_t*}\delta\beta(t) = U(t)^{-1}\delta\beta(t).$$

We now will show that Eq. (6.6) holds. Define $F : O(G) \to O(\mathfrak{g})$ by $F(p) \equiv L_{\pi(p)^{-1}*}p$, 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 $\mathfrak{g} = T_eG$ to T_gG.) 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 $h \in \mathfrak{g}$, let B_h be the associated horizontal vector field on $O(G)$ defined in Definition 3.3. Then $(B_hF)(p) = -D_{L_{\pi(p)^{-1}*}ph}F(p)$.

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

$$dO(t)/dt + D_{\alpha(t)}O(t) = 0 \text{ with } O(0) = L_{g(0)^{-1}*}p, \tag{6.9}$$

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

$$\begin{aligned}(B_hF)(p) &= \frac{d}{dt}\Big|_0 F(p(t)) = \frac{d}{dt}\Big|_0 L_{g(t)^{-1}*}p(t)\\ &= \frac{d}{dt}\Big|_0 L_{g(t)^{-1}*}L_{g(t)*}O(t)\\ &= \frac{d}{dt}\Big|_0 O(t) = -D_{\alpha(0)}O(0)\\ &= -D_{L_{\pi(p)^{-1}*}ph}F(p).\end{aligned}$$

This proves the claim.

Because of the claim, we now know that

$$\begin{aligned} dU(t) = dF(//_t) &= (B_{\delta b(t)}F)(//_t) \\ &= -D_{(L_{\pi(//_t)^{-1}*}//_t\delta b(t))}F(//_t) \\ &= -D_{(L_{\Sigma_t^{-1}*}//_t\delta b(t))}U(t) \\ &= -D_{(U(t)\delta b(t))}U(t), \end{aligned}$$

which in view of Eq. (6.7) proves Eq. (6.6).

To complete the proof we need only verify Eq. (6.8). Now

$$b(t) = \int_0^t U(t)^{-1}\delta\beta(t) = \int_0^t U(t)^{-1}d\beta(t) + V_t, \tag{6.10}$$

where V is the process of bounded variation given by

$$\begin{aligned} V(t) &\equiv \frac{1}{2}\int_0^t d(U(\tau)^{-1})d\beta(\tau) = \\ &= \frac{1}{2}\int_0^t U(\tau)^{-1}D_{d\beta(\tau)}d\beta(\tau) = \frac{1}{2}\int_0^t U(\tau)^{-1}(\sum_{h\in S} D_h h)d\tau. \end{aligned}$$

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 *a priori* that $b(t)$ is also a Brownian motion on $\mathfrak{g}$. In order for these statements to be consistent with Eq. (6.10) is is necessary for $V \equiv 0$, *i.e.* Eq. (6.8) is valid. Q.E.D.

As a bi-product of the proof we have shown that

$$\sum_{h\in S} D_h h = 0. \tag{6.11}$$

We could also verify this equation more directly as follows. For $k \in \mathfrak{g}$ let

$$e^{t\tilde{k}}(g) = R_{e^{tk}}(g) \equiv ge^{tk},$$

then $e^{t\tilde{k}}$ 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 $e^{t\tilde{k}}$ 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 ∇ via,

$$\nabla \cdot \tilde{k} = \mathrm{tr}\nabla\tilde{k} = \sum_{h\in S}\langle\nabla_{\tilde{h}}\tilde{k}, \tilde{h}\rangle = \sum_{h\in S}\langle D_h k, h\rangle = -\sum_{h\in S}\langle k, D_h h\rangle, \tag{6.12}$$

wherein the last equality we have used the metric compatibility of ∇ to conclude that D_h is skew adjoint on $\mathfrak{g}$. Since, as already noted, $\nabla \cdot \tilde{k} = 0$ for all $k \in \mathfrak{g}$, 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*

$$(6.13)\quad E[(\tilde{h}f)(\Sigma_T)] = E\left[f(\Sigma_T)\left\langle U(T)^{-1}h, \int_0^T U^{-1}(t)\left(\dot{l}(t) - \frac{1}{2}l(t)Ric_e\right)d\beta(t)\right\rangle\right] - \frac{1}{2}E\left[f(\Sigma_T)\left\langle U(T)^{-1}h, \int_0^T l(t)U(t)^{-1}J_e dt\right\rangle\right].$$

where Ric_e *is the Ricci tensor restricted to* $\mathfrak{g} = T_eG$ *and* J *is defined in Eq. (4.3).*

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

$$(6.14)\quad E[(\tilde{h}f)(\Sigma_T)] = E\left[\left\langle f(\Sigma_T)//_T^{-1}\tilde{h}(\Sigma_T), \int_0^T\left(\dot{l}(t) - \frac{1}{2}l(t)\mathrm{Ric}_{//_t}\right)db(t) - \frac{1}{2}\int_0^T l(t)J_{//_t}dt\right\rangle\right].$$

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

$$\mathrm{Ric}_{//_t} = \mathrm{Ric}_{L_{\Sigma_t*}U(t)} = U^{-1}(t)\mathrm{Ric}_{L_{\Sigma_t*}}U(t) = U^{-1}(t)\mathrm{Ric}_eU(t),$$

and

$$J_{//_t} = J_{L_{\Sigma_t*}U(t)} = \sum_{h\in S}(L_{\Sigma_t*}U(t))^{-1}(\nabla_{L_{\Sigma_t*}U(t)h}\mathrm{Ric})L_{\Sigma_t*}U(t)h = \sum_{h\in S}U^{-1}(t)(\nabla_{U(t)h}\mathrm{Ric})U(t)h = U^{-1}(t)J_e.$$

Also notice that

$$//_T^{-1}\tilde{h}(\Sigma_T) = U^{-1}(T)L_{\Sigma_T*}^{-1}\tilde{h}(\Sigma_T) = U^{-1}(T)h.$$

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

$$E[(\tilde{h}f)(\Sigma_T)] = E\left[f(\Sigma_T)\left\langle U^{-1}(T)h, \int_0^T\left(\dot{l}(t) - \frac{1}{2}l(t)U^{-1}(t)\mathrm{Ric}_eU(t)\right)U^{-1}(t)d\beta(t)\right\rangle\right] - \frac{1}{2}E\left[\left\langle U^{-1}(T)h, \int_0^T l(t)U(t)^{-1}J_e dt\right\rangle\right],$$

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*

$$(6.15)\quad E[(\tilde{h}f)(\Sigma_T)] = E\left[f(\Sigma_T)\left\langle U(T)^{-1}h, \int_0^T U^{-1}(t)\left(\dot{l}(t) - \frac{1}{2}l(t)Ric_e\right)\overleftarrow{d\beta}(t)\right\rangle\right]$$

$$(6.16)\quad = E\left[f(\Sigma_T)\int_0^T\left\langle U(t,T)h, \left(\dot{l}(t) - \frac{1}{2}l(t)Ric_e\right)\overleftarrow{d\beta}(t)\right\rangle\right],$$

where $\overleftarrow{d\beta}$ denotes the backwards Itô's differential and $U(t,T) \equiv U(t)U(T)^{-1}$ 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 $A(t) = (\dot{l}(\tau) - \frac{1}{2}l(\tau)\text{Ric}_e^*)U(t,T)h$ 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,

$$dU(t,T) + D_{\delta\beta(t)}U(t,T) = 0 \text{ with } U(T,T) = I_{\mathfrak{g}}.$$

From this it follows that $U(t,T)$ may be chosen to be $\sigma(\beta(\tau) - \beta(T) : t \leq \tau \leq T)$-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

$$\begin{aligned} d(U^{-1}(t)\text{Ric}_e)d\beta(t) &= (U^{-1}(t)D_{d\beta(t)}\text{Ric}_e)d\beta(t) \\ &= \sum_{h\in S} U^{-1}(t)(D_h\text{Ric}_e)h\,dt = U^{-1}(t)J\,dt \end{aligned}$$

and

$$\begin{aligned} d(U^{-1}(t))d\beta(t) &= U^{-1}(t)D_{d\beta(t)}d\beta(t) \\ &= \sum_{h\in S} U^{-1}(t)D_h h\; dt = 0. \end{aligned}$$

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 $A : O(M) \to so(T_oM)$ 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 $\hat{\Omega}\langle a,b\rangle$ denote the vertical vector field $\hat{A}$, where $A(u) \equiv \Omega_u\langle a,b\rangle$. Similarly, let $B_{\Theta\langle a,b\rangle}$ denote the horizontal vector field on $O(M)$ defined by $u \in O(M) \to B_{\Theta_u\langle a,b\rangle}(u) \in TO(M)$. 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 $A,C \in so(T_oM)$, then*
1. $[\hat{A},\hat{C}] = \widehat{[A,C]}$, *and*
2. $[B_a,B_b] = -\hat{\Omega}\langle a,b\rangle - B_{\Theta\langle a,b\rangle}$,
3. $[\hat{A},B_a] = B_{Aa}$.

Remark 7.2. – The last commutator formula easily generalizes to

$$[\hat{A}, B_a] = B_{Aa} - \widehat{B_a A} \tag{7.1}$$

when $A : O(M) \to so(T_oM)$ is a non-constant smooth function.

LEMMA 7.3. – *Let* $a \in T_oM$ *and* $A \in so(T_oM)$, *then* $[\mathcal{L}, \hat{A}] = 0$ *and*

$$[\mathcal{L}, B_a] = \sum_{c\in S}\{(B_c\Omega\langle c, a\rangle)\hat{} - 2B_c\hat{\Omega}\langle c, a\rangle\} + B_{\mathrm{Ric}^* a}. \tag{7.2}$$

Proof. – We compute,

$$\begin{aligned}[\mathcal{L}, \hat{A}] &= \sum_{a\in S} B_a[B_a, \hat{A}] + [B_a, \hat{A}]B_a = -\sum_{a\in S}\{B_aB_{Aa} + B_{Aa}B_a\} \\ &= -\sum_{a,c\in S}\{B_aB_c + B_cB_a\}\langle Aa, c\rangle,\end{aligned} \tag{7.3}$$

which is zero, since $\langle Aa, c\rangle$ is skew symmetric in a and c while $\{B_aB_c + B_cB_a\}$ is symmetric. Similarly,

$$\begin{aligned}[\mathcal{L}, B_a] &= \sum_{c\in S}[B_c^2, B_a] = \sum_{c\in S}\{B_c[B_c, B_a] + [B_c, B_a]B_c\} \\ &= -\sum_{c\in S}\{B_c\hat{\Omega}\langle c, a\rangle + \hat{\Omega}\langle c, a\rangle B_c\} - \sum_{c\in S}\{B_cB_{\Theta\langle c,a\rangle} + B_{\Theta\langle c,a\rangle}B_c\} \\ &= -\sum_{c\in S}\{2B_c\hat{\Omega}\langle c, a\rangle + [\hat{\Omega}\langle c, a\rangle, B_c]\} \\ &\quad -\sum_{c\in S}\{\sum_{b\in S}(B_cB_b + B_bB_c)\langle\Theta\langle c, a\rangle, b\rangle + B_{B_c\Theta\langle c,a\rangle}\}.\end{aligned}$$

Now

$$\sum_{b,c\in S}(B_cB_b + B_bB_c)\langle\Theta\langle c, a\rangle, b\rangle = 0,$$

since $(B_cB_b + B_bB_c)$ is symmetric in b and c while $\langle\Theta\langle c, a\rangle, b\rangle$ is anti-symmetric in b and c because of the TSS assumption. Also

$$\begin{aligned}\sum_{c\in S}\{[\hat{\Omega}\langle c, a\rangle, B_c] + B_{B_c\Theta\langle c,a\rangle}\} &= \sum_{c\in S}\{B_{\Omega\langle c,a\rangle c} - (B_c\Omega\langle c, a\rangle)\hat{}\} - B_{\bar{\Theta}a} \\ &= -B_{\mathrm{Ric}\, a} - \sum_{c\in S}(B_c\Omega\langle c, a\rangle)\hat{}\} - B_{\bar{\Theta}a} \\ &= -B_{(\mathrm{Ric}\, a + \bar{\Theta}a)} - \sum_{c\in S}(B_c\Omega\langle c, a\rangle)\hat{} \\ &= -B_{\mathrm{Ric}^* a} - \sum_{c\in S}(B_c\Omega\langle c, a\rangle)\hat{}.\end{aligned}$$

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 $Z : O(M) \to T_oM$ is a smooth function, let

$$B \cdot Z \equiv \sum_{a\in S}\langle B_a Z, a\rangle = \sum_{a\in S} B_a\langle Z, a\rangle. \tag{7.4}$$

COROLLARY 7.4. – *Suppose that $F \in C^\infty(O(M))$ and $Z \in C^\infty(O(M) \to T_oM)$ such that $F(ug) = F(u)$ and $Z(ug) = g^{-1}Z(u)$ for all $g \in O(T_oM)$,* 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,*

$$[\mathcal{L}, B_a]F = B_{Ric^* a}F \tag{7.5}$$

and

$$[\mathcal{L}, B\cdot]Z \equiv \mathcal{L}(B \cdot Z) - B \cdot \mathcal{L}Z = -\langle Z, J\rangle - \sum_{a\in S}\langle B_a Z, Ric\, a\rangle, \tag{7.6}$$

where

$$J \equiv \sum_{a\in S} B_a Ric\, a. \tag{7.7}$$

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

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

where j is the vector field on M such that $j(m) = \sum_{c\in S}(\nabla_c \mathrm{Ric}\,)c$ for all $m \in M$ and any orthonormal basis S of T_mM.

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

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

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

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

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

We then have the following important commutator formula:

$$\begin{aligned}
[\mathcal{L}, B_a]W &= \sum_{c\in S}\{(B_c\Omega\langle c, a\rangle)\hat{} - 2B_c\hat{\Omega}\langle c, a\rangle\}W + B_{\mathrm{Ric}^* a}W \\
&= \sum_{c\in S}\{-\rho_*(B_c\Omega\langle c, a\rangle)W + 2B_c(\rho_*(\Omega\langle c, a\rangle)W)\} + B_{\mathrm{Ric}^* a}W \\
&= \sum_{c\in S}\{\rho_*(B_c\Omega\langle c, a\rangle)W + 2\rho_*(\Omega\langle c, a\rangle)B_cW\} + B_{\mathrm{Ric}^* a}W.
\end{aligned}$$

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

$$[\mathcal{L}, B_a]Z = \sum_{c\in S}\{(B_c\Omega\langle c, a\rangle)Z + 2\Omega\langle c, a\rangle B_c Z\} + B_{\mathrm{Ric}^* a}Z.$$

Hence, we have

$$\begin{aligned}\sum_{a\in S}\langle[\mathcal{L}, B_a]Z, a\rangle &= \sum_{a,c\in S}\{\langle(B_c\Omega\langle c, a\rangle)Z, a\rangle + 2\langle\Omega\langle c, a\rangle B_c Z, a\rangle\} + \sum_{a\in S} B_{\mathrm{Ric}^* a}\langle Z, a\rangle\\ &= \sum_{a,c\in S}\{-\langle Z, (B_c\Omega\langle c, a\rangle)a\rangle - 2\langle B_c Z, \Omega\langle c, a\rangle a\rangle + \langle B_c Z, a\rangle\langle \mathrm{Ric}^* a, c\rangle\}\\ &= -\sum_{c\in S}\{\langle Z, B_c \mathrm{Ric}\, c\rangle + 2\langle B_c Z, \mathrm{Ric}\, c\rangle - \langle B_c Z, \mathrm{Ric}\, c\rangle\}\\ &= -\langle Z, J\rangle - \sum_{c\in S}\langle B_c Z, \mathrm{Ric}\, c\rangle\},\end{aligned}$$

which is Eq. (7.6). Q.E.D.

Acknowledgments

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

REFERENCES

[1] S. AIDA, *On the irreducibility of certain Dirichlet forms on loop spaces over compact homogeneous spaces*, in the Proceedings of the 1994 Taniguchi Symposium in "New Trends in Stochastic Analysis" (K. D. Elworthy, S. Kusuoka, I. Shigekawa Eds.), World Scientific, New Jersey 1997, pp. 3-42.

[2] H. AIRAULT and P. MALLIAVIN, Integration by parts formulas and dilatation vector fields on elliptic probability spaces, *Prob. Theory Related Fields*, *106*, 1996, no. 4, pp. 447-494.

[3] J.-M. BISMUT, Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions. *Z. Wahrsch. Verw. Gebiete*, *56*, no. 4, 1981, pp. 469-505.

[4] J.-M. BISMUT, "*Large Deviations and the Malliavin Calculus*", Birkhauser, Boston/Basel/Stuttgart, 1984.

[5] R. H. CAMERON, The first variation of an indefinite Wiener integral, *Proc. A.M.S.*, *2*, 1951, pp. 914-924.

[6] B. K. DRIVER, YM_2: Continuum expectations, lattice convergence, and lassos, *Comm. Math. Phys.*, *123*, 1989, pp. 575-616.

[7] B. K. DRIVER, A Cameron-Martin type quasi-invariance theorem for Brownian motion on a compact Riemannian manifold, *J. of Func. Anal.*, *110*, 1992, pp. 272-376.

[8] B. K. DRIVER, *The Lie bracket of adapted vector fields on Wiener spaces*, UCSD preprint, 1995. To appear in *Applied Mathematics and Optimization.*

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

[10] B. K. DRIVER, Integration by parts and quasi-invariance for heat kernel measures on loop groups, December 1996, UCSD preprint, 64 pages. To appear in *J. of Funct. Anal.*

[11] B. K. DRIVER and L. GROSS, *Hilbert spaces of holomorphic functions on complex Lie groups*, in the Proceedings of the 1994 Taniguchi Symposium in "New Trends in Stochastic Analysis" (K. D. Elworthy, S. Kusuoka, I. Shigekawa Eds.), World Scientific, New Jersey 1997, pp. 76-106.

[12] R. J. ELLIOTT and M. KOHLMANN, Integration by parts, homogeneous chaos expansions and smooth densities. *Ann. Probab.*, *17*, no. 1, 1989, pp. 194-207.

[13] K. D. ELWORTHY, "Stochastic differential equations on manifolds," London Mathematical Society Lecture note Series 70. Cambridge Univ. Press, London, 1982.

[14] K. D. ELWORTHY and Xue-Mei LI, Formulae for the derivatives of heat semigroups, *J. of Func. Anal.*, *125*, 1994, pp. 252-286.

[15] K. D. ELWORTHY and Xue-Mei LI, A class of integration by parts formulae in stochastic analysis I, pp. 15-30, in "Itô Stochastic Calculus and Probability Theory," N. Ikeda, S. Watanabe, M. Fukushima, and H. Kunita (Eds.), Springer Verlag, Tokyo, 1996.

[16] M. EMERY, "*Stochastic Calculus in Manifolds*," Springer, Berlin/Heidelberg/New York, 1989.

[17] O. ENCHEV and D. W. STROOCK, Towards a Riemannian geometry on the path space over a Riemannian manifold. *J. Funct. Anal.*, *134*, 1995, no. 2, pp. 392-416.

[18] S. FANG, Inégalité du type de Poincaré sur l'espace des chemins riemanniens. (French) [Poincare-type inequality on a Riemannian path space] *C. R. Acad. Sci. Paris Sér. I Math.*, *318*, 1994, no. 3, pp. 257-260.

[19] S. FANG, Stochastic anticipative integrals on a Riemannian manifold. *J. Funct. Anal.*, *131*, 1995, no. 1, pp. 228-253.

[20] S. FANG and P. MALLIAVIN, Stochastic analysis on the path space of a Riemannian manifold. I. Markovian stochastic calculus, *J. Funct. Anal.*, *118*, no. 1, 1993, pp. 249-274.

[21] E. P. HSU, Quasi-invariance of the Wiener measure on the path space over a compact Riemannian manifold. *J. Funct. Anal.*, *134*, no. 2, 1995, pp. 417-450.

[22] E. P. HSU, Flows and quasi-invariance of the Wiener measure on path spaces. Stochastic analysis (Ithaca, NY, 1993), 265-279, *Proc. Sympos. Pure Math.*, *57*, Amer. Math. Soc., Providence, RI, 1995.

[23] E. P. HSU, Inégalités de Sobolev logarithmiques sur un espace de chemins. (French) [Logarithmic Sobolev inequalities on path spaces] *C. R. Acad. Sci. Paris Sér. I Math.*, *320*, no. 8, 1995, pp. 1009-1012.

[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.

[25] N. IKEDA and S. WATANABE, "Stochastic differential equations and diffusion processes," North-Holland Mathematical Library, 24. North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1981.

[26] KOBAYASHI and NOMIZU, "Foundations of differential geometry, Vols. 1.," Interscience Publishers (John Wiley and Sons), New York/London, 1963.

[27] R. LEANDRE, Integration by parts formulas and rotationally invariant Sobolev calculus on free loop spaces, *J. of Geometry and Physics*, *11*, 1993, pp. 517-528.

[28] R. LEANDRE and J. R. NORRIS, Integration by parts Cameron-Martin formulas for the free path space of a compact Riemannian manifold, Warwick Univ. preprint, January 1995.

[29] T. LYONS and Z. QIAN, Stochastic Jacobi fields and vector fields induced by varying area on path spaces, Preprint, 1996 Imperial College of Science, London England.

[30] M.-P. MALLIAVIN and P. MALLIAVIN, Integration on loop groups. I. Quasi invariant measures, Quasi invariant integration on loop groups, *J. of Funct. Anal.*, *93*, 1990, pp. 207-237.

[31] M.-P. MALLIAVIN and P. MALLIAVIN, Integration on loop group III, Asymptotic Peter-Weyl orthogonality, *J. of Funct. Anal.*, *108*, 1992, pp. 13-46.

[32] J. R. NORRIS, Simplified Malliavin calculus, Seminaire de Probabilites XX 1984/85 (ed. par J. Azema et M. Yor), *Lect. Notes in Math.*, *1204*, Springer-Verlag, Berlin, 1986, pp. 101-130.

[33] J. R. NORRIS, Twisted Sheets, *J. Funct. Anal.*, *132*, 1995, pp. 273-334.

[34] D. W. STROOCK, An estimate on the Hessian of the heat kernel, pp. 355-372 in "Itô Stochastic Calculus and Probability Theory," Ikeda, Watanabe, Fukushima, and Kunita (Eds.), Springer-Verlag, Berlin/Heidelberg/New York, 1996.

[35] D. W. STROOCK and J. TURETSKY, Upper bounds on derivatives of the logarithm of the heat kernel, M.I.T. preprint, July 1996.

[36] D. W. STROOCK and J. TURETSKY, Short time behavior of logarithmic derivatives of the heat kernel, M.I.T. preprint, December 1996.

[37] D. W. STROOCK and O. ZEITOUNI, Variations on a theme by Bismut, *Astérisque*, *236*, 1996, pp. 291-301.

[KU] H. KUNITA, *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, Cambridge, 1990.

[M0] P. MALLIAVIN, *Géométrie différentielle stochastique*, Montréal: Presses de l'Université de Montréal, 1978.

[M1] P. MALLIAVIN, *Stochastic calculus of variation and hypoelliptic operators*, Proc. Int. Symp. S.D.E. Kyoto, 1976, pp. 195-264, Wiley and Sons, New York, 1978.

(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