DERIVATIVE AND DIVERGENCE FORMULAE FOR DIFFUSION SEMIGROUPS

ANTON THALMAIER AND JAMES THOMPSON

Mathematics Research Unit, FSTC, University of Luxembourg
6, rue Richard Coudenhove-Kalergi, 1359 Luxembourg, Grand Duchy of Luxembourg

Abstract. For a semigroup P_t generated by an elliptic operator on a smooth manifold M, we use straightforward martingale arguments to derive probabilistic formulae for $P_t(V(f))$, not involving derivatives of f, where V is a vector field on M. For non-symmetric generators, such formulae correspond to the derivative of the heat kernel in the *forward* variable. As an application, these formulae can be used to derive various *shift-Harnack* inequalities.

Introduction

For a Banach space E, $e \in E$ and a Markov operator P on $\mathcal{B}_b(E)$, it is known that certain estimates on $P(\nabla_e f)$ are equivalent to corresponding *shift-Harnack* inequalities. This was proved by F.-Y. Wang in [18]. For example, for $\delta_e \in (0,1)$ and $\beta_e \in C((\delta_e,\infty)\times E;[0,\infty))$, he proved that the derivative-entropy estimate

$$\left|P(\nabla_e f)\right| \le \delta(P(f\log f)-(Pf)\log Pf)+\beta_e(\delta,\cdot)Pf$$

holds for any $\delta \ge \delta_e$ and positive $f \in C_b^1(E)$ if and only if the inequality

$$(Pf)^p \le (P(f^p(re+\cdot)))\exp\left(\int_0^1 \frac{pr}{1+(p-1)s}\beta_e\left(\frac{p-1}{r+r(p-1)s},\cdot+sre\right)ds\right)$$

holds for any $p \ge 1/(1-r\delta_e)$, $r \in (0,1/\delta_e)$ and positive $f \in \mathcal{B}_b(E)$. Furthermore, he also proved that if $C \ge 0$ is a constant then the L^2-derivative inequality

$$\left|P(\nabla_e f)\right|^2 \le CPf^2$$

holds for any non-negative $f \in C_b^1(E)$ if and only if the inequality

$$Pf \le P(f(\alpha e+\cdot))+|\alpha|\sqrt{CPf^2}$$

holds for any $\alpha \in \mathbb{R}$ and non-negative $f \in \mathcal{B}_b(E)$. The objective of this article is to find probabilistic formulae for $P_T(V(f))$ from which such estimates can be derived, for the case in which P_T is the Markov operator associated to a non-degenerate diffusion X_t on a smooth, finite-dimensional manifold M, and V a vector field.

E-mail address: `anton.thalmaier@uni.lu, james.thompson@uni.lu`.
Date: January 13, 2017.
2010 *Mathematics Subject Classification.* 58J65, 60J60, 53C21.
Key words and phrases. Diffusion semigroup, Heat kernel, Gradient estimate, Harnack inequality, Ricci curvature.

In Section 1 we suppose that M is a Riemannian manifold and that the generator of X_t is $\Delta + Z$, for some smooth vector field Z. Any non-degenerate diffusion on a smooth manifold induces a Riemannian metric with respect to which its generator takes this form. The basic strategy is then to use the relation $V(f) = \operatorname{div}(fV) - f \operatorname{div} V$ to reduce the problem to finding a suitable formula for $P_T(\operatorname{div}(fV))$. Such a formula was given in [3] for the case $Z = 0$, which we extend to the general case with Theorem 1.16. In doing so, we do not make any assumptions on the derivatives of the curvature tensor, as occurred in [2]. For an adapted process h_t with paths in the Cameron-Martin space $L^{1,2}([0,T];\mathbb{R})$, with $h_0 = 0$ and $h_T = 1$ and under certain additional conditions, we obtain the formula

$$\begin{aligned} &P_T\left(V(f)\right)(x) \\ &\quad = -\mathbb{E}[f(X_T(x))(\operatorname{div} V)(X_T(x))] \\ &\quad\quad -\frac{1}{2}\mathbb{E}\left[f(X_T(x))\left\langle V(X_T(x)), //_T\Theta_T \int_0^T \left((\operatorname{div} Z)(X_t(x))h_t - \dot{h}_t\right)\Theta_t^{-1}dB_t\right\rangle\right] \end{aligned}$$

where Θ is the $\operatorname{Aut}(T_xM)$-valued process defined by the pathwise differential equation

$$\frac{d}{dt}\Theta_t = -//_t^{-1}\left(\operatorname{Ric}^\sharp + (\nabla_\cdot Z)^* - \operatorname{div} Z\right)//_t\Theta_t$$

with $\Theta_0 = \operatorname{id}_{T_xM}$. Here $//_t$ denotes the stochastic parallel transport associated to $X_t(x)$, whose antidevelopment to T_xM has martingale part B. In particular, B is a diffusion on $\mathbb{R}^n$ generated by the Laplacian; it is a standard Brownian motion sped up by 2, so that $dB_t^i dB_t^j = 2\delta_{ij}\,dt$. Choosing h_t explicitly yields a formula from which estimates then can be deduced, as described in Subsection 1.5.

The problem of finding a suitable formula for $P_T(V(f))$ is dual to that of finding an analogous one for $V(P_Tf)$. A formula for the latter is called the Bismut formula [1] or the Bismut-Elworthy-Li formula, on account of [6]. We provide a brief proof of it in Subsection 1.3, since we would like to compare it to our formula for $P_T(V(f))$. Our approach to these formulae is based on martingale arguments; integration by parts is done at the level of local martingales. Under conditions which assure that the local martingales are true martingales, the wanted formulae are then obtained by taking expectations. They allow for the choice of a finite energy process. Depending on the intended type, conditions are imposed either on the right endpoint, as in the formula for $P_T(V(f))$, or the left endpoint, as in the formula for $V(P_Tf)$. The formula for $P_T(V(f))$ requires non-explosivity; the formula for $V(P_Tf)$ does not. From the latter can be deduced Bismut's formula for the logarithmic derivative in the *backward* variable x of the heat kernel $p_T(x,y)$ determined by

$$(P_Tf)(x) = \int_M f(y)p_T(x,y)\operatorname{vol}(dy), \quad f \in C_b(M).$$

From our formula for $P_T(V(f))$ can be deduced the following formula for the derivative in the *forward* variable y:

$$(\nabla \log p_T(x,\cdot))_y = -\frac{1}{2}\mathbb{E}\left[//_T\Theta_T\int_0^T\left((\operatorname{div} Z)(X_t(x))h_t - \dot{h}_t\right)\Theta_t^{-1}dB_t \,\middle|\, X_T(x) = y\right].$$

In Section 2 we consider the general case in which M is a smooth manifold and X_t a non-degenerate diffusion solving a Stratonovich equation of the form

$$dX_t = A_0(X_t)\,dt + A(X_t)\circ dB_t.$$

We denote by TX_t the derivative (in probability) of the solution flow. Using a similar approach to that of Section 1, and a variety of geometric objects naturally associated to the equation, we obtain, under certain conditions, the formula

$$\begin{aligned} &P_T(V(f)) \\ &= -\sum_{i=1}^m \mathbb{E}[f(X_T)A_i\langle V,A_i\rangle(X_T)] \\ &\quad -\frac{1}{2}\mathbb{E}\Big[f(X_T)\Big\langle V(X_T),\Xi_T\int_0^T \Xi_t^{-1}\Big(\big((\operatorname{trace}\hat\nabla A_0)(X_t)h_t-\dot h_t\big)A(X_t)\,dB_t+2h_tA_0^A\,dt\Big)\Big\rangle\Big] \end{aligned}$$

with

$$\begin{aligned} \Xi_t &= TX_t - TX_t\int_0^t TX_s^{-1}\Big(\big((\breve\nabla A_0)^* + \breve\nabla A_0 + \operatorname{trace}\hat\nabla A_0\big)(\Xi_s)\Big)ds, \\ A_0^A &= \sum_{i=1}^m \big((\breve\nabla A_0)^* + \breve\nabla A_0\big)\big(\breve T(\cdot,A_i)^*(A_i)\big) + [A_0,\breve T(\cdot,A_i)^*(A_i)], \end{aligned}$$

where the operators $\hat\nabla A_0$, $\breve\nabla A_0$ and $\breve T(\cdot,A_i)$ are given at each $x\in M$ and $v\in T_xM$ by

$$\begin{aligned} \hat\nabla_v A_0 &= A(x)\big(d(A^*(\cdot)A_0(\cdot))_x(v) - (dA^*)_x(v,A_0)\big), \\ \breve\nabla_v A_0 &= A(x)d\big(A(\cdot)^*A_0(\cdot)\big)_x(v), \\ \breve T(v,A_i)_x &= A(x)(dA^*)_x(v,A_i). \end{aligned}$$

This formula has the advantage of involving neither parallel transport nor Riemannian curvature, both typically difficult to calculate in terms of A.

1. Intrinsic Formulae

1.1. **Preliminaries.** Let M be a complete and connected n-dimensional Riemannian manifold, ∇ the Levi-Civita connection on M and $\pi\colon \mathrm{O}(M)\to M$ the orthonormal frame bundle over M. Let $E\to M$ be an associated vector bundle with fibre V and structure group $G=\mathrm{O}(n)$. The induced covariant derivative

$$\nabla\colon \Gamma(E)\to\Gamma(T^*M\otimes E)$$

determines the so-called *connection Laplacian* (or *rough Laplacian*) $\square$ on $\Gamma(E)$,

$$\square a = \operatorname{trace}\nabla^2 a.$$

Note that $\nabla^2 a\in\Gamma(T^*M\otimes T^*M\otimes E)$ and $(\square a)_x=\sum_i\nabla^2 a(v_i,v_i)\in E_x$ where v_i runs through an orthonormal basis of T_xM. For $a,b\in\Gamma(E)$ of compact support it is immediate to check that

$$\langle\square a,b\rangle_{L^2(E)} = -\langle\nabla a,\nabla b\rangle_{L^2(T^*M\otimes E)}.$$

In this sense we have $\square=-\nabla^*\nabla$. Let H be the horizontal subbundle of the G-invariant splitting of $T\,\mathrm{O}(M)$ and

$$h\colon \pi^*TM \xrightarrow{\sim} H \hookrightarrow T\,\mathrm{O}(M)$$

the *horizontal lift* of the G-connection; fibrewise this bundle isomorphism reads as

$$h_u\colon T_{\pi(u)}M\xrightarrow{\sim} H_u,\quad u\in\mathrm{O}(M).$$

In terms of the *standard horizontal vector fields* $H_1,\dots,H_n$ on $\mathrm{O}(M)$,

$$H_i(u) := h_u(ue_i),\quad u\in\mathrm{O}(M),$$

Bochner's horizontal Laplacian Δ^{hor}, acting on smooth functions on $\mathrm{O}(M)$, is given as

$$\Delta^{\mathrm{hor}} = \sum_{i=1}^n H_i^2.$$

To formulate the relation between $\square$ and Δ^{hor}, it is convenient to write sections $a \in \Gamma(E)$ as equivariant functions $F_a \colon \mathrm{O}(M) \to V$ via $F_a(u) = u^{-1} a_{\pi(u)}$ where we read $u \in \mathrm{O}(M)$ as an isomorphism $u \colon V \xrightarrow{\sim} E_{\pi(u)}$. Equivariance means that

$$F_a(ug) = g^{-1} F_a(u), \quad u \in \mathrm{O}(M),\ g \in G = \mathrm{O}(n).$$

Lemma 1.1 (see [9], p. 115)**.** *For $a \in \Gamma(E)$ and F_a the corresponding equivariant function on* $\mathrm{O}(M)$*, we have*

$$(H_i F_a)(u) = F_{\nabla_{ue_i} a}(u), \quad u \in \mathrm{O}(M).$$

Hence

$$\Delta^{\mathrm{hor}} F_a = F_{\square a},$$

where as above

$$\square \colon \Gamma(E) \xrightarrow{\nabla} \Gamma(T^*M \otimes E) \xrightarrow{\nabla} \Gamma(T^*M \otimes T^*M \otimes E) \xrightarrow{\mathrm{trace}} \Gamma(E).$$

Proof. Fix $u \in \mathrm{O}(M)$ and choose a curve γ in M such that $\gamma(0) = \pi(u)$ and $\dot\gamma = ue_i$. Let $t \mapsto u(t)$ be the horizontal lift of γ to $\mathrm{O}(M)$ such that $u(0) = u$. Note that $\dot u(t) = h_{u(t)}(\dot\gamma(t))$, and in particular $\dot u(0) = h_u(ue_i) = H_i(u)$. Hence, denoting the parallel transport along γ by $//_\varepsilon = u(\varepsilon)u(0)^{-1}$, we get

$$\begin{aligned}
F_{\nabla_{ue_i} a}(u) &= u^{-1}\left(\nabla_{ue_i} a\right)_{\pi(u)} \\
&= u^{-1} \lim_{\varepsilon \downarrow 0} \frac{//_\varepsilon^{-1} a_{\gamma(\varepsilon)} - a_{\gamma(0)}}{\varepsilon} \\
&= \lim_{\varepsilon \downarrow 0} \frac{u(\varepsilon)^{-1} a_{\gamma(\varepsilon)} - u(0)^{-1} a_{\gamma(0)}}{\varepsilon} \\
&= \lim_{\varepsilon \downarrow 0} \frac{F_a(u(\varepsilon)) - F_a(u(0))}{\varepsilon} \\
&= (H_i)_u F_a \\
&= (H_i F_a)(u).
\end{aligned}$$

$\square$

Now consider diffusion processes X_t on M generated by the operator

$$\mathscr{L} = \Delta + Z$$

where $Z \in \Gamma(TM)$ is a smooth vector field. Such diffusions on M may be constructed from the corresponding horizontal diffusions on $\mathrm{O}(M)$ generated by

$$\Delta^{\mathrm{hor}} + \bar Z$$

where the vector field $\bar Z$ is the horizontal lift of Z to $\mathrm{O}(M)$, i.e. $\bar Z_u = h_u(Z_{\pi(u)})$, $u \in \mathrm{O}(M)$. More precisely, we start from the Stratonovich stochastic differential equation on $\mathrm{O}(M)$,

$$dU_t = \sum_{i=1}^n H_i(U_t) \circ dB_t^i + \bar Z(U_t)\, dt, \quad U_0 = u \in \mathrm{O}(M) \tag{1.1}$$

where B_t is a Brownian motion on $\mathbb{R}^n$ sped up by 2, that is $dB_t^i dB_t^j = 2\delta_{ij}\, dt$. Then for $X_t = \pi(U_t)$, the following equation holds:

$$dX_t = \sum_{i=1}^n U_t e_i \circ dB_t^i + Z(X_t)\, dt, \quad X_0 = x := \pi u. \tag{1.2}$$

The Brownian motion B is the martingale part of the anti-development $\int_U \vartheta$ of X, where ϑ denotes the canonical 1-form ϑ on $\mathrm{O}(M)$, i.e.

$$\vartheta_u(e) = u^{-1} e_{\pi(u)}, \quad e \in T_u \mathrm{O}(M).$$

In particular, for $F \in C^\infty(\mathrm{O}(M))$, resp. $f \in C^\infty(M)$, we have

$$\begin{aligned} d(F \circ U_t) &= \sum_{i=1}^n (H_i F)(U_t) \circ dB_t^i + (\bar{Z}F)(U_t)\, dt \\ &= \sum_{i=1}^n (H_i F)(U_t)\, dB_t^i + \left(\Delta^{\mathrm{hor}} + \bar{Z}\right)(F)(U_t)\, dt, \end{aligned} \tag{1.3}$$

respectively

$$\begin{aligned} d(f \circ X_t) &= \sum_{i=1}^n (df)(U_t e_i) \circ dB_t^i + (Zf)(X_t)\, dt \\ &= \sum_{i=1}^n (df)(U_t e_i)\, dB_t^i + (\Delta + Z)(f)(X_t)\, dt. \end{aligned}$$

Typically, solutions to (1.2) are defined up to some maximal lifetime $\zeta(x)$ which may be finite. Then we have, almost surely,

$$\{\zeta(x) < \infty\} \subset \{X_t \to \infty \text{ as } t \uparrow \zeta(x)\}$$

where on the right-hand side, the symbol ∞ denotes the point at infinity in the one-point compactification of M. It can be shown that the maximal lifetime of solutions to equation (1.1) and to (1.2) coincide, see e.g. [12].

In case of a non-trivial lifetime the subsequent stochastic equations should be read for $t < \zeta(x)$.

Proposition 1.2. *Let $/\!/_t \colon E_{X_0} \to E_{X_t}$ be parallel transport in E along X, induced by the parallel transport on M,*

$$/\!/_t = U_t U_0^{-1} \colon T_{X_0} M \to T_{X_t} M.$$

Then, for $a \in \Gamma(E)$, we have

$$d\left(/\!/_t^{-1} a(X_t)\right) = \sum_{i=1}^n /\!/_t^{-1} \left(\nabla_{U_t e_i} a\right) \circ dB_t^i + /\!/_t^{-1} (\nabla_Z a)(X_t)\, dt,$$

respectively in Itô form,

$$d\left(/\!/_t^{-1} a(X_t)\right) = \sum_{i=1}^n /\!/_t^{-1} \left(\nabla_{U_t e_i} a\right) dB_t^i + /\!/_t^{-1} (\Box a + \nabla_Z a)(X_t)\, dt.$$

More succinctly, the last two equations may be written as

$$d\left(/\!/_t^{-1} a(X_t)\right) = /\!/_t^{-1} \nabla_{\circ dX_t} a,$$

respectively

$$d\left(/\!/_t^{-1} a(X_t)\right) = /\!/_t^{-1} \nabla_{dX_t} \alpha + /\!/_t^{-1} (\Box a)(X_t)\, dt.$$

Proof. We have $//_t^{-1}a(X_t) = U_0U_t^{-1}a(X_t) = U_0F_a(U_t)$. It is easily checked that $\bar{Z}F_a = F_{\nabla_Z a}$. Thus, we obtain from equation (1.3)

$$\begin{aligned} dF_a(U_t) &= \sum_{i=1}^n (H_iF_a)(U_t)\,dB_t^i + \left(\Delta^{\mathrm{hor}}F_a + \bar{Z}F_a\right)(U_t)\,dt \\ &= \sum_{i=1}^n \left(F_{\nabla_{U_te_i}a}\right)(U_t)\,dB_t^i + \left(F_{\Box a} + F_{\nabla_Z a}\right)(U_t)\,dt \\ &= \sum_{i=1}^n U^{-1}\left(\nabla_{U_te_i}a\right)(X_t)\,dB_t^i + U_t^{-1}\left(\Box a + \nabla_Z a\right)(X_t)\,dt. \end{aligned}$$ □

Corollary 1.3. *Fix $T > 0$ and let $a_t \in \Gamma(E)$ solve the equation*

$$\frac{\partial}{\partial t}a_t = \Box a_t + \nabla_Z a_t \quad \text{on } [0,T]\times M.$$

Then

$$//_t^{-1}a_{T-t}(X_t), \quad 0 \le t < T \wedge \zeta(x),$$

is a local martingale.

Proof. Indeed we have

$$d(//_t^{-1}a_{T-t}(X_t)) \overset{\mathrm{m}}{=} //_t^{-1}\underbrace{\left(\Box a_{T-t} + \nabla_Z a_t + \frac{\partial}{\partial t}a_{T-t}\right)}_{=0}(X_t)\,dt = 0,$$

where $\overset{\mathrm{m}}{=}$ denotes equality modulo differentials of local martingales. □

We are now going to look at operators $\mathscr{L}^{\mathscr{R}}$ on $\Gamma(E)$ which differ from $\Box$ by a zero-order term, in other words,

$$\Box - \mathscr{L}^{\mathscr{R}} = \mathscr{R} \quad \text{where } \mathscr{R} \in \Gamma(\mathrm{End}E). \tag{1.4}$$

Thus, by definition, the action $\mathscr{R}_x\colon E_x \to E_x$ is linear for each $x \in M$.

Example 1.4. A typical example is $E = \Lambda^pT^*M$ and $A^p(M) = \Gamma(\Lambda^pT^*M)$ with $p \ge 1$. The *de Rham-Hodge Laplacian*

$$\Delta^{(p)} = -(d^*d + dd^*) : A^p(M) \to A^p(M)$$

then takes the form

$$\Delta^{(p)}\alpha = \Box\alpha - \mathscr{R}\alpha$$

where $\mathscr{R}$ is given by the Weitzenböck decomposition. In the special case $p = 1$, one obtains $\mathscr{R}\alpha = \mathrm{Ric}(\alpha^\sharp, \cdot)$ where $\mathrm{Ric}\colon TM \oplus TM \to \mathbb{R}$ is the Ricci tensor.

Definition 1.5. Fix $x \in M$ and let X_t be a diffusion to $\mathscr{L} = \Delta + Z$, starting at x. Let Q_t be the $\mathrm{Aut}(E_x)$-valued process defined by the following linear pathwise differential equation

$$\frac{d}{dt}Q_t = -Q_t\mathscr{R}_{//_t}, \quad Q_0 = \mathrm{id}_{E_x},$$

where

$$\mathscr{R}_{//_t} := //_t^{-1} \circ \mathscr{R}_{X_t} \circ //_t \in \mathrm{End}(E_x)$$

and $//_t$ is parallel transport in E along X.

Proposition 1.6. *Let $\mathscr{L}^{\mathscr{R}} = \square - \mathscr{R}$ be as in equation* (1.4) *and X_t be a diffusion to $\mathscr{L} = \Delta + Z$, starting at x. Then, for any $a \in \Gamma(E)$,*

$$d\Big(Q_t //_t^{-1} a(X_t)\Big) = \sum_{i=1}^n Q_t //_t^{-1} \Big(\nabla_{U_t e_i} a\Big)\, dB_t^i + Q_t //_t^{-1} (\square a + \nabla_Z a - \mathscr{R} a)(X_t)\, dt.$$

Proof. Let $n_t := //_t^{-1} a(X_t)$. Then

$$\begin{aligned} d(Q_t n_t) &= (dQ_t)\, n_t + Q_t\, dn_t \\ &= -Q_t //_t^{-1} \mathscr{R}_{X_t} //_t^{-1} n_t\, dt + Q_t\, dn_t \\ &= -Q_t //_t^{-1} (\mathscr{R} a)(X_t)\, dt + Q_t\, dn_t. \end{aligned}$$

The claim thus follows from Proposition 1.2. □

Corollary 1.7. *Fix $T > 0$ and let $X_t(x)$ be a diffusion to $\mathscr{L} = \Delta + Z$, starting at x. Suppose that a_t solves*

$$\begin{cases} \dfrac{\partial}{\partial t} a_t = (\square - \mathscr{R} + \nabla_Z)\, a_t \quad \text{on } [0,T] \times M, \\ a_t|_{t=0} = a \in \Gamma(E). \end{cases}$$

Then

$$N_t := Q_t //_t^{-1} a_{T-t}(X_t(x)), \quad 0 \le t < T \wedge \zeta(x), \tag{1.5}$$

is a local martingale, starting at $a_T(x)$. In particular, if $\zeta(x) = \infty$ and if equation (1.5) *is a true martingale on $[0,T]$, we arrive at the formula*

$$a_T(x) = \mathbb{E}[Q_T //_T^{-1} a(X_T(x))], \quad a \in \Gamma(E).$$

Proof. Indeed, we have

$$dN_t \overset{\mathrm{m}}{=} Q_t //_t^{-1} \underbrace{\left((\square + \nabla_Z - \mathscr{R}) a_{T-t} + \frac{\partial}{\partial t} a_{T-t}\right)}_{=0} (X_t)\, dt = 0. \qquad \square$$

Remark 1.8. Note that

$$\frac{d}{dt} Q_t = -Q_t \mathscr{R}_{//_t}, \quad \text{with } Q_0 = \mathrm{id}_{E_x},$$

implies the obvious estimate

$$\|Q_t\|_{\mathrm{op}} \le \exp\left(-\int_0^t \underline{\mathscr{R}}(X_s(x))\, ds\right)$$

where $\underline{\mathscr{R}}(x) = \inf\{\langle \mathscr{R}_x v, w\rangle : v, w \in E_x,\ \|v\| \le 1 \text{ and } \|w\| \le 1\}$.

1.2. **Commutation formulae.** In the sequel, we consider the special case $E = T^*M$. Thus $\Gamma(E)$ is the space of differential 1-forms on M. The results of this section apply to vector fields as well, by identifying vector fields $V \in \Gamma(TM)$ and 1-forms $\alpha \in \Gamma(T^*M)$ via the metric:

$$V \longleftrightarrow V^\flat, \quad \alpha \longleftrightarrow \alpha^\#.$$

Let $Z \in \Gamma(TM)$ be a vector field on M. Then the divergence of Z, denoted by $\operatorname{div} Z \in \mathscr{C}^\infty(M)$, is defined by $\operatorname{div} Z :=\ \operatorname{trace}(v \mapsto \nabla_v Z)$. Therefore

$$(\operatorname{div} Z)(x) = \sum_{i=1}^n \langle \nabla_{v_i} X, v_i\rangle$$

for any orthonormal basis $\{v_i\}_{i=1}^n$ for T_xM. For compactly supported f we have

$$\langle Z, \nabla f\rangle_{L^2(TM)} = -\langle \operatorname{div} Z, f\rangle_{L^2(M)}.$$

The adjoint Z^* of Z is given by the relation

$$Z^* f = -Zf - (\operatorname{div} Z) f, \quad f \in C^\infty(M).$$

If either f or h is compactly supported, this implies

$$\langle Zf, h\rangle_{L^2(M)} = \langle f, Z^* h\rangle_{L^2(M)}.$$

Similarly, for $\alpha \in \Gamma(T^*M)$, we let

$$(\operatorname{div}\alpha)(x) = \operatorname{trace}(T_xM \xrightarrow{\nabla\alpha} T_x^*M \xrightarrow{\#} T_xM).$$

Thus $\operatorname{div} Y = \operatorname{div} Y^\flat$ and $\operatorname{div}\alpha = \operatorname{div}\alpha^\#$. That is, if $\delta = d^*$ denotes the usual codifferential then $\operatorname{div}\alpha = -\delta\alpha$. Finally, we define

$$\operatorname{Ric}_Z(X,Y) := \operatorname{Ric}(X,Y) - \langle \nabla_X Z, Y\rangle, \quad X, Y \in \Gamma(TM).$$

Notation 1.9. For the sake of convenience, we read bilinear forms on M, such as Ric_Z, likewise as sections of $\operatorname{End}(T^*M)$ or $\operatorname{End}(TM)$, e.g.

$$\operatorname{Ric}_Z(\alpha) := \operatorname{Ric}_Z(\alpha^\sharp, \cdot), \quad \alpha \in T^*M,$$
$$\operatorname{Ric}_Z(v) := \operatorname{Ric}_Z(v, \cdot)^\sharp, \quad v \in TM.$$

If there is no risk of confusion, we do not distinguish in notation. In particular, depending on the context, $(\operatorname{Ric}_Z)_{//_t}$ may be a random section of $\operatorname{End}(T^*M)$ or of $\operatorname{End}(TM)$.

Lemma 1.10 (Commutation rules)**.** *Let $Z \in \Gamma(TM)$.*

(1) *For the differential d, we have*

$$d(\Delta + Z) = (\square - \operatorname{Ric}_Z + \nabla_Z) d;$$

(2) *for the codifferential $d^* = -\operatorname{div}$, we have*

$$(\Delta + Z^*) d^* = d^*(\square - \operatorname{Ric}_Z^* + \nabla_Z^*),$$

where the formal adjoint of ∇_Z (acting on 1-forms) is $\nabla_Z^*\alpha = -\nabla_Z\alpha - (\operatorname{div} Z)\alpha$.

Proof. Indeed, for any smooth function f we have

$$\begin{aligned} d(\Delta + Z) f &= d(-d^* df + (df)Z) \\ &= \Delta^{(1)} df + \nabla_Z df + \langle \nabla_{\bullet} Z, \nabla f\rangle \\ &= (\square + \nabla_Z)(df) - \operatorname{Ric}_Z(\cdot, \nabla f) \\ &= (\square - \operatorname{Ric}_Z + \nabla_Z)(df). \end{aligned}$$

The formula in (2) is then just dual to (1). □

1.3. **A formula for the differential.** Now, let $X_t(x)$ be a diffusion to $\Delta + Z$ on M, starting at $X_0(x) = x$, U_t a horizontal lift of X to $\mathrm{O}(M)$ and $B = U_0 \int_U \vartheta$ the martingale part of the anti-development of $X_t(x)$ to T_xM. Let Q_t be the $\operatorname{Aut}(T_x^*M)$-valued process defined by

$$\frac{d}{dt} Q_t = -Q_t (\operatorname{Ric}_Z)_{//_t}$$

with $Q_0 = \mathrm{id}_{T_x^*M}$, let

$$P_t f(x) = \mathbb{E}\Big[\mathbb{1}_{\{t<\zeta(x)\}} f(X_t(x))\Big]$$

be the minimal semigroup generated by $\Delta + Z$ on M, acting on bounded measurable functions f.

Fix $T > 0$ and let ℓ_t be an adapted process with paths in the Cameron-Martin space $L^{1,2}([0,T];T_xM)$. By Corollary 1.7

$$N_t := Q_t //_t^{-1}(dP_{T-t}f), \quad t < T \wedge \zeta(x), \tag{1.6}$$

is local martingale. Therefore

$$N_t(\ell_t) - \int_0^t Q_s //_s^{-1}(dP_{T-s}f)(\dot\ell_s)ds$$

is a local martingale. By integration by parts

$$\int_0^t Q_s //_s^{-1}(dP_{T-s}f)(\dot\ell_s)ds - \frac{1}{2}(P_{T-t}f)(X_t(x)) \int_0^t \langle Q_s^{\mathrm{tr}}(\dot\ell_s), dB_s\rangle$$

is also a local martingale and therefore

$$Q_t //_t^{-1}(dP_{T-t}f)(\ell_t) - \frac{1}{2}(P_{T-t}f)(X_t(x)) \int_0^t \langle Q_s^{\mathrm{tr}}\dot\ell_s, dB_s\rangle \tag{1.7}$$

is a local martingale, starting at $(dP_Tf)(\ell_0)$. Choosing ℓ_t so that (1.7) is a true martingale on $[0,T]$ with $\ell_0 = v$ and $\ell_T = 0$, we obtain the formula

$$(dP_Tf)(v) = -\frac{1}{2}\mathbb{E}\left[\mathbb{1}_{\{T<\zeta(x)\}} f(X_T(x)) \int_0^T \langle Q_s^{\mathrm{tr}}\dot\ell_s, dB_s\rangle\right]. \tag{1.8}$$

For further details, see [14, 15]. Denoting by $p_t(x,y)$ the smooth heat kernel associated to $\Delta + Z$, since formula (1.8) holds for all smooth functions f of compact support, it implies Bismut's formula

$$(d\log p_T(\cdot,y))_x(v) = -\frac{1}{2}\mathbb{E}\left[\int_0^{\tau\wedge T} \langle Q_s^{\mathrm{tr}}\dot\ell_s, dB_s\rangle \Big| X_T(x) = y\right].$$

The argument leading to formula (1.8) is based on the fact that the local martingale (1.7) is a true martingale. Since the condition on ℓ_t is imposed on the left endpoint, this can always be achieved, by taking $\ell_s = 0$ for $s \geq \tau \wedge T$ where τ is the first exit time of some relatively compact neighbourhood of x. No bounds on the geometry are needed; also explosion in finite times of the underlying diffusion can be allowed. For the problem of constructing appropriate finite energy processes ℓ_s with the property $\ell_s = 0$ for $s \geq \tau \wedge T$, see [15], resp. [16, Lemma 4.3].

Imposing in (1.7) however the conditions $\ell_0 = 0$ and $\ell_T = v$ would lead to a formula for

$$\mathbb{E}\left[Q_T //_T^{-1}(df)_{X_T(x)}(v)\right]$$

not involving derivatives of f, which clearly requires strong assumptions. If the local martingale (1.6) is a true martingale, we get the formula

$$(dP_Tf)_x(v) = \mathbb{E}\left[Q_T //_T^{-1}(df)_{X_T(x)}(v)\right].$$

For such a formula to hold, obviously $X_t(x)$ needs to be non-explosive.

1.4. **A formula for the codifferential.** Recall that, according to Lemma 1.10, we have

$$(\Delta + Z + \operatorname{div} Z)\operatorname{div} = \operatorname{div}(\Box + \nabla_Z - \operatorname{Ric}^*_{-Z} + \operatorname{div} Z). \tag{1.9}$$

For a bounded 1-form α suppose α_t satisfies

$$\frac{d}{dt}\alpha_t = \Big(\Box + \nabla_Z - \operatorname{Ric}^*_{-Z} + \operatorname{div} Z\Big)\alpha_t \tag{1.10}$$

with $\alpha_0 = \alpha$, where $\operatorname{div} Z$ acts fibrewise as a multiplication operator, and that Θ_t is the $\operatorname{Aut}(T_xM)$-valued process which solves

$$\frac{d}{dt}\Theta_t = -(\operatorname{Ric}^*_{-Z} - \operatorname{div} Z)_{//_t}\Theta_t$$

with $\Theta_0 = \mathrm{id}_{T_xM}$. Here $\operatorname{Ric}^*_{-Z}$ is the adjoint to Ric_{-Z} acting as endomorphism of T_xM, see Notation 1.9.

Remark 1.11. We have $\Theta_t = Q_t^{\mathrm{tr}}$ if we set $\mathscr{R} := \operatorname{Ric}^*_{-Z} - \operatorname{div} Z \in \operatorname{End}(T^*M)$ and define Q_t via Definition 1.5.

Proposition 1.12. *Fix $T > 0$. Let $X_t(x)$ be a diffusion to $\Delta + Z$ on M, starting at x.*

(i) *Then*

$$(\operatorname{div} \alpha_{T-t})(X_t(x)) \exp\left(\int_0^t (\operatorname{div} Z)(X_s(x))\, ds\right)$$

is a local martingale, starting at $\operatorname{div}\alpha_T$.

(ii) *Suppose h_t is an adapted process with paths in $L^{1,2}([0,T];\mathbb{R})$. Then*

$$\operatorname{div}\alpha_{T-t}h_t + \frac{1}{2}\alpha_{T-t}\left(//_t\Theta_t\int_0^t \Big(\dot h_s - (\operatorname{div} Z)(X_s(x))\, h_s\Big)\Theta_s^{-1}//_s^{-1}dB_s\right) \tag{1.11}$$

is a local martingale, starting at $\operatorname{div}\alpha_T h_0$.

Proof. (i) Taking into account the commutation rule (1.9) and the evolution equation (1.10) of α_t, we get

$$\begin{aligned}\partial_t \operatorname{div}\alpha_t &= \operatorname{div}\partial_t\alpha_t\\ &= \operatorname{div}(\Box + \nabla_Z - \operatorname{Ric}^*_{-Z} + \operatorname{div} Z)\alpha_t\\ &= (\Delta + Z + \operatorname{div} Z)\operatorname{div}\alpha_t.\end{aligned} \tag{1.12}$$

The claim then follows from Itô's formula.

(ii) To verify the second item, set

$$\mathbb{A}_t := \exp\left(\int_0^t (\operatorname{div} Z)(X_s(x))\, ds\right)$$

and define $\ell_t := \mathbb{A}_t^{-1}h_t$. Using the fact that $\alpha_{T-t}(//_t\Theta_t)$ is a local martingale, indeed

$$d(\alpha_{T-t}(//_t\Theta_t)) = \sum_{i=1}^n (\nabla_{U_te_i}\alpha_{T-t})(//_t\Theta_t)\, dB_t^i$$

we obtain

$$\begin{aligned}
&(\operatorname{div}\alpha_{T-t})(X_t(x))\mathbb{A}_t\dot{\ell}_t\,dt\\
&\qquad=\sum_{i=1}^n(\nabla_{U_te_i}\alpha_{T-t})(U_te_i)\,\mathbb{A}_t\dot{\ell}_t\,dt\\
&\qquad=\sum_{i=1}^n\Big(//_t^{-1}\nabla_{U_te_i}\alpha_{T-t}\Big)(U_0e_i)\,\mathbb{A}_t\dot{\ell}_t\,dt\\
&\qquad=\sum_{i=1}^n\Big(\nabla_{U_te_i}\alpha_{T-t}\Big)(//_t\Theta_t\Theta_t^{-1}U_0e_i)\,\mathbb{A}_t\dot{\ell}_t\,dt\\
&\qquad=\frac{1}{2}\Big\langle\sum_{i=1}^n\Big(\nabla_{U_te_i}\alpha_{T-t}\Big)(//_t\Theta_t)\,dB_t^i,\mathbb{A}_t\dot{\ell}_t\Theta_t^{-1}dB_t\Big\rangle\\
&\qquad\overset{\mathrm{m}}{=}\frac{1}{2}d\left(\alpha_{T-t}\Big(//_t\Theta_t\int_0^t\mathbb{A}_s\dot{\ell}_s\Theta_s^{-1}\,dB_s\Big)\right)
\end{aligned}$$

where $\overset{\mathrm{m}}{=}$ denotes equality modulo the differential of a local martingale. By part (i)

$$n_t:=(\operatorname{div}\alpha_{T-t})(X_t(x))\mathbb{A}_t$$

is a local martingale and therefore so is

$$n_t\ell_t-\int_0^t n_s\,d\ell_s.$$

Since

$$\mathbb{A}_t\dot{\ell}_t=\dot{h}_t-(\operatorname{div}Z)(X_t(x))\,h_t$$

the result follows by substitution. □

Remark 1.13. a) Let D^n be an exhausting sequence of M by relatively compact open domains. Following the discussion of [3, Appendix B] and [8, Section III.1] it is standard to show that there is a strongly continuous semigroup P_t^n on compactly supported 1-forms α on D^n generated by $L:=\square+\nabla_Z-\operatorname{Ric}^*_{-Z}+\operatorname{div}Z$ with Dirichlet boundary conditions. In probabilistic terms, $\alpha_t^n(x):=(P_t^n\alpha)(x)$ is easily identified as

$$\alpha_t^n(x)=\mathbb{E}\Big[\mathbb{1}_{\{t<\tau^n(x)\}}\,\alpha(//_t\Theta_t)\Big]$$

where $\tau^n(x)$ is the first exit time of $X_t(x)$ from D^n, when started at $x\in D^n$. As $n\to\infty$, the semigroup α_t^n converges to

$$\alpha_t(x)=\mathbb{E}\Big[\mathbb{1}_{\{t<\zeta(x)\}}\,\alpha(//_t\Theta_t)\Big].\tag{1.13}$$

In particular, α_t solves equation (1.10) on M.

b) Formula (1.13) shows that α_t is bounded in case α is bounded. Choosing the process h in (1.11) in such a way that $h_0=1$ but $h_t=0$ for $t\geq\tau\wedge T$ where τ is the first exit time of $X_t(x)$ of some relatively compact neighbourhood of x, we arrive at the formula

$$(\operatorname{div}\alpha_T)(x)=-\frac{1}{2}\mathbb{E}\left[\mathbb{1}_{\{T<\zeta(x)\}}\alpha\left(//_T\Theta_T\int_0^T\Big(\dot{h}_s-(\operatorname{div}Z)(X_s(x))\,h_s\Big)\Theta_s^{-1}//_s^{-1}dB_s\right)\right].\tag{1.14}$$

Note that the local formula (1.14) doesn't require assumptions, either on the geometry of M or on the drift vector field Z. Indeed, with an appropriate choice of h it is always possible to make (1.11) a true martingale.

Lemma 1.14. *Suppose* Ric_Z *is bounded below, that* $\mathrm{Ric} + (\nabla_{\bullet} Z)^*$, $\operatorname{div} Z$ *and* $\operatorname{div} \alpha$ *are bounded with* h_t *bounded and*

$$\left(\int_0^T |\dot{h}_s|^2 ds\right)^{1/2} \in L^{1+\epsilon}$$

for some $\epsilon > 0$. *Then the local martingale* (1.11) *is a true martingale.*

Proof. Since Ric_Z is bounded below, X_t is non-explosive, by [17, Corollary 2.1.2]. In this case we have $\alpha_t = \mathbb{E}[\alpha(//_t \Theta_t)]$. From equation (1.12) we see that

$$u(t,x) := (\operatorname{div} \alpha_t)(x)$$

solves the heat equation

$$\partial_t u = (\Delta + Z + \operatorname{div} Z) u \tag{1.15}$$

with initial condition $u(0,\cdot) = \operatorname{div} \alpha$. By means of equation (1.14), combined with the bound on $\operatorname{div} Z$ and the other assumptions, we see that $\operatorname{div} \alpha_t$ is a bounded solution to (1.15), which implies

$$\operatorname{div} \alpha_t = \mathbb{E}\left[(\operatorname{div} \alpha)(X_t) \exp\left(\int_0^t (\operatorname{div} Z)(X_s)\, ds\right)\right] \tag{1.16}$$

for all $t \geq 0$. Note that our assumptions control the norms of Θ_t and Θ_t^{-1}. Combined with the assumptions on h this proves that (1.11) is indeed a true martingale. □

Remark 1.15. Equation (1.16) shows that div commutes with the semigroup $P_t^{(1)} \alpha := \alpha_t$ on 1-forms:

$$\operatorname{div} P_t^{(1)} \alpha = P_t^{\operatorname{div} Z} (\operatorname{div} \alpha)$$

where

$$P_t^{\rho} f := \mathbb{E}\left[f(X_t) \exp\left(\int_0^t \rho(X_s)\, ds\right)\right]$$

denotes the Feynman-Kac semigroup on functions to $\Delta + Z$ with scalar potential ρ.

Using the identification of differential forms and vector fields via the metric, we obtain the following result.

Theorem 1.16. *Let* M *be a Riemannian manifold and* Z *a smooth vector field on* M. *Let* $X = X(x)$ *be a diffusion to* $\Delta + Z$ *on* M, *starting at* $X_0(x) = x$, *which is assumed to be non-explosive. Let* $T > 0$ *and* h *be an adapted process with paths in* $L^{1,2}([0,T];\mathbb{R})$ *such that* $h_0 = 0$ *and* $h_T = 1$, *and such that* (1.11) *is a true martingale. Then for all bounded smooth vector fields* V *on* M,

$$\mathbb{E}[(\operatorname{div} V)(X_T(x))] = -\frac{1}{2}\mathbb{E}\left[\left\langle V(X_T(x)), //_T \Theta_T \int_0^T \left((\operatorname{div} Z)(X_t(x)) h_t - \dot{h}_t\right) \Theta_t^{-1} dB_t \right\rangle\right]$$

where Θ *is the* $\mathrm{Aut}(T_x M)$*-valued process defined by the following pathwise differential equation:*

$$\frac{d}{dt}\Theta_t = -\mathrm{Ric}_{//_t} \Theta_t - (\nabla_{\bullet} Z)^*_{//_t} \Theta_t + (\operatorname{div} Z) \Theta_t$$

with $\Theta_0 = \mathrm{id}_{T_x M}$.

Corollary 1.17. *Suppose f is a bounded smooth function and that V is a bounded smooth vector field with* $\operatorname{div} V$ *bounded. Then, under the assumptions of Theorem 1.16, by using the relation* $\operatorname{div}(fV) = Vf + f \operatorname{div} V$*, we get*

$$\begin{aligned} P_T(V(f))(x) &= -\mathbb{E}[f(X_T(x))\,(\operatorname{div} V)(X_T(x))] \\ &\quad - \frac{1}{2}\mathbb{E}\Big[f(X_T(x))\Big\langle V(X_T(x)), //_T\Theta_T \int_0^T \Big((\operatorname{div} Z)(X_t(x))h_t - \dot{h}_t\Big)\,\Theta_t^{-1} dB_t\Big\rangle\Big] \end{aligned}$$

where the right-hand side does not contain any derivatives of f.

Corollary 1.18. *Under the assumptions of Theorem 1.16 we have*

$$(\nabla \log p_T(x,\cdot))_y = -\frac{1}{2}\mathbb{E}\Big[//_T\Theta_T \int_0^T \Big((\operatorname{div} Z)(X_t(x))h_t - \dot{h}_t\Big)\Theta_t^{-1} dB_t \Big| X_T(x) = y\Big]$$

with Θ given as above.

Proof. By Theorem 1.16, for all smooth, compactly supported vector fields V we have

$$\begin{aligned} &P_T(\operatorname{div} V)(x) \\ &= -\frac{1}{2}\int_M \Big\langle V(y), \mathbb{E}\Big[//_T\Theta_T \int_0^T \Big((\operatorname{div} Z)(X_t(x))h_t - \dot{h}_t\Big)\Theta_t^{-1} dB_t \Big| X_T(x) = y\Big]\Big\rangle p_T(x,y)\operatorname{vol}(dy), \end{aligned}$$

but on the other hand

$$\begin{aligned} P_T(\operatorname{div} V)(x) &= \int_M (\operatorname{div} V)(y)\, p_T(x,y)\operatorname{vol}(dy) \\ &= -\int_M (dp_T(x,\cdot))_y\; V(y)\operatorname{vol}(dy) \\ &= -\int_M (d\log p_T(x,\cdot))_y\; V(y)\, p_T(x,y)\operatorname{vol}(dy) \end{aligned}$$

so the result follows. □

1.5. **Shift-Harnack Inequalities.** Suppose Ric_Z is bounded below, that $\mathrm{Ric} + (\nabla_{\bullet} Z)^*$ and $\operatorname{div} Z$ are bounded and that the following formula holds, for all $t > 0$, all $f \in C_b^1(M)$ and all bounded vector fields V with $\operatorname{div} V$ bounded (see Corollary 1.17):

$$\begin{aligned} P_t(V(f))(x) &= -\mathbb{E}[f(X_t(x))\,(\operatorname{div} V)(X_t(x))] \\ &\quad - \frac{1}{2}\mathbb{E}\Big[f(X_t(x))\Big\langle V(X_t(x)), //_t\Theta_t \int_0^t \Big[(\operatorname{div} Z)(X_r(x))\frac{r}{t} - \frac{1}{t}\Big]\Theta_r^{-1} dB_r\Big\rangle\Big]. \end{aligned}$$

Fix $T > 0$. Then, by Jensen's inquality (see [13, Lemma 6.45]), there exist $c, C_1(T) > 0$ such that

$$|P_t(V(f))| \le \delta\,(P_t(f\log f) - P_t f \log P_t f) + \underbrace{\Big(|\operatorname{div} V|_\infty + \delta c + \frac{C_1(T)}{\delta t}|V|_\infty^2\Big)}_{=:\,\alpha_1(\delta,t,V)} P_t f \tag{1.17}$$

for all $\delta > 0$, $t \in (0,T]$ and positive $f \in C_b^1(M)$. Alternatively, by the Cauchy-Schwarz inequality, there exists $C_2(T) > 0$ such that

$$|P_t(V(f))|^2 \le \underbrace{\Big(|\operatorname{div} V|_\infty + \frac{C_2(T)}{\sqrt{t}}|V|_\infty\Big)^2}_{=:\,\alpha_2(t,V)} P_t f^2 \tag{1.18}$$

for all $t \in (0,T]$ and $f \in C_b^1(M)$. These estimates can be used to derive *shift-Harnack inequalities*, as shown by F.-Y. Wang for the case of a Markov operator on a Banach space (see [18, Proposition 2.3]). In particular, suppose $\{F_s\colon s \in [0,1]\}$ is a C^1 family of diffeomorphisms of M with $F_0 = \mathrm{id}_M$. For each $s \in [0,1]$ define a vector field V_s on M by

$$V_s := (DF_s)^{-1}\dot{F}_s$$

and assume V_s and $\operatorname{div} V_s$ are uniformly bounded. Note $\frac{d}{ds}(f \circ F_s) = \nabla_{V_s}(f \circ F_s)$. Fixing $p \geq 1$ and setting $\beta(s) = 1 + (p-1)s$, as in the first part of [18, Proposition 2.3], we deduce from inequality (1.17) that

$$\frac{d}{ds}\log\Big(P_t(f^{\beta(s)} \circ F_s)\Big)^{p/\beta(s)} \geq -\frac{p}{\beta(s)}\alpha_1\left(\frac{\beta'(s)}{\beta(s)}, t, V_s\right)$$

for all $s \in [0,1]$, which when integrated gives the shift-Harnack inequality

$$(P_t f)^p \leq (P_t(f^p \circ F_1))\exp\left(\int_0^1 \frac{p}{\beta(s)}\alpha_1\left(\frac{\beta'(s)}{\beta(s)}, t, V_s\right)ds\right)$$

for each $t \in [0,T]$ and positive $f \in C_b^1(M)$. Alternatively, from inequality (1.18) and following the calculation in the second part of [18, Proposition 2.3], we deduce

$$P_t f \leq P_t(f \circ F_1) + \left(\int_0^1 \alpha_2(t, V_s)ds\right)^{1/2} \sqrt{P_t f^2}$$

for each $t \in [0,T]$ and positive $f \in C_b^1(M)$. The shift F_1 could be given by the exponential of a well-behaved vector field; the shifts considered in [18] are of the form $x \mapsto x + v$, for some v belonging to the Banach space.

2. Extrinsic Formulae

Suppose now that M is an n-dimensional smooth manifold. Suppose A_0 is a smooth vector field and

$$A\colon M \times \mathbb{R}^m \to TM, \quad (x,e) \mapsto A(x)e,$$

a smooth bundle map over M. This means $A(\cdot)e$ is a vector field on M for each $e \in \mathbb{R}^m$, and $A(x)\colon \mathbb{R}^m \to T_xM$ is linear for each $x \in M$

For an $\mathbb{R}^m$-valued Brownian motion B_t, sped up by 2 so that $d[B,B]_t = 2\,\mathrm{id}_{\mathbb{R}^m}dt$, defined on a filtered probability space $(\Omega, \mathscr{F}, \mathbb{P}; (\mathscr{F}_t)_{t\in\mathbb{R}_+})$, satisfying the usual completeness conditions, consider the Stratonovich stochastic differential equation

$$dX_t = A_0(X_t)\,dt + A(X_t) \circ dB_t. \tag{2.1}$$

Given an orthonormal basis $\{e_i\}_{i=1}^m$ of $\mathbb{R}^m$ set $A_i(\cdot) := A(\cdot)e_i$ and $B_t^i := \langle B_t, e_i\rangle$. Then the previous equation can be equivalently written

$$dX_t = A_0(X_t)\,dt + \sum_{i=1}^m A_i(X_t) \circ dB_t^i.$$

There is a partial flow $X_t(\cdot)$, $\zeta(\cdot)$ associated to (2.1) (see [10] for details) such that for each $x \in M$ the process $X_t(x)$, $0 \leq t < \zeta(x)$ is the maximal strong solution to (2.1) with starting point $X_0(x) = x$, defined up to the explosion time $\zeta(x)$; moreover using the notation $X_t(x,\omega) = X_t(x)(\omega)$ and $\zeta(x,\omega) = \zeta(x)(\omega)$, if

$$M_t(\omega) = \{x \in M : t < \zeta(x,\omega)\}$$

then there exists $\Omega_0 \subset \Omega$ of full measure such that for all $\omega \in \Omega_0$:

i) $M_t(\omega)$ is open in M for each $t \geq 0$, i.e. $\zeta(\cdot,\omega)$ is lower semicontinuous on M;

ii) $X_t(\cdot,\omega) : M_t(\omega) \to M$ is a diffeomorphism onto an open subset of M;
iii) The map $s \mapsto X_s(\cdot,\omega)$ is continuous from $[0,t]$ into $C^\infty(M_t(\omega), M)$ with its C^∞-topology, for each $t > 0$.

The solution processes $X = X(x)$ to (2.1) are diffusions on M with generator

$$\mathscr{L} := A_0 + \sum_{i=1}^m A_i^2$$

We will assume that the equation is non-degenerate, which is to say that $A(x) : \mathbb{R}^m \to T_xM$ is surjective for all $x \in M$. Then A induces a Riemannian metric on M, the quotient metric, with respect to which

$$A(x)^* = (A(x)|_{\ker A(x)^\perp})^{-1}$$

and whose inner product $\langle \cdot, \cdot \rangle$ on a tangent space T_xM is given by

$$\langle v, u \rangle = \langle A(x)^* v, A(x)^* u \rangle_{\mathbb{R}^m}.$$

2.1. **A formula for the differential.** Denote by

$$P_t f(x) := \mathbb{E}\left[\mathbb{1}_{\{t<\zeta(x)\}} f(X_t(x)) \right]$$

the minimal semigroup associated to equation (2.1), acting on bounded measurable functions f. In terms of any linear connection $\tilde{\nabla}$ on TM, a solution $TX_t(x)$ to the derivative equation

$$d^{\tilde{\nabla}} TX_t(x) = \tilde{\nabla}_{TX_t(x)} A_0 \, dt + \sum_{i=1}^m \tilde{\nabla}_{TX_t(x)} A_i \circ dB_t^i$$

with $TX_0(x) = \mathrm{id}_{T_xM}$ is the derivative (in probability) at x of the solution flow to (2.1). Our objective will be to find a formula for $P_T(V(f))$ in terms of TX_t. Before doing so, let us briefly derive the corresponding formula for $(dP_T)(v)$. As in Subsection 1.3, let ℓ_t be an adapted process with paths in $L^{1,2}([0,T]; T_{x_0}M)$. By Itô's formula and the Weitzenböck formula (see [4, Theorem 2.4.2])

$$N_t := (dP_{T-t} f)(TX_t(x))$$

is local martingale. Therefore

$$N_t(\ell_t) - \int_0^t (dP_{T-s} f)(TX_s(x)(\dot{\ell}_s)) ds$$

is a local martingale. By integration by parts

$$\int_0^t (dP_{T-s} f)(TX_s(x)(\dot{\ell}_s)) ds - \frac{1}{2} (P_{T-t} f)(X_t(x)) \int_0^t \langle TX_s(x)(\dot{\ell}_s), A(X_s(x)) \, dB_s \rangle$$

is also a local martingale and therefore

$$(dP_{T-t} f)(TX_t(x)\ell_t) - \frac{1}{2} (P_{T-t} f)(X_t(x)) \int_0^t \langle TX_s(x)\dot{\ell}_s, A(X_s(x)) \, dB_s \rangle \tag{2.2}$$

is a local martingale, starting at $(dP_T f)(\ell_0)$. Choosing ℓ_t so that (2.2) is a true martingale with $\ell_0 = v$ and $\ell_T = 0$, we obtain the formula

$$(dP_T f)(v) = -\frac{1}{2} \mathbb{E}\left[\mathbb{1}_{\{T<\zeta(x)\}} f(X_T(x)) \int_0^T \langle TX_s(x)\dot{\ell}_s, A(X_s(x)) \, dB_s \rangle \right]. \tag{2.3}$$

This formula is well-known; it is the one given by [14, Theorem 2.4]. Formula (1.8) can be obtained from it by filtering. Furthermore, it as always possible to choose such ℓ_t, as in Subsection 1.3. Now denote by $p_t(x,y)$ the smooth heat kernel associated to (2.1) such that

$$P_t f(x) = \int_M f(y) p_t(x,y)\,\mathrm{vol}(dy)$$

where $\mathrm{vol}(dy)$ denotes integration with respect to the induced Riemannian volume measure. Since formula (2.3) holds for all smooth functions f of compact support, we deduce from it the Bismut formula

$$(d \log p_T(\cdot,y))_x(v) = -\frac{1}{2}\mathbb{E}\left[\int_0^{\tau\wedge T} \langle TX_s(x)\dot{\ell}_s, A(X_s(x))\,dB_s\rangle\rangle \Big| X_T(x)=y\right],$$

the original version of which was given in [1] for compact manifolds. The version stated here is [14, Corollary 2.5], the non-local version having been earlier given in [6].

2.2. **Induced linear connections.** There are a number of linear connections naturally associated to the map A. Firstly, there is the Levi-Civita connection ∇ for the induced metric. Secondly, there is the *Le Jan-Watanabe connection*, which is given by the push forward under A of the flat connection on $\mathbb{R}^m$. Its covariant derivative $\breve{\nabla}$ is defined by

$$\breve{\nabla}_v U = A(x) d\left(A(\cdot)^* U(\cdot)\right)_x(v) \tag{2.4}$$

for a vector field U and $v \in T_xM$. Like the Levi-Civita connection, it is adapted to the induced metric. In fact, all metric connections on TM arise in this way. In addition to the properties of $\breve{\nabla}$ summarized below, further details of it can be found in [4, 5, 7]. It has the property that if $e \in \ker A(x)^\perp$ then $\breve{\nabla}_v A_e = 0$ for all $v \in T_xM$, where by A_e we mean the section $x \mapsto A(x)e$. It therefore satisfies the Le Jan-Watanabe property

$$\sum_{i=1}^m \breve{\nabla}_{A_i} A_i = 0.$$

To any linear connection $\tilde{\nabla}$ on TM one can associate an adjoint connection $\tilde{\nabla}'$ by

$$\tilde{\nabla}'_v U = \tilde{\nabla}_v U - \tilde{T}(v,U)$$

for v a vector and U a smooth vector field, where $\tilde{T}$ denotes the torsion tensor of $\tilde{\nabla}$. The adjoint of the Le Jan-Watanabe connection will be denoted by $\hat{\nabla}$. It therefore satisfies

$$\hat{\nabla}_v U = \breve{\nabla}_v U - \breve{T}(v,U)$$

or equivalently $\breve{\nabla}_v U = \hat{\nabla}_v U - \hat{T}(v,U)$, where $\breve{T}$ and $\hat{T}$ denote the torsion tensors of $\breve{\nabla}$ and $\hat{\nabla}$, respectively; these antisymmetric tensors satisfy $\breve{T} = -\hat{T}$. By [4, Proposition 2.2.3] the torsion can be written in terms of A by

$$\breve{T}(v,u)_x = A(x)(dA^*)_x(v,u) \tag{2.5}$$

where dA^* denotes the exterior derivative of the $\mathbb{R}^m$-valued 1-form $A^* : TM \to \mathbb{R}^m$. The adjoint connection can therefore be written in terms of A by

$$\hat{\nabla}_v U = A(x)\left(d\left(A^*(\cdot)U(\cdot)\right)_x(v) - (dA^*)_x(v,U)\right).$$

Besides torsion, we will also encounter several expressions involving curvature, including

$$\breve{\mathrm{Ric}} := \sum_{i=1}^m \breve{R}(\cdot, A_i)A_i$$

where $\breve{R}$ denotes the curvature tensor of $\breve{\nabla}$. In particular, [4, Lemma 2.4.3] states for a smooth 1-form ϕ that

$$\sum_{i=1}^{m} L_{A_i} L_{A_i} \phi = \operatorname{trace} \hat{\nabla}^2 \phi - \phi(\breve{\mathrm{Ric}}) \tag{2.6}$$

where L denotes Lie differentiation.

2.3. **Induced differential operators.** With respect to the metric induced by A, we set $\delta := d^*$. For a 1-form ϕ, the codifferential δ satisfies

$$\delta\phi = -\sum_{i=1}^{m} \nabla_{A_i}\phi(A_i) \tag{2.7}$$

but this relation does not hold with ∇ replaced by $\hat{\nabla}$. Nonetheless, for the divergence of a smooth vector field U we do have

$$\operatorname{div} U = \sum_{i=1}^{m} \langle \nabla_{A_i} U, A_i \rangle = \sum_{i=1}^{m} \langle \hat{\nabla}_{A_i} U, A_i \rangle = \operatorname{trace} \hat{\nabla} U \tag{2.8}$$

by the adaptedness of $\breve{\nabla}$.

Lemma 2.1. *For any smooth vector field U, 1-form ϕ and linear connection $\tilde{\nabla}$ with adjoint $\tilde{\nabla}'_{\bullet} U = \tilde{\nabla}_{\bullet} U - \tilde{T}(\,\cdot\,, U)$ we have*

$$(U + \operatorname{div} U)\delta\phi = -\delta\left(\tilde{\nabla}^*_U + (\tilde{\nabla}' U)^*\right)\phi.$$

Proof. As a linear connection, $\tilde{\nabla}$ satisfies

$$L_U \phi = \tilde{\nabla}_U \phi + \phi(\tilde{\nabla}' U).$$

Since d commutes with Lie differentiation, we thus have

$$dUf = L_U df = \tilde{\nabla}_U df + df(\tilde{\nabla}' U) = \tilde{\nabla}_U df + (\tilde{\nabla}' U) df.$$

By duality this implies

$$U^* \delta\phi = \delta\left(\tilde{\nabla}^*_U + (\tilde{\nabla}' U)^*\right)\phi$$

and therefore

$$(U + \operatorname{div} U)\delta\phi = -\delta\left(\tilde{\nabla}^*_U + (\tilde{\nabla}' U)^*\right)\phi$$

since $U^* = -U - \operatorname{div} U$. □

With respect to the induced metric, the formal adjoint ∇^*_U of the differential operator ∇_U acting on 1-forms is given by

$$\nabla^*_U = -\nabla_U - \operatorname{div} U.$$

More generally, we have the following lemma.

Lemma 2.2. *For any smooth vector field U and metric connection $\tilde{\nabla}'$ with adjoint $\tilde{\nabla}$ we have*

$$\tilde{\nabla}^*_U = -\tilde{\nabla}_U - \operatorname{div} U.$$

Proof. Denoting by μ_g the Riemannian volume density, the divergence of a vector field U satisfies $L_U\mu_g = (\operatorname{div} U)\mu_g$ and thus for compactly supported 1-forms ϕ, ψ we have

$$\begin{aligned} L_U(\langle\phi,\psi\rangle\mu_g) &= \langle\tilde{\nabla}'_U\phi,\psi\rangle\mu_g + \langle\phi,\tilde{\nabla}'_U\psi\rangle\mu_g + (\operatorname{div} U)\langle\phi,\psi\rangle\mu_g \\ &= \langle\tilde{\nabla}_U\phi,\psi\rangle\mu_g + \langle\phi,\tilde{\nabla}_U\psi\rangle\mu_g + \langle\phi(\tilde{T}'(U,\cdot)),\psi\rangle\mu_g \\ &\quad + \Big\langle\phi,\psi(\tilde{T}'(U,\cdot))\Big\rangle\mu_g + (\operatorname{div} U)\langle\phi,\psi\rangle\mu_g \end{aligned}$$

from which the result follows, since $\tilde{T}'$ is antisymmetric and $\int_M L_U(\langle\phi,\psi\rangle\mu_g) = 0$, by Stokes' theorem. □

The map A also induces a differential operator $\hat{\delta}$, mapping 1-forms to functions by

$$\hat{\delta}\phi := -\sum_{i=1}^m \iota_{A_i} L_{A_i}\phi.$$

Since $L_{A_i}\phi = \iota_{A_i} d\phi + d(\iota_{A_i}\phi)$, the generator $\mathscr{L}$ can be expressed in terms of $\hat{\delta}$ by

$$\mathscr{L} = L_{A_0} - (\hat{\delta}d + d\hat{\delta}). \tag{2.9}$$

Clearly $\hat{\delta}^2 = 0$, so to find an analogue of the second commutation rule in Lemma 1.10 for $\hat{\delta}$ and $\mathscr{L}$ it suffices to calculate the Lie derivative of $\hat{\delta}$ in the direction A_0. This is the main objective of the remainder of this section. Note that $\hat{\delta}$ need not agree with the codifferential δ. For any smooth vector field U and linear connection $\tilde{\nabla}$ with adjoint $\tilde{\nabla}'$ we have

$$L_U\phi = (\tilde{\nabla}_U\phi) + \phi(\tilde{\nabla}' U) \tag{2.10}$$

and therefore

$$\hat{\delta}\phi = -\sum_{i=1}^m (\hat{\nabla}_{A_i}\phi)(A_i) - \sum_{i=1}^m \phi(\breve{\nabla}_{A_i}A_i) = -\sum_{i=1}^m (\hat{\nabla}_{A_i}\phi)(A_i) \tag{2.11}$$

or alternatively

$$\hat{\delta}\phi = -\sum_{i=1}^m (\breve{\nabla}_{A_i}\phi)(A_i) - \sum_{i=1}^m \phi(\hat{\nabla}_{A_i}A_i) = -\sum_{i=1}^m (\breve{\nabla}_{A_i}\phi)(A_i) \tag{2.12}$$

by the Le Jan-Watanabe property and the fact that $\breve{T}(A_i, A_i) = 0$. Applying (2.10) to the Levi-Civita connection gives

$$\hat{\delta}\phi = -\sum_{i=1}^m (\nabla_{A_i}\phi)(A_i) - \sum_{i=1}^m \phi(\nabla_{A_i}A_i)$$

and so by (2.7) we have

$$\hat{\delta}\phi = \delta\phi - \phi\left(\sum_{i=1}^m \nabla_{A_i}A_i\right) \tag{2.13}$$

which expresses the difference of the operators δ and $\hat{\delta}$.

Lemma 2.3. *For any smooth vector field U and 1-form ϕ we have*

$$\Big(U + \operatorname{trace}\hat{\nabla}U\Big)\hat{\delta}\phi = \hat{\delta}\Big(\hat{\nabla}_U - (\breve{\nabla}U)^* + \operatorname{trace}\hat{\nabla}U\Big)\phi + \phi(U^A)$$

where the vector field U^A is defined by

$$U^A := -\sum_{i=1}^m \Big((\breve{\nabla}U)^* + \breve{\nabla}U\Big)(\nabla_{A_i}A_i) - \sum_{i=1}^m [U, \nabla_{A_i}A_i].$$

Proof. By Lemmas 2.1 and 2.2 we have

$$(U + \operatorname{div} U)\delta\phi = \delta(\hat{\nabla}_U + \operatorname{div} U - (\breve{\nabla} U)^*)\phi.$$

By (2.13) we have

$$\begin{aligned}(U + \operatorname{div} U)\delta\phi = (U + \operatorname{div} U)\hat{\delta}\phi + (\operatorname{div} U)\phi(\nabla_{A_i} A_i) \\ + (\hat{\nabla}_U \phi)(\nabla_{A_i} A_i) + \phi(\hat{\nabla}_U \nabla_{A_i} A_i)\end{aligned}$$

and

$$\begin{aligned}\delta\Big(\hat{\nabla}_U + \operatorname{div} U - (\breve{\nabla} U)^*\Big)\phi = \hat{\delta}\Big(\hat{\nabla}_U + \operatorname{div} U - (\breve{\nabla} U)^*\Big)\phi + (\hat{\nabla}_U \phi)(\nabla_{A_i} A_i) \\ + \Big((\operatorname{div} U - (\breve{\nabla} U)^*)\phi\Big)(\nabla_{A_i} A_i).\end{aligned}$$

Rearranging, the result follows by equation (2.8). □

Note that the vector field A_0^A appears to depend on the Levi-Civita connection via the sum of the vector fields $\nabla_{A_i} A_i$. It is clear that all other objects appearing in the definition of A_0^A can be calculated explicitly in terms of A and A_0, by formula (2.4). The following lemma, combined with formula (2.5), shows that the sum of the vector fields $\nabla_{A_i} A_i$ can also be expressed directly in terms of A.

Lemma 2.4. *We have*

$$\sum_{i=1}^m \nabla_{A_i} A_i = -\sum_{i=1}^m \breve{T}(\cdot, A_i)^*(A_i)$$

where $\breve{T}$ denotes the torsion of the Le Jan-Watanabe connection.

Proof. Suppressing the summation over i, the Le Jan-Watanabe property implies

$$\nabla_{A_i} A_i = \breve{\nabla}_{A_i} A_i - \breve{K}(A_i, A_i) = -\breve{K}(A_i, A_i)$$

where $\breve{K}$ denotes the contorsion tensor of $\breve{\nabla}$. The contorsion tensor measures the extent to which a metric connections fails to be the Levi-Civita connection, vanishing if the connection is torsion free. It is discussed in [9] and [11]. The components of $\breve{K}$ satisfy $\breve{K}^i{}_{jj} = \breve{T}_j{}^i{}_j$, which is to say

$$\breve{K}(A_i, A_i) = (\breve{T}(\cdot, A_i)^\flat)(A_i)^\sharp,$$

where $\flat$ and $\sharp$ are the musical isomorphisms associated to the induced metric. This implies

$$\langle \breve{K}(A_i, A_i), U\rangle = \langle \breve{T}(U, A_i), A_i\rangle$$

for all smooth vector fields U, and therefore

$$\breve{K}(A_i, A_i) = \breve{T}(\cdot, A_i)^*(A_i)$$

as required. □

Consequently

$$A_0^A = \sum_{i=1}^m \Big((\breve{\nabla} A_0)^* + \breve{\nabla} A_0\Big)\Big(\breve{T}(\cdot, A_i)^*(A_i)\Big) + [A_0, \breve{T}(\cdot, A_i)^*(A_i)]. \tag{2.14}$$

2.4. **Commutation formula.** We have, in summary, the following commutation rule, extending formula (1.9).

Proposition 2.5. *For any smooth* 1*-form* ϕ *we have*

$$\Big(\mathscr{L}+\operatorname{trace}\hat{\nabla}A_0\Big)\hat{\delta}\phi=\hat{\delta}\Big(\operatorname{trace}\hat{\nabla}^2+\hat{\nabla}_{A_0}-\breve{\operatorname{Ric}}-(\breve{\nabla}A_0)^*+\operatorname{trace}\hat{\nabla}A_0\Big)\phi+\phi(A_0^A)$$

where the vector field A_0^A *is given by* (2.14)*.*

Proof. The claim follows from Lemmas 2.3 and 2.4 and the relations (2.6) and (2.9). □

Finally, note that for a smooth function f, the codifferential δ satisfies

$$\langle df,\phi\rangle=f\delta(\phi)-\delta(f\phi).$$

We will need an analogous formula for $\hat{\delta}$, as given by the following lemma.

Lemma 2.6. *For any smooth function* f *we have*

$$\langle df,\phi\rangle=f\hat{\delta}(\phi)-\hat{\delta}(f\phi).$$

Proof. Suppressing notationally the summation over i, we have

$$\begin{aligned}\hat{\delta}(f\phi)&=-\iota_{A_i}L_{A_i}(f\phi)\\&=-\iota_{A_i}\Big(\iota_{A_i}d(f\phi)+d(\iota_{A_i}f\phi)\Big)\\&=-\iota_{A_i}\Big(\iota_{A_i}(df\wedge\phi+fd\phi)+\phi(A_i)df+fd(\phi(A_i))\Big)\\&=-\iota_{A_i}\iota_{A_i}(df\wedge\phi)-\phi(A_i)df(A_i)+f\hat{\delta}(\phi)\\&=-\langle df,\phi\rangle+f\hat{\delta}(\phi)\end{aligned}$$

since $\iota_{A_i}\iota_{A_i}(df\wedge\phi)=0$. □

Now we are in a position to deduce formulae for the induced differential operator in terms of the derivative flow TX_t.

2.5. **A formula for the induced differential operator.** We must now assume equation (2.1) is complete, which is to say $\zeta(x)=\infty$, almost surely. For a bounded smooth 1-form α suppose α_t satisfies

$$\partial_t\alpha_t=(\operatorname{trace}\hat{\nabla}^2+\hat{\nabla}_{A_0}-\breve{\operatorname{Ric}}-(\breve{\nabla}A_0)^*+\operatorname{trace}\hat{\nabla}A_0)\alpha_t$$

with $\alpha_0=\alpha$ and that $\Xi_t(x):T_xM\to T_{X_t(x)}M$ solves the covariant Itô equation

$$d^{\hat{\nabla}}\Xi_t(x)=-\Big(\breve{\operatorname{Ric}}+(\breve{\nabla}A_0)^*+\operatorname{trace}\hat{\nabla}A_0\Big)(\Xi_t(x))\,dt+\sum_{i=1}^m\breve{\nabla}_{\Xi_t(x)}A_i\,dB_t^j$$

along the paths of $X_t(x)$ with $\omega_0=\mathrm{id}_{T_xM}$. Fixing $T>0$, by Itô's formula we have

$$\begin{aligned}d(\alpha_{T-t}(\Xi_t(x)))&=\sum_{i=1}^m\hat{\nabla}_{A_i}\alpha_{T-t}(\Xi_t(x))\,dB_t^i+\hat{\nabla}_{A_0}\alpha_{T-t}(\Xi_t(x))\,dt+\partial_t\alpha_{T-t}(\Xi_t(x))\,dt\\&\quad+\operatorname{trace}\hat{\nabla}^2\alpha_{T-t}(\Xi_t(x))\,dt+\alpha_{T-t}(d^{\hat{\nabla}}\Xi_t(x))\\&=\sum_{i=1}^m((\hat{\nabla}_{A_i}\alpha_{T-t})\cdot+\alpha_{T-t}(\breve{\nabla}_{\bullet}A_i))(\Xi_t(x))\,dB_t^i.\end{aligned}\tag{2.15}$$

It follows that $\alpha_{T-t}(\Xi_t(x))$ is a local martingale, starting at α_T. Furthermore, according to equation (26) in [5], for the derivative process $TX_t(x)$ we have

$$d^{\hat{\nabla}} TX_t(x) = -\breve{\mathrm{Ric}}(TX_t(x))\,dt + \breve{\nabla}_{TX_t(x)}A_0\,dt + \sum_{i=1}^m \breve{\nabla}_{TX_t(x)}A_i\,dB_t^i$$

and therefore, by the variation of constants formula, we have

$$\Xi_t(x) = TX_t(x) - TX_t(x)\int_0^t TX_s(x)^{-1}\Big(((\breve{\nabla}A_0)^* + \breve{\nabla}A_0 + \mathrm{trace}\,\hat{\nabla}A_0)(\Xi_s(x))\Big)\,ds.$$

Thus it is possible to calculate $\Xi_t(x)$ without using the parallel transport implicit in the original equation. Moreover, if the vector field A_0 vanishes then $\Xi_t(x)$ is given precisely by the derivative process $TX_t(x)$.

Proposition 2.7. *Suppose h_t is an adapted process with paths in $L^{1,2}([0,T];\mathbb{R})$. Then*

$$\begin{aligned}&\hat{\delta}\alpha_{T-t}h_t - \int_0^t h_s\alpha_{T-s}(A_0^A)\,ds \\ &\quad+ \frac{1}{2}\alpha_{T-t}\bigg(\Xi_t(x)\int_0^t\Big(\dot{h}_s - (\mathrm{trace}\,\hat{\nabla}A_0)(X_s(x))h_s\Big)\Xi_s(x)^{-1}A(X_s(x))\,dB_s\bigg)\end{aligned} \tag{2.16}$$

is a local martingale, starting at $\hat{\delta}\alpha_T h_0$, where the vector field A_0^A is given by (2.14).

Proof. Set

$$\mathbb{A}_t := \exp\bigg(\int_0^t(\mathrm{trace}\,\hat{\nabla}A_0)(X_s(x))\,ds\bigg)$$

and define $\ell_t := \mathbb{A}_t^{-1}h_t$. By equation (2.15), integration by parts and formula (2.11), we have, suppressing the summation over i, that

$$\begin{aligned}&d\bigg(\alpha_{T-t}\Big(\Xi_t(x)\frac{1}{2}\int_0^t \mathbb{A}_s\dot{\ell}_s\Xi_s(x)^{-1}A(X_s(x))\,dB_s\Big)\bigg)\\ &\overset{\mathrm{m}}{=} \frac{1}{2}\Big(((\hat{\nabla}_{A_i}\alpha_{T-t})\cdot + \alpha_{T-t}(\breve{\nabla}_\cdot A_i))(\Xi_t(x))\,dB_t^i\Big)\Big(\mathbb{A}_t\dot{\ell}_t\Xi_t(x)^{-1}A_j(X_t(x))\,dB_t^j\Big)\\ &= \Big((\hat{\nabla}_{A_i}\alpha_{T-t})A_i + \alpha_{T-t}(\breve{\nabla}_{A_i}A_i)\Big)\mathbb{A}_t\dot{\ell}_t\,dt\\ &= (\hat{\nabla}_{A_i}\alpha_{T-t})A_i\mathbb{A}_t\dot{\ell}_t\,dt\\ &= -(\hat{\delta}\alpha_{T-t})\mathbb{A}_t\dot{\ell}_t\,dt\end{aligned} \tag{2.17}$$

where $\overset{\mathrm{m}}{=}$ denotes equality modulo the differential of a local martingale. By Proposition 2.5 and Itô's formula we have

$$d(\mathbb{A}_t\hat{\delta}\alpha_{T-t}) \overset{\mathrm{m}}{=} \mathbb{A}_t\hat{\delta}\partial_t\alpha_{T-t}\,dt + \mathbb{A}_t(\mathscr{L} + \mathrm{trace}\,\hat{\nabla}A_0)\hat{\delta}\alpha_{T-t}\,dt = \mathbb{A}_t\alpha_{T-t}(A_0^A)\,dt$$

which implies

$$n_t := \mathbb{A}_t\hat{\delta}\alpha_{T-t} - \int_0^t \mathbb{A}_s\alpha_{T-s}(A_0^A)ds$$

is a local martingale, starting at $\hat{\delta}\alpha_T$. This implies

$$\begin{aligned}d(n_t\ell_t) &\overset{\mathrm{m}}{=} n_t\dot{\ell}_t\,dt\\ &= (\hat{\delta}\alpha_{T-t})\mathbb{A}_t\dot{\ell}_t\,dt - \dot{\ell}_t\int_0^t \mathbb{A}_s\alpha_{T-s}(A_0^A)\,dsdt.\end{aligned}$$

Substituting the definition of n_t into the left-hand side and performing integration by parts to the second term on the right-hand side implies

$$\hat{\delta}\alpha_{T-t}h_t - \int_0^t (\hat{\delta}\alpha_{T-s})\mathbb{A}_s\dot{\ell}_s\,ds - \int_0^t h_s\alpha_{T-s}(A_0^A)\,ds \tag{2.18}$$

is another local martingale. Since

$$\dot{\ell}_t = \mathbb{A}_t^{-1}\left(\dot{h}_t - (\operatorname{trace}\hat{\nabla}A_0)(X_t(x))h_t\right),$$

substituting formula (2.17) into the second term in (2.18) completes the proof. □

Theorem 2.8. *Suppose h_t is any adapted process with paths in $L^{1,2}([0,\infty);\mathbb{R})$ such that $h_0 = 0$ and $h_T = 1$ and that α is a bounded smooth 1-form. Suppose* (2.1) *is complete and that the local martingales $\alpha_{T-t}(\Xi_t)$ and* (2.16) *are true martingales. Then*

$$P_T(\hat{\delta}\alpha) = \frac{1}{2}\mathbb{E}\left[\alpha\left(\Xi_T\int_0^T \Xi_t^{-1}\left(\left((\operatorname{trace}\hat{\nabla}A_0)(X_t)h_t - \dot{h}_t\right)A(X_t)\,dB_t + 2h_tA_0^A\,dt\right)\right)\right].$$

Proof. By (2.15) we have

$$\alpha_{T-t}(\Xi_t) = \alpha(\Xi_T) - \int_t^T \left((\hat{\nabla}_{A_i}\alpha_{T-s})\cdot + \alpha_{T-s}(\breve{\nabla}_{\cdot}A_i)\right)(\Xi_t)\,dB_t^i$$

and therefore

$$\mathbb{E}\left[\int_0^T \alpha_{T-t}(\Xi_t h_t \Xi_t^{-1}A_0^A)\,dt\right] = \mathbb{E}\left[\alpha\left(\Xi_T\int_0^T h_t\Xi_t^{-1}A_0^A\,dt\right)\right]$$

since $\alpha_{T-t}(\Xi_t)$ is assumed to be a martingale. The result now follows from Proposition 2.7, by taking expectations. □

In analogue to Lemma 1.14, an integrability assumption on h plus suitable bounds on $\breve{\nabla}A_0$, $\operatorname{trace}\hat{\nabla}A_0$, A_0^A and $\hat{\delta}\alpha$ and on the moments of TX_t and TX_t^{-1} would be sufficient to guarantee that $\alpha_{T-t}(\Xi_t)$ and (2.16) are true martingales.

Corollary 2.9. *Suppose f is a bounded smooth function. Suppose V is a bounded smooth vector field with $\sum_{i=1}^m A_i\langle V,A_i\rangle$ bounded. Then, under the assumptions of Theorem 2.8 with $\alpha = fV^\flat$, we have*

$$\begin{aligned} &P_T(V(f))\\ &\quad= -\sum_{i=1}^m \mathbb{E}\left[f(X_T)A_i\langle V,A_i\rangle(X_T)\right]\\ &\qquad - \frac{1}{2}\mathbb{E}\left[f(X_T)\left\langle V(X_T),\Xi_T\int_0^T \Xi_t^{-1}\left(\left((\operatorname{trace}\hat{\nabla}A_0)(X_t)h_t - \dot{h}_t\right)A(X_t)\,dB_t + 2h_tA_0^A\,dt\right)\right\rangle\right] \end{aligned}$$

with

$$\Xi_t = TX_t - TX_t\int_0^t TX_s^{-1}\left((\breve{\nabla}A_0)^* + \breve{\nabla}A_0 + \operatorname{trace}\hat{\nabla}A_0\right)(\Xi_s)\,ds,$$

$$A_0^A = \sum_{i=1}^m \left((\breve{\nabla}A_0)^* + \breve{\nabla}A_0\right)\left(\breve{T}(\cdot,A_i)^*(A_i)\right) + [A_0,\breve{T}(\cdot,A_i)^*(A_i)],$$

where the operators $\hat{\nabla}A_0$, $\breve{\nabla}A_0$ and $\breve{T}(\cdot,A_i)$ are given at each $x \in M$ and $v \in T_xM$ by

$$\begin{aligned}\hat{\nabla}_v A_0 &= A(x)\left(d\left(A^*(\cdot)A_0(\cdot)\right)_x(v) - (dA^*)_x(v,A_0)\right),\\ \breve{\nabla}_v A_0 &= A(x)d\left(A(\cdot)^*A_0(\cdot)\right)_x(v),\\ \breve{T}(v,A_i)_x &= A(x)(dA^*)_x(v,A_i).\end{aligned}$$

Proof. This follows from Theorem 2.8. In particular, Lemma 2.6 implies

$$V(f) = f\hat{\delta}(V^\flat) - \hat{\delta}(fV^\flat)$$

while formula (2.12), the Le Jan-Watanabe property and the adaptedness of $\breve{\nabla}$ imply

$$\hat{\delta}(V^\flat) = -\sum_{i=1}^m \langle \breve{\nabla}_{A_i} V, A_i\rangle = -\sum_{i=1}^m A_i\langle V, A_i\rangle. \qquad \square$$

Note that if (2.1) is a gradient system then $\mathscr{L} = \Delta + A_0$ and A_0^A vanishes and

$$\sum_{i=1}^m A_i\langle V, A_i\rangle = \operatorname{div} V.$$

In this case, since $\operatorname{trace}\hat{\nabla}A_0 = \operatorname{div} A_0$, Corollary 2.9 yields the unfiltered version of Corollary 1.17.

Corollary 2.10. *Under the assumptions of Corollary 2.9 we have*

$$\begin{aligned}&\left(d\log p_T(x,\cdot)\right)_y(v)\\ &\quad= -\Big\langle v, \sum_{i=1}^m \breve{T}(\cdot,A_i)^*(A_i)(y)\Big\rangle\\ &\qquad+\frac{1}{2}\Big\langle v, \mathbb{E}\Big[\Xi_T\int_0^T \Xi_t^{-1}\Big(\Big((\operatorname{trace}\hat{\nabla}A_0)(X_t)h_t - \dot{h}_t\Big)A(X_t)\,dB_t + 2h_tA_0^A\,dt\Big)\Big|X_T(x) = y\Big]\Big\rangle\end{aligned}$$

for all $v \in T_yM$ where the various terms appearing in the right-hand side can be calculated as in Corollary 2.9.

Proof. Since Corollary 2.9 holds for all smooth functions f and vector fields V of compact support, and since by Lemma 2.6

$$f\hat{\delta}(V^\flat) - \hat{\delta}(fV^\flat) = V(f) = f\delta(V^\flat) - \delta(fV^\flat),$$

the result follows from equation (2.13), Lemma 2.4 and Corollary 2.9. $\square$

Example 2.11. Consider the special case $M = \mathbb{R}^n$. Denote by $q_T(x,y)$ the smooth density of $X_T(x)$ with respect to the standard n-dimensional Lebesgue measure. Recall that $p_T(x,y)$ denotes the density with respect to the induced Riemannian measure. It follows that

$$q_T(x,y) = p_T(x,y)\rho^{1/2}(y)$$

where $\rho(y)$ denotes the absolute value of the determinant of the matrix

$$\{\langle A^*\partial_i, A^*\partial_j\rangle_{\mathbb{R}^m}(y)\}_{i,j=1}^n$$

in which $\{\partial_i\}_{i=1}^n$ denotes the standard basis of vector fields on $\mathbb{R}^n$. Consequently

$$\left(d\log q_T(x,\cdot)\right)_y(v) = \left(d\log p_T(x,\cdot)\right)_y(v) + \left(d\log\rho^{1/2}(\cdot)\right)_y(v)$$

with the first term on the right-hand side given, in terms of the induced metric, by Corollary 2.10.

Acknowledgment. This work has been supported by the Fonds National de la Recherche Luxembourg (FNR) under the OPEN scheme (project GEOMREV O14/7628746).

References

1. Jean-Michel Bismut, *Large deviations and the Malliavin calculus*, Progress in Mathematics, vol. 45, Birkhäuser Boston, Inc., Boston, MA, 1984. MR 755001
2. Ana Bela Cruzeiro and Xicheng Zhang, *Bismut type formulae for diffusion semigroups on Riemannian manifolds*, Potential Anal. **25** (2006), no. 2, 121–130. MR 2238939
3. Bruce K. Driver and Anton Thalmaier, *Heat equation derivative formulas for vector bundles*, J. Funct. Anal. **183** (2001), no. 1, 42–108. MR 1837533
4. K. D. Elworthy, Y. Le Jan, and Xue-Mei Li, *On the geometry of diffusion operators and stochastic flows*, Lecture Notes in Mathematics, vol. 1720, Springer-Verlag, Berlin, 1999. MR 1735806
5. K. D. Elworthy, Yves Le Jan, and Xue-Mei Li, *Concerning the geometry of stochastic differential equations and stochastic flows*, New trends in stochastic analysis (Charingworth, 1994), World Sci. Publ., River Edge, NJ, 1997, pp. 107–131. MR 1654511
6. K. D. Elworthy and X.-M. Li, *Formulae for the derivatives of heat semigroups*, J. Funct. Anal. **125** (1994), no. 1, 252–286. MR 1297021
7. K. David Elworthy, Yves Le Jan, and Xue-Mei Li, *The geometry of filtering*, Frontiers in Mathematics, Birkhäuser Verlag, Basel, 2010. MR 2731662
8. Klaus-Jochen Engel and Rainer Nagel, *One-parameter semigroups for linear evolution equations*, Graduate Texts in Mathematics, vol. 194, Springer-Verlag, New York, 2000, With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt. MR 1721989
9. Shoshichi Kobayashi and Katsumi Nomizu, *Foundations of differential geometry. Vol I*, Interscience Publishers, a division of John Wiley & Sons, New York-London, 1963. MR 0152974
10. Hiroshi Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990. MR 1070361
11. Mikio Nakahara, *Geometry, topology and physics*, second ed., Graduate Student Series in Physics, Institute of Physics, Bristol, 2003. MR 2001829
12. Ichiro Shigekawa, *On stochastic horizontal lifts*, Z. Wahrsch. Verw. Gebiete **59** (1982), no. 2, 211–221. MR 650613
13. Daniel W. Stroock, *An introduction to the analysis of paths on a Riemannian manifold*, Mathematical Surveys and Monographs, vol. 74, American Mathematical Society, Providence, RI, 2000. MR 1715265
14. Anton Thalmaier, *On the differentiation of heat semigroups and Poisson integrals*, Stochastics Stochastics Rep. **61** (1997), no. 3-4, 297–321. MR 1488139
15. Anton Thalmaier and Feng-Yu Wang, *Gradient estimates for harmonic functions on regular domains in Riemannian manifolds*, J. Funct. Anal. **155** (1998), no. 1, 109–124. MR 1622800
16. ———, *A stochastic approach to a priori estimates and Liouville theorems for harmonic maps*, Bull. Sci. Math. **135** (2011), no. 6-7, 816–843. MR 2838103
17. Feng-Yu Wang, *Analysis for diffusion processes on Riemannian manifolds*, Advanced Series on Statistical Science & Applied Probability, vol. 18, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2014. MR 3154951
18. ———, *Integration by parts formula and shift Harnack inequality for stochastic equations*, Ann. Probab. **42** (2014), no. 3, 994–1019. MR 3189064