Intertwined Diffusions by Examples

Xue-Mei Li

Mathematics Institute, the University of Warwick, U.K.

Abstract

We discuss the geometry induced by pairs of diffusion operators on two states spaces related by a map from one space to the other. This geometry led us to an intrinsic point of view on filtering. This will be explained plainly by examples, in local coordinates and in the metric setting. This article draws largely from the books [11, 13] and aim to have a comprehensive account of the geometry for a general audience.

1 Introduction

Let p be a differentiable map from a manifold N to M which intertwines a diffusion operator $\mathcal{B}$ on N with another diffusion operator, $\mathcal{A}$ on M, that is $(\mathcal{A}f)\circ p = \mathcal{B}(f\circ p)$ for a given function f from M to **R**. Suppose that $\mathcal{A}$ is elliptic and f is smooth. It is stated in [13] this intertwining pair of operators determine a unique horizontal lifting map $\mathfrak{h}$ from TM to TN which is induced by the symbols of $\mathcal{A}$ and $\mathcal{B}$ and the image of the lifting map determines a subspace of the tangent space to N and is called the associated horizontal tangent space and denoted by H. The condition that $\mathcal{A}$ is elliptic can be replaced by cohesiveness, that is, the symbol $\sigma^{\mathcal{A}}: T_x^*M \to T_xM$ has constant non-zero rank and $\mathcal{A}$ is along the image of $\sigma^{\mathcal{A}}$. If $\mathcal{A}$ is of the form, $\mathcal{A} = \frac{1}{2}\sum_{i=1}^m L_{X^i}L_{X^i} + L_{X^0}$, it is cohesive if span$\{X^1(x), \dots, X^m(x)\}$ are of constant rank and contains $X^0(x)$. Throughout this paper we assume that $\mathcal{A}$ is cohesive.

For simplicity we assume that the $\mathcal{B}$-diffusion does not explode. The pair of intertwining operators induces the splitting of TN in the case that $\mathcal{B}$ is elliptic or the splitting of $Tp^{-1}[Im(\sigma^{\mathcal{A}})] = \text{ker}(T_u p)\oplus H_u$. Hence a diffusion operator $\mathcal{A}$ in Hörmander form has a horizontal lift $\mathcal{A}^H$, operator on N, through the horizontal lift of the defining corresponding vector fields. For operators not in Hörmander form an intrinsic definition of horizontal lift can also be defined by the lift of its symbols and another associated operator δ^A from the space of differential forms to the space of functions and such that $\delta^{\mathcal{A}}(df) = \mathcal{A}f$. In this case the diffusion operator $\mathcal{B}$ splits and $\mathcal{B} = \mathcal{A}^H + B^V$ where B^V acts only on the vertical bundle, which leads to computation of the conditional distribution of the $\mathcal{B}$ diffusion given a $\mathcal{A}$ diffusion. We describe this in a number of special cases.

This work was inspired by an observation for gradient stochastic flows. Let

$$dx_t = X(x_t)\circ dB_t + X_0(x_t)dt$$

be a gradient stochastic differential equations (SDEs). As usual (B_t) is an $\mathbf{R}^m$ valued Brownian motion. The bundle map $X : \mathbf{R}^m \times M \to TM$ is induced by an isometric embedding map $f : M \to \mathbf{R}^m$. Define $Y(x) := df(x) : T_xM \to \mathbf{R}^m$ and

$$\langle X(x)e, v\rangle := \langle e, Y(x)(v)\rangle.$$

Then ker $X(x)$ is the normal bundle νM and [ker $X(x)]^\perp$ corresponds to the tangential bundle. It was observed by Itô that the solution is a Brownian motion, that is the infinitesimal generator of the solutions is $\frac{1}{2}\Delta$. It was further developed in [8] that if we choose an orthonormal basis $\{e_i\}$ of $\mathbf{R}^m$ and define the vector fields $X_i(x) = X(x)(e)$ then the SDE now written as

$$dx_t = \sum_{i=1}^m X_i(x_t) \circ dB_t^i + X_0(x_t)dt \tag{1.1}$$

and the Itô correction term $\sum \nabla X^i(X^i)$ vanishes. In [18] this observation is used to prove a Bismut type formula for differential forms related to gradient Brownian flow, in [20] to obtain an effective criterion for strong 1-completeness, and in [19] to obtain moment estimates for the derivative flow $T\xi_t$ of the gradient SDEs. The key observation was that for each i either ∇X_i or X_i vanishes and if $T\xi_t(v)$ is the derivative flow for the SDE, $T\xi_t(v)$ is in fact the derivative in probability of the solution $\xi_t(x)$ at x in the direction v satisfying

$$/\!/_t d(/\!/_t^{-1} v_t) = \sum_{i=1}^m \nabla X_i(v_t) \circ dB_t^i + \nabla X_0(v_t)dt.$$

where $/\!/_t(\sigma) : T_{\sigma_0}M \to T_{\sigma_t}M$ denotes the stochastic parallel translation corresponding to the Levi-Civita connection along a path σ which is defined almost surely for almost all continuous paths. Consider the Girsanov transform

$$B_t \to B_t + \int_0^t \sum_i \frac{\langle \nabla_{v_s} X_i, v_s\rangle_{x_s}}{|v_s|_{x_s}^2} e_i ds$$

and $\int_0^t \sum_i \frac{\langle \nabla_{v_s} X_i, v_s\rangle_{x_s}}{|v_s|_{x_s}^2} e_i ds = \int_0^t \frac{(\nabla_{v_s} X)^*(v_s)}{|v_s|_{x_s}^2} ds$. Let $\tilde{x}_t$ and $\tilde{v}_t$ be the corresponding solutions to the above two SDEs then $\mathbf{E}|v_t|_{x_t}^p = \mathbf{E}|\tilde{v}_t|_{\tilde{x}_t}^p G_t$ where G_t is the Girsanov density. It transpires that the transformation does not change (1.1). Since $|v_t|^p = |v_0|^p G_t e^{a_t^p(x_t)}$, where a_t is a term only depending on x_t not on v_t, see equation (18) in [20] $\mathbf{E}|v_t|^p = \mathbf{E}e^{a_t^p(\tilde{x}_t)} = \mathbf{E}e^{a_t^p(x_t)}$. In summary the exponential martingale term in the formula for $|v_t|^p$ can be considered as the Radon-Nikodym derivative of a new measure given by a Cameron-Martin transformation on the path space and this Cameron-martin transformation has no effect on the x-process.

Letting $\mathcal{F}_s^x = \sigma\{\xi_s(x) : 0 \le s \le t\}$, $\mathbf{E}\{/\!/_t^{-1} v_t | \mathcal{F}_t^x\}$ satisfies [9],

$$\frac{d}{dt} /\!/_t^{-1} W_t = -\frac{1}{2} /\!/_t^{-1}(\sigma)\mathrm{Ric}^\#(W_t)dt$$

where $\mathrm{Ric}_x : T_xM \to T_xM$ is the linear map induced by the Ricci tensor. The process W_t is called damped stochastic parallel translation and this observation allows us to give pointwise bounds on the conditional expectation of the derivative flow. Together with an intertwining formula

$$dP_t(v) = \mathbf{E} df(T\xi_t(v)),$$

this gives an intrinsic probabilistic representation for $dP_t f = \mathbf{E} df(W_t)$, and leads to

$$|\nabla P_t f|(x) \le |P_t(\nabla f)|_{L^p}(x)(\mathbf{E}|W_t|^q)^{\frac{1}{q}}(x)$$

and $\nabla |P_t f|(x) \le |df|_{L^\infty} \mathbf{E}|W_t|(x)$ which in the case of the Ricci curvature is bounded below by a positive constant leads to:

$$|\nabla P_t|(x) \le e^{-Ct} |P_t(\nabla f)|_{L^p}(x)$$

and

$$\nabla |P_t f|(x) \le |df|_{L^\infty} e^{-Ct}$$

respectively.

If the Ricci curvature is bounded below by a function ρ, one has the following pointwise bound on the derivative of the heat semigroup:

$$|\nabla P_t|(x) \le |P_t(\nabla f)|_{L^p}(x) \left(\mathbf{E} e^{-q/2 \int_0^t \rho(x_s) ds}\right)^{\frac{1}{q}}.$$

See [21] for an application, and [6], [22], [7], [24] for interesting work associated with differentiation of heat semi-groups.

It turns out that the discussion for the gradient SDE are not particular to the gradient system. Given a cohesive operator the same consideration works provided that the linear connection, equivalently stochastic parallel transport or horizontal lifting map, we use is the correct one.

To put the gradient SDE into context we introduce a diffusion generator on GLM, the general linear frame bundle of M. Let η_t^i be the partial flow of X_i and let X_i^G be the vector field corresponding to the flow $\{T\eta_t^i(u) : u \in GLM\}$. Let

$$\mathcal{B} = \frac{1}{2} \sum L_{X_i^G} L_{X_i^G} + L_{X_0^G}.$$

Then $\mathcal{B}$ is over $\mathcal{A}$, the generator of SDE (1.1). The symbol of $\mathcal{A}$ is $X^*(x)X(x)$ and likewise there is a similar formulation for $\sigma^{\mathcal{B}}$ and $h_u = X^{GL}(u)Y(\pi(u))$ where $Y(x)$ is the partial inverse of $X(x)$ and $X^G(e) = \sum X_i^G \langle e, e_i \rangle$.

Open Question. Let W_d be the Wasserstein distance on the space of probability measures on M associated to the Riemannian distance function, show that if

$$W_d(P_t^* \mu, P_t^* \nu) \le e^{ct} W_d(\mu, \nu)$$

the same inequality holds true for the Riemannian covering space of M. Note that if this inequality is obtained by an estimate through lower bound on the Ricci curvature, the same inequality holds on the universal covering space. It would be interesting to see a direct transfer of the inequality from one space to the other. On the other hand if e^{cT} is replaced by Ce^{ct} we do not expect the same conclusion.

2 Horizontal lift of vectors and operators

Let $p : N \to M$ be a smooth map and $\mathcal{B}$, $\mathcal{A}$ intertwining diffusions, that is

$$\mathcal{B}(f \circ p) = \mathcal{A}f \circ p$$

for all smooth function $f : M \to \mathbf{R}$, with semi-groups Q_t and P_t respectively. Instead of intertwining we also say that $\mathcal{B}$ is over $\mathcal{A}$.

Note that for some authors intertwining may refer to a more general concept for operators: $\mathcal{A}D = D(\mathcal{B}+k)$, where k is a constant and D an oprator. For example if Δ^q is the Laplace-Beltrami operator on differential q-forms over a Riemannian manifold $d\Delta = \Delta^1 d$. The usefulness of such relation comes largely from the relation between their respective eigenfunctions. For h a smooth positive function, the following relation $(\Delta - 2L_{\nabla h})(e^h) = e^h(\Delta + V)$ relates to h-transform and links a diffusion operator $\Delta - 2L_{\nabla h}$ with $L+V$ for a suitable potential function V. See [1] for further discussion.

It follows that

$$\frac{\partial}{\partial t}(P_t f \circ p) = \frac{\partial}{\partial t}(P_t f) \circ p = \mathcal{A}(P_t f) \circ p = \mathcal{B}(P_t f \circ p).$$

Since $P_t f \circ p = f \circ p$ at $t = 0$ and $P_t f \circ p$ solves $\frac{\partial}{\partial t} = \mathcal{B}$, we have the intertwining relation of semi-groups:

$$P_t f \circ p = Q_t(f \circ p). \tag{2.1}$$

The $\mathcal{B}$ diffusion u_t is seen to satisfy Dynkin's criterion for $p(u_t)$ to be a Markov process. This intertwining of semi-groups has come up in other context. The relation $VP_t = Q_t V$, where V is a Markov kernel from N to M, is relates to this one when a choice of an inverse to p is made. For example take M to be a smooth Riemannian manifold and N the orthonormal frame bundle. One would fix a frame for each point in M. Note that the law of the Horizontal Brownian motion has been shown to be the law of the Brownian motion and independent of the initial frame [8].

The symbol of an operator $\mathcal{L}$ on a manifold M is a map from $T^*M \times T^*M \to \mathbf{R}$ such that for $f, g : M \to \mathbf{R}$,

$$\sigma^{\mathcal{L}}(df, dg) = \frac{1}{2}\left[\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f\right].$$

If $\mathcal{L} = \frac{1}{2}a_{ij}\frac{\partial^2}{\partial x_j \partial x_j} + b_k \frac{\partial}{\partial x_k}$ is an elliptic operator on $\mathbf{R}^n$ its symbol is (a_{ij}) which induces a Riemannian metric $(g_{ij}) = (a_{ij})^{-1}$ on $\mathbf{R}^n$.

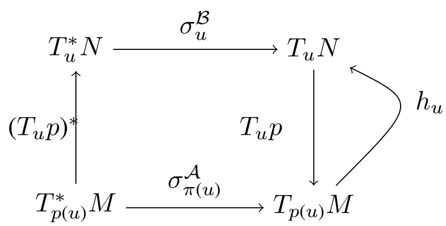
For the intertwining diffusions: $p^*\sigma^{\mathcal{B}} = \sigma^A$, or

$$T_u p \circ \sigma^{\mathcal{B}}_u((T_u p)^*) = \sigma^{\mathcal{A}}_{p(u)},$$

if the symbols are considered as linear maps from the cotangent to the tangent spaces. We stress again that throughout this article we assume that $\mathcal{A}$ is **cohesive**, i.e. $\sigma^{\mathcal{A}}$ has constant rank and $\mathcal{A}$ is along the distribution $E = Im[\sigma^{\mathcal{A}}]$.

There is a unique horizontal lifting map [13] such that, now with the symbols considered as linear maps from the cotangent space to the tangent spaces

$$\mathfrak{h}_u \circ \sigma^{\mathcal{A}}_{p(u)} = \sigma^{\mathcal{B}}_u (T_u p)^*.$$



Let H_u to be the image of $\mathfrak{h}_u$, called the **horizontal distribution**. They consists of image of differential forms of the form $\phi(Tp-)$ for $\phi \in T^*M$ by $\sigma^{\mathcal{B}}$. Note that this cannot be reduced to the case of $\mathcal{A}$ being elliptic because E_x may not give rise to a submanifold of M.

If an operator $\mathcal{L}$ has the Hörmander form representation

$$\mathcal{L} = \frac{1}{2}\sum_{j=1}^{m} L_{X^j}L_{X^j} + L_{X^0} \tag{2.2}$$

Define $X(x) : \mathbf{R}^m \to T_xM$ by $X(x) = \sum X^i(x)e_i$ for $\{e_i\}$ an orthonormal basis of $\mathbf{R}^m$. Then

$$\sigma_x^{\mathcal{L}} = \frac{1}{2}X(x)X(x)^* : T_x^*M \to T_xM.$$

In the elliptic case, $\sigma^{\mathcal{L}}$ induces a Riemannian metric and $X^*(x)\phi = Y(x)\phi^{\#}$.

An operator $\mathcal{L}$ is **along a distribution** $S := \{S_x : x \in M\}$, where each S_x is a subspace of T_xM, if $\mathcal{L}\psi = 0$ whenever $\psi_x(S_x) = \{0\}$. The horizontal lifts of tangent vectors induce a horizontal lift of the operator which is denoted as $\mathcal{A}^H$. To define a horizontal lift of a diffusion operator intrinsically, we introduced an operator $\delta^{\mathcal{L}}$. If M is endowed with a Riemannian metric let $\mathcal{L} = \Delta$ be the Laplace-Beltrami operator, this is d^*, the L^2 adjoint of d the differential operator d. Then $d^*(f\phi) = fd^*\phi + \iota_{\nabla f}(\phi)$ for ϕ a differential 1-form and f a function, using the Riemannian metric to define the gradient operator, and $\Delta = d^*d$.

For a general diffusion operator it was shown in [13] that there is a unique linear operator $\delta^{\mathcal{L}} : C^{r+1}T^*M \to C^r(M)$ determined by $\delta^{\mathcal{L}}(df) = \mathcal{L}f$ and $\delta^{\mathcal{L}}(f\phi) = f\delta^{\mathcal{L}}(\phi) + df\sigma^{\mathcal{L}}(\phi)$. If $\mathcal{L}$ has the representation (2.2),

$$\delta^{\mathcal{L}} = \frac{1}{2}\sum_{j=1}^{m} L_{X^j}\iota_{X^j} + \iota_{X^0}.$$

Here ι is the interior product, $\iota_v\phi := \phi(v)$. The symbol of the operator now plays the role of the Riemannian metric. For $\mathcal{B}$ over $\mathcal{A}$,

$$\delta^{\mathcal{B}}(p^*(df)) = p^*(\delta^{\mathcal{A}}df).$$

There are many operators over $\mathcal{A}$ and only one of which, $\mathcal{A}^H$, is horizontal. An operator $\mathcal{L}$ is **horizontal (respectively vertical)** if it is along the horizontal or the vertical distribution. An operator $\mathcal{B}$ is vertical if and only if $\mathcal{B}(f \circ p) = 0$ for all f and $\mathcal{B} - \mathcal{A}^H$ is a vertical operator.

The foundation of the noise decomposition theorem in [13] depends on the following decomposition of operator $\mathcal{B}$, when $\mathcal{A}$ is cohesive,

$$\mathcal{B} = \mathcal{A}^H + (\mathcal{B} - \mathcal{A}^H) \tag{2.3}$$

and it can be proven that $\mathcal{B} - \mathcal{A}^H$ is a vertical operator.

2.1 In metric form

Note that $\sigma^{\mathcal{A}}$ gives rise to a positive definite bilinear form on T^*M:

$$\langle \phi, \psi \rangle_x = \phi(x)(\sigma^{\mathcal{A}}_x(\psi(x)))$$

and this induces an inner product on E_x:

$$\langle u, v \rangle_x = (\sigma^{\mathcal{A}}_x)^{-1}(u)(v).$$

For an orthonormal basis $\{e_i\}$ of E_x, let $e_i^* = (\sigma^{\mathcal{A}}_x)^{-1}(e_i)$. Then $e_j^* \sigma^{\mathcal{A}}(e_i^*) = (\sigma^{\mathcal{A}}_x)^{-1}(e_j)(e_i) = \langle e_j, e_i \rangle$ and hence

$$\langle \phi, \psi \rangle_x = \sum_i \langle e_j, e_i \rangle \phi(e_i)\psi(e_j) = \sum_i \phi(e_i)\psi(e_i).$$

Likewise the symbol $\sigma^{\mathcal{A}^H}$ induces an inner product on T^*N with the property that $\langle \phi \circ Tp, \psi \circ Tp \rangle = \langle \phi, \psi \rangle$ and a metric on $H \subset TN$ which is the same as that induced by $\mathfrak{h}$ from TM. Note that $\sigma^{\mathcal{B}} = \sigma^{\mathcal{A}^H} + \sigma^{\mathcal{B}^V}$, where $\mathcal{B}^V$ is the vertical part of $\mathcal{B}$, and $Im[\sigma^{\mathcal{B}^V}] \cap H = \{0\}$. Let μ be an invariant measure for $\mathcal{A}^H$ and $\mu_M = p_*(\mu)$ the pushed forward measure which is an invariant measure for $\mathcal{A}$.

If $\mathcal{A}$ is symmetric,

$$\begin{aligned} \int_M \langle df, dg \rangle \; \mu_M(dx) &= \int \sigma^{\mathcal{A}}(df, dg) \; \mu_M(dx) \\ &= \frac{1}{2} \int [\mathcal{A}(fg) - f(\mathcal{A}g) - g(\mathcal{A}f)] \; \mu_M(dx) \\ &= -\int_M f\mathcal{A}g \; d\mu_M(x). \end{aligned}$$

Hence $\mathcal{A} = -d^*d$ and

$$\delta^{\mathcal{A}} = -d^*$$

for d^* the L^2 adjoint. Similarly we have an L^2 adjoint on N and $\mathcal{A}^H = -d^*d$. For a 1-form ϕ on M,

$$\int_N \langle \phi \circ Tp, d(g \circ p) \rangle d\mu_N = \int \langle d^*(\phi \circ Tp), g \circ p \rangle d\mu_N = \int \langle \mathbf{E}\{d^*(\phi \circ Tp)|p\}, g \circ p \rangle d\mu_N$$

Hence $\mathbf{E}\{d^*(\phi \circ Tp)|p\} = (d^*\phi) \circ p$. Since for $u+v \in H \oplus \text{ker}[Tp]$, $h \circ Tp(u+v) = u$, every differential form ψ on N induces a form $\phi = \psi \circ \mathfrak{h}$ such that $\psi = \phi(T\pi)$ when restricted to H, hence $\mathbf{E}\{d^*\psi)|p\} = (d^*(\psi \circ h)) \circ p$.

2.2 On the Heisenberg group

A Lie group is a group G with a manifold structure such that the group multiplication $G \times G \to G$ and taking inverse are smooth. Its tangent space at the identity $\mathfrak{g}$ can be identified with left invariant vector fields on G, $X(a) = TL_a X(e)$ and we denote A^* the left invariant vector field with value A at the identity. The tangent space T_aG at a can be identified with $\mathfrak{g}$ by the derivative TL_a of the left translation map. Let $\alpha_t = \exp(tA)$ be the solution flow to the left invariant vector field TL_aA whose value at 0 is the identity then it is also the flow for the corresponding right invariant vector field: $\dot\alpha_s = \frac{d}{dt}|_{t=s} \exp^{(t-s)A}\exp^{sA} = TR_{\alpha_s}A$. Then $u_t = a\exp(tA)$ is the solution flow through a.

Consider the Heisenberg group G whose elements are $(x, y, z) \in \mathbf{R}^3$ with group product

$$(x_1, y_1, z_1)(x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2 + \frac{1}{2}(x_1y_2 - x_2y_1).$$

The Lie bracket operation is $[(a, b, c),(a', b', c')] = (0, 0, ab' - a'b)$. Note that for $X, Y \in \mathfrak{g}$, $e^X e^Y = e^{X+Y+\frac{1}{2}[X,Y]}$. If $A = (a, b, c)$, then $A^* = (a, b, c + \frac{1}{2}(xb - ya))$. Consider the projection $\pi : G \to \mathbf{R}^2$ where $\pi(x, y, z) = (x, y)$. Let

$$X_1(x, y, z) = (1, 0, -\frac{1}{2}y), \quad X_2(x, y, z) = (0, 1, \frac{1}{2}x), \qquad X_3(x, y, z) = (0, 0, -1)$$

be the left invariant vector fields corresponding to the standard basis of $\mathfrak{g}$. The vector spaces $H_{(x,y,z)} = \text{span}\{X_1, X_2\} = \{(a, b, \frac{1}{2}(xb - ya))\}$ are of rank 2. They are the horizontal tangent spaces associated to the Laplacian $\mathcal{A} = \frac{1}{2}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})$ on $\mathbf{R}^2$ and the left invariant Laplacian $\mathcal{B} := \frac{1}{2}\sum_{i=1}^3 L_{X_i}L_{X_i}$ on G. The vertical tangent space is $\{(0, 0, c)\}$ and there is a a horizontal lifting map from $T_{(x,y)}\mathbf{R}^2$:

$$h_{(x,y,z)}(a, b) = (a, b, \frac{1}{2}(xb - ya)).$$

The horizontal lift of $\mathcal{A}$ is the hypo-elliptic diffusion operator $\mathcal{A}^H = \frac{1}{2}\sum_{i=1}^2 L_{X_i}L_{X_i}$ and the horizontal lift of a 2-dimensional Brownian motion, the horizontal Brownian motion, has its third component the Levy area. In fact for almost surely all continuous path $\sigma : [0, T] \to M$ with $\sigma(0) = 0$ we have the horizontal lift curve :

$$\tilde\sigma(t) = \left(\sigma^1(t), \sigma^2(t), \frac{1}{2}\int_0^t \left(\sigma^1(t) \circ d\sigma^2(t) - \sigma^2(t) \circ d\sigma^1(t)\right)\right).$$

The hypoelliptic semi-group Q_t in $\mathbf{R}^3$ and the heat semigroup P_t satisfies $Q_t(f \circ \pi) = e^{\frac{1}{2}t\Delta}f \circ \pi$ and $d(e^{\frac{1}{2}t\Delta}f) = Q_t(df \circ \pi) \circ \mathfrak{h}$.

2.3 The local coordinate formulation

Let M be a smooth Riemannian manifold and $\pi : P \to M$ a principal bundle with group G acting on the right, of which we are mainly interested in the case when P is the general linear frame bundle of M or the orthonormal frame bundle with G the special general linear group or the special orthogonal group of $\mathbf{R}^n$. For $A \in \mathfrak{g}$, the

Lie algebra of G, the action of the one parameter group $\exp(tA)$ on P induces the fundamental vector field A^* on P. Let VTP be the vertical tangent bundle consisting of tangent vectors in the kernel of the projection $T\pi$ so the fundamental vector fields are tangent to the fibres and $A \mapsto A^*(u)$ is a linear isomorphism from $\mathfrak{g}$ to VT_uP. At each point a complementary space, called the horizontal space, can be assigned in a right invariant way: $HT_{ua}P = (R_a)_* HT_uP$.

For the general linear group $GL(n)$ its Lie algebra is the vector space of all n by n matrices and the value at a of the left invariant vector field A^* is aA. The Lie bracket is just the matrix commutator, $[A,B] = AB - BA$. Every finite dimensional Lie group is homomorphic to a matrix Lie group by the adjoint map. For $a \in G$, the tangent map to the conjugation $\phi : g \in G \mapsto aga^{-1} \in G$ induces the adjoint representation $\text{ad}(a) : G \to GL(\mathfrak{g};\mathfrak{g})$. For $X \in \mathfrak{g}$, $\phi_* X^*(g) = TL_a TR_{a^{-1}}(X(a^{-1}ga)) = TR_{a^{-1}}X(ga) = (R_{a^{-1}})_* X(g)$ and is left invariant so $\text{ad}(a)(A) = TR_{a^{-1}}X^*(a)$. The Lie bracket of two left invariant vector fields $[X^*, Y^*] = \lim_{t\to 0} \frac{1}{t}(\exp(tY)_* X^* - X^*) = \lim_{t\to 0} \frac{1}{t}(R_{e^{tY})})_* X^* - X^*)$ is again a left invariant vector field and this defines a Lie bracket on $\mathfrak{g}$ by $[X,Y]^* = [X^*, Y^*]$. The Lie algebra homomorphism induced by $a \mapsto ad(a)$ is denoted by $\text{Ad} : \mathfrak{g} \to \mathfrak{g}l(n, \mathbf{R})$ is given by $\text{Ad}_X(Y) = [X, Y]$. A tangent vector at $a \in G$ can be represented in a number of different ways, notably by the curves of the form $a\exp(tA)$, $\exp(tB)a$. The Lie algebra elements are related by $B = aAa^{-1} = \text{ad}(a)A$ and $\frac{d}{dt}|_{t=0} A\exp(TB)A^{-1} = \text{ad}(A)B$ so $A\exp(tB)A^{-1} = \exp(t\text{ad}(A)B)$. The left invariant vector fields provides a parallelism of TG and there is a canonical left invariant 1-form on G, $\omega_g(TL_g(v)) = G_e(v)$, determined by $\theta(A^*) = A$.

The collection of left invariant vector fields on TP forms also an algebra and the map $A \to A^*$ is a Lie-algebra isomorphism. A horizontal subspace of the tangent space to the principal bundle TP is determined by the kernel of a connection 1-form ω, which is a $\mathfrak{g}$-value differential 1-form on P such that (i) $\omega(A^*) = A$, for all $A \in \mathfrak{g}$, and (2) $(R_a)^*\omega = ad(a^{-1})\omega(-)$. Here A^* refers to the TP valued left invariant vector field. The first condition means that the connection 1-form restricts to an isomorpism from VTP to $\mathfrak{g}$ and the second is a compatibility condition following from that the fundamental vector field corresponding to $ad(a^{-1})A$ is $(R_a)_* A^*$. The kernel of ω is right invariant since $\omega_{ua}(TR_aV) = (R_a)^*\omega(V) = ad(a)\,\omega_u(V)$ for any $V \in T_uP$.

In a local chart $\pi^{-1}(U)$ with U an open set of M and $u \in \pi^{-1}(U) \to (\pi(u), \phi(u))$ the chart map where $\phi(ua) = \phi(u)a$, the connection map satisfies $\omega_{(x,a)}(0, B^*) = B$ for B^* the left invariant vector field of G corresponding to $B \in \mathfrak{g}$ and

$$\omega_{(x,a)}(v, B^*(a)) = ad(a^{-1})(M_x v) + B$$

where $M_x : T_xM \to \mathfrak{g}$ is a linear map varying smoothly with x. The trivial connection for a product manifold $M \times G$ would correspond to a choice of M_x with M_x identically zero and so the horizontal vectors are of the form $(v, 0)$. The horizontal tangent space at (x, a) is the linear space generated by

$$H_{(x,a)}TP = \{(v, -TR_a(M_x v)), \qquad v \in T_xU, a \in G\}.$$

Given a connection on P, for every differentiable path σ_t on M, through each frame u_0 over σ_0 there is a unique u_t which projects down to σ_t on M given by $\omega(\dot{u}_t) = 0$.

In local coordinates $u_t = (\sigma_t, g_t)$, $\omega(\dot{u}_t) = \mathrm{ad}(g_t^{-1})\omega_{(\sigma_t,e)}(\dot{\sigma}_t, TR_{g_t}^{-1}\dot{g}_t)$ and $g_t{}^{-1}\dot{g}_t + ad(g_t^{-1})M_{\sigma_t}\dot{\sigma}_t = 0$. If u_t is a lift of x_t then $u_t \circ g$ is the horizontal lift of x_t through $u_0 g$ so $u_t : \pi^{-1}(\sigma_0) \to \pi^{-1}(\sigma_t)$ is an isomorphism. This formulation works for continuous paths. Consider the path of continuous paths over M and a Brownian motion measure. For almost surely all continuous paths σ_t a horizontal curve exists, as solution to the stochastic differential equation in Stratnovitch form:

$$dg_t = -M_{\sigma_t}(e_i)(g_t) \circ d\sigma_t^i.$$

Here (e_i) is an orthonormal basis of $\mathbf{R}^n$, and the $M_\cdot(e_i)$'s are matrices in $\mathfrak{g}$ and the solution u_t induces a transformation from the fibre at σ_0 to the fibre at σ_t.

2.4 The orthonormal frame bundle

Let $N = OM$ be the orthonormal frame bundle with π the natural projection to a Riemannian manifold M and an right invariant Riemannian metric. Let $\mathcal{A} = \Delta$ be the Laplacian on M and $\mathcal{B}$ the Laplacian on N. We may choose the Laplacian $\mathcal{B}$ to be of the form $\frac{1}{2}L_{A_i^*}L_{A_i^*} + \frac{1}{2}L_{H_i}L_{H_i}$ where A_i are fundamental vector fields and $\{H_i\}$ the standard horizontal vector fields. The horizontal lifting map $\mathfrak{h}_u$ is: $v \in TM \mapsto (v, 0)$. We mention two Hörmander form representation for the horizontal lift. The first one consists of of horizontal lifts of vector fields that defines $\mathcal{A}$. The second one is more canonical. Let $\{B(e), e \in \mathbf{R}^n\}$ be the standard horizontal vector fields on OM determined by $\theta(B(e)) = e$ where θ is the canonical form of OM, that is $T\pi[B(e)(u)] = u(e)$. Take an orthonormal basis of $\mathbf{R}^n$ and obtaining never vanishing vector fields $H_i =: B(e_i)$, then $\mathcal{A}^H = \sum L_{H_i}L_{H_i}$ and $\mathcal{A}^H$ is called the horizontal Laplacian. The two heat semigroups Q_t, upstairs, and P_t intertwine: $Q_t(f \circ \pi) = P_t f \circ \pi$. Let us observe that if Q_t^H is the semigroup corresponding to horizontal Laplacian $\mathcal{A}^H$, since $dQ_t f \circ T\pi$ annihilates the vertical bundle, $Q_t^H(f \circ \pi) = Q_t(f \circ \pi)$ and Q_t^H restricts to a semigroup on the set of bounded measurable functions of the form $f \circ \pi$.

Denote by the semi-group corresponding to the Laplace-Beltrami operators by the same letters with the supsctipt one indicates the semi-group on 1-forms, then $dP_t f = P_t^1 d$ and $dQ_t = Q_t^1 d$, which follows from that the exterior differentiation d and the Laplace-Beltrami operator commute. Now

$$d(P_t f \circ \pi) = d(P_t f) \circ T\pi = P_t^1(df) \circ T\pi$$

Similarly $d(Q_t(f \circ \pi)) = Q_t^1(df \circ T\pi)$. Now we represent Q_t by the horizontal diffusion which does not satisfy the commutation relation: $d\mathcal{A}^H \neq \mathcal{A}^H d$ in general. Let $\bar{W}_t$ be the solution to a differential equation involving the Weitzenbock curvature operator, see Proposition 3.4.5 in [13], $\bar{W}_t /\!/_t = W_t$ where $\frac{d}{dt}\bar{W}_t = -\frac{1}{2}u_t^{-1}\,\mathrm{Ric}^\#(u_t\bar{W}_t)$.

$$d(Q_t^H f \circ \pi)(\mathfrak{h}v) = d(P_t f)(v) = \mathbf{E}df(\bar{W}_t u_t \circ u_0^{-1}(v)),$$

the formula as we explained in the introduction, after conditioning the derivative flow. Note also that $d(Q_t^H(f \circ \pi)) = d(P_t f \circ \pi) = dP_t f \circ T\pi$ and

$$d(P_t f)(-) = Q_t(df \circ T\pi)(\mathfrak{h}-).$$

3 Examples

3.1 Diffusions on the Euclidean Space

Take the example that $N=\mathbf{R}^2$ and $M=\mathbf{R}$. Any elliptic diffusion operators on M is of the form $a(x)\frac{d^2}{dx^2}$ and a diffusion operator on N is of the form $\mathcal{B}=a(x,y)\frac{d^2}{dx^2}+d(x,y)\frac{d^2}{dy^2}+c(x,y)\frac{d^2}{dxdy}$ with $4ad>c^2$ and $a>0$. Now $\mathcal{B}$ is over $\mathcal{A}$ implies that $a(x,y)=a(x)$ for all y. If a,b,c are constants, a change of variable of the form $x=u$ and $y=(c/2\sqrt{a})u+v$ transforms $\mathcal{B}$ to $a^2\frac{\partial^2}{\partial u^2}+(d-c^2/4a)\frac{\partial^2}{\partial v^2}$. In this local coordinates $\mathcal{B}$ and $\mathcal{A}$ have a trivial projective relation. In general we may seek a diffeomorphism $\Phi:(x,y)\mapsto(u,v)$ so that Φ intertwines $\mathcal{B}$ and $\tilde{B}$ where $\tilde{B}$ is the sum of $a^2\frac{\partial^2}{\partial u^2}$ and an operator of the form $\frac{\partial^2}{\partial v^2}$. This calculation is quite messy. However according to the theory in [13], the horizontal lifting map

$$v\mapsto\sigma^{\mathcal{B}}(Tp)^*(\sigma^{\mathcal{A}})^{-1}(v)=\sigma^{\mathcal{B}}(\frac{v}{a},0)^T=\begin{pmatrix}a&\frac{c}{2}\\\frac{c}{2}&d\end{pmatrix}\begin{pmatrix}\frac{v}{a}\\0\end{pmatrix}=(v,\frac{c}{2a}v).$$

where $p:(x,y)\to x$ and Tp is the derivative map and $(Tp)^*$ is the corresponding adjoint map. Hence the lifting of $\mathcal{A}$, as the square of the lifting $\sqrt{a}\frac{d}{dx}$ gives $\sqrt{a}(\frac{d}{dx}+\frac{c}{2a}\frac{d}{dy})$ and resulting the completion of the square procedure and the splitting of $\mathcal{B}$:

$$\mathcal{B}=a\left(\frac{d}{dx}+\frac{c}{2a}\frac{d}{dy}\right)^2+(d-\frac{c^2}{4a})\frac{d^2}{dy^2}.$$

This procedure trivially generalises to multidimensional case $p:\mathbf{R}^{n+p}\to\mathbf{R}^n$ with $p(x,y)=x$. If $\pi:\mathbf{R}^N\to\mathbf{R}^m$ is a surjective smooth map not necessarily of the form $p(x,y)=x$ we may try to find two diffeomorphisms ψ on $\mathbf{R}^N$ and ϕ on $\mathbf{R}^m$ and so that $p=\phi\circ\pi\psi^{-1}$ if of simple form. The diffusion operators $\mathcal{B}$ and $\mathcal{A}$ induce two operators $\tilde{B}$ and $\tilde{\mathcal{A}}$. If $\mathcal{B}$ and $\mathcal{A}$ are intertwining then so are $\tilde{B}$ and $\tilde{A}$. Indeed from

$$\begin{aligned}\tilde{\mathcal{B}}(gp)(y)) &= \mathcal{B}(gp\circ\psi)(\psi^{-1}(y))=\mathcal{B}(g\circ\phi\pi)(\psi^{-1}(y))\\ &= \mathcal{A}(g\circ\phi)(\pi\psi^{-1}(y))=\mathcal{A}(g\circ\phi)(\phi^{-1}p(y))=\tilde{\mathcal{A}}g(p(y)).\end{aligned}$$

This transformation is again not necessary because of the for-mentioned theorem.

In general, [13], if $p:\mathbf{R}^n\times\mathbf{R}\to\mathbf{R}^n$ is the trivial projection and $\mathcal{B}$ is defined by

$$\mathcal{B}g(x,y)=\sigma^{ij}(x)\frac{\partial^2 g}{\partial x_i\partial x_j}+\sum b^k(x,y)\frac{\partial^2 g}{\partial y\partial x_k}+c(x,y)\frac{\partial^2 g}{\partial y^2}$$

with $a=(a_{ij})$ symmetric positive definite and of constant rank, $[b(x,y)]^Tb(x,y)\le c(x,y)a(x)$, there is a horizontal lift induced by $\mathcal{B}$ and $\sigma^{ij}(x)\frac{\partial^2 g}{\partial x_i\partial x_j}$ given by

$$h_{(x,y)}(v)=(v,\langle a(x)^{-1}b,v\rangle).$$

Or even more generally if $p:\mathbf{R}^{m+p+q}\to\mathbf{R}^{m+p}$ with $\mathcal{A}$ a $(m+p)\times(m+p)$ matrix and B a $(m+p)\times q$ matrix and C a $q\times q$ matrix with each column of $B(x,y)$ in the image of $\mathcal{A}$, the horizontal lift map is $h_{(x,y)}(v)=(v,B^T(x,y)A^{-1}v)$.

3.2 The SDE example and the associated connection

Consider SDE (1.1). For each $y \in M$, define the linear map $X(y)(e) : \mathbf{R}^m \to T_yM$ by $X(y)(e) = \sum_{i=1}^m X_i(y)\langle e, e_i\rangle$. Let $Y(y) : T_yM \to [\ker X(y)]^\perp$ be the right inverse to $X(y)$. The symbol of the generator $\mathcal{A}$ is $\sigma_y^{\mathcal{A}} = \frac{1}{2}X(y)X(y)^*$, which induces a Riemannian metric on the manifold in the elliptic case, and a sub-Riemannian metric in the case of $\sigma^{\mathcal{A}}$ being of constant rank .

This map X also induces an affine connection $\breve{\nabla}$, which we called the LW connection, on the tangent bundle which is compatible with the Riemannian metric it induced as below. If $v \in T_{y_0}M$ is a tangent vector and $U \in \Gamma TM$ a vector field,

$$(\breve{\nabla}_v U)(y_0) = X(y_0)D(Y(y)U(y))(v).$$

At each point $y \in M$ the linear map

$$X(y) : \mathbf{R}^m = \ker X(y) \oplus [\ker X(y)]^\perp \to T_yM$$

induces a direct sum decomposition of $\mathbf{R}^m$. The connection defined above is a metric connection with the property that

$$\breve{\nabla}_v X(e) \equiv 0, \qquad \forall e \in [\ker X(y_0)]^\perp, v \in T_{y_0}M.$$

This connection is the adjoint connection by the induced diffusion pair on the general linear frame bundle mentioned earlier. See [11] where it is stated any metric connection on M can be defined through an SDE, using Narasimhan and Ramanan's universal connection.

3.3 The sphere Example

Consider the inclusion $i : S^n \to \mathbf{R}^{n+1}$. The tangent space to T_xS^n for $x \in S^n$ is of the form:

$$T_xS^n = \{v : \langle x, v\rangle = 0\}, \qquad \langle u, v\rangle_x = \langle u, v\rangle_{\mathbf{R}^{n+1}}.$$

Let P_x be the orthogonal projection of $\mathbf{R}^n$ to T_xS^n:

$$P_x : e \in \mathbf{R}^{n+1} \mapsto e - \langle e, x\rangle \frac{x}{|x|^2}.$$

This induces the vector fields $X_i(x) = P_x(e_i)$ and the gradient SDE

$$dx_t = \sum_{i=1}^m P_{x_t}(e_i) \circ dB_t^i.$$

For a vector field $U \in \Gamma TS^n$ on S^n and a tangent vector $v \in T_xS^n$, define the Levi-Civita connection as following:

$$\begin{aligned} \nabla_v U &:= P_x((DU)_x(v)) \\ &= (DU)_x(v) - \langle (DU)_x(v), x\rangle \frac{x}{|x|^2}. \end{aligned}$$

The term

$$\langle (DU)_x(v), x\rangle \frac{x}{|x|^2}$$

is actually tensorial since $\langle (DU)_x(v), x\rangle = \langle U, v\rangle$ and hence defines the Christoffel symbols Γ_{ij}^k, where

$$\nabla_{e_i} e_j = \Gamma_{ij}^k, \qquad (\nabla_v U)^k = D_v U^\nu + \Gamma_{ij}^k v_i u_j.$$

Solution to gradient SDE are BMs since $\nabla_{X_i} X_i = 0$ as observed by Itô. From tensorial property, get Gauss and Weingarten's formula,

$$\begin{array}{lll} (DU)_x(v) & = \nabla_v U + \alpha_x(Z(x), v), & v \in T_xM, U \in \Gamma TM \\ (D\xi)_x(v) & = -A(\xi(x), v) + [(D\xi)_x(v)]^\nu, & \xi \in \nu M \end{array}$$

For $e \in \mathbf{R}^m$, write $e = P_x(e) + e^\nu(x)$ and obtain

$$D_v[P_x(e)] + D_v[e^\nu] = 0.$$

Take the tangential part of all terms in the above equation to see that

$$\text{if } e \in [\ker X(x_0)]^\perp, \quad \nabla_v[P_x(e)] = A(v, e^\nu(x_0)) = 0.$$

3.4 The pairs of SDEs example and decomposition of noise

In general if we have $p : N \to M$ and the bundle maps $\tilde X : N \times \mathbf{R}^m \to TN$ and $X : M \times \mathbf{R}^m \to TM$ are p-related: $Tp\tilde X(u) = X(p(u))$, let $y_t = p(u_t)$ for u_t the solution to

$$du_t = \tilde X(u_t) \circ dB_t + \tilde X_0(u_t)dt.$$

Then y_t satisfies

$$dy_t = X(y_t) \circ dB_t + X_0(y_t)dt.$$

Consider the orthogonal projections at each $y \in M$,

$$\begin{array}{rll} K^\perp(y): & \mathbf{R}^m \to [\ker X(y)]^\perp, & K^\perp(y) := Y(y)X(y) \\ K(y): & \mathbf{R}^m \to \ker[X(y)], & K(y) := I - Y(y)X(y). \end{array}$$

Then

$$dy_t = X(y_t)K^\perp(y_t) \circ dB_t + X_0(y_t)dt \tag{3.1}$$

where the term $K^\perp(y_t) \circ dB_t$ captures the noise in y_t.

To find the conditional law of y_t we express the SDE for u_t use the term $K^\perp(y_t) \circ dB_t$. For a suitable stochastic parallel translation [13] that preserves the splitting of $\mathbf{R}^m$ as the kernel and orthogonal kernel of $X(y)$, define two independent Brownian motions

$$\begin{aligned} B_t^\perp &:= \int_0^s /\!/_t^{-1} K^\perp(p(u_t)) dB_t \\ \beta_s &:= \int_0^s /\!/_t^{-1} K(p(u_t)) \circ dB_t. \end{aligned}$$

Assume now the parallel translation on [ker $X(x)^\perp$ is that given in section 3.2. Since $dx_t = X(x_t)K^\perp(x_t) \circ dB_t + X_0(x_t)dt$, the following filtrations are equal:

$$\sigma\{x_u : 0 \le u \le s\} = \sigma\{B_u^\perp : 0 \le u \le s\}.$$

The horizontal lifting map induced by the pair $(\mathcal{A}, \mathcal{B})$ is given as following:

$$\mathfrak{h}_u(v) = \tilde{X}(u)Y(\pi(u)v), \qquad u \in T_{p(u)}M,$$

From which we obtain the horizontal lift $X^H(u)$ of the bundle map X:

$$X^H(u) = \tilde{X}(u)K^\perp(p(u))$$

and it follows that

$$\begin{aligned} du_t &= \tilde{X}(u_t)K^\perp(p(u_t)) \circ dB_t + \tilde{X}(u_t)K(p(u_t)) \circ dB_t + \tilde{X}_0(u_t)dt \\ &= X^H(u_t) \circ dB_t + \tilde{X}(u_t)K(p(u_t)) \circ dB_t + \tilde{X}_0(u_t)dt \\ &= h_{u_t} \circ dx_t + \tilde{X}(u_t)K(p(u_t)) \circ dB_t + (\tilde{X}_0 - X_0^H)(u_t)dt \end{aligned}$$

If this equation is linear in u_t it is possible to compute the conditional expectation of u_t with respect to $\sigma\{x_u : 0 \le u \le s\}$ as in the derivative flow case (section 2.8 below). This discussion is continued at the end of the article.

3.5 The diffeomorphism group example

If M is a compact smooth manifold and X is smooth we may consider an equation on the space of smooth diffeomorphisms Diff(M). Define $\tilde{X}(f)(x) = X(f(x))$ and $\tilde{X}_0(f)(x) = X_0(f(x))$ and consider the SDE on Diff(M):

$$df_t = \tilde{X}(f_t) \circ dB_t + \tilde{X}_0(f_t)dt$$

with $f_0(x) = x$. Then $f_t(x)$ is solution to $dx_t = X(x_t) \circ dB_t$ with initial point x.

Fix $x_0 \in M$, we have a map $\theta : \text{Diff}(M) \to M$ given by $\theta(f) = f(x_0)$. Let $\mathcal{B} = \frac{1}{2} L_{\tilde{X}_i} L_{\tilde{X}_i}$ and $\mathcal{A} = \frac{1}{2} L_{X_i} L_{X_i}$. Then

$$h_f(v)(x) = \tilde{X}(f)\left(Y(f(x_0))v\right)(x) = X(f(x))(Y(f(x_0))v).$$

3.6 The twist effect

Consider the polar coordinates in $\mathbf{R}^n$, with the origin removed. Consider the conditional expectation of a Brownian motion W_t on $\mathbf{R}^n$ on $|W_t|$ where $|W_t|$, and n-dimensional Bessel Process, $n > 1$, lives in $\mathbf{R}_+$. For $n = 2$ we are in the situation that $p : \mathbf{R}^2 \to \mathbf{R}$ given by $p : (r, \theta) \mapsto r$. The $\mathcal{B}$ and $\mathcal{A}$ diffusion are the Laplacians, $\mathcal{A}^H = \frac{\partial^2}{\partial r^2}$. The map $p(r, \theta) = r^2$ would result the lifting map $v\frac{\partial}{\partial x} \mapsto (\frac{v}{2r}, 0) = \frac{v}{2r}\frac{\partial}{\partial r}$.

At this stage we note that if B_t is a one dimensional Brownian motion, l_t the local time at 0 of B_t and $Y_t = |B_t| + \ell_t$, a 3-dimensional Bessel process starting from 0. There is the following beautiful result of Pitman:

$$E\{f(|B_t|)|\sigma(Y_s : s \le t)\} = \int_0^1 f(xY_t)dx = Vf(Y_t)$$

where V is the Markov kernel: $V(x,dz) = \frac{\mathbf{1}_{0\le z\le x}}{x}dz$ [23, 4].

A second example, [13], which demonstrates the twist effect is on the product space of the circle. Let $p : S^1\times S^1 \to S^1$ be the projection on the first factor. For $0 < \alpha < \frac{\pi}{4}$, define the diffusion operator on $S^1 \times S^1$:

$$\mathcal{B} = \frac{1}{2}(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}) + \tan\alpha \frac{\partial^2}{\partial x \partial y}.$$

and the diffusion operator $\mathcal{A} = \frac{1}{2}\frac{\partial^2}{\partial x^2}$ on S^1. Then

$$\begin{aligned}\mathcal{B}^V &= \frac{1}{2}(1-(\tan\alpha)^2)\frac{\partial^2}{\partial y^2}\\ \mathcal{A}^H &= \frac{1}{2}(\frac{\partial^2}{\partial x^2} + (\tan\alpha)^2\frac{\partial^2}{\partial y^2}) + \tan\alpha\frac{\partial^2}{\partial x\partial y}.\end{aligned}$$

4 Applications

4.1 Parallel Translation

Let $P = GLM$, the space of linear frames on M with an assignment of metrics on the fibres. The connection on P is said to be metric if the parallel translation preserves the metric on the fibres. A connection on P reduces to a connection on the sub-bundle of oriented orthonormal frame bundles OM, i.e. the horizontal lifting belongs to OM if and only if it is metric. Let $F = P \times \mathbf{R}^n/\sim$ be the associated vector bundle determined by the equivalent relation $[u,e] \sim [ug, g^{-1}e]$ hence the vector bundle is $\{ue\}$ where $e \in \mathbf{R}^n, u \in P$. A section of F corresponds to a vector field over M. A parallel translation is induced on TM in the obvious way and given a connection on P let $H(e)$ be the standard horizontal vector field such that $H(e)_u$ is the horizontal lift through u of the vector $u(e)$. If $e \neq 0$, $H(e)$ are never vanishing vector fields such that $TR_a(H(e)) = H(a^{-1}e)$. The fundamental vector fields generated by a basis of $\mathfrak{g}l(n,\mathbf{R})$ and $H(e_i)$ for e_i a basis of $\mathbf{R}^n$ forms a basis of TP at any point and gives a global parallelism on TP.

If we have a curve σ_t with $\sigma_0 = x$ and $\dot\sigma_0 = v$,

$$\nabla_v Y = \lim_{\epsilon\to 0}\frac{1}{\epsilon}[/\!/_\epsilon^{-1}Y(\sigma_\epsilon) - Y(x)].$$

Alternatively $\nabla_X Y(x) = u_0(\tilde{X}f)$ where $\tilde{X}$ is a horizontal lift of X, $f : P \to \mathbf{R}^n$ is defined by $f(u) = u^{-1}[Y(\pi(u))]$ and

$$\tilde{X}f(u_0) = \lim_{h\to 0}\frac{1}{h}(u_h^{-1}Y(\sigma_h) - u_0^{-1}Y(x))$$

for u_h a horizontal lift of x_t starting from u_0. Note that the linear maps $M_x(e)$ which defines the connection form on TP are skew symmetric in the case of $P = OM$, and determines the Christoffel symbols. A vector field Y is horizontal along a curve σ_t if $\nabla_{\dot\sigma(t)}Y = \lim_{h\to 0}\frac{1}{h}[/\!/_h^{-1}Y(\sigma_h) - Y(x)] = 0$. Define the curvature form to be the 2-form $\Omega(-,-) := d\omega(P_h-, P_h-)$ where P_h is the projection to the horizontal space. Then the horizontal part of the Lie bracket of two horizontal vector fields X, Y is the horizontal lift of $[\pi(X), \pi(Y)]$ and its vertical part is determined by $\omega([X,Y]) = -2\Omega(X,Y)$.

The horizontal lift map u_t can also be thought of solutions to:

$$du_t = \sum H(e_i)(u_t) \circ \, d\sigma_t.$$

In fact if $\dot{v}_t$ is the horizontal lift of $\dot{\sigma}_t$, $\dot{v}_t = \sum_{i=1}^n \langle \dot{\sigma}_t, e_i \rangle H(e_i)(\tilde{\sigma}_t)$. Now $/\!/_t(\sigma)$ is not a solution to a Markovian equation, the pair $(/\!/_t(\sigma), u_t)$ is. In local coordinates for v_t^i the ith component of $/\!/_t(\sigma)(v)$, $v \in T_{\sigma_0}M$,

$$dv_t^k = -\Gamma_{i,j}^k(\sigma_t) v_t^j \circ d\sigma_t^i. \tag{4.1}$$

If σ_t is the solution of the SDE $dx_t^k = X_i^k(x_t) \circ dB_t^i + X_0^k(x_t)dt$ then

$$dv_t^k = -\Gamma_{i,j}^k(x_t) v_t^j X_i^k(x_t) \circ dB_t^i - \Gamma_{i,j}^k(x_t) v_t^j X_0^k(x_t) dt.$$

4.2 How does the choice of connection help in the case of the derivative flow?

One may wonder why a choice of a linear connection removes a martingale term in a SDE? The answer is that it does not and what it does is the careful choice of a matrix which transforms the original objects of interest. Recall the differentiation formula:

$$d(P_t f)(v) = \mathbf{E} df(X_t^v)$$

where for each t, X_t^v is a vector field with $X(x) = v$. The choice of X_t^v is by no means unique. Both the derivative flows and the damped parallel translations are valid choices and the linear connection which is intrinsic to the SDE leads to the correct choice. To make this plain let us now consider $\mathbf{R}^n$ as a trivial manifold with the non-trivial Riemannian metric and affine connection induced by X. In components, let U_i be functions on $\mathbf{R}^n$ and $U = (U_1, \dots, U_n)$ and $x_0, v \in \mathbf{R}^n$,

$$(\breve{\nabla}_v U)_k(x_0) = (DU_k)_{x_0}(v) + \sum_j \langle X(x_0) D(Y(x)(e_j), e_k \rangle (v) U_j e_k.$$

The last term determines the Christoffel symbols, c.f. [15].

Given a vector field along a continuous curve there is the stochastic covariant differentiation defined for almost surely all paths, given by $\hat{D}V_t = \hat{/\!/}_t \frac{d}{dt} (\hat{/\!/}_t)^{-1} V_t$ where $\hat{/\!/}_t$ is the stochastic parallel translation using the connection $\hat{\nabla}$, the adjoint connection to $\breve{\nabla}$ to take into account of the torsion effect. Alternatively

$$(\hat{D}V_t)^k = \frac{d}{dt} V_t^k + \Gamma_{ji}^k(\sigma_t) V_t^j \circ d\sigma_t^i.$$

The derivative flow $V_t = T\xi_t(v_0)$ satisfies the SDE:

$$\hat{D}V_t = \breve{\nabla} X_j(V_t) \circ dB_t^j + \breve{\nabla} X_0(V_t) dt.$$

Let $\bar{V}_t = \mathbf{E}\{V_t | \, x_s : 0 \le s \le T\}$. Then

$$\hat{D}\bar{V}_t = -\frac{1}{2}(\breve{\mathrm{Ric}})^{\#}(\bar{V}_t)dt + \nabla X_0(\bar{V}_t)dt.$$

In the setting of the Wiener space Ω and $\mathcal{I} = \xi_\cdot(x_0)$ the Itô map, let $V_t = T\mathcal{I}_t(h)$ for h a Cameron Martin vector then

$$\hat{D}V_t = \breve{\nabla}X_j(V_t) \circ dB_t^j + \breve{\nabla}X_0(V_t)dt + X(x_t)(\dot{h}_t)dt$$

and the corresponding conditional expectation of the vector field V_t satisfies

$$\hat{D}\bar{V}_t = -\frac{1}{2}(\breve{\mathrm{Ric}})^{\#}(\bar{V}_t)dt + \breve{\nabla}X_0(\bar{V}_t)dt + X(x_t)(\dot{h}_t)dt.$$

This means, $\hat{/\!/}_t^{-1}\bar{V}_t$ is differentiable in t and hence a Cameron-Martin vector and $\bar{V}_t$ is the induced Bismut-tangent vector by parallel translation.

4.3 A word about the stochastic filtering problem

Consider the filtering problem for a one dimensional signal process $x(t)$ transmitted through a noise channel

$$\begin{aligned} dx_t &= \alpha(x_t)dt + \sigma \, dW_t \\ dy_t &= \beta(x_t)dt + \sqrt{a}dB_t \end{aligned}$$

where B_t and W_t are independent Brownian motions. The problem is to find the probability density of $x(t)$ conditioned on the observation process $y(t)$ which is closely associated to the following horizontal lifting problem.

Let $\mathcal{B}$ and $\mathcal{A}$ be intertwined diffusion operators. Consider the martingale problems on the path spaces, $C_{u_0}N$ and $C_{y_0}M$, on N and M respectively. Let u_t and y_t be the canonical process on N and on M, assumed to exist for all time, so that for $f \in C_c^\infty(M)$ and $g \in C_c^\infty(N)$

$$\begin{aligned} M_t^{df,\mathcal{A}} : &= f(y_t) - f(y_0) - \int_0^t \mathcal{A}f(y_s)ds \\ M_t^{dg,\mathcal{B}} : &= g(u_t) - g(u_0) - \int_0^t \mathcal{B}g(u_s)ds \end{aligned}$$

are martingales. For a $\sigma\{y_s : 0 \le s \le t\}$-predictable T^*M-valued process ϕ_t which is along y_t we could also define a local martingale $M_t^{\phi,\mathcal{A}}$ by

$$\langle M_t^{\phi,\mathcal{A}}, M_t^{df,\mathcal{A}} \rangle = 2\int_0^t df(\sigma^{\mathcal{A}}(\phi))(y_s)ds.$$

It is also denoted by

$$M_t^{\phi,\mathcal{A}} \equiv \int_0^t \phi_s d\{y_s\}.$$

The conditional law of u_t given y_t is given by integration against function f from N to **R**, define

$$\pi_t f(u_0)(\sigma) = \mathbf{E}\Big\{ f(u_t) | p(u.) = \sigma \Big\}. \tag{4.2}$$

This conditional expectation is defined for $P^{\mathcal{A}}_{p(u_0)}$, the $\mathcal{A}$ diffusion measures, almost surely all σ and extends to $\phi_t \circ h_{u_t}$ for ϕ_t as before and h the horizontal lifting map. The following is from Theorem 4.5.1 in [13].

Theorem 4.1 If f is C^2 with $\mathcal{B}f$ and $\sigma^{\mathcal{B}}(df, df) \circ \mathfrak{h}$ bounded, then

$$\pi_t f(u_0) = f(u_0) + \int_0^t \pi_s(\mathcal{B}f)(u_0) ds + \int_0^t \pi_s(df \circ h_{u.})(u_0) d\{y_s\}. \tag{4.3}$$

To see this holds, taking conditional expectation of the following equation:

$$f(u_t) = f(u_0) + \int_0^t \mathcal{B}f(u_s) ds + M_t^{df, \mathcal{B}}$$

and use the following theorem, Proposition 4.3.5 in [13],

$$\mathbf{E}\{ M_t^{df,\mathcal{B}} | p(u.) = x.\} = M_t^{\mathbf{E}\{df \circ h_{u_s} | p(u.) = x.\}, \mathcal{A}}.$$

In the case that $p : M \times M_0 \to M$ is the trivial projection of the product manifold to M, let $\mathcal{A}$ be a cohesive diffusion operator on M, $\mathcal{L}$ the diffusion generator on M_0 and $u_t = (y_t, x_t)$ a $\mathcal{B}$ diffusion. If x_t is a Markov process with generator $\mathcal{L}$ and $\mathcal{B}$ a coupling of $\mathcal{L}$ and $\mathcal{A}$, by which we mean that $\mathcal{B}$ is intertwined with $\mathcal{L}$ and $\mathcal{A}$ by the projections p_i to the first or the second coordinates, there is a bilinear $\Gamma^{\mathcal{B}} : T^*M \times T^*M_0 \to \mathbf{R}$ such that

$$\mathcal{B}(g_1 \otimes g_2)(x, y) = (\mathcal{L}g_1)(x) g_2(y) + g_1(x)(\mathcal{A}g_2)(y) + \Gamma^{\mathcal{B}}((dg_1)_x, (dg_2)_y) \tag{4.4}$$

where $g_1 \otimes g_2 : M \times M_0 \to \mathbf{R}$ denotes the map $(x, y) \mapsto g_1(x) g_2(y)$ and g_1, g_2 are C^2. In fact $\Gamma^{\mathcal{B}}((dg_1)_x, (dg_2)_y) = \sigma^{\mathcal{B}}_{(x,y)}(d\tilde{g}_1, d\tilde{g}_2)$ where $\tilde{g}_i = g(p_i)$. Then $\sigma^{\mathcal{B}} : T^*M_1 \times T^*M_2 \to TM_1 \times TM_2$ is of the following form. For $\ell_1 \in T_x^* M_1$, $\ell_2 \in T_y^* M_2$

$$\sigma^B_{(x,y)}(\ell_1, \ell_2) = \begin{pmatrix} \sigma_x^L & \sigma^{1,2}_{(x,y)} \\ \sigma^{2,1}_{(x,y)} & \sigma_x^{\mathcal{A}} \end{pmatrix} \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix}.$$

The horizontal lifting map is given by

$$v \mapsto (v, \alpha \circ (\sigma^{\mathcal{A}})^{-1}(v))$$

where $\alpha : T_x M^* \to T_y M_0$ are defined by

$$\ell_2(\alpha(\ell_1)) = \frac{1}{2} \Gamma^{\mathcal{B}}(\ell_1, \ell_2).$$

In the theorem above take $1 \otimes f$ to see that $\pi_s \mathcal{B}(1 \otimes f)$ reduces to $\mathcal{L}f$ and the filtering equation is:

$$\pi_t f(x_0) = f(x_0) + \int_0^t \pi_s(\mathcal{L}f)(x_0)ds + \int_0^t \pi_s(df(\alpha \circ (\sigma^{\mathcal{A}})^{-1}))(x_0)d\{y_s\}.$$

The case of non-Markovian observation when the non-Makovian factor is introduced through the drift equation for the noise process y_t can be dealt with through a Girsanov transformation. See [13] for detail. Finally we note that the field of stochastic filtering is vast and deep and we did not and would not attempt to give historical references as they deserve. However we would like to mention a recent development [5] which explore the geometry of the signal-observation system. See also [16], [17], [14] and recent work of T. Kurtz .

Acknowledgement. This article is based on the books [11, 13] and could and should be considered as joint work with K. D. Elworthy and Y. LeJan and I would like to thank them for helpful discussions. However any short comings and errors are my sole responsibility. This research is supported by an EPSRC grant (EP/E058124/1).

References

[1] A. Anderson and R. Camporesi. Intertwining Operators for Solving Differential Equations, with Applications to Symmetric Spaces. Commun. Math. Phys. 130, 61-82 (1990)

[2] M. Arnaudon, H. Plank and A. Thalmaier, A Bismut type formula for the Hessian of heat semigroups. C. R. Acad. Sci. Paris 336 (2003) 661-666.

[3] J.-M. Bismut. *Large deviations and the Malliavin calculus*, volume 45 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1984.

[4] P. Carmona, F. Petit, and M. Yor. Beta-gamma random variables and intertwining relations between certain Markov processes. Revista Matematica Iberoamericana, Vol 14, No. 2 1998.

[5] D. Crisan, M. Kouritzin, and J. Xiong. Nonlinear filtering with signal dependent observation noise. *Electron. J. Probab.*, 14:no. 63, 1863–1883, 2009.

[6] B. K. Driver and A. Thalmaier. Heat Equation Derivative Formulas for Vector Bundles, J. Funct. Anal., 183(1), 42-108, 2001.

[7] B. K. Driver and T. Melcher, Hypoelliptic heat kernel inequalities on the Heisenberg group. Journal of Functional Analysis, 221 (2) , 340-365, 2005.

[8] K. D. Elworthy. Geometric aspects of diffusions on manifolds. In P. L. Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour XV-XVII, 1985-1987. Lecture Notes in Mathematics 1362*, volume 1362, pages 276–425. Springer-Verlag, 1988.

[9] K. D. Elworthy and M. Yor. Conditional expectations for derivatives of certain stochastic flows. In J. Azéma, P.A. Meyer, and M. Yor, editors, *Sem. de Prob. XXVII. Lecture Notes in Maths. 1557*, pages 159–172. Springer-Verlag, 1993.

[10] K. D. Elworthy, Y. LeJan, and Xue-Mei Li. Concerning the geometry of stochastic differential equations and stochastic flows. In *'New Trends in stochastic Analysis', Proc. Taniguchi Symposium, Sept. 1994, Charingworth, ed. K. D. Elworthy and S. Kusuoka, I. Shigekawa*. World Scientific Press, 1996.

[11] K. D. Elworthy, Y. LeJan, and Xue-Mei Li. *On the geometry of diffusion operators and stochastic flows, Lecture Notes in Mathematics 1720*. Springer, 1999.

[12] K. D. Elworthy, Y. Le Jan, and Xue-Mei Li. Equivariant diffusions on principal bundles. In *Stochastic analysis and related topics in Kyoto*, volume 41 of *Adv. Stud. Pure Math.*, pages 31–47. Math. Soc. Japan, Tokyo, 2004.

[13] K. D. Elworthy, Yves Le Jan, and Xue-Mei Li. The Geometry of Filtering. To appear in Frontiers in Mathematics Series, Birkhauser.

[14] H. Kunita. Cauchy problem for stochastic partial differential equations arising in nonlinear filtering theory. *Systems Control Lett.*, 1(1):37–41, 1981/82.

[15] Y. LeJan and S. Watanabe. Stochastic flows of diffeomorphisms. In *Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, 32,*, pages 307–332. North-Holland, Amsterdam, 1984.

[16] J. Lázaro-Camí and J. Ortega. Reduction, reconstruction, and skew-product decomposition of symmetric stochastic differential equations. *Stoch. Dyn.*, 9(1):1–46, 2009.

[17] M. Liao. Factorization of diffusions on fibre bundles. *Trans. Amer. Math. Soc.*, 311(2):813–827, (1989).

[18] Xue-Mei Li, Stochastic Flows on Noncompact manifolds (1992). University Of Warwick , Thesis. In http://www.xuemei.org/Xue-Mei.Li-thesis.pdf

[19] Xue-Mei Li, Stochastic differential equations on non-compact manifolds: moment stability and its topological consequences. Probab. Theory Relat. Fields.100, 4, 417-428. 1994.

[20] Xue-Mei Li, Strong p-completeness of stochastic differential equations and the existence of smooth flows on non-compact manifolds. Probab. Theory Relat. Fields. 100, 4, 485-511. 1994.

[21] Xue-Mei Li, On extensions of Myers' theorem. Bull. London Math. Soc. 27, 392-396. 1995.

[22] J. R. Norris, Path integral formulae for heat kernels and their derivatives. Probab. Theory Related Fields 94 (1993), pp. 525541.

[23] J. Pitman. One dimensional Brownian motion and the three dimensional Bessel process. Adv. Appl. Probab. 7 (1975), 511-526.

[24] A. Thalmaier and F. Wang, Gradient estimates for harmonic functions on regular domains in Riemannian manifolds, Journal of Functional Analysis 155 (1998) 109-124.