Chapter 2

Geometry of subelliptic diffusions

Anton Thalmaier[1]

Contents

1 Introduction

In the same way as a vector field on a differentiable manifold induces a flow, second-order differential operators induce stochastic flows with similar properties. In this sense, Brownian motion on a Riemannian manifold appears as the stochastic flow associated to the Laplace–Beltrami operator. The new feature of stochastic flows is that the flow curves depend on a random parameter and behave irregularly as functions of time [36]. This irregularity reveals an irreversibility of time which is inherent in stochastic phenomena.

Subelliptic diffusions are stochastic flows to canonical second-order differential operators associated with sub-Riemannian structures and corresponding horizontal distributions. A common feature of these operators is their lack of ellipticity. Typically they degenerate along a subbundle of the tangent bundle.

[1]anton.thalmaier@uni.lu
Mathematics Research Unit, FSTC, University of Luxembourg, 6 rue Richard Coudenhove-Kalergi, L-1359 Luxembourg, Grand-Duchy of Luxembourg.

2 Stochastic flows

Let M be a differentiable manifold of dimension n and denote by

$$TM \xrightarrow{\pi} M$$

its *tangent bundle*. In particular, we have

$$TM = \dot\cup_{x\in M} T_xM, \quad \pi|T_xM = x.$$

The space of smooth sections of TM is denoted by

$$\begin{aligned}\Gamma(TM) &= \{A\colon M \to TM \text{ smooth} \mid \pi\circ A = \mathrm{id}_M\}\\ &= \{A\colon M \to TM \text{ smooth} \mid A(x)\in T_xM \text{ for all } x\in M\}\end{aligned}$$

and constitutes the *vector fields* on M. As usual, we identify vector fields on M and $\mathbb{R}$-derivations on $C^\infty(M)$ as

$$\begin{aligned}\Gamma(TM) \mathrel{\widehat{=}} \bigl\{&A\colon C^\infty(M)\to C^\infty(M)\ \mathbb{R}\text{-linear}\\ &\mid A(fg) = fA(g) + gA(f), \quad \forall f,g\in C^\infty(M)\bigr\},\end{aligned}$$

where a vector field $A\in\Gamma(TM)$ is considered an $\mathbb{R}$-derivation via

$$A(f)(x) := df_x\, A(x) \in \mathbb{R}, \quad x\in M, \tag{2.1}$$

using the *differential* $df_x\colon T_xM \to \mathbb{R}$ of f at x.

There is a dynamical point of view to vector fields on manifolds: it associates to each vector field a dynamical system given by the flow of the vector field.

2.1 Flow of a vector field. Given a vector field $A\in\Gamma(TM)$, for each $x\in M$ we consider the smooth curve $t\mapsto x(t)$ in M with the properties

$$x(0) = x \quad\text{and}\quad \dot x(t) = A(x(t)).$$

We write $\phi_t(x) := x(t)$. In this way, we obtain for each $A\in\Gamma(TM)$ the corresponding *flow to* A given by

$$\begin{cases}\frac{d}{dt}\phi_t = A(\phi_t),\\ \phi_0 = \mathrm{id}_M.\end{cases} \tag{2.2}$$

System (2.2) means that for any $f\in C_c^\infty(M)$ (space of compactly supported smooth functions on M) the following conditions hold:

$$\begin{cases}\frac{d}{dt}(f\circ\phi_t) = A(f)\circ\phi_t,\\ f\circ\phi_0 = f.\end{cases} \tag{2.3}$$

Indeed, by the chain rule along with definition (2.1), we have for each $f \in C_c^\infty(M)$,

$$\frac{d}{dt}(f \circ \phi_t) = (df)_{\phi_t} \frac{d}{dt}\phi_t = (df)_{\phi_t} A(\phi_t) = A(f)(\phi_t).$$

In integrated form, for each $f \in C_c^\infty(M)$, conditions (2.3) can be written as

$$f \circ \phi_t(x) - f(x) - \int_0^t A(f)(\phi_s(x))\, ds = 0, \quad t \geq 0,\ x \in M. \tag{2.4}$$

As usual, the curve

$$\phi_{\bullet}(x)\colon\ t \mapsto \phi_t(x)$$

is called the *flow curve* (or *integral curve*) to A starting at x.

Remark 2.1. Defining $P_t f := f \circ \phi_t$, we observe that $\frac{d}{dt} P_t f = P_t(A(f))$, in particular,

$$\left.\frac{d}{dt}\right|_{t=0} P_t f = A(f). \tag{2.5}$$

In other words, from knowledge of the flow ϕ_t, the underlying vector field A can be recovered by taking the derivative at zero as in Eq. (2.5).

2.2 Flow to a second-order differential operator. Now let L be a second-order partial differential operator (PDO) on M, e.g., of the form

$$L = A_0 + \sum_{i=1}^r A_i^2, \tag{2.6}$$

where $A_0, A_1, \dots, A_r \in \Gamma(TM)$ for some $r \in \mathbb{N}$. Note that $A_i^2 = A_i \circ A_i$ is understood as a composition of derivations, i.e.,

$$A_i^2(f) = A_i(A_i(f)), \quad f \in C^\infty(M).$$

Example 2.2. Let $M = \mathbb{R}^n$ and consider

$$A_0 = 0 \quad \text{and} \quad A_i = \frac{\partial}{\partial x_i} \text{ for } i = 1, \dots, n.$$

Then $L = \Delta$ is the classical Laplace operator on $\mathbb{R}^n$.

Alternatively, we may consider partial differentiable operators L on M which locally in a chart (h, U) can be written as

$$L|U = \sum_{i=1}^n b_i \partial_i + \sum_{i,j=1}^n a_{ij} \partial_i \partial_j, \tag{2.7}$$

where $b \in C^\infty(U, \mathbb{R}^n)$ and $a \in C^\infty(U, \mathbb{R}^n \otimes \mathbb{R}^n)$ such that $a_{ij} = a_{ji}$ for all i, j (a symmetric). Here we use the notation $\partial_i = \frac{\partial}{\partial h_i}$.

Motivated by the example of a flow to a vector field (vector fields can be seen as first-order differential operators) we want to investigate the question of whether an analogous concept of flow exists for second-order PDOs.

Question 1. Is there a notion of a flow to L if L is a second-order PDO given by (2.6) or (2.7)?

Definition 2.3. Let $(\Omega, \mathscr{F}, \mathbb{P}; (\mathscr{F}_t)_{t\geq 0})$ be a filtered probability space, i.e., a probability space equipped with an increasing sequence of sub-σ-algebras $\mathscr{F}_t$ of $\mathscr{F}$. An adapted continuous process

$$X_{\bullet}(x) \mathrel{\widehat{=}} (X_t(x))_{t\geq 0}$$

on $(\Omega, \mathscr{F}, \mathbb{P}; (\mathscr{F}_t)_{t\geq 0})$ taking values in M, is called a *flow process* to L (or an L-*diffusion*) with starting point x if $X_0(x) = x$ and if, for all test functions $f \in C_c^\infty(M)$, the process

$$N_t^f(x) := f(X_t(x)) - f(x) - \int_0^t (Lf)(X_s(x))\, ds, \quad t \geq 0, \tag{2.8}$$

is a martingale, i.e.,

$$\mathbb{E}^{\mathscr{F}_s} \underbrace{\left[f(X_t(x)) - f(X_s(x)) - \int_s^t (Lf)(X_r(x))\, dr \right]}_{= N_t^f(x) - N_s^f(x)} = 0 \quad \text{for all } s \leq t.$$

Note that, by definition, flow processes to a second-order PDO depend on an additional random parameter $\omega \in \Omega$. For each $t \geq 0$, $X_t(x) \equiv (X_t(x, \omega))_{\omega \in \Omega}$ is an $\mathscr{F}_t$-measurable random variable. The defining equation (2.4) for flow curves translates to the martingale property of (2.8), i.e., the flow curve condition (2.4) holds only under conditional expectations. The theory of martingales gives a rigorous meaning to the idea of a process without systematic drift [59].

Remark 2.4. Since $N_0^f(x) = 0$, we get from the martingale property of $N^f(x)$ that

$$\mathbb{E}\big[N_t^f(x)\big] = \mathbb{E}\big[N_0^f(x)\big] = 0.$$

Hence, defining $P_t f(x) := \mathbb{E}\,[f(X_t(x))]$, we observe that

$$P_t f(x) = f(x) + \int_0^t \mathbb{E}\,[(Lf)(X_s(x))]\, ds,$$

and thus

$$\frac{d}{dt} P_t f(x) = \mathbb{E}\,[(Lf)(X_t(x))] = P_t(Lf)(x);$$

in particular,

$$\frac{d}{dt}\bigg|_{t=0} \mathbb{E}\,[f(X_t(x))] \equiv \frac{d}{dt}\bigg|_{t=0} P_t f(x) = Lf(x).$$

The last formula shows that as for deterministic flows we can recover the operator L from its stochastic flow process. To this end however, we have to average over all possible trajectories starting from x.

For background on stochastic flows we refer to the monograph of Kunita [36].

Example 2.5 (Brownian motion). Let $M = \mathbb{R}^n$ and $L = \frac{1}{2}\Delta$, where Δ is the Laplacian on $\mathbb{R}^n$. Let $X \equiv (X_t)$ be a Brownian motion on $\mathbb{R}^n$ starting at the origin. By Itô's formula [52], for $f \in C^\infty(\mathbb{R}^n)$, we have

$$\begin{aligned} d(f \circ X_t) &= \sum_{i=1}^n \partial_i f(X_t)\, dX_t^i + \frac{1}{2} \sum_{i,j=1}^n \partial_i \partial_j f(X_t)\, dX_t^i dX_t^j \\ &= \langle (\nabla f)(X_t), dX_t \rangle + \frac{1}{2} (\Delta f)(X_t)\, dt. \end{aligned}$$

Thus, for each $f \in C_c^\infty(\mathbb{R}^n)$,

$$f(X_t) - f(X_0) - \int_0^t \frac{1}{2} (\Delta f)(X_s)\, ds, \quad t \geq 0$$

is a martingale. This means that the process

$$X_t(x) := x + X_t$$

is an L-diffusion to $\frac{1}{2}\Delta$ in the sense of Definition 2.3.

Remarks 2.6. As for deterministic flows, we have to deal with the problem that stochastic flows may explode in finite times.

(1) We allow $X_{\bullet}(x)$ to be defined only up to some stopping time $\zeta(x)$, i.e.,

$$X_{\bullet}(x) \mid [0, \zeta(x)[,$$

where

$$\{\zeta(x) < \infty\} \subset \left\{ \lim_{t \uparrow \zeta(x)} X_t(\omega) = \infty \text{ in } \hat{M} := M \mathbin{\dot\cup} \{\infty\} \right\} \quad \mathbb{P}\text{-a.s.} \tag{2.9}$$

Here $\hat{M}$ denotes the one-point compactification of M. A stopping time $\zeta(x)$ with property (2.9) is called a (maximal) *lifetime* for the process $X_\bullet(x)$ starting at x. In equivalent terms, let $U_n \subset M$ be open, relatively compact subsets exhausting M in the sense that

$$U_n \subset \bar{U}_n \subset U_{n+1} \subset \cdots, \quad \bar{U}_n \text{ compact}, \quad \text{and } \cup_n U_n = M.$$

Then we have $\zeta(x) = \sup_n \tau_n(x)$ for the maximal lifetime of $X_\bullet(x)$, where $\tau_n(x)$ is the family of stopping times (*first exit times* of U_n) defined by

$$\tau_n(x) := \inf\{t \geq 0\colon\ X_t(x) \notin U_n\}.$$

(2) For $f \in C^\infty(M)$ (not necessarily compactly supported), the process $N^f(x)$ will in general only be a *local* martingale [52], i.e., there exist stopping times $\tau_n \uparrow \zeta(x)$ such that

$$\forall n \in \mathbb{N}, \quad \big(N^f_{t\wedge\tau_n}(x)\big)_{t\geq 0} \text{ is a (true) martingale.}$$

(3) The following two statements are equivalent (the proof will be given later):

   (a) The process
   $$f(X_\bullet(x)) = (f(X_t(x)))_{t\geq 0}$$
   is of locally bounded variation for all $f \in C_c^\infty(M)$.

   (b) The operator L is of first order, i.e., L is a vector field (in which case the flow is deterministic).

In other words, flow processes have "nice paths" (for instance, paths of bounded variation) if and only if the corresponding operator is first order (i.e., a vector field).

2.3 What are L-diffusions good for? Before discussing the problem of how to construct L-diffusions, we want to study some implications to indicate the usefulness and power of this concept. In the following two examples we assume only existence of an L-diffusion to a given operator L.

A. (*Dirichlet problem*) Let $\varnothing \neq D \subsetneq M$ be an open, connected, relatively compact domain, $\varphi \in C(\partial D)$ and let L be a second-order PDO on M. The *Dirichlet problem* (DP) is the problem to find a function $u \in C(\bar{D}) \cap C^2(D)$ such that

$$\begin{cases} Lu = 0 \text{ on } D, \\ u|_{\partial D} = \varphi. \end{cases} \tag{DP}$$

Suppose that there is an L-diffusion $(X_t(x))_{t\geq 0}$. We choose a sequence of open domains $D_n \uparrow D$ such that $\bar{D}_n \subset D$, and for each n, we consider the *first exit time* of D_n,

$$\tau_n(x) = \inf\{t \geq 0,\ X_t(x) \notin D_n\}.$$

Then $\tau_n(x) \uparrow \tau(x)$, where

$$\tau(x) = \sup_n \tau_n(x) = \inf\{t \geq 0,\ X_t(x) \notin D\}.$$

Now assume that u is a solution to (DP). We may choose test functions $u_n \in C_c^\infty(M)$ such that $u_n|D_n = u|D_n$ and $\operatorname{supp} u_n \subset D$. Then, by the property of an L-diffusion,

$$N_t(x) := u_n(X_t(x)) - u_n(x) - \int_0^t (Lu_n)(X_r(x))\,dr$$

is a martingale. We suppose that $x \in D_n$. Then

$$\begin{aligned} N_{t\wedge\tau_n(x)}(x) &= u_n(X_{t\wedge\tau_n(x)}(x)) - u_n(x) - \int_0^{t\wedge\tau_n(x)} \underbrace{(Lu_n)(X_r(x))}_{=0}\,dr \qquad (2.10)\\ &= u(X_{t\wedge\tau_n(x)}(x)) - u(x) \end{aligned}$$

is also a martingale (here we used that the integral in (2.10) is 0 since $Lu_n = Lu = 0$ on D_n). Thus we get

$$\mathbb{E}\left[N_{t\wedge\tau_n(x)}(x)\right] = \mathbb{E}\left[N_0(x)\right] = 0,$$

which shows that for each $n \in \mathbb{N}$,

$$u(x) = \mathbb{E}\left[u(X_{t\wedge\tau_n(x)}(x))\right]. \qquad (2.11)$$

From Eq. (2.11) we may conclude by dominated convergence and since $\tau_n(x) \uparrow \tau$ that

$$u(x) = \lim_{n\to\infty} \mathbb{E}\left[u(X_{t\wedge\tau_n(x)}(x))\right] = \mathbb{E}\left[\lim_{n\to\infty} u(X_{t\wedge\tau_n(x)}(x))\right] = \mathbb{E}\left[u(X_{t\wedge\tau(x)}(x))\right].$$

We now make the *hypothesis* that $\tau(x) < \infty$ a.s. (the process exits the domain D in finite time). Then

$$\begin{aligned} u(x) &= \lim_{t\to\infty} \mathbb{E}\left[u(X_{t\wedge\tau(x)}(x))\right] = \mathbb{E}\left[\lim_{t\to\infty} u(X_{t\wedge\tau(x)}(x))\right]\\ &= \mathbb{E}\left[u(X_{\tau(x)}(x))\right] = \mathbb{E}\left[\varphi(X_{\tau(x)}(x))\right], \end{aligned}$$

where for the last equality we used the boundary condition $u|\partial D = \varphi$. Note that by passing to the image measure $\mu_x := \mathbb{P} \circ X_{\tau(x)}(x)^{-1}$ on the boundary we get

$$\mathbb{E}\left[\varphi(X_{\tau(x)}(x))\right] = \int_{\partial D} \varphi(z)\,\mu_x(dz).$$

Notation 2.7. The measure μ_x, defined on Borel sets $A \subset \partial D$,

$$\mu_x(A) = \mathbb{P}\left\{X_{\tau(x)}(x) \in A\right\},$$

is called an *exit measure* from the domain D of the diffusion $X_t(x)$. It represents the probability that the process X_t, when started at x in D, exits the domain D through the boundary set A.

Conclusions. From the discussion of the Dirichlet problem above we can make the following two observations.

(a) (Uniqueness) Under the hypothesis

$$\boxed{\tau(x) < \infty \text{ a.s.} \quad \text{for all } x \in D}$$

we have uniqueness of the solutions to the Dirichlet problem (DP). It will be shown later that this hypothesis concerns nondegeneracy of the operator L.

(b) (Existence) Under the hypothesis

$$\boxed{\tau(x) \longrightarrow 0 \ \text{ if } D \ni x \longrightarrow a \in \partial D}$$

we have

$$\mathbb{E}\left[\varphi(X_{\tau(x)}(x))\right] \longrightarrow \varphi(a), \qquad \text{if } D \ni x \longrightarrow a \in \partial D.$$

Thus one may define $u(x) := \mathbb{E}\left[\varphi(X_{\tau(x)}(x))\right]$. It can be shown then that u is L-harmonic on D if it is twice differentiable; thus under the hypothesis in (b), u will then satisfy the boundary condition and hence solve (DP). The hypothesis in (b) is obviously a regularity condition on the boundary ∂D.

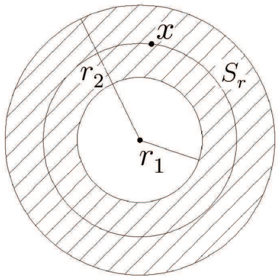
Note that in the arguments above we nowhere used the explicit form of the operator L nor of the domain D. We used only the general properties of a stochastic flow process associated to the given operator L. For a more complete discussion of the Dirichlet problem, see [54, 13].

Examples 2.8.

(1) Let $M = \mathbb{R}^2 \backslash \{0\}$ and $D = \left\{x \in \mathbb{R}^2 : r_1 < |x| < r_2\right\}$ with $0 < r_1 < r_2$. Consider the operator

$$L = \frac{1}{2}\frac{\partial^2}{\partial \vartheta^2},$$

where ϑ denotes the angle when passing to polar coordinates on M. If u is a solution of (DP), then $u + v(r)$ is a solution of (DP) as well, for any radial function $v(r)$ satisfying $v(r_1) = v(r_2) = 0$. Hence, uniqueness of solutions fails.



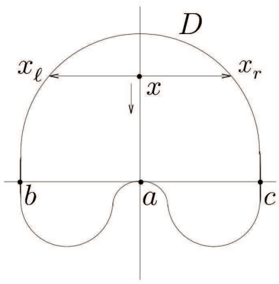
Note: For $x \in D$ with $|x| = r$, let $S_r = \{x \in \mathbb{R}^2 : |x| = r\}$. Then, the flow process $X_\bullet(x)$ to L is easily seen to be a (one-dimensional) Brownian motion on S_r. In particular,

$$\tau(x) = +\infty \text{ a.s.}$$

(2) Let $M = \mathbb{R}^2$ and consider the operator

$$L = \frac{1}{2} \frac{\partial^2}{\partial x_1^2}$$

on a domain D in $\mathbb{R}^2$ of the following shape:



Then, for $x = (x_1, x_2) \in D$, the flow process $X_\bullet(x)$ starting at x is a (one-dimensional) Brownian motion on $\mathbb{R} \times \{x_2\}$. In other words, flow processes move on horizontal lines. In particular, when started at $x \in D$, the process can exit only at two points (e.g., x_ℓ and x_r in the picture). Letting x vertically approach a, by symmetry of the one-dimensional Brownian motion, we see that there exists a solution of (DP) if and only if

$$\varphi(a) = \frac{\varphi(b) + \varphi(c)}{2}.$$

B. (*Heat equation*) Let L be a second-order PDO on M and fix $f \in C(M)$. The *heat equation* on M with initial condition f concerns the problem of finding a real-valued function $u = u(t,x)$ defined on $\mathbb{R}_+ \times M$ such that

$$\begin{cases} \dfrac{\partial u}{\partial t} = Lu \text{ on }]0,\infty[\times M, \\ u|_{t=0} = f. \end{cases} \tag{HE}$$

Suppose now that there is an L-diffusion $X_\bullet(x)$. It is straightforward to see that the "time-space process" $(t, X_t(x))$ will then be an $\hat{L}$-diffusion for the parabolic operator

$$\hat{L} = \frac{\partial}{\partial t} + L$$

with starting point $(0,x)$. By definition, this means that for all $\varphi \in C^2(\mathbb{R}_+ \times M)$,

$$d\varphi(t, X_t(x)) - \big(\hat{L}\varphi\big)(t, X_t(x))\, dt \overset{\mathrm{m}}{=} 0$$

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

From now on we assume nonexplosion of the L-diffusion. In other words, we adopt the *hypothesis* that $\zeta(x) = +\infty$ a.s. for all $x \in M$, i.e.,

$$\boxed{\mathbb{P}\left\{X_t(x) \in M, \quad \forall t \geq 0\right\} = 1, \ \forall x \in M.}$$

Suppose now that u is a *bounded* solution of (HE). We fix $t \geq 0$ and consider the restriction $u \mid [0,t] \times M$. Then

$$u(t-s, X_s(x)) - u(t,x) - \int_0^s \left[\left(\frac{\partial}{\partial r} + L\right) u(t-r,\cdot)\right](X_r(x))\, dr, \quad 0 \leq s < t$$

is a local martingale. In other words, fixing $t > 0$, we have for $0 \leq s < t$,

$$\begin{aligned} u(t-s, X_s(x)) = u(t,x) + \int_0^s \underbrace{\left(\frac{\partial}{\partial r} + L\right) u(t-r,\cdot)(X_r(x))dr}_{=0,\ \text{since } u \text{ solves (HE)}} \\ + (\text{local martingale})_s. \end{aligned} \tag{2.12}$$

Since the integral in (2.12) vanishes, we see that the local martingale term in (2.12) is actually a bounded local martingale (since $u(t-s, X_s(x)) - u(t,x)$ is bounded) and hence a true martingale (equal to 0 at time 0). Using the martingale property we first take expectations and then pass to the limit as $s \uparrow t$ to obtain

$$u(t,x) = \mathbb{E}\left[u(t-s, X_s(x))\right] \to \mathbb{E}\left[u(0, X_t(x))\right] = \mathbb{E}\left[f(X_t(x))\right], \quad \text{as } s \uparrow t, \tag{2.13}$$

where for the limit in (2.13) we have used dominated convergence (recall that u is bounded).

Conclusion. Under the hypothesis $\zeta(x) = +\infty$ for all $x \in M$, we have uniqueness of (bounded) solutions to the heat equation (HE). Solutions are necessarily of the form

$$\boxed{u(t,x) = \mathbb{E}\left[f(X_t(x))\right].}$$

Interpretation. The solution $u(t,x)$ at time t and at point x can be constructed as follows: Run an L-diffusion process starting from x up to time t, apply the initial condition f to the obtained random position $X_t(x)$ at time t, and average over all possible paths.

2.4 Γ-operators and quadratic variation.

Definition 2.9. Let $L\colon C^\infty(M) \to C^\infty(M)$ be a linear mapping (for instance, a second-order PDO). The Γ-*operator associated to* L (“*l'operateur carré du champ*”) is the bilinear map

$$\Gamma\colon C^\infty(M) \times C^\infty(M) \to C^\infty(M) \text{ given as}$$

$$\boxed{\Gamma(f,g) := \tfrac{1}{2}\left[L(fg) - fL(g) - gL(f)\right].}$$

Example 2.10. Let L be a second-order PDO on M without constant term (i.e., $L1 = 0$). Suppose that in a local chart (h,U) for M the operator L is written as

$$L|C_U^\infty(M) = \sum_{i,j=1}^n a_{ij}\,\partial_i\partial_j + \sum_{i=1}^n b_i\,\partial_i,$$

where $C_U^\infty(M) = \{f \in C^\infty(M):\ \operatorname{supp} f \subset U\}$ and $\partial_i = \frac{\partial}{\partial h_i}$. Then

$$\Gamma(f,g) = \sum_{i,j=1}^n a_{ij}(\partial_i f)(\partial_j g), \quad \forall f,g \in C_U^\infty(M).$$

For instance, in the special case that $M = \mathbb{R}^n$ and $L = \Delta$, we find

$$\Gamma(f,f) = \|\nabla f\|^2.$$

Remark 2.11. Let L be a second-order PDO. Then the following equivalence holds:

$$\Gamma(f,g) = 0, \quad \forall f,g \in C^\infty(M) \quad \text{if and only if} \quad L \text{ is of first order, i.e., } L \in \Gamma(TM).$$

For instance, if $L = A_0 + \sum_{i=1}^r A_i^2$, then

$$\Gamma(f,g) = \sum_{i=1}^r A_i(f)A_i(g),$$

and in particular,

$$\Gamma \equiv 0 \quad \text{if and only if} \quad A_1 = A_2 = \cdots = A_r = 0.$$

Remark 2.12. A continuous real-valued stochastic process $(X_t)_{t\ge 0}$ is called a *semimartingale* if it can be decomposed as

$$X_t = X_0 + M_t + A_t, \tag{2.14}$$

where M is a local martingale and A an adapted process of locally bounded variation (with $M_0 = A_0 = 0$). The representation of a semimartingale X as in (2.14) (Doob–Meyer decomposition) is unique: If $\mathscr{M}_0$ denotes the class of local martingales starting from 0 and $\mathscr{A}_0$ is the class of adapted processes with paths of locally bounded variation starting from the origin, then $\mathscr{M}_0 \cap \mathscr{A}_0 = 0$.

Definition 2.13. Let X be a continuous adapted process taking values in a manifold M. Then X is called *semimartingale on M* if

$$f(X) \equiv (f(X_t))_{t\ge 0}$$

is a real-valued semimartingale for all $f \in C^\infty(M)$.

Remark 2.14. If X has maximal lifetime ζ, i.e.,

$$\{\zeta < \infty\} \subset \left\{\lim_{t\uparrow\zeta} X_t = \infty \text{ in } \hat{M} = M\dot{\cup}\{\infty\}\right\} \quad \text{a.s.,}$$

then $f(X)$ is well defined as a process globally on $\mathbb{R}_+$ for all $f \in C_c^\infty(M)$ (with the convention $f(\infty) = 0$). For $f \in C^\infty(M)$, in general,

$$f(X) \equiv (f(X_t))_{t<\zeta}$$

is only a semimartingale with lifetime ζ.

Proposition 2.15. *Let $L\colon C^\infty(M) \to C^\infty(M)$ be an $\mathbb{R}$-linear map and X be a semimartingale on M such that for all $f \in C^\infty(M)$,*

$$N_t^f := f(X_t) - f(X_0) - \int_0^t Lf(X_r)\,dr$$

is a continuous local martingale (of the same lifetime as X) (i.e., $d(f(X)) - Lf(X)\,dt \overset{\mathrm{m}}{=} 0$, where $\overset{\mathrm{m}}{=}$ denotes equality modulo differentials of local martingales).

Then, for all $f, g \in C^\infty(M)$, the quadratic variation $[f(X), g(X)]$ of $f(X)$ and $g(X)$ is given by

$$d\,[f(X), g(X)] \equiv d\big[N^f, N^g\big] = 2\,\Gamma(f,g)(X)\,dt.$$

In particular, $\Gamma(f,f)(X) \geq 0$ a.s.

Proof. Let $f \in C^\infty(M, \mathbb{R}^r)$ and $\phi \in C^\infty(\mathbb{R}^r)$. Writing as above $\overset{\mathrm{m}}{=}$ for equality modulo differentials of local martingales, we have

$$d(\phi \circ f)(X) \overset{\mathrm{m}}{=} L(\phi \circ f)(X)\,dt. \tag{2.15}$$

Developing the left-hand side in Eq. (2.15) by Itô's formula (the function ϕ is applied to the semimartingale $f \circ X$), we get

$$\begin{aligned} &d(\phi(f(X))) \\ &\quad = \sum_{i=1}^r (D_i\phi)(f \circ X)\,d(f^i \circ X) + \frac{1}{2}\sum_{i,j=1}^r (D_iD_j\phi)(f \circ X)\,d[f^i(X), f^j(X)] \\ &\quad \overset{\mathrm{m}}{=} \sum_{i=1}^r (D_i\phi)(f \circ X)\,(Lf^i)(X)\,dt + \frac{1}{2}\sum_{i,j=1}^r (D_iD_j\phi)(f \circ X)\,d[f^i(X), f^j(X)], \end{aligned}$$

where $D_i = \partial/\partial x_i$. By equating the drift parts we find

$$\left[L(\phi \circ f) - \sum_{i=1}^r ((D_i\phi) \circ f)\,(Lf^i)\right](X)\,dt = \frac{1}{2}\sum_{i,j=1}^r (D_iD_j\phi)(f \circ X)\,d[f^i(X), f^j(X)].$$

Taking now $r = 2$ and considering the special case $\phi(x,y) = xy$, we get with $f = (f^1, f^2)$,

$$\left[L(f^1f^2) - f^1L(f^2) - f^2L(f^1)\right](X)\,dt = d\left[f^1(X), f^2(X)\right].$$

Since $\left[L(f^1f^2) - f^1L(f^2) - f^2L(f^1)\right](X) = 2\Gamma(f^1, f^2)(X)$, this completes the proof. □

Lemma 2.16. *For an $\mathbb{R}$-linear map $L\colon C^\infty(M) \to C^\infty(M)$ the following statements are equivalent:*

(i) L is a second-order PDO (without constant term).

(ii) L satisfies the second-order chain rule, i.e., for all $f \in C^\infty(M, \mathbb{R}^r)$ and $\phi \in C^\infty(\mathbb{R}^r)$,

$$L(\phi \circ f) = \sum_{i=1}^r (D_i\phi \circ f)(Lf^i) + \sum_{i,j=1}^r (D_iD_j\phi \circ f)\,\Gamma(f^i, f^j).$$

Proof. (i) $\Rightarrow$ (ii): Write L in local coordinates as

$$L \mid C_U^\infty(M) = \sum_{i,j=1}^n a_{ij}\,\partial_i\partial_j + \sum_{i=1}^n b_i\,\partial_i$$

and use that $\Gamma(f, g) = \sum_{i,j=1}^n a_{ij}\,\partial_i f\,\partial_j g$.

(ii) $\Rightarrow$ (i): Determine the action of L on functions φ written in local coordinates (h, U) via

$$L(\varphi)|U = L(\varphi \circ h^{-1} \circ h) \equiv L(\phi \circ f),$$

where $\phi = \varphi \circ h^{-1}$ and $f = h$. Details are left as an exercise to the reader. □

Corollary 2.17. *Let $L\colon C^\infty(M) \to C^\infty(M)$ be an $\mathbb{R}$-linear mapping. Suppose that for each $x \in M$ there exists a semimartingale X on M such that $X_0 = x$ and such that for each $f \in C^\infty(M)$,*

$$f(X_t) - f(x) - \int_0^t Lf(X)\,dr$$

is a local martingale. Then L is necessarily a PDO of order at most 2.

In addition, X has "nice" trajectories (e.g., in the sense that $[f(X), f(X)] = 0$ for all $f \in C^\infty(M)$) if and only if L is first order.

Proof. As in the proof of Proposition 2.15, for all $f \in C^\infty(M, \mathbb{R}^r)$ and $\phi \in C^\infty(\mathbb{R}^r)$, we have

$$\left[L(\phi \circ f) - \sum_{i=1}^r (D_i\phi \circ f)(Lf^i) + \sum_{i,j=1}^r (D_iD_j\phi \circ f)\,\Gamma(f^i, f^j)\right](X) = 0,$$

so that L is a second-order PDO by Lemma 2.16. The second claim uses

$$d[f(X), g(X)] = 2\,\Gamma(f, g)(X)\,dt, \qquad f, g \in C^\infty(M).$$ □

3 Construction of stochastic flows

Flows to vector fields are classically constructed as solutions of ordinary differential equations on manifolds. In the same way, stochastic flows can be constructed as solutions to stochastic differential equations (SDEs) on manifolds. We start by recalling same basic facts about stochastic differential equations on $\mathbb{R}^n$.

3.1 Stochastic differential equations on Euclidean space.

Example 3.1 (SDEs on $\mathbb{R}^n$). Given $\beta\colon \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ and in addition a function

$$\sigma\colon \mathbb{R}_+ \times \mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^r, \mathbb{R}^n) \equiv \mathrm{Matr}(n \times r, \mathbb{R}),$$

let B be a Brownian motion on $\mathbb{R}^r$. Now one wants to find a continuous semimartingale Y on $\mathbb{R}^n$ such that

$$\boxed{dY_t = \beta(t, Y_t)\,dt + \sigma(t, Y_t) dB_t}$$

in the sense of Itô, i.e.,

$$Y_t = Y_0 + \int_0^t \beta(s, Y_s)\,ds + \int_0^t \sigma(s, Y_s)\,dB_s. \tag{3.1}$$

In Eq. (3.1) the first term describes the "systematic part" (*drift term*) in the evolution of Y, whereas the second integral represents the "fluctuating part" (*diffusion term*).

Definition 3.2. An $\mathbb{R}^n$-valued stochastic process $(Y_t)_{t\geq 0}$ is called an *Itô process* if it has a representation as

$$Y_t = Y_0 + \int_0^t K_s\,ds + \int_0^t H_s\,dB_s,$$

where

- Y_0 is $\mathscr{F}_0$-measurable;
- K_s and H_s are adapted processes taking values in $\mathbb{R}^n$ and $\mathrm{Hom}(\mathbb{R}^r, \mathbb{R}^n)$ respectively;
- $\mathbb{E}\big[\int_0^t |K_s|\,ds\big] < \infty$ and $\mathbb{E}\big[\int_0^t H_s^2\,ds\big] < \infty$ for each $t \geq 0$.

Proposition 3.3. *Let $\beta\colon \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma\colon \mathbb{R}_+ \times \mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^r, \mathbb{R}^n)$ be continuous functions. For a continuous semimartingale Y on $\mathbb{R}^n$, defined up to some predictable stopping time τ (i.e., there exists a sequence of stopping times $\tau_n < \tau$ with $\tau_n \uparrow \tau$), the following conditions are equivalent:*

(a) Y is a solution of the SDE

$$dY_t = \beta(t, Y_t)\,dt + \sigma(t, Y_t)\,dB_t \quad \textit{on } [0, \tau[\,, \tag{3.2}$$

i.e.,

$$Y_t = Y_0 + \int_0^t \beta(s, Y_s)\,ds + \int_0^t \sigma(s, Y_s)\,dB_s, \quad \forall 0 \leq t < \tau \textit{ a.s.}$$

(b) For all $f \in C^\infty(\mathbb{R}^n)$,

$$d(f \circ Y) = (Lf)(t,Y)\,dt + \sum_{k=1}^{n}\sum_{i=1}^{r} \sigma_{ki}(t,Y)\,D_k f(Y)\,dB_i \quad \text{on } [0,\tau[,$$

where

$$L = \sum_{k=1}^{n} \beta_k\,D_k + \frac{1}{2}\sum_{k,\ell=1}^{n} (\sigma\sigma^*)_{k\ell} D_k D_\ell,$$

where σ^* is a transpose of σ, and $(\sigma\sigma^*)_{k\ell} = \sum_{i=1}^r \sigma_{ki}\sigma_{\ell i}$. In particular, every solution of (3.2) is an L-diffusion on $[0,\tau[$ in the sense that

$$d(f \circ Y) - Lf(t,Y)\,dt = d(\text{local martingale}) \text{ on } [0,\tau[.$$

Proof. (a) $\Rightarrow$ (b) Let Y be a solution of SDE (3.2). Then

$$dY^k\,dY^\ell \equiv d[Y^k,Y^\ell] = (\sigma\sigma^*)_{k\ell}(t,Y)\,dt,$$

where $[Y^k,Y^\ell]$ represents quadratic covariation of Y^k and Y^ℓ. By Itô's formula we get

$$\begin{aligned} d(f\circ Y) &= \sum_{k=1}^{n} D_k f(Y)\Big(\beta_k(t,Y)\,dt + \sum_{i=1}^{r}\sigma_{ki}(t,Y)\,dB^i\Big) \\ &\quad + \frac{1}{2}\sum_{k,\ell=1}^{n} D_k D_\ell f(Y)\,\underbrace{(\sigma\sigma^*)_{k\ell}(t,Y)\,dt}_{=d[Y^k,Y^\ell]} \\ &= Lf(t,Y)\,dt + \sum_{k=1}^{n}\sum_{i=1}^{r}\sigma_{ki}(t,Y)\,D_k f(t,Y)\,dB_i \\ &= Lf(t,Y)\,dt + d(\text{local martingale}). \end{aligned}$$

(b) $\Rightarrow$ (a) Take $f(x) = x_\ell$. Then $D_k f = \delta_{k\ell}$ and $Lf = \beta_\ell$, thus

$$dY^\ell = \beta_\ell(t,Y)\,dt + \sum_{i=1}^{r}\sigma_{\ell i}(t,Y)dB^i \quad \text{for each } \ell = 1,\dots,n.$$

This shows that Y solves SDE (3.2) on $[0,\tau[$. $\square$

Proposition 3.4 (Itô SDE on $\mathbb{R}^n$; case of *global Lipschitz conditions*). *Let Z be a continuous semimartingale on $\mathbb{R}^r$ and*

$$\alpha\colon \mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^r,\mathbb{R}^n) \quad (= \mathrm{Matr}(n\times r;\mathbb{R}))$$

such that

$$\exists L > 0, \quad |\alpha(y) - \alpha(z)| \leq L|y - z|, \quad \forall y, z \in \mathbb{R}^n \quad \textit{(global Lipschitz conditions)}.$$

Then, for each $\mathscr{F}_0$-measurable $\mathbb{R}^n$-valued random variable x_0, there exists a unique continuous semimartingale $(X_t)_{t\geq 0}$ on $\mathbb{R}^n$ such that

$$dX = \alpha(X)\, dZ \quad \textit{and} \quad X_0 = x_0. \tag{3.3}$$

Uniqueness holds in the following sense: Suppose that Y is another continuous semimartingale such that $dY = \alpha(Y)\, dZ$ and $Y_0 = x_0$; then $X_t = Y_t$ for all t a.s.

Proof. The proof is standard in stochastic analysis; see for instance [51] or [30]. □

Proposition 3.5 (Itô SDEs on $\mathbb{R}^n$: the case of *local Lipschitz coefficients*). *Let Z be a continuous semimartingale on $\mathbb{R}^r$ and let*

$$\alpha : \mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^r, \mathbb{R}^n)$$

be locally Lipschitz, i.e., for each compact $K \subset \mathbb{R}^n$ there exists a constant $L_K > 0$ such that

$$\forall y, z \in K, \quad |\alpha(y) - \alpha(z)| \leq L_K|y - z|.$$

Then, for any x_0 $\mathscr{F}_0$-measurable, there exists a unique maximal solution $X|[0, \zeta[$ of the SDE

$$dX = \alpha(X)\, dZ, \quad X_0 = x_0.$$

Uniqueness holds in the sense that if $Y|[0, \xi[$ is another solution and $y_0 = x_0$, then $\xi \leq \zeta$ a.s. and $X|[0, \xi[= Y$.

Proof. The proof is reduced to Proposition 3.4 by a standard truncation method. We briefly sketch the argument, since it will be used several times in the sequel. Let $B(0, R) = \{x \in \mathbb{R}^n : |x| \leq R\}$ where $R = 1, 2, \ldots$ and choose test functions $\phi_R \in C_c^\infty(\mathbb{R}^n)$ such that $\phi_R|B(0, R) \equiv 1$. For $R > 0$ consider the "truncated SDE"

$$dX^R = \alpha^R(X^R)\, dZ, \quad X_0^R = x_0, \tag{3.4}$$

where $\alpha^R := \phi_R\, \alpha$ is now global Lipschitz. By Proposition 3.4 there is a unique solution X^R to (3.4). Then

$$X|[0, \tau_R[:= X^R|[0, \tau_R[$$

is well defined by uniqueness, where

$$\tau_R = \inf\left\{t \geq 0 : X_t^R \notin B(0, R)\right\}.$$

This finally defines X on the stochastic interval $[0, \zeta[$ where $\zeta = \sup_R \tau_R$. Uniqueness of X is deduced from the uniqueness of $X|[0, \tau_R[$. □

Example 3.6. Consider the following Itô SDE on $\mathbb{R}^n$:

$$dX = \underbrace{\beta(X)}_{n\times 1}\, dt + \underbrace{\sigma(X)}_{n\times r}\,\underbrace{dB}_{r\times 1}\,, \tag{3.5}$$

where B is Brownian motion on $\mathbb{R}^r$. Then the space-time process $Z_t = (t, B_t)$ is a semimartingale on $\mathbb{R}^{r+1}$ and SDE (3.5) can be written as

$$dX = \begin{pmatrix} \beta(X) \\ \sigma(X) \end{pmatrix} \begin{pmatrix} dt \\ dB \end{pmatrix} = \alpha(X)\, dZ,$$

where $\alpha(X) := \binom{\beta(X)}{\sigma(X)}$. Thus, under a local Lipschitz condition on the coefficients β and σ, the SDE

$$dX = \beta(X)\, dt + \sigma(X)\, dB \tag{3.6}$$

has a unique strong solution for every given initial condition x_0. By Proposition 3.3, maximal solutions of Eq. (3.6) are L-diffusions to the operator

$$L = \sum_{i=1}^{n} \beta_i \partial_i + \frac{1}{2} \sum_{i,j=1}^{n} (\sigma\sigma^*)_{ij} \partial_i \partial_j,$$

where $\partial_i = \frac{\partial}{\partial x_i}$ is the derivative in direction i.

3.2 Stratonovich differentials.

Definition 3.7. For continuous real-valued semimartingales X and Y let

$$X \circ dY := X\, dY + \frac{1}{2} d[X, Y]$$

be the *Stratonovich differential*. Here $X\, dY$ is the usual Itô differential and $d[X, Y] = dX\, dY$ the differential of the quadratic covariation of X and Y. The integral

$$\int_0^t X \circ dY = \int_0^t X\, dY + \frac{1}{2}[X, Y]_t \tag{3.7}$$

is called the *Stratonovich integral* of X with respect to Y.

Formula (3.7) gives the relation between the Stratonovich integral and the usual Itô integral. Since Stratonovich integrals can always be converted back to Itô integrals, their use in our context will be only formal and for the sake of convenient notation.

Remark 3.8. We have the following properties of the Stratonovich differential and Stratonovich integrals.

(1) Associativity: $X \circ (Y \circ dZ) = (XY) \circ dZ$, i.e.,

$$X \circ d\left(\int_0^{\bullet} Y \circ dZ\right) = (XY) \circ dZ.$$

Indeed, we have

$$\begin{aligned} X \circ (Y \circ dZ) &= X \circ d\left(\int_0^{\bullet} Y \circ dZ\right) \\ &= X\, d\left(\int_0^{\bullet} Y \circ dZ\right) + \frac{1}{2} dX\, d\left(\int_0^{\bullet} Y \circ dZ\right) \\ &= X(Y\, dZ) + \tfrac{1}{2}\, X\, dY\, dZ + \tfrac{1}{2} dX\left(Y\, dZ + \tfrac{1}{2}\, dY\, dZ\right) \\ &= (XY)dZ + \tfrac{1}{2}\,(X\, dY + Y\, dX + dX\, dY)dZ \\ &= (XY)dZ + \tfrac{1}{2}\, d(XY)dZ \\ &= (XY) \circ dZ. \end{aligned}$$

(2) Product rule: $d(XY) = X \circ dY + Y \circ dX$.

Proof. By Itô's formula we have

$$d(XY) = X\, dY + Y\, dX + dX\, dY = X \circ dY + Y \circ dX. \qquad \square$$

Proposition 3.9 (Itô–Stratonovich formula). *Let X be a continuous $\mathbb{R}^n$-valued semimartingale and $f \in C^3(\mathbb{R}^n)$. Then*

$$d(f \circ X) = \sum_{i=1}^{n} (D_i f)(X) \circ dX^i \equiv \langle \nabla f(X), \circ\, dX \rangle. \tag{3.8}$$

Proof. By Itô's formula, we have

$$d(D_i f(X)) = \sum_{k=1}^{n} (D_i D_k f)(X)\, dX^k + \frac{1}{2} \sum_{k,\ell=1}^{n} (D_i D_k D_\ell f)(X)\, dX^k dX^\ell.$$

Hence we get

$$\begin{aligned} \sum_{i=1}^{n} (D_i f)(X) \circ dX^i &= \sum_{i=1}^{n} (D_i f)(X)\, dX^i + \frac{1}{2} \sum_{i=1}^{n} d(D_i f(X)) dX^i \\ &= \sum_{i=1}^{n} (D_i f)(X)\, dX^i + \frac{1}{2} \sum_{i,k=1}^{n} (D_i D_k f(X))\, dX^k dX^i \\ &= d(f \circ X). \end{aligned} \qquad \square$$

Formula (3.8) shows the main advantage of the Stratonovich differential: it converts Itô's formula into the usual chain rule of classical analysis. Hence, at least formally, classical differential calculus can be applied in calculations involving Stratonovich differentials.

Proposition 3.10. *Let $\beta\colon \mathbb{R}_+\times\mathbb{R}^n \to \mathbb{R}^n$ be continuous, $\sigma\colon \mathbb{R}_+\times\mathbb{R}^n \to \mathrm{Hom}(\mathbb{R}^r, \mathbb{R}^n)$ be C^1. Furthermore, let B be a Brownian motion on $\mathbb{R}^r$. For a semimartingale Y on $\mathbb{R}^n$ (defined up to some predictable stopping time τ) the following conditions are equivalent:*

(i) The semimartingale Y is a solution of the Stratonovich SDE

$$dY = \beta(t,Y)\,dt + \sigma(t,Y)\circ dB, \tag{3.9}$$

i.e.,

$$Y_t = Y_0 + \int_0^t \beta(s,Y_s)\,ds + \int_0^t \sigma(s,Y_s)\circ dB_s, \quad \text{for } 0 \le t < \tau \text{ a.s.}$$

(ii) For all $f \in C^\infty(\mathbb{R}^n)$,

$$d(f\circ Y) = (Lf)(t,Y)\,dt + \sum_{k=1}^r (A_k f)(t,Y)dB^k \quad \text{on } [0,\tau[,$$

where

$$L = A_0 + \frac{1}{2}\sum_{i=1}^r A_k^2,$$

with the vector fields $A_i \in \Gamma(T\mathbb{R}^n)$ defined as

$$A_0 = \sum_{i=1}^n \beta_i D_i, \quad A_k = \sum_{i=1}^n \sigma_{ik} D_i, \quad k = 1,\dots,r. \tag{3.10}$$

Proof. (i) ⇒ (ii) By the Itô–Stratonovich formula (Proposition 3.9) we have

$$\begin{aligned}
d(f\circ Y) &= \sum_{i=1}^n (D_i f)(Y)\circ dY^i \\
&= \sum_{i=1}^n (D_i f)(Y)\,\beta_i(t,Y)\,dt + \sum_{i=1}^n (D_i f)(Y)\left(\sum_{k=1}^r \sigma_{ik}(t,Y)\circ dB^k\right) \\
&= (A_0 f)(t,Y)\,dt + \sum_{k=1}^r (A_k f)(t,Y)\circ dB^k \\
&= (A_0 f)(t,Y)\,dt + \sum_{k=1}^r (A_k f)(t,Y)\,dB_k + \frac{1}{2}\sum_{k=1}^r d\big((A_k f)(t,Y)\big)dB^k.
\end{aligned}$$

Since

$$d(A_k f(t,Y)) = \partial_t (A_k f)(t,Y)\,dt + (A_0 A_k f)(t,Y)\,dt + \sum_{\ell=1}^{r} (A_\ell A_k f)(t,Y) \circ dB^\ell,$$

we observe that

$$d(A_k f(t,Y))\,dB^k = (A_k^2 f)(t,Y)\,dt,$$

and hence

$$d(f \circ Y) = \underbrace{\left((A_0 f)(t,Y) + \frac{1}{2} \sum_{k=1}^{r} (A_k^2 f)(t,Y) \right)}_{=(Lf)(t,Y)} dt + \sum_{k=1}^{r} (A_k f)(t,Y)\,dB^k.$$

(ii) $\Rightarrow$ (i) It is sufficient to take $f(x) = x_\ell$. □

Corollary 3.11. *Solutions to the Stratonovich SDE*

$$dY = \beta(t,Y)\,dt + \sigma(t,Y) \circ dB$$

define L-diffusions for the operator

$$L = A_0 + \frac{1}{2} \sum_{i=1}^{r} A_i^2 \quad \textit{with } A_0, A_1, \ldots, A_r \textit{ as in Eq. } (3.10),$$

in the sense that

$$d(f \circ Y) - (Lf)(t,Y)\,dt \overset{\mathrm{m}}{=} 0,$$

for all $f \in C^\infty(\mathbb{R}^n)$.

3.3 Stochastic differential equations on manifolds.

In this section we describe the construction of L-diffusions as solutions of stochastic differential equations on manifolds [18, 27].

Definition 3.12. Let M be a differentiable manifold, $\pi\colon TM \to M$ its tangent bundle, and E a finite-dimensional vector space (without restrictions $E = \mathbb{R}^r$). A *stochastic differential equation on* M is a pair (A, Z) where

(1) Z is a semimartingale taking values in E;

(2) $A\colon M \times E \to TM$ is a smooth homomorphism of vector bundles over M, i.e.,

$$(x,e) \longmapsto A(x)e := A(x,e),$$

$$\begin{tikzcd} M \times E \arrow[r, "A"'] \arrow[d, "\mathrm{pr}_1"'] & TM \arrow[d, "\pi"] \\ M \arrow[r, "\mathrm{id}"'] & M. \end{tikzcd}$$

Remark 3.13. Formally the homomorphism A may be considered as section $A \in \Gamma(E^* \otimes TM)$. In particular, we have

$$\begin{cases} \forall x \in M \text{ fixed}, & A(x) \in \mathrm{Hom}(E, T_x M), \\ \forall e \in E \text{ fixed}, & A(\cdot)e \in \Gamma(TM). \end{cases}$$

Notation 3.14. For the SDE (A, Z) we also write

$$\boxed{dX = A(X) \circ dZ}$$

or

$$\boxed{dX = \sum_{i=1}^{r} A_i(X) \circ dZ^i,}$$

where $A_i = A(\cdot)e_i \in \Gamma(TM)$ and $e_1, \dots, e_r$ is a basis of E.

Definition 3.15. Let (A, Z) be an SDE on M and let $x_0 : \Omega \to M$ be $\mathscr{F}_0$-measurable. An adapted continuous process $X|[0, \zeta[\equiv (X_t)_{t<\zeta}$ taking values in M, defined up to the stopping time ζ, is called a solution to the SDE

$$dX = A(X) \circ dZ \tag{3.11}$$

with initial condition $X_0 = x_0$ if, for all $f \in C_c^\infty(M)$, the following conditions are satisfied:

(i) $f \circ X$ is a semimartingale.

(ii) For any stopping time τ such that $0 \le \tau < \zeta$, we have

$$f(X_\tau) = f(X_0) + \int_0^\tau (df)_{X_s} A(X_s) \circ dZ_s. \tag{3.12}$$

We call X a *maximal solution* of the SDE (3.11) if

$$\{\zeta < \infty\} \subset \left\{ \lim_{t \uparrow \zeta} X_t = \infty \text{ in } \hat{M} = M \,\dot{\cup}\, \{\infty\} \right\} \text{ a.s.}$$

Note: The integral in (3.12) is defined using

$$E \xrightarrow{A(x)} T_x M \xrightarrow{(df)_x} \mathbb{R}, \quad x \in M.$$

Remark 3.16. We adopt the convention $X_t(\omega) := \infty$ for $\zeta(\omega) \le t < \infty$ and $f(\infty) = 0$ for $f \in C_c^\infty(M)$. Then we may write, for all $t \ge 0$,

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t (df)_{X_s} A(X_s) \circ dZ_s \\ &= f(X_0) + \sum_{i=1}^r \int_0^t (df)_{X_s} A_i(X_s) \circ dZ_s^i \\ &= f(X_0) + \sum_{i=1}^r \int_0^t (A_i f)(X_s) \circ dZ_s^i \quad \text{with } A_i = A(\cdot)e_i. \end{aligned}$$

Example 3.17. Let $E = \mathbb{R}^{r+1}$ and $Z = (t, Z^1, \dots, Z^r)$, where $(Z^1, \dots, Z^r)$ is a Brownian motion on $\mathbb{R}^r$. Denote the standard basis of $\mathbb{R}^{r+1}$ by $(e_0, e_1, \dots, e_r)$. Letting

$$A\colon M \times E \to TM$$

be a homomorphism of vector bundles over M, we consider the vector fields

$$A_i := A(\cdot)e_i \in \Gamma(TM), \quad i = 0, 1, \dots, r.$$

Then the SDE

$$\boxed{dX = A(X) \circ dZ} \tag{3.13}$$

may be written as

$$\boxed{dX = A_0(X)\,dt + \sum_{i=1}^r A_i(X) \circ dZ^i}$$

and for each $f \in C_c^\infty(M)$ we have

$$\begin{aligned} d(f \circ X) &= (df)_X A(X) \circ dZ \\ &= \sum_{i=0}^r (df)_X A(X) e_i \circ dZ^i \\ &= \sum_{i=0}^r (df)_X A_i(X) \circ dZ^i \\ &= \sum_{i=0}^r (A_i f)(X) \circ dZ^i \end{aligned}$$

$$
\begin{aligned}
&= (A_0 f)(X)\, dt + \sum_{i=1}^{r} (A_i f)(X) \circ dZ^i \\
&= (A_0 f)(X)\, dt + \sum_{i=1}^{r} \left[(A_i f)(X)\, dZ^i + \frac{1}{2} d\big((A_i f)(X)\big)\, dZ^i \right].
\end{aligned}
$$

Taking into account that

$$
d\big((A_i f)(X)\big) = \sum_{j=1}^{r} (A_j A_i f)(X)\, dZ^j + d(\text{terms of bounded variation}),
$$

we see that

$$
d\big((A_i f)(X)\big)\, dZ^i = (A_i^2 f)(X)\, dt,
$$

where we used that $dZ^i dZ^j = \delta_{ij}\, dt$ for $1 \le i, j \le r$. Hence we get

$$
\begin{aligned}
d(f \circ X) &= (A_0 f)(X)\, dt + \frac{1}{2} \sum_{j=1}^{r} (A_i^2 f)(X)\, dt + \sum_{i=1}^{r} (A_i f)(X)\, dZ^i \\
&= (L f)(X)\, dt + \sum_{i=1}^{r} (A_i f)(X)\, dZ^i.
\end{aligned}
$$

Corollary 3.18. *Let $L = A_0 + \frac{1}{2} \sum_{i=1}^{r} A_i^2$ and let X be a solution to Eq. (3.13). Then, for all $f \in C_c^\infty(M)$,*

$$
d(f \circ X) - (L f)(X)\, dt \overset{\mathrm{m}}{=} 0,
$$

where $\overset{\mathrm{m}}{=}$ denotes equality modulo differentials of martingales. In other words, maximal solutions to the SDE

$$
\boxed{dX = A(X) \circ dZ}
$$

are L-diffusions to the operator $L = A_0 + \frac{1}{2} \sum_{i=1}^{r} A_i^2$.

Theorem 3.19 (SDE: Existence and uniqueness of solutions; $M = \mathbb{R}^n$). *Let (A, Z) be an SDE on $M = \mathbb{R}^n$ and x_0 an $\mathscr{F}_0$-measurable random variable taking values in $\mathbb{R}^n$. Then there exists a unique maximal solution X (with maximal lifetime $\zeta > 0$ a.s.) of the SDE*

$$
dX = A(X) \circ dZ \tag{3.14}
$$

with initial condition $X_0 = x_0$. Uniqueness holds in the following sense: If $Y|[0, \xi[$ is another solution of (3.14) to the same initial condition, then $\xi \le \zeta$ a.s. and $X|[0, \xi[= Y$ a.s.

Proof. As in the proof of Proposition 3.5 let $B(0,R)=\{x\in\mathbb{R}^n\colon |x|\le R\}$, where $R=1,2,\dots$, and choose test functions $\phi_R\in C_c^\infty(\mathbb{R}^n)$ such that $\phi_R|B(0,R)\equiv 1$. Since

$$A\in\Gamma(\mathrm{Hom}(\mathbb{R}^r,TM)),$$

we have for each $x\in\mathbb{R}^n$ the linear map

$$A(x)\colon \mathbb{R}^r\to T_xM.$$

In this way A gives rise to a smooth map $\mathbb{R}^n\to\mathrm{Matr}(n\times r;\mathbb{R})$.

Consider now the "truncated SDE"

$$dX^R=A^R(X^R)\circ dZ, \tag{3.15}$$

where $A^R=\phi_R A$. By Proposition 3.4, the truncated SDE (3.15) has a unique global solution X^R with initial condition $X_0^R=x_0$, i.e., for each R there exists a continuous $\mathbb{R}^n$-valued semimartingale $(X_t^R)_{t\ge 0}$ satisfying $X_0^R=x_0$ such that (3.15) holds in the Itô–Stratonovich sense. In terms of the stopping times

$$\tau_R:=\inf\left\{t\ge 0: X_t^R\notin B(0,R)\right\},$$

we have for $R<R'$,

$$X^{R'}|[0,\tau_R[\,=X^R|[0,\tau_R[\quad \text{a.s.}$$

Hence a stochastic process X (with lifetime $\zeta=\lim_{R\uparrow\infty}\tau_R$) is well defined via

$$X|[0,\tau_R[\,=X^R|[0,\tau_R[.$$

For each $f\in C_c^\infty(\mathbb{R}^n)$ such that $\mathrm{supp}\,(f)\subset B(0,R)$ (with R sufficiently large), we have

$$\begin{aligned}
d(f\circ X)&=d(f\circ X^R)\\
&=\sum_{k=1}^n (D_kf(X^R))\circ d(X^R)^k \quad \text{(using the Itô–Stratonovich formula)}\\
&=\langle\nabla f(X^R),\circ\, dX^R\rangle\\
&=\langle\nabla f(X^R),\phi_R(X^R)\,A(X^R)\circ dZ\rangle\\
&=\langle\nabla f(X),A(X)\circ dZ\rangle\\
&=\sum_{i=1}^r\langle\nabla f(X),A_i(X)\circ dZ^i\rangle\\
&=\sum_{i=1}^r (df)_X A_i(X)\circ dZ^i\\
&=(df)_X A(X)\circ dZ.
\end{aligned}$$

Hence, X is the unique solution to Eq. (3.14) with initial condition $X_0 = x_0$. Note that X is a solution of $dX = A(X) \circ dZ$ in the Itô–Stratonovich sense (in $\mathbb{R}^n$) if and only if, for all $f \in C_c^\infty(\mathbb{R}^n)$,

$$d(f \circ X) = (df)_X A(X) \circ dZ. \qquad \square$$

Theorem 3.20 (SDE: Existence and uniqueness of solutions; general case). *Let (A, Z) be an SDE on a differentiable manifold M and let $x_0 \colon \Omega \to M$ be $\mathscr{F}_0$-measurable. There exists a unique maximal solution $X|[0, \zeta[$ (where $\zeta > 0$ a.s.) of the SDE*

$$dX = A(X) \circ dZ$$

with initial condition $X_0 = x_0$. Uniqueness holds in the sense that if $Y|[0, \xi[$ is another solution with $Y_0 = x_0$, then $\xi \leq \zeta$ a.s. and $X|[0, \xi[= Y$ a.s.

We shall reduce Theorem 3.20 to Theorem 3.19 via embedding the manifold M into a high-dimensional Euclidean space.

Whitney's embedding theorem. *Each manifold M of dimension n can be embedded into $\mathbb{R}^{n+k}$ as a closed submanifold (for k sufficiently large, e.g., $k = n+1$), i.e.,*

$$M \hookrightarrow \iota(M) \subset \mathbb{R}^{n+k},$$

where $\iota : M \to \iota(M)$ is a diffeomorphism and $\iota(M) \subset \mathbb{R}^{n+k}$ is a closed submanifold.

Proof (of Theorem 3.20). We choose a Whitney embedding (in general not intrinsic)

$$M \underset{\text{diffeom.}}{\overset{\iota}{\hookrightarrow}} \iota(M) \subset \mathbb{R}^{n+k}$$

and identify M and $\iota(M)$; in particular, for each $x \in M$ the tangent space T_xM is then a linear subspace of $\mathbb{R}^{n+k}$ according to

$$T_xM \overset{d\iota_x}{\hookrightarrow} T_x\mathbb{R}^{n+k} \equiv \mathbb{R}^{n+k}.$$

Vector fields $A_1, \ldots, A_r \in \Gamma(TM)$ can be extended to vector fields

$$\bar{A}_1, \ldots, \bar{A}_r \in \Gamma(T\mathbb{R}^{n+k}) \equiv C^\infty(\mathbb{R}^{n+k}; \mathbb{R}^{n+k}) \quad \text{with } \bar{A}_i|M = A_i,$$

i.e., $\bar{A}_i \circ \iota = d\iota \circ A_i$. Hence a given bundle map

$$A : M \times \mathbb{R}^r \to TM, \quad (x, z) \mapsto A(x)z = \sum_{i=1}^r A_i(x)z^i$$

has a continuation

$$\bar{A} : \mathbb{R}^{n+k} \times \mathbb{R}^r \to \mathbb{R}^{n+k} \times \mathbb{R}^{n+k}, \quad (x, z) \mapsto \bar{A}(x)z = \sum_{i=1}^{r} \bar{A}_i(x)z^i.$$

In place of the original SDE

$$dX = A(X) \circ dZ \ \text{ on } M, \tag{$*$}$$

the idea is to consider the SDE

$$dX = \bar{A}(X) \circ dZ \ \text{ on } \mathbb{R}^{n+k}. \tag{$\bar{*}$}$$

First of all, it is clear that any solution of $(*)$ in M provides a solution of $(\bar{*})$ in $\mathbb{R}^{n+k}$. More precisely, if X is a solution to $(*)$ with starting value $X_0 = x_0$, then $\bar{X} := \iota \circ X$ solves equation $(\bar{*})$ with starting value $\bar{X}_0 = \iota \circ x_0$. Indeed, if $\bar{f} \in C_c^\infty(\mathbb{R}^{n+k})$, then $f := \bar{f}|M = \bar{f} \circ \iota \in C_c^\infty(M)$, and we have

$$\begin{aligned} d(\bar{f} \circ \bar{X}) = d(f \circ X) &= \sum_{i=1}^{r} (df)_X \, A_i(X) \circ dZ^i \\ &= \sum_{i=1}^{r} (d\bar{f})_{\bar{X}} \, (d\iota)_X \, A_i(X) \circ dZ^i \\ &= \sum_{i=1}^{r} (d\bar{f})_{\bar{X}} \, \bar{A}_i(\iota \circ X) \circ dZ^i \\ &= \sum_{i=1}^{r} (d\bar{f})_{\bar{X}} \, \bar{A}_i(\bar{X}) \circ dZ^i. \end{aligned}$$

This implies, in particular, uniqueness of solutions to $(*)$, since Eq. $(\bar{*})$ has a unique solution to a given initial condition.

To establish the existence of solutions to $(*)$ we first remark that any test function $f \in C_c^\infty(M)$ has a continuation $\bar{f} \in C_c^\infty(\mathbb{R}^{n+k})$ such that $\bar{f}|M \equiv \bar{f} \circ \iota = f$. We make the following important observation:

Each solution $X|[0, \zeta[$ of $(\bar{*})$ in $\mathbb{R}^{n+k}$ with $X_0 = x_0$ which stays on M for $t < \zeta$ (where x_0 is an M-valued $\mathscr{F}_0$-measurable random variable) gives a solution of $(*)$.

Hence, to complete the proof it is sufficient to show the following lemma. □

Lemma 3.21. *If $X|[0, \zeta[$ is the maximal solution of $(\bar{*})$ in $\mathbb{R}^{n+k}$ with $X_0 = x_0$, then*

$$\{t < \zeta\} \subset \{X_t \in M\}, \quad \textit{for all } t \ \textit{a.s.}$$

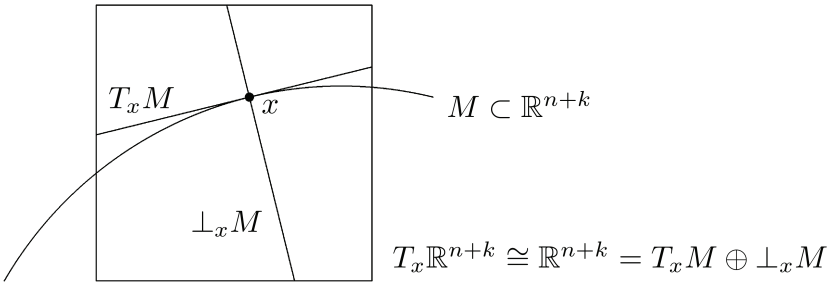
Observe that it is enough to verify Lemma 3.21 for one specific continuation $\bar{A}$ of A.

Proof (of Lemma 3.21). Let

$$\perp M = \left\{(x,v) \in M \times \mathbb{R}^{n+k} \middle| v \in (T_x M)^\perp\right\}$$

be the normal bundle of M and consider M embedded into $\perp M$ as a zero section:

$$M \hookrightarrow \perp M, \quad x \mapsto (x,0).$$



Fact: There is a smooth function $\varepsilon\colon M \to]0,\infty[$ such that the map

$$\begin{aligned}\tau_\varepsilon(M) := \big\{(x,v) \in \perp M\colon\ |v| < \varepsilon(x)\big\} &\xrightarrow{\cong} \bigcup_{x\in M} \big\{y \in \mathbb{R}^{n+k} : |y-x| < \varepsilon(x)\big\},\\ (x,v) &\longmapsto x+v\end{aligned}$$

is a diffeomorphism from the tubular neighborhood $\tau_\varepsilon(M)$ of M of radius ε onto the indicated part in $\mathbb{R}^{n+k}$. This follows from the local inversion theorem since the given map has full rank along the zero section of $\perp M$.

Note that both

$$\pi\colon \tau_\varepsilon(M) \to M, \quad (x,v) \mapsto x \quad \text{and} \quad \operatorname{dist}^2(\cdot, M)\colon \tau_\varepsilon(M) \to \mathbb{R}, \quad (x,v) \mapsto |v|^2$$

are smooth maps.

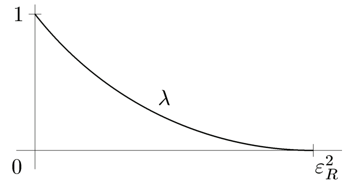
Now letting $R > 0$ be sufficiently large such that

$$M \cap B(0, R+1) \neq \varnothing,$$

then

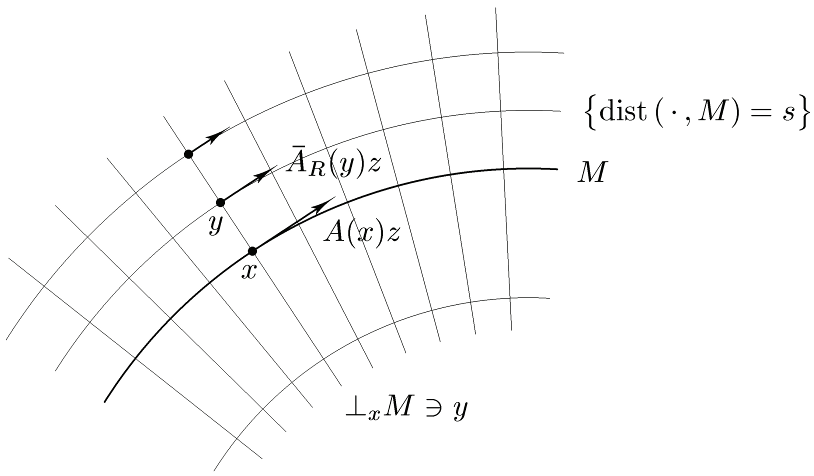
$$\varepsilon_R = \inf\{\varepsilon(x) \mid x \in M \cap B(0,R+1)\} > 0.$$

We choose a decreasing smooth function $\lambda : [0, \infty[\to [0, 1]$ of the form



and a test function $0 \le \varphi \in C_c^\infty(\mathbb{R}^{n+k})$ such that $\varphi|B(0, R) \equiv 1$ and $\operatorname{supp}(\varphi) \subset B(0, R + 1)$. Consider the map

$$\begin{aligned}
&\bar{A}^R : \mathbb{R}^{n+k} \times \mathbb{R}^r \to \mathbb{R}^{n+k} \times \mathbb{R}^{n+k},\\
&\bar{A}^R(y, z) := \begin{cases} \varphi(y)\, \lambda(\operatorname{dist}^2(y, M))\, A(\pi(y))\, z & \text{if } y \in \tau_\varepsilon(M),\\ 0 & \text{if } y \notin \tau_\varepsilon(M).\end{cases}
\end{aligned}$$



Let X be the solution of

$$dX = \bar{A}^R(X) \circ dZ, \quad X_0 = x_0.$$

Consider the test function $f \in C_c^\infty(\mathbb{R}^{n+k})$ given as

$$f(y) = \varphi(y)\, \lambda(\operatorname{dist}^2(y, M)).$$

Then

$$\begin{aligned}
d(f \circ X) &= (df)_X \bar{A}^R(X) \circ dZ\\
&= \langle \nabla f(X),\, \bar{A}^R(X) \circ dZ\rangle\\
&= 0 \quad \text{on } [0, \tau_R[,
\end{aligned}$$

where $\tau_R := \inf\{t \geq 0 \colon X_t \notin B(0,R)\}$. Indeed, f is constant on each submanifold of the form
$$\{\mathrm{dist}(\cdot, M) = s\} \cap B(0,R), \quad s < \varepsilon_R,$$
whereas $\bar{A}^R(y,z)$ is tangent to such submanifolds. Thus, for all $y \in B(0,R)$ and $z \in \mathbb{R}^r$,
$$\nabla f(y) \perp \bar{A}^R(y)z.$$
Hence, for any solution X of $(\bar{*})$, we obtain that
$$f(X) \equiv \text{constant on } [0, \tau_R[\ \text{ a.s.}$$
Since R is arbitrary, this completes the proof of the lemma. □

Solutions to an SDE on M of the type (3.11) are by definition semimartingales on M as defined above: A continuous adapted process X with values in M is a *semimartingale* on M if, for each $f \in C_c^\infty(M)$, the composition $f \circ X$ provides a continuous real-valued semimartingale. It is easy to see that each M-valued semimartingale can be obtained as the solution of an SDE on M.

Theorem 3.22 (Manifold-valued semimartingales as solutions of an SDE). *Every semimartingale on a manifold M is given as the solution of an SDE of type* (3.11).

Proof. Let X be an arbitrary semimartingale on M. Without loss of generality (after an eventual change of time), we may assume that X has infinite lifetime. Choosing a Whitney embedding $\iota \colon M \hookrightarrow \mathbb{R}^{n+k}$ we may consider the semimartingale $Z := \iota \circ X$ taking values in $E := \mathbb{R}^{n+k}$. Let $A \colon M \times E \to TM$ be the bundle homomorphism which is fiberwise the orthogonal projection $A(x) \colon \mathbb{R}^{n+k} \to T_xM$ of $\mathbb{R}^{n+k}$ onto $T_xM \subset T_x\mathbb{R}^{n+k} = \mathbb{R}^{n+k}$. We show that X solves the equation
$$dX = A(X) \circ dZ.$$
Let $f \in C_c^\infty(M)$ be given. We choose a continuation $\bar{f} \in C_c^\infty(\mathbb{R}^{n+k})$, where $\bar{f} \circ \iota = f$ such that $\bar{f}$ is constant locally about M on the normal subspaces $\perp_x M$ (this is $\bar{f}(y) = f(x)$ for $y \in \perp_x M$ sufficiently small). Now let $x \in M$ and $z \in \mathbb{R}^{n+k}$. By decomposing $z = z_0 + z^\perp$, where $z_0 \in T_xM$ and $z^\perp \in \perp_x M$, we obtain
$$(df)_x\, A(x)z = (d\bar{f})_{\iota(x)}\, (d\iota)_x\, A(x)z = (d\bar{f})_{\iota(x)}\, z_0 = (d\bar{f})_{\iota(x)}\, z.$$
But then
$$\begin{aligned} d(f \circ X) = d(\bar{f} \circ \iota \circ X) &= \sum_{i=1}^{n+k} \big(D_i \bar{f}\big)\, (\iota \circ X) \circ dZ^i \\ &= \sum_{i=1}^{n+k} (df)_X A(X) e_i \circ dZ^i = (df)_X A(X) \circ dZ, \end{aligned}$$
which gives the claim. □

4 Some probabilistic formulas for solutions of PDEs

Let L be a second-order partial differentiable operator on M, e.g., M is a general differentiable manifold and L is given in so-called Hörmander form as

$$L = A_0 + \frac{1}{2}\sum_{i=1}^{r} A_i^2. \tag{4.1}$$

For $x \in M$, let $X_t(x)$ be an L-diffusion, starting from x at time $t = 0$, i.e., $X_0(x) = x$. Recall that $X_t(x)$ can be constructed as the solution to the SDE on M,

$$\begin{cases} dX = A_0(X)\,dt + \sum_{i=1}^{r} A_i(X) \circ dB^i, \\ X_0 = x, \end{cases}$$

where B denotes Brownian motion on $\mathbb{R}^r$. Sometimes one starts with a partial differentiable operator L on M which locally in a chart (h, U) is written as

$$L|U = \sum_{i=1}^{n} b_i \partial_i + \sum_{i,j=1}^{n} (\sigma\sigma^*)_{ij}\, \partial_i \partial_j, \tag{4.2}$$

where $b \in C^\infty(U, \mathbb{R}^n)$ and $a \in C^\infty(U, \mathbb{R}^r \otimes \mathbb{R}^n)$ (using the notation $\partial_i = \frac{\partial}{\partial h_i}$). It is straightforward to rewrite such an operator in Hörmander form (4.1) and then to construct an L-diffusion by solving a Stratonovich SDE.

In the special case $M = \mathbb{R}^n$ and

$$L = \sum_{i=1}^{n} b_i \partial_i + \frac{1}{2}\sum_{i,j=1}^{n} (\sigma\sigma^*)_{ij} \partial_i \partial_j,$$

an L-diffusion can be constructed directly as the solution of the Itô SDE on $\mathbb{R}^n$:

$$\begin{cases} dX = b(X)\,dt + \sigma(X)\,dB, \\ X_0 = x, \end{cases}$$

where B is again a Brownian motion on $\mathbb{R}^r$.

4.1 Feynman–Kac formula.

Let L be as in Eq. (4.1). Suppose that the lifetime of $X_t(x)$ is infinite a.s. for all $x \in M$.

Proposition 4.1 (Feynman–Kac formula). *Let $f\colon M \to \mathbb{R}$ be continuous and bounded and let $V\colon M \to \mathbb{R}$ be continuous and bounded above, i.e., $V(x) \le K$*

for some constant $K > 0$. Let $u\colon \mathbb{R}_+ \times M \to \mathbb{R}$ be a bounded solution of the following "initial value problem"

$$\begin{cases} \frac{\partial}{\partial t} u = Lu + Vu, \\ u|_{t=0} = f, \end{cases}$$

i.e.,

$$\begin{cases} \frac{\partial}{\partial t} u(t,\cdot) = Lu(t,\cdot) + V(\cdot)\, u(t,\cdot), \\ u(0,\cdot) = f(\cdot). \end{cases}$$

Then the solution u is given by the formula

$$u(t,x) = \mathbb{E}\left[\exp\left(\int_0^t V(X_s(x))\, ds\right) f(X_t(x))\right].$$

Remark 4.2. Operators of the form $H = L + V$ (where V is the multiplication operator by V) are called *Schrödinger operators*, for instance, $H = \frac{1}{2}\triangle + V$. The function V is called *potential*. If H is (essentially) self-adjoint, then

$$u(t,\cdot) = e^{tH} f$$

by *semigroup theory*.

Proof (of Proposition 4.1). Fix $t > 0$ and consider the process $Y_s := A_s Z_s$, where

$$\begin{cases} A_s := \exp\left(\int_0^s V(X_r(x))\, dr\right), \\ Z_s := u(t-s, X_s(x)). \end{cases}$$

We will show that $(Y_s)_{0\le s\le t}$ is a martingale in our setting.

Indeed, first note that by Itô's formula,

$$dZ_s = \Big(\partial_s u(t-s,\cdot) + Lu(t-s,\cdot)\Big)(X_s(x))\, ds + dN_s,$$

where N_s is a local martingale. Thus, since A_s is of bounded variation, we have

$$\begin{aligned} dY_s &= Z_s dA_s + A_s dZ_s \\ &= Z_s A_s V(X_s(x))\, ds + A_s\Big(\partial_s u(t-s,\cdot) + Lu(t-s,\cdot)\Big)(X_s(x))\, ds + A_s dN_s \\ &= A_s \underbrace{\left(-\partial_t u + Lu + Vu\right)}_{=0}(t-s, X_s(x))\, ds + A_s dN_s. \end{aligned}$$

Hence $(Y_s)_{0\le s\le t}$ is a local martingale, and as it is bounded, $(Y_s)_{0\le s\le t}$ is a true martingale. In particular, by taking expectations we obtain

$$\begin{aligned} u(t,x) = \mathbb{E}\,[Y_0] = \mathbb{E}[Y_t] &= \mathbb{E}\left[\exp\left(\int_0^t V(X_r(x))\,dr\right) u(0,X_t(x))\right] \\ &= \mathbb{E}\left[\exp\left(\int_0^t V(X_r(x))\,dr\right) f(X_t(x))\right]. \end{aligned}$$

□

4.2 Elliptic boundary value problems. Let L be a second-order partial differential operator on a differential manifold M, e.g.,

$$L = A_0 + \frac{1}{2}\sum_{i=1}^r A_i^2 \quad \text{on a differential manifold } M, \text{ or}$$

$$L = \sum_{i=1}^n b_i \partial_i + \frac{1}{2}\sum_{i,j=1}^n (\sigma\sigma^*)_{ij}\partial_i\partial_j \quad \text{in local coordinates on } M.$$

Remark 4.3 (Ellipticity).

(1) The *diffusion vector fields* $A_1, \dots, A_r$ define for each $x \in M$ a linear map

$$A(x)\colon\ \mathbb{R}^r \to T_xM, \quad z \mapsto \sum_{i=1}^r A_i(x) z_i.$$

The operator

$$L = A_0 + \frac{1}{2}\sum_{i=1}^r A_i^2$$

is called *elliptic* on some subset $D \subset M$ if the map $A(x)$ is *surjective* for each $x \in D$.

(2) Similarly, an operator of the type

$$L = \sum_{i=1}^n b_i \partial_i + \frac{1}{2}\sum_{i,j=1}^n (\sigma\sigma^*)_{ij}\partial_i\partial_j$$

is called *elliptic* on some subset $D \subset M$ if the linear map

$$\sigma(x)\colon\ \mathbb{R}^r \to \mathbb{R}^n, \quad z \mapsto \underbrace{\sigma(x)}_{n\times r}\, z$$

is *surjective* for each $x \in D$.

It is easily checked that both notions of ellipticity are compatible.

Note 4.4. The following conditions are equivalent:

$$\begin{aligned}\sigma(x) \text{ is surjective} &\iff \sigma^*(x) \text{ is injective}\\ &\iff a(x) := \sigma(x)\sigma^*(x) \text{ is invertible}\\ &\iff \langle a(x)v, v\rangle > 0, \quad \forall 0 \neq v \in \mathbb{R}^n.\end{aligned}$$

Example 4.5 (Expected hitting time of a boundary). Let $\varnothing \neq D \subsetneq M$ be some open, relatively compact domain with boundary ∂D. Suppose that there exists a solution $u \in C^2(D) \cap C(\bar{D})$ to the problem

$$\begin{cases} Lu = -1 & \text{on } D,\\ u|_{\partial D} = 0. \end{cases} \tag{4.3}$$

(For instance, if L is elliptic on $\bar{D}$ and the boundary ∂D is smooth, it is well known in classical PDE theory that such a solution exists.)

Let $X_t(x)$ be an L-diffusion such that $X_0(x) = x$ and define

$$\tau_D(x) = \inf\{t > 0 : X_t(x) \in \partial D\}\,.$$

Then, for each $x \in D$,

$$\boxed{u(x) = \mathbb{E}\big[\tau_D(x)\big].}$$

In particular, we see that $u > 0$ on D.

Proof. For $x \in D$, let $X_t = X_t(x)$ and $\tau_D = \tau_D(x)$. We know that the process

$$u(X_{t\wedge\tau_D}) - u(x) - \int_0^{t\wedge\tau_D} Lu(X_s)\,ds, \quad t \geq 0$$

is a martingale (starting at 0), and hence

$$\mathbb{E}[u(X_{t\wedge\tau_D})] - u(x) = \mathbb{E}\bigg[\int_0^{t\wedge\tau_D} \underbrace{Lu(X_s)}_{=-1}\,ds\bigg].$$

This shows that

$$\mathbb{E}[t \wedge \tau_D] = u(x) - \mathbb{E}[u(X_{t\wedge\tau_D})]. \tag{4.4}$$

Recall that u is bounded, since $u \in C(\bar{D})$ with $\bar{D}$ compact, and hence by Beppo Levi,

$$\mathbb{E}[\tau_D] = \lim_{t\to\infty} \mathbb{E}[t \wedge \tau_D] < +\infty.$$

Thus, by letting $t \uparrow +\infty$ in (4.4), we obtain

$$\mathbb{E}[\tau_D] = u(x) - \mathbb{E}[u(X_{\tau_D})] = u(x),$$

where we used that $u|\partial D = 0$. □

Corollary 4.6. *If the smooth boundary value problem* (4.3) *has a solution, then* $\mathbb{E}[\tau_D(x)] < \infty$*, and hence* $\tau_D(x) < \infty$ *a.s., for all* $x \in D$*. Thus* L*-diffusions starting at any point* $x \in D$ *eventually hit* ∂D *with probability* 1.

Remark 4.7. The property of an L-diffusion of hitting the boundary with probability 1 is a "nondegeneracy" condition on the operator L. We demonstrate this in the following simple example on $\mathbb{R}^n$.

Example 4.8. Consider an operator of the form

$$L = \sum_{i=1}^{n} b_i \partial_i + \frac{1}{2} \sum_{i,j=1}^{n} (\sigma\sigma^*)_{ij} \partial_i \partial_j \quad \text{on } \mathbb{R}^n, \quad a_{ij} = (\sigma\sigma^*)_{ij},$$

and let $D \subset \mathbb{R}^n$ be relatively compact. Suppose that L is nondegenerate in the following weak sense: For some $1 \leq \ell \leq n$ there holds

$$\boxed{\min_{x \in \bar{D}} a_{\ell\ell}(x) > 0.}$$

Then $\mathbb{E}[\tau_D(x)] < \infty$ for any $x \in D$.

Proof. Set

$$A := \min_{x \in \bar{D}} a_{\ell\ell}(x) \quad \text{and} \quad B := \max_{x \in \bar{D}} |b(x)|.$$

For constants $\mu, \nu > 0$ consider the smooth function

$$h(x) = -\mu e^{\nu x_\ell}, \quad x \in D.$$

Then, choosing $\nu > 2B/A$ and taking $K = \min_{x \in \bar{D}} x_\ell$, we get

$$\begin{aligned} -Lh(x) &= \mu e^{\nu x_\ell} \left(\frac{\nu^2}{2} a_{\ell\ell}(x) + \nu b_\ell(x) \right) \\ &\geq \frac{1}{2} \mu \nu A\, e^{\nu x_\ell} \left(\nu - \frac{2B}{A} \right) \\ &\geq \frac{1}{2} \nu \mu A e^{\nu K} \left(\nu - \frac{2B}{A} \right) \\ &\geq 1 \quad \text{for } \mu \text{ sufficiently large.} \end{aligned}$$

Thus

$$Lh \leq -1 \quad \text{on } D.$$

As above, we may proceed as follows. The process

$$N_t^h := h(X_{t \wedge \tau_D}) - h(x) - \int_0^{t \wedge \tau_D} Lh(X_s)\, ds, \quad t \geq 0$$

is a martingale (where again $X_t = X_t(x)$ and $\tau_D = \tau_D(x)$). By taking expectations we obtain

$$h(x) - \mathbb{E}[h(X_{t\wedge\tau_D})] = -\mathbb{E}\left[\int_0^{t\wedge\tau_D} \underbrace{Lh(X_s)}_{\le -1}\, ds\right] \ge \mathbb{E}[t\wedge\tau_D].$$

Hence,

$$\begin{aligned}\mathbb{E}[\tau_D] &= \mathbb{E}\left[\liminf_{t\to\infty} t\wedge\tau_D\right]\\ &\le \liminf_{t\to\infty} \mathbb{E}[t\wedge\tau_D]\\ &\le 2\max_{y\in\bar D} |h(y)| < \infty,\end{aligned}$$

which shows the claim. □

Definition 4.9 (Generalized Dirichlet problem). Let $\varnothing \neq D \subsetneq M$ be an open and relatively compact domain and let L be a second-order PDO on M as above. Assume $g, k \in C(\bar D)$, $k \ge 0$ and $\varphi \in C(\partial D)$ are given. The generalized Dirichlet problem consists in finding $u \in C^2(D) \cap C(\bar D)$ such that

$$\begin{cases} -Lu + ku = g \text{ on } D,\\ u|_{\partial D} = \varphi. \end{cases} \tag{GDP}$$

Theorem 4.10 (Stochastic representation of solutions to the GDP). *Assume that u solves (GDP). For $x \in D$, let $X_t(x)$ be an L-diffusion, starting from x, and assume that*

$$\mathbb{E}[\tau_D(x)] < \infty, \quad \textit{for all } x \in D.$$

Then

$$u(x) = \mathbb{E}\left[\varphi(X_{\tau_D}) \exp\left\{-\int_0^{\tau_D} k(X_s)\, ds\right\} + \int_0^{\tau_D} g(X_s) \exp\left\{-\int_0^s k(X_r)\, dr\right\} ds\right],$$

where $\tau_D = \tau_D(x)$ and $X_t = X_t(x)$.

Proof. Consider the semimartingale

$$N_t := u(X_t) \exp\left\{-\int_0^t k(X_s)\, ds\right\} + \int_0^t g(X_s) \exp\left\{-\int_0^s k(X_r)\, dr\right\} ds.$$

We find that

$$\begin{aligned} dN_t &= \exp\left\{-\int_0^t k(X_s)\, ds\right\} \left[d\big(u(X_t)\big) - u(X_t)k(X_t)\, dt + g(X_t)\, dt\right]\\ &\overset{\mathrm{m}}{=} \exp\left\{-\int_0^t k(X_s)\, ds\right\} \left[(Lu)(X_t)\, dt - u(X_t)k(X_t)\, dt + g(X_t)\, dt\right] = 0, \end{aligned}$$

where as before the symbol $\overset{\mathrm{m}}{=}$ denotes equality modulo differentials of (local) martingales. Thus, the process

$$(N_{t\wedge\tau_D})_{t\geq 0}$$

is a martingale. In particular, by dominated convergence, we get

$$u(x) = \mathbb{E}[N_0] = \mathbb{E}[N_{t\wedge\tau_D}] \to \mathbb{E}[N_{\tau_D}],$$

and thus

$$u(x) = \mathbb{E}\left[u(X_{\tau_D})\exp\left\{-\int_0^{\tau_D} k(X_s)\,ds\right\} + \int_0^{\tau_D} g(X_s)\exp\left\{-\int_0^s k(X_s)\,dr\right\}ds\right].$$

Since $u|_{\partial D} = \varphi$, we have $u(X_{\tau_D}) = \varphi(X_{\tau_D})$ which gives the claim. □

We shall consider the result of Theorem 4.10 in some special cases.

(I) *Classical Feynman–Kac formula*. Consider the boundary value problem of finding $u \in C^2(D) \cap C(\bar{D})$ such that

$$\begin{cases} -Lu + ku = g \text{ on } D, \\ u|_{\partial D} = 0. \end{cases}$$

Its solution is given by

$$u(x) = \mathbb{E}\left[\int_0^{\tau_D(x)} g(X_t(x))\exp\left\{-\int_0^t k(X_r(x))\,dr\right\}dt\right], \quad x \in D. \tag{4.5}$$

In particular, if $k \equiv 0$ then

$$u(x) = \mathbb{E}\left[\int_0^{\tau_D(x)} g(X_t(x))\,dt\right] \quad \text{(Green's kernel).}$$

Note that $-Lu = g$ is equivalent to $u = -L^{-1}g$. Thus the Green kernel gives an inverse to $-L$.

(II) *Classical Dirichlet problem*. Consider the problem of finding $u \in C^2(D) \cap C(\bar{D})$ such that

$$\begin{cases} Lu = 0 \text{ on } D, \\ u|_{\partial D} = \varphi. \end{cases} \tag{DP}$$

If $X_t(x)$ is an L-diffusion, then

$$u(x) = \mathbb{E}\left[\varphi(X_{\tau_D}(x))\right] = \int_{\partial D} \varphi\,d\mu_x,$$

where the exit measure μ_x is given by

$$\mu_x(B) := \mathbb{P}\{X_{\tau_D}(x) \in B\}, \quad B \subset \partial D \text{ measurable}.$$

Note that $u(x) = \int_{\partial D} \varphi \, d\mu(x)$ makes sense also for boundary functions φ which are just bounded and measurable.

Example 4.11. Assume that $\partial D = A \cup B$, where $A \cap B = \emptyset$. In physics, a solution $u \in C^2(D) \cap C(\bar{D})$ to the Dirichlet problem

$$\begin{cases} Lu = 0 \text{ on } D, \\ u|A = 1, \\ u|B = 0 \end{cases}$$

is called the *equilibrium potential* for the capacitor (A, B). Let $\varphi|_{\partial D}$ be defined as

$$\varphi(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in B. \end{cases}$$

Then

$$u(x) = \mathbb{E}\left[\varphi(X_{\tau_D}(x))\right] = \mathbb{P}\big\{\tau_A(x) < \tau_B(x)\big\},$$

where

$$\begin{aligned} \tau_A(x) &= \inf\{t > 0,\ X_t(x) \in A\}, \\ \tau_B(x) &= \inf\{t > 0,\ X_t(x) \in B\}. \end{aligned}$$

Thus $u(x)$ corresponds to *the probability that an L-diffusion, starting from x, hits A before hitting B*.

4.3 Parabolic boundary value problems. Let $D \subset M$ be an open and relatively compact domain. Consider a second-order PDO L on M and let $(X_t(x))_{t \geq 0}$ be an L-diffusion. Let $T > 0$ and V be a measurable function on D such that

$$\mathbb{E}\left[\exp\left(\int_0^{T \wedge \tau_D(x)} V_-(X_s(x)) \, ds\right)\right] < \infty, \quad \forall x \in D,$$

where $V_- := (-V) \vee 0$ denotes the negative part of V and $\tau_D(x) = \inf\{s \geq 0 : X_t(x) \in \partial D\}$. Furthermore, let $f, g \in C(\bar{D})$ and $\varphi \in C(\partial D)$.

Problem. Find a solution to the following parabolic boundary value problem:

$$\begin{cases} \frac{\partial}{\partial t}u = Lu - Vu + g & \text{on } [0,T]\times D, \\ u(t,\cdot)|_{\partial D} = \varphi & \text{for } t \in [0,T], \\ u|_{t=0} = f. \end{cases} \tag{BVP}$$

Note that necessarily $f|_{\partial D} = \varphi$.

Theorem 4.12. *Every solution* $u \in C^2([0,T]\times D) \cap C([0,T]\times \bar{D})$ *of (BVP) is of the form*

$$u(t,x) = \mathbb{E}\left[f(X_{t\wedge\tau_D}) \exp\left(-\int_0^{t\wedge\tau_D} V(X_s)\, ds\right) + \int_0^{t\wedge\tau_D} g(X_s) \exp\left(-\int_0^s V(X_r)\, dr\right) ds \right],$$

where $X_t = X_t(x)$ *and* $\tau_D = \tau_D(x)$.

Proof. For $0 < t_0 \le T$, we check by Itô's formula that

$$\begin{aligned} N_t &:= u(t_0 - t, X_t) \exp\left(-\int_0^t V(X_s) ds\right) \\ &\quad + \int_0^t g(X_s) \exp\left(-\int_0^s V(X_r) dr\right) ds, \quad t \le t_0 \wedge \tau_D \end{aligned}$$

is a martingale. Then it suffices to evaluate $u(t_0,x) = \mathbb{E}[N_0] = \mathbb{E}[N_{t_0\wedge\tau_D}]$, which gives the claim. □

In the discussion of this section we have restricted ourselves to representation formulas for solutions to elliptic–parabolic equations of second order. For establishing the existence of solutions by probabilistic methods the reader may consult [54].

5 Stochastic calculus on manifolds

5.1 Quadratic variation and integration of 1-forms. In this section we give canonical constructions related to continuous semimartingales on a manifold M, including the quadratic variation of continuous semimartingales with respect to bilinear forms on TM and the integral of 1-forms on M along semimartingales; see [19] for more details.

The following technical lemma on continuous processes is well known (e.g., see [24]) and very useful for a localization in space of continuous adapted processes, besides the usual localization in time (through a sequence of stopping times).

Lemma 5.1. *Let $(V_k)_{k\in\mathbb{N}}$ be a countable covering of M by open sets V_k and X be a continuous adapted M-valued process. Then there exists a nondecreasing sequence $(\tau_n)_{n\geq 0}$ of stopping times with $\tau_0 = 0$ and $\sup_n \tau_n = \infty$, such that on each of the intervals $[\tau_n, \tau_{n+1}] \cap \big(\mathbb{R}_+ \times \{\tau_n < \tau_{n+1}\}\big)$, the process X takes values in only one of the V_k.*

Given a filtered probability space $(\Omega, \mathscr{F}, \mathbb{P}; (\mathscr{F}_t)_{t\in\mathbb{R}_+})$, we denote by $\mathscr{S}$ the vector space of real-valued continuous semimartingales:

$$\mathscr{S} = \mathscr{M}_0 \oplus \mathscr{A},$$

where $\mathscr{M}_0$ denotes the space of continuous local martingales starting at 0 and $\mathscr{A}$ the space of continuous adapted processes pathwise locally of bounded variation.

Lemma 5.2. *Let M be an arbitrary differentiable manifold. There exist finitely many functions $h^1, \dots, h^\ell \in C^\infty(M)$ such that the following properties hold:*

(i) Each function $f \in C^\infty(M)$ factorizes through $(h^1, \dots, h^\ell)$ as $f = \bar f \circ (h^1, \dots, h^\ell)$ for some $\bar f \in C^\infty(\mathbb{R}^\ell)$.

(ii) Each section $b \in \Gamma(T^*M \otimes T^*M)$ can be written as $b = \sum_{i,j=1}^\ell b_{ij}\, dh^i \otimes dh^j$ with functions $b_{ij} \in C^\infty(M)$.

(iii) Each differential form $\alpha \in \Gamma(T^*M)$ can be written as $\alpha = \sum_{i=1}^\ell \alpha_i\, dh^i$ with functions $\alpha_i \in C^\infty(M)$.

Proof. We represent M via a Whitney embedding $h\colon M \hookrightarrow \mathbb{R}^\ell$ as a closed submanifold of some $\mathbb{R}^\ell$. Then there exists a differentiable partition $(\varphi_\lambda)_{\lambda\in\Lambda}$ of the unity on M and a family $(I_\lambda)_{\lambda\in\Lambda}$ of subsets $I_\lambda \subset \{1, \dots, \ell\}$ with the following property: For each $\lambda \in \Lambda$ the $(h^i)_{i\in I_\lambda}$ define a chart for M on some open neighborhood of $\operatorname{supp}(\varphi_\lambda)$.

Part (i) is evident: One defines $\bar f | h(M)$ through $f = \bar f \circ h$ and extends $\bar f$ constantly along the normal subspaces $\perp_x M$ to an open neighborhood of $M \cong h(M)$. Then, one may smooth $\bar f$ by multiplication with a function identical to 1 locally about $h(M)$ and vanishing outside a suitable larger neighborhood. For part (ii), note that $\varphi_\lambda\, b = \sum_{i,j=1}^\ell b_{ij}^\lambda\, dh^i \otimes dh^j$ with $b_{ij}^\lambda \in C^\infty(M)$ such that $\operatorname{supp}(b_{ij}^\lambda) \subset \operatorname{supp}(\varphi_\lambda)$ and $b_{ij}^\lambda := 0$ for $\{i, j\} \not\subset I_\lambda$, but then

$$b = \sum_{i,j=1}^{\ell} b_{ij}\, dh^i \otimes dh^j, \quad \text{where } b_{ij} := \sum_\lambda b_{ij}^\lambda.$$

The proof of part (iii) is analogous to (ii). □

Theorem 5.3. *Let X be an M-valued semimartingale. There exists a unique linear mapping*

$$\Gamma(T^*M \otimes T^*M) \to \mathscr{A}, \quad b \mapsto \int b(dX, dX),$$

such that for all $f, g \in C^\infty(M)$,

$$df \otimes dg \mapsto [f(X), g(X)], \tag{5.1}$$

$$f\, b \mapsto \int f(X)\, b(dX, dX). \tag{5.2}$$

Here, by definition $b(dX, dX) := d\int b(dX, dX)$. Recall that $[f(X), g(X)]$ in condition (5.1) *denotes the quadratic covariation process of $f(X)$ and $g(X)$.*

Definition 5.4. The process $\int b(dX, dX)$ is called the *integral of b along X* or the *b-quadratic variation of X*. The random variable giving its value at time t is usually written as $\int_0^t b(dX, dX)$.

Proof (of Theorem 5.3). By Lemma 5.2(ii) each section $b \in \Gamma(T^*M \otimes T^*M)$ can be represented as $b = \sum b_{ij}\, dh^i \otimes dh^j$. We define

$$\int b(dX, dX) := \sum \int (b_{ij} \circ X)\, d[h^i(X), h^j(X)]. \tag{5.3}$$

Then uniqueness is obvious; to prove existence it remains to show that (5.3) is well defined. To this end assume that

$$b = \sum_{\text{finite}} u_\nu\, df^\nu \otimes dg^\nu = 0.$$

We need to check that

$$\sum_\nu u_\nu(X)\, d[f^\nu(X), g^\nu(X)] = 0$$

as well. Without loss of generality, by means of Lemma 5.1, we may assume that h is already a global chart for M. According to Lemma 5.2(i), we write $u_\nu = \bar{u}_\nu \circ h$, $f^\nu = \bar{f}^\nu \circ h$, and $g^\nu = \bar{g}^\nu \circ h$ in terms of appropriate extensions $\bar{u}_\nu, \bar{f}^\nu, \bar{g}^\nu \in C^\infty(\mathbb{R}^\ell)$. Defining $\bar{X} = h \circ X$, the claim then follows from the following calculation:

$$\begin{aligned}
\sum_\nu u_\nu(X)\, d[f^\nu(X), g^\nu(X)] &= \sum_\nu \bar{u}_\nu(\bar{X})\, d[\bar{f}^\nu(\bar{X}), \bar{g}^\nu(\bar{X})] \\
&= \sum_{i,j} \sum_\nu \bar{u}_\nu(\bar{X})\, (D_i \bar{f}^\nu)(\bar{X})\, (D_j \bar{g}^\nu)(\bar{X})\, d[\bar{X}^i, \bar{X}^j] \\
&= \sum_{i,j} \Big(\sum_\nu u_\nu\, df^\nu \otimes dg^\nu \Big) \left(\left(\frac{\partial}{\partial h^i} \right)_X, \left(\frac{\partial}{\partial h^j} \right)_X \right) d[\bar{X}^i, \bar{X}^j] = 0.
\end{aligned}$$

□

Corollary 5.5. *The b-quadratic variation $\int b(dX, dX)$ depends only on the symmetric part of b. In particular, $\int b(dX, dX) = 0$ if b is antisymmetric.*

Proof. Defining $\bar{b}(A, B) := b(B, A)$, the assignment $b \mapsto \int \bar{b}(dX, dX)$ has the defining properties (5.1) and (5.2) as well. □

Theorem 5.6 (Pullback formula for the b-quadratic variation). *Let $\phi\colon M \to N$ be a differentiable map and $b \in \Gamma(T^*N \otimes T^*N)$. Let $\phi^*b \in \Gamma(T^*M \otimes T^*M)$ be the pullback of b via ϕ, i.e.,*

$$(\phi^* b)_p(u, v) := b_{\phi(p)}(d\phi_p\, u, d\phi_p\, v), \quad u, v \in T_pM,\ p \in M.$$

Then, for any semimartingale X on M,

$$\int (\phi^* b)\,(dX, dX) = \int b\big(d(\phi \circ X), d(\phi \circ X)\big). \tag{5.4}$$

Proof. The left-hand side of (5.4) obviously has the defining properties for the b-quadratic variation of the image process $\phi \circ X$. □

We now turn to the problem of integrating 1-forms on M along M-valued semimartingales.

Theorem 5.7. *Let X be a semimartingale taking values in M. There is a unique linear mapping*

$$\Gamma(T^*M) \to \mathscr{S}, \quad \alpha \mapsto \int \alpha(\circ\, dX) \equiv \int_X \alpha$$

such that, for all $f \in C^\infty(M)$,

$$df \;\mapsto\; f(X) - f(X_0), \tag{5.5}$$

$$f\,\alpha \;\mapsto\; \int f(X) \circ \alpha(\circ\, dX). \tag{5.6}$$

In (5.6) the integral means the Stratonovich integral of the process $f(X)$ with respect to the semimartingale $\int \alpha(\circ\, dX)$. Thus, in other words, $f(X) \circ \alpha(\circ\, dX) \equiv f(X) \circ d\big(\int \alpha(\circ\, dX)\big)$.

Definition 5.8 (Stratonovich integral of 1-forms along semimartingales). The process $\int \alpha(\circ\, dX)$ is called the *Stratonovich integral of α along X*. We also write $\int_X \alpha$ instead of $\int \alpha(\circ\, dX)$.

Proof (of Theorem 5.7). By Lemma 5.2(iii) each differential form $\alpha \in \Gamma(T^*M)$ can be represented as $\alpha = \sum_i \alpha_i \, dh^i$ with functions $\alpha_i \in C^\infty(M)$. We define

$$\int_X \alpha := \sum_i \int \alpha_i(X) \circ d(h^i(X)). \tag{5.7}$$

Uniqueness is again obvious; it is thus sufficient to show that formula (5.7) is well defined. To this end, we have to verify that if $\alpha = \sum_{\text{finite}} u_\nu \, df^\nu = 0$ then

$$\sum_\nu u_\nu(X) \circ d(f^\nu(X)) = 0$$

holds as well. Proceeding as in the proof of Theorem 5.7, without loss of generality, we may assume again that h is already a global chart for M. But then we have

$$\begin{aligned}\sum_\nu u_\nu(X) \circ d(f^\nu(X)) &= \sum_\nu \bar{u}_\nu(\bar{X}) \circ d(\bar{f}^\nu(\bar{X})) \\ &= \sum_i \sum_\nu \bar{u}_\nu(\bar{X}) \circ \left[D_i \bar{f}^\nu(\bar{X}) \circ d\bar{X}^i\right] \\ &= \sum_i \left(\left(\sum_\nu u_\nu \, df^\nu\right)\left(\tfrac{\partial}{\partial h^i}\right)_X\right) \circ d\bar{X}^i = 0,\end{aligned}$$

which gives the claim. □

Example 5.9. In the special case of a deterministic C^1 curve X in M, say $X_t = x(t)$, which is trivially a semimartingale, we obtain

$$\int_X \alpha = \int \alpha\big(\dot{x}(t)\big) \, dt, \qquad \alpha \in \Gamma(T^*M). \tag{5.8}$$

Indeed, the right-hand side of (5.8) obviously has the defining properties of $\int_X \alpha$.

Theorem 5.10 (Pullback formula for the Stratonovich integral of a 1-form). *Let $\phi\colon M \to N$ be a differentiable map and $\alpha \in \Gamma(T^*N)$. Then, for any semimartingale X on M, we have*

$$\int_X \phi^*\alpha = \int_{\phi \circ X} \alpha. \tag{5.9}$$

Proof. The left-hand side of Eq. (5.9) satisfies the defining properties for the Stratonovich integral of α along $\phi \circ X$. By uniqueness we therefore have equality. □

Remark 5.11. Let $\alpha, \beta \in \Gamma(T^*M)$. Then $\alpha \otimes \beta \in \Gamma(T^*M \otimes T^*M)$ and for the quadratic covariation process of $\int_X \alpha$ and $\int_X \beta$ we have the formula

$$\left[\int_X \alpha, \int_X \beta\right] = \int (\alpha \otimes \beta)(dX, dX).$$

5.2 Martingales and Brownian motions. The aim of this section is to introduce martingales and Brownian motions on manifolds. This task requires additional geometric structures on the manifolds: linear connections and Riemannian metrics. These results will then be extended later to the setting of sub-Riemannian geometry where the metric is defined only on a subbundle of TM.

Notation 5.12. Let $\pi\colon TM \to M$ be the tangent bundle over M. A *linear connection* in TM, or equivalently a *covariant derivative on* TM, is an $\mathbb{R}$-linear mapping

$$\nabla\colon \Gamma(TM) \to \Gamma(T^*M \otimes TM) \tag{5.10}$$

satisfying the product rule $\nabla(fX) = df \otimes X + f\,\nabla X$, for all $X \in \Gamma(TM)$ and $f \in C^\infty(M)$. Alternatively, (5.10) may be written as a mapping

$$\Gamma(TM) \times \Gamma(TM) \to \Gamma(TM), \quad (A, X) \mapsto \nabla_A X \equiv (\nabla X)A,$$

which is $C^\infty(M)$-linear in the first variable and derivative in the second variable. For $f \in C^\infty(M)$, we have the *second fundamental form* (or *Hessian*) of f defined as

$$\nabla df \equiv \operatorname{Hess} f \in \Gamma(T^*M \otimes T^*M), \quad (\nabla df)(A, B) = ABf - (\nabla_A B)f.$$

The bilinear form

$$(A, B) \mapsto (\nabla df)(A, B)$$

is symmetric for each $f \in C^\infty(M)$ if and only if the connection ∇ is torsion-free, i.e., if for all $A, B \in \Gamma(TM)$,

$$T(A, B) \equiv \nabla_A B - \nabla_B A - [A, B] = 0.$$

Definition 5.13 (∇-martingale). Let M be a manifold and ∇ be a linear connection in TM. An M-valued semimartingale X defined on some filtered probability space $(\Omega, \mathscr{F}, \mathbb{P}; (\mathscr{F}_t)_{t\ge 0})$ is called ∇-*martingale* if for each $f \in C^\infty(M)$,

$$d(f \circ X) - \tfrac{1}{2}\,(\nabla df)(dX, dX) \overset{\mathrm{m}}{=} 0, \tag{5.11}$$

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

Since $(\nabla df)(dX, dX)$ depends only on the symmetric part of ∇df, one may always assume that the linear connection ∇ is torsion-free. Symmetrization of the connection does not change the class of ∇-martingales.

Example 5.14. In the special case of $M = \mathbb{R}^n$ equipped with the canonical linear connection $\nabla_{D_i} D_j = 0$, we have

$$(\nabla df)(D_i, D_j) = D_i D_j f,$$

and hence ∇-martingales in the sense of Definition 5.13 coincide with the usual class of continuous local martingales on $\mathbb{R}^n$. Indeed, according to Itô's formula, a continuous $\mathbb{R}^n$-valued semimartingale X is a local martingale if and only if

$$d(f \circ X) - \frac{1}{2} \sum_{i,j} (D_i D_j f)(X)\, d[X^i, X^j] \overset{\mathrm{m}}{=} 0,$$

for all $f \in C^\infty(\mathbb{R}^n)$. This is exactly condition (5.11) of Definition 5.13.

Remark 5.15 (Martingales as solutions of SDEs). Let ∇ be a linear connection on TM which without loss of generality is torsion-free. Let $A_0 \in \Gamma(TM)$ and $A \in \Gamma\big(\mathrm{Hom}(M \times \mathbb{R}^r, TM)\big)$, and suppose that X is a solution to the SDE

$$dX = A_0(X)\, dt + A(X) \circ dZ. \tag{5.12}$$

Here Z may be an arbitrary continuous $\mathbb{R}^r$-valued semimartingale. Then for $f \in C^\infty(M)$ we have

$$d(f \circ X) = \big(A_0 f\big)(X)\, dt + \sum_{i=1}^r \big(A_i f\big)(X)\, dZ^i + \frac{1}{2} \sum_{i,j=1}^r \big(A_i A_j f\big)(X)\, d[Z^i, Z^j],$$

where $A_i = A(\cdot)e_i \in \Gamma(TM)$ for $i = 1, \dots, r$. Since $(\nabla df)(A_i, A_j) = A_i A_j f - (\nabla_{A_i} A_j) f$ and since on the other hand,

$$(\nabla df)(dX, dX) = \sum_{i,j=1}^r (\nabla df)(A_i, A_j)(X)\, d[Z^i, Z^j],$$

we obtain

$$\begin{aligned} d(f \circ X) - \frac{1}{2} (\nabla df)(dX, dX) = (A_0 f)(X)\, dt + \sum_{i=1}^r \big(A_i f\big)(X)\, dZ^i \\ + \frac{1}{2} \sum_{i,j=1}^r \big(\nabla_{A_i} A_j f\big)(X)\, d[Z^i, Z^j]. \end{aligned}$$

Denoting the drift of the semimartingale Z by Z^{drift}, we obtain that X is a ∇-martingale if

$$\big(A_0 f\big)(X)\, dt + \sum_{i=1}^r \big(A_i f\big)(X)\, d(Z^{\mathrm{drift}})^i + \frac{1}{2} \sum_{i,j=1}^r \big(\nabla_{A_i} A_j f\big)(X)\, d[Z^i, Z^j] = 0$$

for any $f \in C^\infty(M)$. In the special case when Z is a Brownian motion on $\mathbb{R}^r$ we find that solutions X to the SDE (5.12) are ∇-martingales if

$$A_0 = -\frac{1}{2} \sum_{i=1}^r \nabla_{A_i} A_i.$$

Definition 5.16 (Riemannian quadratic variation). Let $(M, g) = (M, \langle\cdot,\cdot\rangle)$ be a Riemannian manifold and X be a semimartingale taking values in M. The process

$$[X, X] := \int g(dX, dX) = \int \langle dX, dX\rangle \tag{5.13}$$

is called a *Riemannian quadratic variation* of X.

Theorem 5.17 (Lévy's characterization of M-valued Brownian motions). *Let (M, g) be a Riemannian manifold and ∇ be the Levi-Civita connection. For a semimartingale X of maximal lifetime and taking values in M, the following conditions are equivalent:*

(i) X is a Brownian motion on (M, g), i.e., for any $f \in C^\infty(M)$ the real-valued process

$$f \circ X - \frac{1}{2}\int \Delta f \circ X\, dt$$

is a local martingale; here $\Delta f = \operatorname{trace} \nabla df \in C^\infty(M)$ denotes the Laplace–Beltrami operator on M.

(ii) X is a ∇-martingale such that $[f(X), f(X)] = \int \|\nabla f\|^2(X)\, dt$ for every $f \in C^\infty(M)$.

(iii) X is a ∇-martingale such that $\int b(dX, dX) = \int (\operatorname{trace} b)(X)\, dt$ for every $b \in \Gamma(T^*M \otimes T^*M)$.

In particular, for the Riemannian quadratic variation (5.13) of X, we then have

$$\int_0^t g(dX, dX) = t \dim M.$$

Proof. (1) To prove (ii) $\Longleftrightarrow$ (iii) we verify that for X the following two conditions are equivalent:

(a) $[f(X), f(X)] = \int \|\nabla f\|^2(X)\, dt$.

(b) $\int b(dX, dX) = \int (\operatorname{trace} b)(X)\, dt$ for every $b \in \Gamma(T^*M \otimes T^*M)$.

Indeed, for $f, h \in C^\infty(M)$ we have

$$\begin{aligned}\operatorname{trace}(df \otimes dh) &= \sum_i (df \otimes dh)(e_i, e_i) = \sum_i (df)(e_i)\,(dh)(e_i)\\ &= \sum_i \langle \nabla f, e_i\rangle\,\langle \nabla h, e_i\rangle = \langle \nabla f, \nabla h\rangle.\end{aligned}$$

The implication (b) ⇒ (a) is then the special case that $b = df \otimes df$. To verify the direction (a) ⇒ (b), first note that by polarization (a) implies

$$[f(X), h(X)] = \int \langle \nabla f \circ X, \nabla h \circ X \rangle \, dt$$

for $f, h \in C^\infty(M)$. Thus $[f \circ X, h \circ X] = \int (df \otimes dh)(dX, dX) = \int \operatorname{trace}(df \otimes dh)(X)\, dt$. By means of the uniqueness part of Theorem 5.3, we get

$$\int b(dX, dX) = \int (\operatorname{trace} b)(X)\, dt$$

for any bilinear form $b \in \Gamma(T^*M \otimes T^*M)$.

(2) (iii) ⇒ (i): Part (1) applied to the given ∇-martingale X shows that $b(dX, dX) = (\operatorname{trace} b)(X)\, dt$ for bilinear forms $b \in \Gamma(T^*M \otimes T^*M)$; thus in particular for $b = \nabla df$,

$$d(f \circ X) \overset{\mathrm{m}}{=} \tfrac{1}{2} \nabla df(dX, dX) = \tfrac{1}{2}(\Delta f)(X)\, dt.$$

(3) (i) ⇒ (ii): Now let X be a Brownian motion on M. According to $\nabla df^2 = 2\,(f\, \nabla df + df \otimes df)$ we first note that $\Delta(f^2) = 2f\, \Delta f + 2\, \|\nabla f\|^2$, and thus

$$d(f^2 \circ X) \overset{\mathrm{m}}{=} \tfrac{1}{2}(\Delta f^2)(X)\, dt = (f\, \Delta f)(X)\, dt + \|\nabla f\|^2(X)\, dt.$$

On the other hand, by means of Itô's formula,

$$d(f^2 \circ X) = 2\, f(X)\, d(f \circ X) + d[f(X), f(X)] \overset{\mathrm{m}}{=} f(X)\, (\Delta f)(X)\, dt + d[f(X), f(X)].$$

Uniqueness of the Doob–Meyer decomposition implies

$$[f(X), f(X)] = \int \|\nabla f\|^2(X)\, dt.$$

Finally, once again by means of part (1), the last formula gives

$$\nabla df(dX, dX) = (\operatorname{trace} \nabla df)(X)\, dt = (\Delta f)(X)\, dt$$

from which we conclude that X is a ∇-martingale. □

On $\mathbb{R}^n$ with the canonical Euclidean metric, Brownian motions in the sense of Lévy's characterization coincide with the usual class of $\mathbb{R}^n$-valued Brownian motions.

Theorem 5.18 (M-valued Brownian motions as solutions of an SDE). *Let (M, g) be a Riemannian manifold and ∇ be the Levi-Civita connection on M. Consider the SDE*

$$dX = A_0(X)\, dt + A(X) \circ dB \tag{5.14}$$

with $A_0 \in \Gamma(TM)$ and $A \in \Gamma\big(\operatorname{Hom}(M \times \mathbb{R}^r, TM)\big)$; here B is a Brownian motion on $\mathbb{R}^r$. Then maximal solutions to (5.14) are Brownian motions on (M, g) if the two subsequent conditions are satisfied:

(i) $A_0 = -\frac{1}{2}\sum_i \nabla_{A_i} A_i$ *with* $A_i \equiv A(\cdot)e_i$ *for* $i = 1, \ldots, r$.

(ii) The map $A(x)^*\colon T_xM \to \mathbb{R}^r$ *is an isometric embedding for every* $x \in M$, *i.e.,* $A(x)\,A(x)^* = \mathrm{id}_{T_xM}$, *where* $A(x)^*$ *is the adjoint to* $A(x) \in \mathrm{Hom}(\mathbb{R}^r, T_xM)$.

Proof. Let X be a solution to Eq. (5.14) and assume that conditions (i) and (ii) are satisfied. According to Remark 5.15 condition (i) guarantees that X is a ∇-martingale. In addition, we have for $f \in C^\infty(M)$,

$$d(f \circ X) \overset{\mathrm{m}}{=} \frac{1}{2}\sum_{i=1}^{r}(\nabla df)(A_i, A_i)(X)\,dt.$$

It is thus sufficient to verify that

$$\sum_i(\nabla df)(A_i, A_i) = \Delta f.$$

This is however a straightforward consequence of condition (ii). □

Remark 5.19. Conditions (i) and (ii) of Theorem 5.18 can always be satisfied for r sufficiently large. For instance, let $M \hookrightarrow \mathbb{R}^r$ be a Whitney embedding. Then T_xM can be seen as a subspace $\mathbb{R}^r$ for each $x \in M$. Defining $A \in \Gamma(\mathrm{Hom}(M \times \mathbb{R}^r, TM))$ fiberwise as an orthogonal projection $A(x)\colon \mathbb{R}^r \to T_xM$ onto T_xM and setting $A_0 = -\frac{1}{2}\sum_i \nabla_{A_i} A_i$, then every solution to the SDE (5.14) (with a given initial condition) is a Brownian motion on (M, g). The drawback of this construction is that to a given Riemannian manifold (M, g) there is no *canonical* choice of the coefficients A_0 and A; there is however a canonical SDE on the orthonormal frame bundle $\mathrm{O}(TM)$ over M such that its solutions project to Brownian motions on (M, g). We deal with this construction in the next subsection.

Theorem 5.20 (Brownian motions on submanifolds of $\mathbb{R}^n$). *Let* M *be a submanifold of* $\mathbb{R}^n$ *endowed with the induced Riemannian metric. Consider the SDE*

$$dX = A(X) \circ dB, \tag{5.15}$$

where B *is a Brownian motion on* $\mathbb{R}^n$ *and*

$$A \in \Gamma(\mathrm{Hom}(M \times \mathbb{R}^n, TM)), \quad (x, v) \mapsto A(x)v,$$

such that $A(x)\colon \mathbb{R}^n \to T_xM$ *is the orthogonal projection onto* T_xM. *Then every solution of* (5.15), *for some specified initial condition, gives a Brownian motion on* (M, g).

Proof. In terms of the vector fields $A_i \equiv A(\cdot)e_i \in \Gamma(TM)$, $i = 1, \ldots, n$, it is sufficient by Theorem 5.18 to verify that $\sum_i \nabla_{A_i} A_i = 0$. This is however a straightforward calculation. □

5.3 Parallel transport and stochastically moving frames. The fundamental observation that diffusion processes on a manifold M can be horizontally lifted via a connection to the frame bundle over M goes back to the pioneering work of Malliavin, Eells, and Elworthy. Conversely, solving SDEs on the frame bundle and projecting the solution down to the manifold M allows canonical constructions of diffusion processes on M.

Intuitively this procedure corresponds to a "rolling without slipping" of the manifold along the trajectories of a continuous $\mathbb{R}^n$-valued semimartingale. It allows us to construct, for each semimartingale in T_xM, its *stochastic development* on M, together with a notion of parallel transport along the paths of the obtained process. Clearly this method requires a connection on M. The problem that in sub-Riemannian geometry, typically only "partial connections" are canonically given, will be addressed in the next subsection.

Notation 5.21. Let M be an n-dimensional differentiable manifold and denote by $P = \mathrm{L}(TM)$ its frame bundle. Then $\pi\colon P \to M$ is a G-principal bundle with $G = \mathrm{GL}(n;\mathbb{R})$. The fiber P_x consists of the linear isomorphisms $u\colon \mathbb{R}^n \to T_xM$ where $u \in P_x$ is identified with the $\mathbb{R}$-basis

$$(u_1, \dots, u_n) := (ue_1, \dots, ue_n).$$

A linear connection in TM induces canonically a *G-connection* in P given as a G-invariant differentiable splitting h of the following exact sequence of vector bundles over P:

$$\begin{tikzcd} 0 \arrow[r] & \ker d\pi \arrow[r] & TP \arrow[r, "d\pi"] & \pi^*TM \arrow[r] \arrow[l, dashed, bend left, "h"] & 0. \end{tikzcd}$$

This splitting induces a decomposition of TP:

$$TP = V \oplus H := \ker d\pi \oplus h(\pi^*TM).$$

G-invariance of the splitting means that $H_{ug}{=}(dR_g)H_u$ for each $u \in P$, where $R_g u := u\,g$ denotes the right action of $g{\in}G$. For $u{\in}P$, we call H_u the *horizontal space at u* and $V_u{=}\{v{\in}T_uP\colon\ (d\pi)v = 0\}$ the *vertical space at u*. The bundle isomorphism

$$h\colon\ \pi^*TM \xrightarrow{\sim} H \hookrightarrow TP \tag{5.16}$$

is called the *horizontal lift* of the G-connection; fiberwise it reads $h_u : T_{\pi(u)}M \xrightarrow{\sim} H_u$.

By means of the G-connection in P each vector field $X \in \Gamma(TP)$ decomposes into a horizontal and a vertical part:

$$X = \operatorname{hor} X + \operatorname{vert} X.$$

Definition 5.22 (Connection form). Each $u \in P$ defines an embedding $I_u\colon G \hookrightarrow P$, $g \mapsto ug$. Its differential at the unit element $e \in G$,

$$\iota_u \equiv (dI_u)_e\colon T_eG \to T_uP, \quad A \longmapsto \hat{A}(u), \tag{5.17}$$

gives an identification $\kappa_u\colon \mathfrak{g} \xrightarrow{\sim} V_u$ of the Lie algebra $\mathfrak{g} = T_eG$ of G with the vertical fiber V_u at u. The vertical vector field $\hat{A} \in \Gamma(TP)$ on P defined by (5.17) is called the *standard-vertical vector field* to $A \in \mathfrak{g}$. The $\mathfrak{g}$-valued 1-form $\omega \in \Gamma(T^*P \otimes \mathfrak{g})$ on P, defined by

$$\omega_u(X_u) := \kappa_u^{-1}(\operatorname{vert} X)_u, \quad X \in \Gamma(TP), \tag{5.18}$$

is called the *connection form* of the G-connection.

Note that for the frame bundle $\pi\colon \mathrm{L}(TM) \to M$ over M we have $\mathfrak{g} = \mathrm{GL}(n;\mathbb{R})$. In the case that M is a Riemannian manifold it is natural to consider the orthonormal frame bundle $\pi\colon \mathrm{O}(TM) \to M$ over M with structure group $G = \mathrm{O}(n;\mathbb{R})$. The fiber P_x then consists of the linear isometries $u\colon \mathbb{R}^n \to T_xM$. As above a metric connection on TM then gives rise to a G-invariant splitting $TP = V \oplus H$. The connection form then takes its values in the Lie algebra $\mathfrak{g}$ of skew symmetric $n \times n$ matrices.

In the sequel we deal with the two cases of G-principal bundles: $P = \mathrm{L}(TM)$ over a manifold M with $G = \mathrm{GL}(n;\mathbb{R})$ and $P = \mathrm{O}(TM)$ over a Riemannian manifold M with $G = \mathrm{O}(n;\mathbb{R})$. In addition to the $\mathfrak{g}$-valued connection form (see Definition 5.22) we have the *canonical* 1-*form*

$$\vartheta \in \Gamma(T^*P \otimes \mathbb{R}^n), \qquad \vartheta_u(X_u) := u^{-1}(d\pi X_u), \quad u \in P \text{ and } X \in \Gamma(TP), \tag{5.19}$$

where as usual we read $u \in P$ as a linear isomorphism (resp. isometry), $u\colon \mathbb{R}^n \xrightarrow{\sim} T_{\pi(u)}M$.

Remark 5.23. The frame bundles $P = \mathrm{L}(TM)$ with M a manifold (resp. $P = \mathrm{O}(TM)$ with M a Riemannian manifold), considered as manifolds, are parallelizable, i.e., the tangent bundles $T\,\mathrm{L}(TM) \to \mathrm{L}(TM)$ and $T\,\mathrm{O}(TM) \to \mathrm{O}(TM)$ are trivial.

Proof. Indeed a G-connection in P decomposes $TP = V \oplus H$. A canonical trivialization for TP is given as follows: the vertical subbundle V is trivialized by the standard-vertical vector fields $\hat{A}$ to A, where A runs through a basis of $\mathfrak{g}$; the horizontal subbundle H is trivialized by the *standard-horizontal vector fields* $L_1, \ldots, L_n$ in $\Gamma(TP)$ defined by

$$L_i(u) := h_u(ue_i).$$

For any $u \in P$, then

$$(\hat{A}(u),\ L_i(u):\ A \in \text{basis for } \mathfrak{g},\ i = 1, \dots, n)$$

is a basis for $T_u P = V_u \oplus H_u$ which is obvious from the isomorphisms $\mathfrak{g} \xrightarrow{\sim} V_u$, $A \mapsto \hat{A}(u)$, and $h_u\colon T_{\pi(u)}M \xrightarrow{\sim} H_u$. □

Remark 5.24. The standard-vertical (resp., standard-horizontal) vector fields are determined by the relations

$$\vartheta(\hat{A}) = 0 \text{ and } \vartheta(L_i) = e_i \quad (\text{resp. } \omega(\hat{A}) = A \text{ and } \omega(L_i) = 0).$$

The canonical second-order partial differential operator $\Delta^{\mathrm{hor}} := \sum_i L_i^2$ is called the *horizontal Laplacian* on $\mathrm{L}(TM)$ (resp. $\mathrm{O}(TM)$).

Definition 5.25 (Horizontal lift of an M-valued semimartingale). For any P-valued semimartingale U the Stratonovich integral $\int_U \omega$ (defined componentwise with respect to a basis of $\mathfrak{g}$) gives a semimartingale taking values in the Lie algebra $\mathfrak{g}$. We call U *horizontal* if $\int_U \omega = 0$ a.s. For an M-valued semimartingale X, a semimartingale U taking values in P is called the *horizontal lift of* X, if U is horizontal and if $\pi \circ U = X$ a.s.

Remark 5.26. Definition 5.25 generalizes the classical notion of horizontal lift for M-valued differentiable curves: a curve $t \mapsto u(t)$ over $t \mapsto x(t)$is called *horizontal* if $\pi \circ u = x$ and $\omega(\dot{u}) = 0$.

For the remainder of this subsection we deal with the following situation: either M will be a differentiable manifold equipped with a torsion-free connection, or M will be a Riemannian manifold equipped with the Levi-Civita connection.

Definition 5.27 (Anti-development of an M-valued semimartingale). Let X be an M-valued semimartingale and U a horizontal lift of X taking values in $P = \mathrm{L}(TM)$ (resp. $\mathrm{O}(TM)$). The $\mathbb{R}^n$-valued semimartingale

$$Z = \int_U \vartheta \equiv \int \vartheta(\circ dU)$$

is called the *anti-development* of X into $\mathbb{R}^n$ (with respect to the initial frame U_0). In terms of the standard basis of $\mathbb{R}^n$ we have $Z \equiv (Z^1, \dots, Z^n)$ where $Z^i = \int_U \vartheta^i$.

Theorem 5.28. *Let X be an M-valued semimartingale, U a horizontal lift of X to $P = \mathrm{L}(TM)$ (resp. $\mathrm{O}(TM)$), and Z an anti-development of X into $\mathbb{R}^n$. The following statements hold:*

(i) $\int_U \sigma = \sum_{i=1}^n \int \sigma(U)\, L_i(U) \circ dZ^i$ *for each differential form* $\sigma \in \Gamma(T^*P)$.

(ii) $\int_X \alpha = \sum_{i=1}^n \int \alpha(X)\, Ue_i \circ dZ^i$ *for each differential form* $\alpha \in \Gamma(T^*M)$.

In particular, $d(f\circ U)=\sum_{i=1}^n (L_i f)(U)\circ dZ^i$ *for each function* $f\in C^\infty(P)$*, or in short,*

$$dU = \sum_{i=1}^n L_i(U)\circ dZ^i, \tag{5.20}$$

as well as $d(f\circ X)=\sum_{i=1}^n (Ue_i)(f)\circ dZ^i$ *for each function* $f\in C^\infty(M)$*, or in short,*

$$dX = U\circ dZ. \tag{5.21}$$

Proof. The additional claims follow from (i) and (ii) with $\sigma = df$ where $f\in C^\infty(P)$ (resp. $\alpha = df$ where $f\in C^\infty(M)$).

To (i): According to Theorem 5.7 it is sufficient that the right-hand side of (i) has the defining properties of $\int_U \sigma$. For $f\in C^\infty(P)$ we have to show that $d(f\circ U) = \sum_i (df)(U)\, L_i(U)\circ dZ^i \equiv \sum_i (L_i f)(U)\circ dZ^i$, which is equivalent to

$$f\circ U - f\circ U_0 = \int_U \sigma, \quad \text{where } \sigma\in\Gamma(T^*P),\ \sigma_u := \sum_i (L_i f)(u)\,\vartheta_u^i. \tag{5.22}$$

But we observe that $\sum_i (L_i f)(u)\,\vartheta_u^i = (df)_u\circ \mathrm{pr}_{H_u}$; indeed for $A\in T_uP$ we have

$$\begin{aligned}\sum_i (L_i f)(u)\,\vartheta_u^i(A) &= \sum_i (df)_u\, L_i(u)\,\vartheta_u^i(A)\\ &= \sum_i (df)_u\, h_u(ue_i)\,\big(u^{-1}(d\pi)_u A\big)^i\\ &= (df)_u\, h_u(u\,u^{-1}(d\pi)_u A)\\ &= (df)_u\, h_u\big((d\pi)_u A\big)\\ &= \big((df)_u\circ \mathrm{pr}_{H_u}\big)(A).\end{aligned}$$

On the other side, we have $(df\circ \mathrm{pr}_V)_u = (df)_u\,\kappa_u\,\omega_u = d(f\circ I_u)_e\,\omega_u$. But U is horizontal and hence $\int_U df\circ \mathrm{pr}_V = 0$ which shows that

$$f\circ U - f\circ U_0 = \int_U df = \int_U df\circ \mathrm{pr}_H + \int_U df\circ\mathrm{pr}_V = \int_U df\circ \mathrm{pr}_H = \int_U \sigma.$$

The second defining property of the Stratonovich integral is obvious.

To (ii): It is sufficient to show that

$$d(f\circ X) = \sum_i (df)(X)\, Ue_i\circ dZ^i \equiv \sum_i (Ue_i)(f)\circ dZ^i$$

holds for each function $f \in C^\infty(M)$. With part (i), using that $(d\pi)_u\, L_i(u) = ue_i$, we obtain

$$\begin{aligned} d(f\circ\pi\circ U) &= \sum_i d(f\circ\pi)(U)\, L_i(U)\circ dZ^i \\ &= \sum_i (df)\big(\pi(U)\big)\,(d\pi)(U)\, L_i(U)\circ dZ^i = \sum_i (df)(X)\, Ue_i\circ dZ^i\,, \end{aligned}$$

which shows the claim. □

Theorem 5.29. *Let X be an M-valued semimartingale, U a horizontal lift of X to $P = \mathrm{L}(TM)$ (resp. $\mathrm{O}(TM)$), and Z an anti-development of X into $\mathbb{R}^n$. Then*

(i) $\int a(dU, dU) = \sum_{i,j=1}^n \int a(U)\,\big(L_i(U), L_j(U)\big)\, d[Z^i, Z^j]$ *for* $a \in \Gamma(T^*P\otimes T^*P)$;

(ii) $\int b(dX, dX) = \sum_{i,j=1}^n \int b(X)\,\big(Ue_i, Ue_j\big)\, d[Z^i, Z^j]$ *for* $b \in \Gamma(T^*M\otimes T^*M)$.

Proof. It is again sufficient to consider the special case $a = d\varphi_1\otimes d\varphi_2$ where $\varphi_1, \varphi_2 \in C^\infty(P)$ (resp. $b = df_1\otimes df_2$ where $f_1, f_2 \in C^\infty(M)$). Then the statements follow from Remark 5.11. □

Theorem 5.30 (Existence of horizontal lifts to M-valued semimartingales). *Let P be a G-principal bundle over a manifold M endowed with a G-connection. Let x_0 be an M-valued random variable and u_0 a P-valued random variable over x_0, i.e., $\pi\circ u_0 = x_0$ a.s. Then to each M-valued semimartingale X with $X_0 = x_0$ there is exactly one horizontal lift U to P with $U_0 = u_0$ a.s.*

Proof. See [53] or [24, Chapter 7]. The existence part is straightforward. According to Theorem 3.22, the semimartingale X can be realized as the solution of a Stratonovich SDE of the form

$$dX = \sum_{i=1}^{\ell} A_i(X)\circ dZ^i, \quad X_0 = x_0, \tag{5.23}$$

where Z is an $\mathbb{R}^\ell$-valued semimartingale for some ℓ. Let $\bar{A}_i \in \Gamma(TP)$ be the horizontal lift of $A_i \in \Gamma(TM)$, i.e., $\bar{A}_i(u) = h_u(A_i(\pi u))$ for $u \in P$, and consider the "horizontally lifted SDE" on P:

$$dU = \sum_{i=1}^{\ell} \bar{A}_i(U)\circ dZ^i, \quad U_0 = u_0. \tag{5.24}$$

It is clear that solutions to (5.24) are canonical candidates for the wanted horizontal lift. Indeed, we have $d(\pi\circ U) = \sum_i (d\pi)_U\, \bar{A}_i(U)\circ dZ^i \equiv \sum_i A_i(\pi\circ U)\circ dZ^i$ with $\pi\circ U_0 = x_0$, and hence $\pi\circ U = X$ by the uniqueness of solutions to (5.23). On the other hand, we have $\int_U \omega = \sum_i \int \omega(U)\,\bar{A}_i(U)\circ dZ^i = 0$. It remains to verify that U and X have identical lifetimes. □

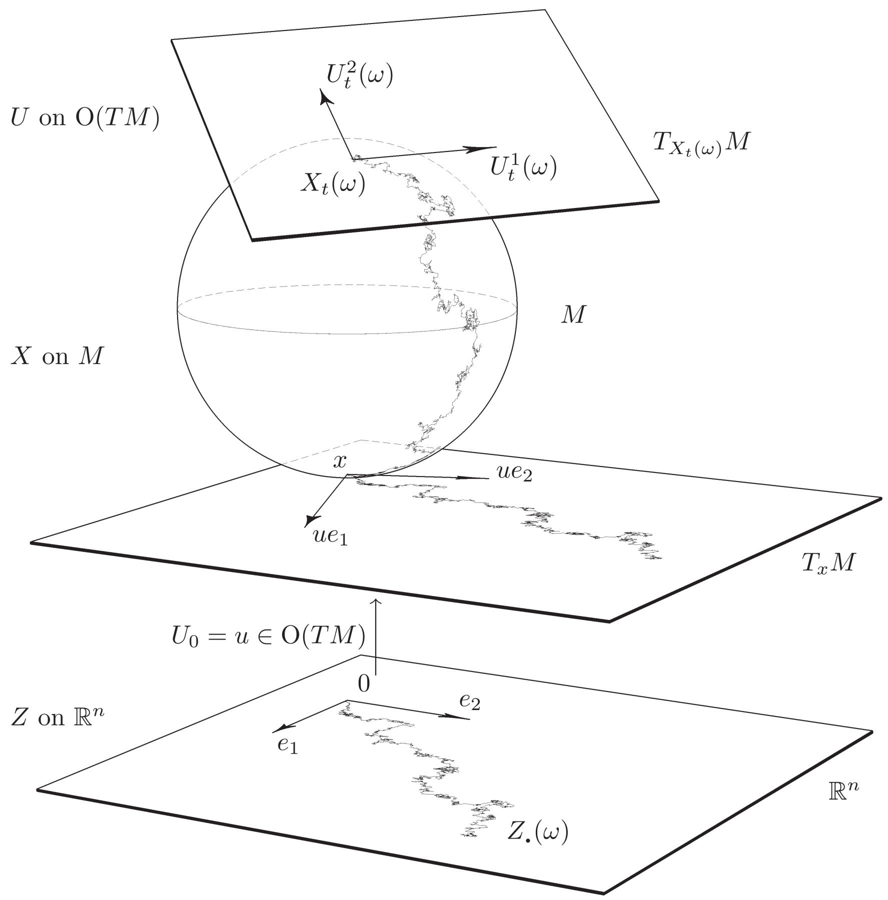
We want to summarize the theory developed so far. Let M be a differentiable manifold equipped with a torsion-free connection, or a Riemannian manifold with the Levi-Civita connection. For a semimartingale X on M we defined its horizontal lift U to $P = \mathrm{L}(TM)$ (resp. $\mathrm{O}(TM)$), and its anti-development Z into $\mathbb{R}^n$. Then (modulo choice of initial conditions $X_0 = x$, $U_0 = u$) each of the three processes X, U, Z determines the two others.
Indeed, we have

(a) Z determines U as the solution to the SDE $dU = \sum_{i=1}^n L_i(U) \circ dZ^i$ with $U_0 = u$;

(b) U determines X via $X = \pi \circ U$;

(c) X determines Z as $Z = \int_U \vartheta$ where U is the unique horizontal lift of X to P with $U_0 = u$.

Typically, one starts with Z on $\mathbb{R}^n$ to determine X on M (the *stochastic development* of Z). The frame U moves along X by parallel transport.



In the deterministic special case of a differentiable curve $Z\colon t \mapsto z(t)$ in $\mathbb{R}^n$, stochastic development reduces to the canonical Cartan development of $z(t)$.

Example 5.31 (Cartan development). The *Cartan development* of an $\mathbb{R}^n$-valued curve $t \mapsto z(t)$ is the construction of curves $x\colon t \mapsto x(t) \in M$ and $u\colon t \mapsto u(t) \in P$ (where $P = \mathrm{L}(TM)$, resp. $P = \mathrm{O}(TM)$ in the Riemannian case) such that $u(\cdot)$ lies above $x(\cdot)$ and such that

(i) $\dot{x} = u\,\dot{z}$, or in equivalent notation $dx(t) = u(t)\,dz(t)$;

(ii) u is parallel along x, i.e., $\nabla_D u \equiv (\nabla_D u^1, \dots, \nabla_D u^n) = 0$, where $D = \partial/\partial t$.

Condition (ii) means that $u(\cdot)$ is a horizontal curve; thus $\dot{u} \in H_u \equiv h_u(T_{\pi(u)}M)$, and hence $\dot{u} = h_u(\dot{x}) = h_u(u\dot{z})$ by using (i). Since $h_u(u\dot{z}) = \sum_i h_u(ue_i)\,\dot{z}^i = \sum_i L_i(u)\,\dot{z}^i$, conditions (i) and (ii) are seen to be equivalent to

$$du = \sum_i L_i(u)\,dz^i.$$

Definition 5.32 (Parallel transport along a semimartingale). Let M be a differentiable manifold equipped with a torsion-free connection, or a Riemannian manifold with the Levi-Civita connection. Let X be a semimartingale on M and U an arbitrary horizontal lift of X to $\mathrm{L}(TM)$ (resp. $\mathrm{O}(TM)$). For $0 \le s \le t$, let $//_{s,t} := U_t \circ U_s^{-1}$ be given by

$$\begin{tikzcd} T_{X_s}M \arrow[rr, "\sim"'] & & T_{X_t}M \\ & \mathbb{R}^n. \arrow[ul, "U_s"] \arrow[ur, "U_t"'] & \end{tikzcd}$$

The isomorphisms (resp. isometries in the Riemannian case)

$$//_{0,t}\colon T_{X_0}M \to T_{X_t}M$$

are called the *stochastic parallel transport along X*.

Theorem 5.33 (Geometric Itô formula). *Let M be a differentiable manifold equipped with a linear connection ∇ (without restriction ∇ torsion-free). Let X be an M-valued semimartingale, U a horizontal lift of X to $\mathrm{L}(TM)$, and $Z = \int_U \vartheta$ the corresponding anti-development of X into $\mathbb{R}^n$. For each $f \in C^\infty(M)$ the following formula holds:*

$$d(f \circ X) = \sum_{i=1}^n (df)(X)\,(Ue_i)\,dZ^i + \frac{1}{2}\sum_{i,j=1}^n \big(\nabla df\big)(X)\,(Ue_i, Ue_j)\,d[Z^i, Z^j], \quad (5.25)$$

or in abbreviated form (see Theorem 5.29),

$$d(f \circ X) = (df)(U\,dZ) + \frac{1}{2}\nabla df(dX, dX). \tag{5.26}$$

Proof. From $dU = \sum_i L_i(U) \circ dZ^i$ we first see that

$$\begin{aligned} d(f \circ X) = d(f \circ \pi \circ U) &= \sum_i L_i(f \circ \pi)(U) \circ dZ^i \\ &= \sum_i L_i(f \circ \pi)(U)\,dZ^i + \frac{1}{2}\sum_{i,j} L_i L_j(f \circ \pi)(U)\,d[Z^i, Z^j], \end{aligned}$$

where $L_i(f \circ \pi)(u) = d(f \circ \pi)_u L_i(u) = (df)_{\pi(u)}(d\pi)_u h_u(ue_i) = (df)_{\pi(u)}(ue_i)$. A straightforward calculation, however, shows that

$$L_i L_j(f \circ \pi)(u) = \nabla df(ue_i, ue_j),$$

from where formula (5.25) results. □

Remark 5.34. Let M be a Riemannian manifold with its Levi-Civita connection. Denoting by $\Delta^{\text{hor}} = \sum_i L_i^2$ the horizontal Laplacian on $\mathrm{O}(TM)$ and by Δ the Laplace–Beltrami operator on M, then for each $f \in C^\infty(M)$ the following relation holds:

$$\Delta^{\text{hor}}(f \circ \pi) = (\Delta f) \circ \pi.$$

Proof. Indeed, for $u \in \mathrm{O}(TM)$, we have

$$\sum_i L_i{}^2(f \circ \pi)(u) = \sum_i \nabla df(ue_i, ue_i) = (\operatorname{trace}\nabla df)\pi(u) = (\Delta f) \circ \pi(u). \qquad \square$$

Theorem 5.35. *Let M be a differentiable manifold equipped with a torsion-free linear connection ∇. Let X be an M-valued semimartingale and U_0 an $\mathrm{L}(TM)$-valued $\mathscr{F}_0$-measurable random variable such that $\pi \circ U_0 = X_0$ a.s.; furthermore let $Z = \int_U \vartheta$ be the anti-development of X into $\mathbb{R}^n$ with respect to the initial frame U_0.*

(i) Then X is a ∇-martingale on M if and only if Z is a local martingale on $\mathbb{R}^n$.

(ii) If ∇ is the Levi-Civita connection to some Riemannian metric g on M and if U_0 takes its values in $\mathrm{O}(TM)$, then X is a Brownian motion on (M, g) if and only if Z is a Brownian motion on $\mathbb{R}^n$ (more precisely, a Brownian motion on $\mathbb{R}^n$ stopped at the lifetime ζ of X).

Proof. (i) According to Definition 5.13, X is a ∇-martingale if

$$d(f \circ X) - \tfrac{1}{2}(\nabla df)(dX, dX) \overset{\mathrm{m}}{=} 0,$$

for functions $f \in C^\infty(M)$. By means of the geometric Itô formula (Theorem 5.33) this means that

$$\sum_i (df)(X)\,(Ue_i)\,dZ^i \overset{\mathrm{m}}{=} 0$$

for any $f \in C^\infty(M)$, which is easily seen to be equivalent to the condition that Z is a local martingale.

(ii) According to Theorem 5.17, X is a Brownian motion on (M, g) if

$$d(f \circ X) - \tfrac{1}{2}(\Delta f \circ X)\,dt \overset{\mathrm{m}}{=} 0,$$

for all $f \in C^\infty(M)$. According to formula (5.25), clearly if Z is a Brownian motion $\mathbb{R}^n$, then X will be a Brownian motion on (M, g). Conversely, if X is a Brownian motion on (M, g), then by Lévy's characterization of M-valued Brownian motions (Theorem 5.17), X is a ∇-martingale, and thus Z a local martingale by part (i). On the other hand, we have $Z^i = \int_U \vartheta^i$, where $\vartheta_u^i = \langle d\pi(\,\cdot\,), ue_i\rangle = \pi^*\langle\cdot, ue_i\rangle$. We may calculate the quadratic variation of Z using Remark 5.11 as follows:

$$\begin{aligned} d[Z^i, Z^j] &= d\Big[\textstyle\int_U \vartheta^i, \int_U \vartheta^j\Big] = (\vartheta^i \otimes \vartheta^j)\,(dU, dU) \\ &= \pi^*\big(\langle\cdot, Ue_i\rangle \otimes \langle\cdot, Ue_j\rangle\big)\,(dU, dU) \\ &= \big(\langle\cdot, Ue_i\rangle \otimes \langle\cdot, Ue_j\rangle\big)\,(dX, dX) \\ &= \mathrm{trace}\big(\langle\cdot, Ue_i\rangle \otimes \langle\cdot, Ue_j\rangle\big)(X)\,dt = \delta_{ij}\,dt. \end{aligned}$$

By means of Lévy's characterization for Brownian motions on $\mathbb{R}^n$ we see that Z is a Brownian motion. □

Theorem 5.35 provides a canonical construction of Brownian motions on Riemannian manifolds. One obtains Brownian motions on (M, g) with starting point $x \in M$ as a stochastic development of a Brownian motion B on $\mathbb{R}^n$ as follows. Choose $u \in \mathrm{O}(TM)$ such that $\pi(u) = x$ and solve the SDE

$$dU = \sum_{i=1}^n L_i(U) \circ dB^i, \quad U_0 = u.$$

According to Theorem 5.35, then $X = \pi \circ U$ will be a Brownian motion on (M, g) starting from $X_0 = x$.

Remark 5.36. Let X be an M-valued semimartingale with starting point $x \in M$. The anti-development Z of X into $\mathbb{R}^n$ (see Definition 5.27) required the choice of a frame u above x,

$$Z = \int_U \vartheta, \quad U_0 = u.$$

Considering the anti-development of X into T_xM, i.e.,

$$Z' = U_0 \int_U \vartheta,$$

makes the notion intrinsic. Then we have the formula

$$dZ' = U_0 U_t^{-1} \circ dX = //_{0,t}^{-1} \circ dX.$$

5.4 Subelliptic diffusions and sub-Riemannian Brownian motions. In this subsection we want to adapt the results developed so far from the Riemannian to the sub-Riemannian setting.

A *sub-Riemannian structure* on a differentiable manifold M is a pair $(\mathcal{H}, g)$, where $\mathcal{H}$ is a subbundle of TM and g is a positive-definite metric tensor defined only on $\mathcal{H}$. Any sub-Riemannian structure induces a vector bundle morphism

$$\sharp\colon T^*M \to TM,$$

determined by the properties $\sharp(T^*M) = \mathcal{H}$ and $q(v) = g(v, \sharp q)$ for any $q \in T^*M$ and $v \in \mathcal{H}$. The kernel of $\sharp$ is the subbundle $\operatorname{Ann}(\mathcal{H}) \subseteq T^*M$ of elements of T^*M vanishing on $\mathcal{H}$. Then the so-called co-metric g^* on T^*M, defined by

$$g^*(q_1, q_2) = q_1(\sharp q_2), \quad q_1, q_2 \in T_x^*M,\ x \in M,$$

degenerates along $\operatorname{Ann}(\mathcal{H})$. It is obvious that sub-Riemannian structures on M and co-metrics degenerating along a subbundle of T^*M are equivalent structures.

Definition 5.37. Let $(\mathcal{H}, g)$ be a sub-Riemannian structure on M. A continuous semimartingale X taking values in M is called *horizontal*, or a *sub-Riemannian diffusion*, if

$$\int \alpha(\circ\, dX) = 0, \quad \text{for all } \alpha \in \Gamma(\operatorname{Ann}(\mathcal{H})).$$

Here $\int \alpha(\circ\, dX) \equiv \int_X \alpha$ denotes the Stratonovich integral of α along X.

Remark 5.38. Note that if X is a horizontal semimartingale then $\int_X \beta$ is well defined for $\beta \in \Gamma(\mathcal{H}^*)$. The same holds true for $\int b(dX, dX)$ if $b \in \Gamma(\mathcal{H}^* \otimes \mathcal{H}^*)$ is a bilinear form on $\mathcal{H}^*$. In particular, the *sub-Riemannian quadratic variation* of X,

$$[X, X] = \int g(dX, dX), \tag{5.27}$$

is well defined.

As we saw in Theorem 3.22, a continuous semimartingale X taking values in M can always be obtained as the solution of an SDE of the type $dX = \sum_i A_i(X) \circ dZ^i$. Then obviously X is horizontal if the vector fields A_i are horizontal in the sense that $A_i \in \Gamma(\mathcal{H})$.

To define horizontal martingales in the sub-Riemannian setting we need to specify a connection ∇. To this end it is enough to have a so-called *partial connection* on $\mathcal{H}$ (see [32] and [20, Section 2]),

$$\Gamma(\mathcal{H}) \times \Gamma(\mathcal{H}) \to \Gamma(\mathcal{H}), \quad (A, B) \mapsto \nabla_A B,$$

and correspondingly the *partial Hessian* of a function $f \in C^\infty(M)$,

$$\nabla df \equiv \operatorname{Hess} f \in \Gamma(\mathcal{H}^* \otimes \mathcal{H}^*), \quad (\nabla df)(A, B) = ABf - (\nabla_A B)f, \quad f \in C^\infty(M).$$

Definition 5.39. Let $(\mathcal{H}, g)$ be a sub-Riemannian structure on M and ∇ a partial connection on $\mathcal{H}$. A continuous semimartingale X taking values in M is called a *horizontal martingale* if H is horizontal and for any $f \in C^\infty(M)$,

$$d(f \circ X) - \tfrac{1}{2}(\nabla df)(dX, dX) \overset{\mathrm{m}}{=} 0.$$

Remark 5.40. (a) Often partial connections are induced from (full) connections $\tilde\nabla$ on M in terms of a projection $p\colon TM \to \mathcal{H}$ as

$$\nabla_A B = p(\tilde\nabla_A B), \quad A, B \in \Gamma(\mathcal{H}).$$

For instance, one may extend the metric g from $\mathcal{H}$ to a full Riemannian metric $\tilde g$ on TM (this is a common procedure in the case of sub-Riemannian structures related to Riemannian foliations); then

$$\nabla_A B = \mathrm{pr}_{\mathcal{H}}(\tilde\nabla_A B), \quad A, B \in \Gamma(\mathcal{H}) \tag{5.28}$$

(where $\tilde\nabla$ is the Levi-Civita connection to $\tilde g$ on M and $\mathrm{pr}_{\mathcal{H}}$ is the orthogonal projection of TM onto $\mathcal{H}$) defines a partial connection on $\mathcal{H}$ which is, moreover, *metric*, i.e., $\nabla_A g = 0$, for all $A \in \Gamma(\mathcal{H})$. Note that (5.28) is the horizontal part of the so-called *Bott connection* on M; see [57, Chapter 5].

(b) More generally, it is straightforward to show the following result. Given a projection $p\colon TM \to \mathcal{H}$, there exists a unique partial connection on $\mathcal{H}$ which is metric and has the property

$$\nabla_A B - \nabla_B A - p[A, B] = 0.$$

This is actually the connection in (5.28) defined relative to any Riemannian metric $\tilde g$ such that p is the orthogonal projection.

Theorem 5.17 is now easily adapted to the sub-Riemannian setting. Given a partial connection ∇ on $\mathcal{H}$ which is metric (i.e., $\nabla_A g = 0$, for all $A \in \Gamma(\mathcal{H})$) we define the *sub-Laplacian* $\Delta_{\mathcal{H}}$ relative to $(\mathcal{H}, g, \nabla)$ as

$$\Delta_{\mathcal{H}} f = \operatorname{trace}_{\mathcal{H}} \nabla df, \quad f \in C^\infty(M).$$

If the partial connection ∇ is as in Remark 5.40(b), then $\Delta_{\mathcal{H}}$ coincides with the sub-Laplacian relative to the complement $\mathcal{V} = \ker p$ as defined in [21, Section 2.2].

Theorem 5.41 (Lévy's characterization of sub-Riemannian Brownian motions on M). *Let $(\mathcal{H}, g)$ be a sub-Riemannian structure on M and ∇ a partial metric connection on $\mathcal{H}$. For a horizontal semimartingale X of maximal lifetime on M the following conditions are equivalent:*

(i) X is a sub-Riemannian Brownian motion on M, i.e., for any $f \in C^\infty(M)$, the real-valued process

$$f \circ X - \frac{1}{2} \int (\Delta_{\mathcal{H}} f) \circ X \, dt$$

is a local martingale.

(ii) X is a ∇-martingale such that $[f(X), f(X)] = \int g^*(df, df)(X_t)\, dt$ for every $f \in C^\infty(M)$.

(iii) X is a ∇-martingale such that $\int b(dX, dX) = \int (\operatorname{trace}_{\mathcal{H}} b)(X)\, dt$ for every $b \in \Gamma(\mathcal{H}^* \otimes \mathcal{H}^*)$.

In particular, for the sub-Riemannian quadratic variation (5.27) of X, we then have

$$\int_0^t g(dX, dX) = t \dim \mathcal{H}.$$

Analogously to Theorem 5.18, we can construct sub-Riemannian Brownian motions on M as solutions to SDEs.

Theorem 5.42 (Sub-Riemannian Brownian motions as solutions of an SDE on M). *Let $(\mathcal{H}, g)$ be a sub-Riemannian structure on M and ∇ a partial metric connection on $\mathcal{H}$. Consider an SDE of the type*

$$dX = A_0(X)\, dt + A(X) \circ dB \tag{5.29}$$

with $A_0 \in \Gamma(\mathcal{H})$ and $A \in \Gamma\big(\operatorname{Hom}(M \times \mathbb{R}^r, \mathcal{H})\big)$; the driving process B is a Brownian motion on $\mathbb{R}^r$ (for some r).

Then maximal solutions to (5.29) are sub-Riemannian Brownian motions M if the two following conditions are satisfied:

(i) $A_0 = -\frac{1}{2}\sum_i \nabla_{A_i} A_i$ *with* $A_i \equiv A(\cdot)e_i$ *for* $i = 1, \ldots, r$.

(ii) *The map* $A(x)^*\colon \mathcal{H}_x \to \mathbb{R}^r$ *is an isometric embedding for every* $x \in M$, *i.e.,* $A(x)\,A(x)^* = \mathrm{id}_{\mathcal{H}_x}$, *where* $A(x)^*$ *is the adjoint to* $A(x) \in \mathrm{Hom}(\mathbb{R}^r, \mathcal{H}_x)$.

The problem of defining sub-Riemannian Brownian motions and corresponding random walk approximations has recently been addressed in [15].

The results of Section 5.3 easily carry over to the case of horizontal martingales and sub-Riemannian Brownian motions. Instead of $\mathrm{L}(TM)$ (resp. $\mathrm{O}(TM)$), we work with the G-principal bundle $P = \mathrm{L}(\mathcal{H})$ of frames in $\mathcal{H}$ (resp. $P = \mathrm{O}(\mathcal{H})$ of orthonormal frames in $\mathcal{H}$), where now $G = \mathrm{GL}(k;\mathbb{R})$ (resp. $G = \mathrm{O}(k;\mathbb{R})$), and $k = \dim \mathcal{H}$. In other words,

$$P_x = \left\{u\colon \mathbb{R}^k \to \mathcal{H}_x \mid u \text{ linear isomorphisms (resp. } u \text{ linear isometry)}\right\}, \quad x \in M.$$

A partial connection ∇ on $\mathcal{H}$ (resp. a metric partial connection ∇ on $\mathcal{H}$), now induces a G-invariant subbundle $H \subset TP$ such that

$$\pi_*\colon\ H_u \xrightarrow{\ \sim\ } \mathcal{H}_{\pi(u)},$$

where π is the projection $P \to M$. In terms of the horizontal lift of this G-connection,

$$h\colon\ \pi^*\mathcal{H} \xrightarrow{\ \sim\ } H \hookrightarrow TP,$$

we have the standard-horizontal vector fields

$$L_i \in \Gamma(TP), \quad L_i(u) = h_u(ue_i),\ u \in P,\ i = 1, \ldots, k.$$

The $\mathfrak{g}$-valued connection form ω and the $\mathbb{R}^k$-valued canonical 1-form ϑ are defined as in the Riemannian case, but for a partial connection they are given only on $H \oplus V$ with $V = \ker d\pi$, and no longer globally on TP; in other words,

$$\omega \in \Gamma((H^* \oplus V^*) \otimes \mathfrak{g}) \quad \text{and} \quad \vartheta \in \Gamma((H^* \oplus V^*) \otimes \mathbb{R}^k).$$

One can now define stochastic developments of $\mathbb{R}^k$-valued semimartingales according to

$$\begin{aligned} dU &= \sum_{i=1}^k L_i(U) \circ dZ^i, \quad U_0 = u, \\ X &= \pi(U), \end{aligned}$$

as we did in Section 5.3. The resulting processes X will be horizontal semimartingales on M. Horizontal lifts of such semimartingales X to $P = \mathrm{L}(\mathcal{H})$ can be established as in the Riemannian case, for instance, by representing X as the solution to an SDE on M with vector fields $A_i \in \Gamma(\mathcal{H})$ and solving the "horizontally lifted" SDE on P (see the proof of Theorem 5.30).

Theorem 5.43 (Geometric Itô formula for horizontal diffusions). *Let $(\mathcal{H}, g)$ be a sub-Riemannian structure on M and ∇ a partial connection on $\mathcal{H}$. Let X be an M-valued horizontal semimartingale, U a horizontal lift of X to $P = \mathrm{L}(\mathcal{H})$ and $Z = \int_U \vartheta$ the corresponding anti-development of X into $\mathbb{R}^k$. For each $f \in C^\infty(M)$ the following formula holds:*

$$d(f \circ X) = \sum_{i=1}^{k} (df)(X)\,(Ue_i)\,dZ^i + \frac{1}{2}\sum_{i,j=1}^{k}\big(\nabla df\big)(X)\,(Ue_i, Ue_j)\,d[Z^i, Z^j], \tag{5.30}$$

or in abbreviated form,

$$d(f \circ X) = (df)(U\,dZ) + \tfrac{1}{2}\nabla df(dX, dX). \tag{5.31}$$

This finally gives the following sub-Riemannian version of Theorem 5.35.

Theorem 5.44. *Let $(\mathcal{H}, g)$ be a sub-Riemannian structure on M and ∇ a partial connection on $\mathcal{H}$. Let X be an M-valued horizontal semimartingale and U_0 an $\mathrm{L}(\mathcal{H})$-valued $\mathscr{F}_0$-measurable random variable such that $\pi \circ U_0 = X_0$ a.s.; furthermore, let $Z = \int_U \vartheta$ be the anti-development of X into $\mathbb{R}^k$ with respect to the initial frame U_0.*

(i) Then X is a ∇-martingale on M if and only if Z is a local martingale on $\mathbb{R}^k$.

(ii) If ∇ is a metric partial connection on $\mathcal{H}$ and if U_0 takes its values in $\mathrm{O}(\mathcal{H})$, then X is a sub-Riemannian Brownian motion on M if and only if Z is a Brownian motion on $\mathbb{R}^k$ (more precisely, a Brownian motion on $\mathbb{R}^k$ stopped at the lifetime ζ of X).

Following Remark 5.36 we have the following remark.

Remark 5.45. Let ∇ be a partial connection on $\mathcal{H}$ and let be X an M-valued horizontal semimartingale with starting point $x \in M$. Let Z be the anti-development of X into $\mathcal{H}_x$,

$$Z = \int /\!/_{0,t}^{-1} \circ dX.$$

(a) Then X is a ∇-martingale on M if and only if its anti-development Z into $\mathcal{H}_x$ is a local martingale on $\mathcal{H}_x$.

(b) If ∇ is a metric partial connection on $\mathcal{H}$, then X is a sub-Riemannian Brownian motion on M if and only if its anti-development Z into $\mathcal{H}_x$ is a Brownian motion on $\mathcal{H}_x$.

Here $/\!/_{0,t}\colon \mathcal{H}_{X_0} \xrightarrow{\sim} \mathcal{H}_{X_t}$ denotes the stochastic parallel transport of horizontal tangent vectors along X. Recall that the $/\!/_{0,t}$ are linear isomorphisms for a partial connection, and isometries for a metric partial connection.

6 Control theory and support theorems

6.1 Control systems. Consider a Stratonovich SDE on M of the type

$$dX = A_0(X)dt + \sum_{i=1}^{r} A_i(X) \circ dB^i, \tag{6.1}$$

driven by a Brownian motion $B = (B^1, \dots, B^r)$ on $\mathbb{R}^r$.

Definition 6.1. Solutions X to SDE (6.1) are called *hypoelliptic diffusions* if the vector fields $A_1, \dots, A_r$ are *bracket generating* in the sense that

$$\dim \operatorname{Lie}(A_1, \dots, A_r)(x) = \dim M, \quad \text{for all } x \in M. \tag{6.2}$$

To the SDE (6.1) we associate the control system

$$\dot{x}(t) = A_0(x(t)) + \sum_{i=1}^{r} A_i(x(t))u^i(t), \tag{6.3}$$

where the control $u = u(\cdot)$ lies in

$$\mathcal{U} = \{u\colon \mathbb{R}_+ \longrightarrow \mathbb{R}^r \text{ piecewise constant}\}; \tag{6.4}$$

see for instance [56]. In the space $\mathcal{U}$ of controls we could have equally taken u piecewise smooth or piecewise continuous with values in $\mathbb{R}^r$.

We denote by

- $X_t(x)$ the solution to SDE (6.1) with starting point $X_0 = x$; and by
- $\phi_t(x, u)$ the solution to the control system (6.3) with initial condition $x(0) = x$ and $u = u(\cdot) \in \mathcal{U}$.

For simplicity, in the remainder of Section 6, all vector fields of the form

$$A_0 + \sum_{i=1}^{r} A_i u^i, \quad u \in \mathbb{R}^r \text{ fixed} \quad (\text{"frozen vector fields"})$$

are assumed to be complete.

We consider the following orbits:

$$O^+(x) := \{y \in M : \ y = \phi_t(x, u),\ t \geq 0,\ u = u(\cdot) \in \mathcal{U}\} \quad \text{"forward orbit"},$$
$$O_t^+(x) := \{y \in M : \ y = \phi_t(x, u),\ u = u(\cdot) \in \mathcal{U}\} \quad \text{"forward orbit at time } t\text{"}.$$

We call the control system (6.3)

- *completely controllable* if $O^+(x) = M$ for each $x \in M$;
- *strongly controllable* if $O_t^+(x) = M$ for each $t > 0$ and each $x \in M$;
- *completely accessible* if $O^+(x)$ has nonvoid interior for each $x \in M$;
- *strongly accessible* if $O_t^+(x)$ has nonvoid interior for each $t > 0$ and each $x \in M$.

Remark 6.2. Geometric control theory characterizes properties of control problems in terms of Lie-algebra conditions on the vector fields $A_0, \dots, A_r$. For example, for system (6.3),

(1) complete accessibility holds if $\dim \mathrm{Lie}(A_0, A_1, \dots, A_r)(x) = \dim M$ for each $x \in M$;

(2) strong accessibility holds if $\dim \mathrm{Lie}\big(A_0 + \frac{\partial}{\partial t}, A_1, \dots, A_r\big)(t, x) = \dim M + 1$ for each $t > 0$ and $x \in M$;

(3) strong controllability holds if $\dim \mathrm{Lie}(A_1, \dots, A_r) = \dim M$ for each $x \in M$.

See for instance [2, 16, 31].

6.2 Support theorems. The famous support theorem of Stroock–Varadhan (1972) establishes a bridge between the theory of SDEs and control theory, more precisely, between Eq. (6.1) and Eq. (6.3).

Induced by $X_{\bullet}(x) \colon \Omega \longrightarrow C(\mathbb{R}_+, M)$, we have the following measures:

$$\mathbb{P}_x := \mathbb{P} \circ X_{\bullet}(x)^{-1} \quad \text{probability measure on } C_x(\mathbb{R}_+, M);$$
$$\mathbb{P}_{t,x} := \mathbb{P} \circ X_t(x)^{-1} \quad \text{probability measure on } M.$$

Here $C_x(\mathbb{R}_+, M)$ denotes the space of continuous trajectories $\mathbb{R}_+ \to M$ starting from x at time 0.

Theorem 6.3 (Support theorem; Stroock–Varadhan [55]). *For the supports of the probability measures $\mathbb{P}_x$ (resp. $\mathbb{P}_{t,x}$), the following properties hold:*

(I) Path space: On $C_x(\mathbb{R}_+, M)$ we have

$$\operatorname{supp} \mathbb{P}_x = \overline{\big\{\phi_{\bullet}(x, u) \colon u \in \mathcal{U}\big\}}.$$

(II) *State space: On M we have*

$$\operatorname{supp} \mathbb{P}_{t,x} = \bar{O}_t^+(x),$$
$$\operatorname{supp} G_\lambda(x,\cdot) = \bar{O}^+(x),$$

where

$$G_\lambda(x,\cdot) = \int_0^\infty e^{-\lambda t}\mathbb{P}_{t,x}(\cdot)\,dt, \quad \lambda > 0$$

denotes Green's measure *with exponent λ on M.*

Proof. The support theorem is proved by approximating the driving Brownian motion B through its piecewise linear polygonal approximation

$$B_t^\pi = (t_{i+1} - t_i)^{-1}\left[(t_{i+1} - t)B_{t_i} + (t - t_i)B_{t_{i+1}}\right], \quad t_i \le t \le t_{i+1},$$

for partitions

$$\pi : 0 = t_0 < t_1 < t_2 < \cdots .$$

See Stroock–Varadhan [55], Kunita [33], and Ichihara–Kunita [28, 29] for technical details. □

Corollary 6.4. *Suppose that the vector fields $A_1, \ldots, A_r$ are bracket generating in the sense that condition* (6.2) *holds. Then*

$$\operatorname{supp} \mathbb{P}_x = C_x(\mathbb{R}_+, M) \quad \text{and} \quad \operatorname{supp} \mathbb{P}_{t,x} = M.$$

Proof. See Remark 6.2 above, as well as Stroock–Varadhan [55]. □

Remark 6.5. For stochastic representations of solutions to classical boundary value problems on a relatively compact open domain D related to the Hörmander-type operator

$$L = A_0 + \frac{1}{2}\sum_{i=1}^r A_i^2$$

(see Sections 4.2 and 2.3) the following "finite exit time condition" has been crucial.

(a) *Finite exit time condition.* For each $x \in D$, the solution X_t to SDE (6.1) with starting point $X_0 = x$ exits D in finite time almost surely.

In terms of the associated control system, a sufficient condition for (a) to hold is given by the following escape condition.

(b) The domain D is said to satisfy the *escape condition* if, for each $x \in D$, there is a control $u = u(\cdot) \in \mathcal{U}$ such that the path $t \mapsto \phi_t(x,u)$ in $C_x(\mathbb{R}_+, M)$ escapes from $\bar{D}$ (i.e., there exists a $T > 0$ such that $\phi_T(x,u) \notin \bar{D}$).

The proof that the escape condition implies the finite exit time condition proceeds along the lines of the support theorem; see [54].

7 Stochastic flows of diffeomorphisms

We consider again an SDE on M of the type

$$dX = A(X) \circ dZ, \tag{7.1}$$

where $Z = (t, B^1, \dots, B^r)$ with $B = (B^1, \dots, B^r)$ a Brownian motion on $\mathbb{R}^r$. In equivalent form, Eq. (7.1) can be written as

$$dX = A_0(X)\, dt + \sum_{i=1}^{r} A_i(X) \circ dB^i, \tag{7.2}$$

where the vector fields $A_i = A(\cdot)e_i \in \Gamma(TM)$ are taken with respect to the standard basis $(e_0, e_1, \dots, e_r)$ of $\mathbb{R}^{r+1}$.

Let $(X_t(\cdot), \zeta(\cdot))$ be the partial flow to

$$L = A_0 + \frac{1}{2}\sum_{i=1}^{r} A_i^2,$$

in the sense that for each $x \in M$, the process $X_t(x)$ has maximal lifetime $\zeta(x)$ and solves SDE (7.1). For $t \geq 0$ fixed, we then have the random set

$$M_t(\omega) = \{x \in M\colon\ t < \zeta(x)(\omega)\}, \quad \omega \in \Omega.$$

Theorem 7.1. *The following assertions hold $\mathbb{P}$-almost surely (in $\omega \in \Omega$):*

(i) $M_t(\omega)$ *is an open subset of* M *for each* $t \geq 0$, *i.e.,* $\zeta(\cdot)(\omega)$ *is lower semicontinuous on* M.

(ii) *For each* $t \geq 0$, *the map*

$$X_t(\cdot)(\omega)\colon M_t(\omega) \longrightarrow R_t(\omega)$$

is a diffeomorphism onto an open subset $R_t(\omega)$ *of* M.

(iii) *The path map* $s \longmapsto X_s(\cdot)(\omega)$ *is continuous from* $[0, t]$ *into* $C^\infty(M_t(\omega), M)$ *with its* C^∞*-topology.*

Proof. See Kunita's theory of stochastic flows [36]. □

Remark 7.2. Under "mild" growth conditions (see [36] for precise statements) on the vector fields $A_0, \dots, A_r$ and their derivatives (which are trivially fulfilled if M is compact), we have almost surely

$$X_t(\cdot) \in \mathrm{Diff}(M), \quad \text{for all } t.$$

7.1 Tangent flows and pullback of vector fields under stochastic flows.

Proposition 7.3. *In the situation of a partial flow to the SDE*

$$dX = \sum_{i=0}^{r} A_i(X) \circ dZ^i \tag{7.3}$$

we consider the "tangent flow" $X_{t*} := TX_t$, defined as the differential of the map $x \mapsto X_t(x)$,

$$T_x M \to T_{X_t(x)} M, \quad v \longmapsto X_{t*}v, \quad x \in M_t(\omega).$$

The tangent map

$$U_t := X_{t*}$$

solves the (formally) differentiated SDE (7.3), i.e.,

$$dU = \sum_{i=0}^{r} (DA_i)_X \, U \circ dZ^i, \tag{7.4}$$

where $(DA_i)_X = T_X A_i \equiv T_{\pi(U)} A_i$. In addition, the inverse tangent flow $U'_t = X_{t*}^{-1}$ solves the SDE

$$dU' = -\sum_{i=0}^{r} U'(DA_i)_X \circ dZ^i. \tag{7.5}$$

Proof. These are standard formulas in the theory of SDEs and are checked in a straightforward way using stochastic calculus; see [35, 36]. □

We now come to a crucial notion: the pullback of a vector field V on M under a stochastic flow $x \mapsto X_t(x)$. More precisely, for $V \in \Gamma(TM)$ we consider the (random) vector field $X_{t*}^{-1}V$ on M_t defined as

$$(X_{t*}^{-1}V)_x = (T_x X_t)^{-1} V_{X_t(x)} \in T_x M, \quad x \in M_t.$$

In other words, we have

$$(X_{t*}^{-1}V)(f) = V(f \circ X_t^{-1}) \circ X_t, \quad f \in C^\infty(M).$$

Lemma 7.4. *The pullback vector field $X_{t*}^{-1}V$ satisfies the equation*

$$d(X_{t*}^{-1}V) = \sum_{i=0}^{r} X_{t*}^{-1}[A_i, V] \circ dZ_t^i.$$

In the special form of SDE (7.2) this means

$$d(X_{t*}^{-1}V) = X_{t*}^{-1}[A_0, V]\, dt + \sum_{i=1}^{r} X_{t*}^{-1}[A_i, V] \circ dB_t^i.$$

Proof. For instance, see [34, Section 5]. □

Corollary 7.5. *Suppose that the vector field V commutes with $A_0, \dots, A_r$. Then we have $X_{t*}^{-1}V = V$.*

Remark 7.6. There are analogous formulas for the pushforward vector fields $X_{t*}V$ on R_t, e.g.,

$$d(X_{t*}V) = \sum_{i=0}^{r} [X_{t*}A_i, V] \circ dZ_t^i,$$

respectively,

$$d(X_{t*}V) = [X_{t*}A_0, V]\,dt + \sum_{i=1}^{r} [X_{t*}A_i, V] \circ dB_t^i.$$

7.2 Malliavin's covariance matrix.

Definition 7.7 (Malliavin's covariance matrix). Suppose that an SDE of the type

$$dX = A_0(X)\,dt + \sum_{i=1}^{r} A_i(X) \circ dB^i$$

is given. For $t > 0$, the tensor

$$C_t(x) = \sum_{i=1}^{r} \int_0^t (X_{s*}^{-1}A_i)_x \otimes (X_{s*}^{-1}A_i)_x\,ds \in T_xM \otimes T_xM, \quad x \in M_t \tag{7.6}$$

defines a smooth (random) section of the bundle $TM \otimes TM$ over M_t. This section is usually called *Malliavin's covariance matrix*.

Malliavin's covariance matrix is at the heart of the so-called Malliavin calculus, also known as stochastic calculus of variations [47, 50]. In the sequel we use different notions of writing Malliavin's covariance matrix (7.6).

Notation 7.8. Putting together the diffusion vector fields $A_1, \dots, A_r$ into a bundle map $A\colon M \times \mathbb{R}^r \to TM$ over M, we have

$$(X_{s*}^{-1}A)_x : \mathbb{R}^r \to T_xM, \quad z \mapsto \sum_{i=1}^{r} (X_{s*}^{-1}A_i)_x z^i. \tag{7.7}$$

(Note that the drift vector field A_0 is not included.) Considering the dual map to Eq. (7.7),

$$(X_{s*}^{-1}A)_x^* :\ T_x^*M \longrightarrow (\mathbb{R}^r)^* \equiv \mathbb{R}^r,$$

we may read Malliavin's covariance matrix (7.6) as

$$C_t(x) = \int_0^t (X_{s*}^{-1}A)_x (X_{s*}^{-1}A)_x^*\,ds \in \mathrm{Hom}(T_xM, T_xM), \quad x \in M_t.$$

Example 7.9. On $\mathbb{R}^2$ consider the SDE

$$dX_t = A_0(X_t)\,dt + A_1(X_t)\circ dB_t^1, \quad X_0 = x = (x^1, x^2), \tag{7.8}$$

where $A_0 = x^1\frac{\partial}{\partial x^2}$ and $A_1 = \frac{\partial}{\partial x^1}$. Obviously SDE (7.8) can be written as

$$dX_t^1 = dB_t^1, \quad dX_t^2 = X_t^1\,dt, \quad (X_0^1, X_0^2) = (x^1, x^2),$$

and so we have an explicit expression for the solution as

$$\begin{cases} X_t^1 = x^1 + B_t^1, \\ X_t^2 = x^2 + x^1 t + \int_0^t B_s^1\,ds. \end{cases}$$

Thus

$$X_{t*} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \quad \text{and} \quad X_{t*}^{-1} = \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}.$$

For Malliavin's covariance matrix we get

$$C_t(x) = \begin{pmatrix} t & -t^2/2 \\ -t^2/2 & t^3/3 \end{pmatrix}.$$

Note that $C_t(x)$ is independent of x and invertible for $t > 0$. SDE (7.8) is degenerate in the sense that A_1 does not span $T_x\mathbb{R}^2$, but observe that $[A_0, A_1] = \frac{\partial}{\partial x^2}$. It is easy to see that the random vector (X_t^1, X_t^2) has a Gaussian distribution with covariance

$$\begin{pmatrix} t & t^2/2 \\ t^2/2 & t^3/3 \end{pmatrix}.$$

For $t > 0$ the covariance is nonsingular, and hence (X_t^1, X_t^2) has a smooth Gaussian density function with respect to two-dimensional Lebesgue measure.

8 Stochastic flows and hypoellipticity

The purpose of this section is to sketch a probabilistic proof of Hörmander's hypoellipticity theorem. We follow some of the arguments in Bismut [14].

Consider a second-order PDO in Hörmander form,

$$L = A_0 + \frac{1}{2}\sum_{i=1}^r A_i^2, \tag{8.1}$$

on a differentiable manifold M with smooth vector fields $A_0, \dots, A_r$. For simplicity, we assume again that all vector fields of the form

$$A_0 + \frac{1}{2} \sum_{i=1}^{r} A_i u^i, \qquad u \in \mathbb{R}^r$$

are complete.

We denote by $\mathscr{D}'(M)$ the space of distributions on M. Recall that an operator L of type (8.1) is called *hypoelliptic* if $u \in \mathscr{D}'(M)$ and $Lu|U \in C^\infty(U)$, where $U \subset M$ is open, implying that $u|U \in C^\infty(U)$.

Our goal is to show hypoellipticity of operator (8.1) under a certain Hörmander-type nondegeneracy.

8.1 Hypoellipticity under Hörmander conditions. Consider the following two canonical measures on M:

$$\mathbb{P}_t(x, dy) := \mathbb{P}\{X_t(x) \in dy\}, \quad \text{and} \tag{8.2}$$

$$G_\lambda(x, dy) := \int_0^\infty e^{-\lambda t}\, \mathbb{P}\{X_t(x) \in dy\}\, dt, \quad \lambda > 0. \tag{8.3}$$

Remark 8.1. In Section 4, these measures were used for stochastic representation formulas of classical PDEs.

(i) Recall that every bounded solution $u(t, x)$ to the initial value problem

$$\frac{\partial}{\partial t} u = Lu, \quad u|_{t=0} = f$$

can be represented as

$$u(t, x) = \int P_t(x, dy)\, f(y) = E[f \circ X_t(x)].$$

(ii) According to the Feynman–Kac formula (4.5), solutions to

$$(\lambda - L)u = f$$

have a representation as

$$u(x) = \int G_\lambda(x, dy)\, f(y), \quad x \in M.$$

In this sense, the operator G_λ defines the inverse to $\lambda - L$, formally $G_\lambda = (\lambda - L)^{-1}$.

Choosing a smooth volume measure, vol, on M, we now come to the following fundamental question.

Problem 8.2. When do measures like $P_t(x,\cdot)$ or $G_\lambda(x,\cdot)$ have densities with respect to vol?

Definition 8.3. To the vector fields $A_0,\dots,A_r$ defining the operator

$$L = A_0 + \frac{1}{2}\sum_{i=1}^{r} A_i^2,$$

we associate several important Lie algebras [28, 29, 5].

- On M,

$$\begin{aligned} \mathscr{L} &:= \operatorname{Lie}(A_0, A_1, \dots, A_r),\\ \mathscr{B} &:= \operatorname{Lie}(A_1, \dots, A_r),\\ \mathscr{I} &:= \text{ideal in } \mathscr{L} \text{ generated by } \mathscr{B}. \end{aligned}$$

- On $M\times\mathbb{R}$,

$$\hat{\mathscr{L}} := \operatorname{Lie}\left(A_0 + \frac{\partial}{\partial t}, A_1, \dots, A_r\right).$$

By definition we have $\mathscr{B}\subset\mathscr{I}\subset\mathscr{L}$.

In terms of these Lie algebras we consider the following Hörmander conditions ($n=\dim M$):

$$\dim\mathscr{L}(x) = n \qquad \text{at each point } x \text{ of } M, \tag{H$_0$}$$

$$\dim\hat{\mathscr{L}}(t,x) = n+1 \qquad \text{at each point } (x,t) \text{ of } M\times\mathbb{R}. \tag{H$_1$}$$

Hypothesis (H$_0$) means that

$$\operatorname{Lie}(A_0, A_1, \dots, A_r)(x) = T_xM, \quad \text{for all } x\in M,$$

whereas (H$_1$) is equivalent to

$$\begin{aligned} &\operatorname{Lie}\bigl(A_1,\dots,A_r,\ [A_i,A_j]_{0\le i,j\le r},\\ &\qquad\qquad [A_i,[A_j,A_k]]_{0\le i,j,k\le r},\dots\bigr)(x) = T_xM, \quad \text{for all } x\in M. \end{aligned}$$

This last condition can be equivalently stated as

$$\dim\mathscr{I}(x) = n \quad \text{at each point } x \text{ of } M.$$

Theorem 8.4 (Hörmander (1967) [26]).

(1) Under hypothesis (H_0) *the operator* L *is hypoelliptic on* M.

(2) Under hypothesis (H_1) *the space-time operator* $L + \frac{\partial}{\partial t}$ *is hypoelliptic on* $M \times \mathbb{R}$.

It can be shown that Hörmander's condition (H_0) is necessary for hypoellipticity for operators L with analytic coefficients. Such is not the case for smooth vector fields $A_0, A_1, \ldots, A_r$. The stochastic approach allows us to derive sharper criteria for hypoellipticity that allow Hörmander's condition to fail on tiny subsets of M; see for instance [11].

Remark 8.5. Let L^* be the formal adjoint operator to L (with respect to the chosen volume measure). It can be written as

$$L^* = \frac{1}{2}\sum_{i=1}^{r} A_i^2 + \tilde{A}_0 + a, \quad \text{where } A_0 + \tilde{A}_0 \in \mathscr{B}.$$

Thus we have the following equivalences:

(i) L satisfies (H_0) if and only if $L^* - a$ satisfies (H_0).

(ii) $L + \frac{\partial}{\partial t}$ satisfies (H_1) if and only if $L^* + \frac{\partial}{\partial t} - a$ satisfies (H_1).

Corollary 8.6. *We may consider the measures* $G_\lambda(x, \cdot)$ *and* $P_t(x, \cdot)$ *as distributions as follows:*

$$G_\lambda(x, dy) \in \mathscr{D}'(M \times M), \quad P_t(x, dy) \in \mathscr{D}'(]0, t[\times M \times M).$$

Denoting by Λ *the diagonal in* $M \times M$*, the following equations hold in the weak sense:*

$$(\lambda - L_x)G_\lambda = 1\!\!1_\Lambda, \qquad (\lambda - L_y^*)G_\Lambda = 1\!\!1_\Lambda,$$

$$\left(\frac{\partial}{\partial t} - L_x\right)P_t(x, dy) = 0, \qquad \left(\frac{\partial}{\partial t} - L_y^*\right)P_t(x, dy) = 0.$$

By means of Hörmander's Theorem 8.4 we obtain the following:

(a) Suppose that condition (H_0) *holds. Then the operator* L *is hypoelliptic and there exists a function* $g_\lambda \in C^\infty((M \times M) \setminus \Delta)$ *such that*

$$G_\lambda(x, dy) = g_\lambda(x, y)\,\mathrm{vol}(dy).$$

(b) Suppose that condition (H_1) *holds. Then the operator* $\frac{\partial}{\partial t} - L$ *is hypoelliptic and there exists a function* $p_t(x, y)$ *in* $C^\infty(]0, \infty[\times M \times M)$ *such that*

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

In the sequel, to avoid technical problems, we assume that $A_0, A_1, \dots, A_r \in \Gamma(TM)$, along with their derivatives, satisfy some growth conditions. Such conditions will be necessary below to make some quantities well defined. To this end we choose a Riemannian metric on M; the volume form $\mathrm{vol}(dy)$ will be taken with respect to this metric.

Standing Hypothesis. Assume that the vector fields $A_0, A_1, \dots, A_r$ are smooth with bounded derivatives of all orders.

Remark 8.7. This hypothesis is far from being necessary, but it guarantees that solutions to the Eqs. (7.3)–(7.5) lie in any L^p space ($1 \le p < \infty$) uniformly over compact time intervals.

The following theorem gives a probabilistic approach to Hörmander's hypoellipticity theorem; see [14, 38, 39, 40, 49, 50], as well as Malliavin's original work [44, 45, 46].

Theorem 8.8. *For $x \in M$, let $X \equiv X_\bullet(x)$ be the solution to the Stratonovich SDE*

$$dX = A_0(X)\,dt + \sum_{i=1}^{r} A_i(X) \circ dB^i, \qquad \text{with initial condition } X_0 = x. \tag{8.4}$$

Suppose that for each $t > 0$ the following two conditions hold true:

(i) The bilinear form

$$C_t(x) := \sum_{i=1}^{r} \int_0^t (X_{s*}^{-1} A_i)_x \otimes (X_{s*}^{-1} A_i)_x \, ds \quad \text{on } T_x^*M \otimes T_x^*M$$

is almost surely nondegenerate.

(ii) In terms of the inner product on T_xM and reading $C_t(x) \in \mathrm{Hom}(T_xM, T_xM)$, we have

$$|C_t(x)^{-1}| \in L^p \quad \text{for each } p \ge 1. \tag{8.5}$$

Then there exists a function $p_t(x,y)$ in $C^\infty(]0,t[\times M \times M)$ such that

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

Remark 8.9. Thus proving Hörmander's parabolic result will come down to showing that under hypothesis (H_1), conditions (i) and (ii) of Theorem 8.8 are satisfied. We will sketch the essential steps of the proof in the remainder of this section.

The idea underlying the probabilistic approach is the following. The measure $P_t(x,dy)$ is the image of the Wiener measure under the mapping $X_t(x)\colon \omega \mapsto X_t(x,\omega)$. Since Wiener measure has a well-understood analytic structure, if this map were “smooth” then regularity properties of $P_t(x, dy)$ could be obtained by integration by parts on Wiener space. The goal of Malliavin calculus is to overcome the difficulty that the map $X_t(x)$, however, is most pathological from the standpoint of classical analysis or standard calculus. See [37, 25, 10] for survey articles along these lines.

Theorem 8.10. *Suppose that* (H_1) *holds true, i.e.,*

$$\dim \operatorname{Lie}\big(A_1, \dots, A_r,\ [A_i, A_j]_{0\le i,j\le r},\ [A_i,[A_j,A_k]]_{0\le i,j,k\le r}, \dots\big)(x) = n \quad \text{for each } x \in M. \tag{H$_1$}$$

Then, for each $x \in M$ *and each* $t > 0$, *almost surely,*

$$C_t(x) = \sum_{i=1}^r \int_0^t (X_{s*}^{-1}A_i)_x \otimes (X_{s*}^{-1}A_i)_x\, ds \in T_xM \otimes T_xM$$

defines a nondegenerate symmetric bilinear form on T_x^*M.

Proof. We fix $x \in M$ and let

$$\begin{aligned}
\mathscr{G}_s &:= \operatorname{span}\Big((X_{s*}^{-1}A_i)_x :\ i = 1, \dots, r\Big) \subset T_xM,\\
\mathscr{U}_t &:= \operatorname{span}\Big(\bigcup_{s\le t} \mathscr{G}_s\Big), \quad t > 0,\\
\mathscr{U}_t^+ &:= \bigcap_{s>t} \mathscr{U}_s.
\end{aligned}$$

Then (by the 0/1-law of Blumenthal) $\mathscr{U}_0^+$ is almost surely a fixed (deterministic) linear subspace of T_xM. We have to show that almost surely

$$\mathscr{U}_0^+ = T_xM.$$

Suppose that $\mathscr{U}_0^+ \subsetneq T_xM$. Then the stopping time

$$\sigma := \inf\big\{t > 0 : \mathscr{U}_t \ne \mathscr{U}_0^+\big\}$$

is almost surely strictly positive. Let $\xi \in T_xM$ such that $\xi \perp \mathscr{U}_0^+$. Then, in particular, $\xi \perp \mathscr{U}_t$ for all $t < \sigma$. In other words, we have for each $i = 1, \dots, r$,

$$\langle \xi, (X_{t*}^{-1}A_i)_x\rangle = 0, \quad \text{for any } t < \sigma.$$

However, for any $V \in \Gamma(TM)$, we know that

$$\begin{aligned} d(X_{s*}^{-1}V)_x &= \Big(X_{s*}^{-1}[A_0, V]\Big)_x\, ds + \sum_{j=1}^{r} \Big(X_{s*}^{-1}[A_j, V]\Big)_x \circ dB_s^j \\ &= \Big(X_{s*}^{-1}[A_0, V]\Big)_x\, ds + \sum_{j=1}^{r} \Big(X_{s*}^{-1}[A_j, V]\Big)_x\, dB_s^j \\ &\quad + \sum_{j=1}^{r} \Big(X_{s*}^{-1}\, \big[A_j, [A_j, V]\big]\Big)_x\, ds. \end{aligned}$$

Thus, taking $V = A_i$, where $1 \le i \le r$, we have for $t < \sigma$,

$$\begin{aligned} \underbrace{\langle \xi, (X_{t*}^{-1} A_i)_x \rangle}_{=0} = \underbrace{\langle \xi, A_i(x) \rangle}_{=0} &+ \int_0^t \langle \xi, (X_{s*}^{-1}[A_0, A_i])_x \rangle\, ds \\ &+ \sum_{j=1}^{r} \int_0^t \langle \xi, (X_{s*}^{-1}[A_j, A_i])_x \rangle\, dB_s^j \\ &+ \sum_{j=1}^{r} \int_0^t \langle \xi, (X_{s*}^{-1}[A_j, [A_j, A_i]])_x \rangle\, ds. \end{aligned} \tag{8.6}$$

By uniqueness of the Doob–Meyer decomposition, canceling the martingale part in Eq. (8.6), we first obtain

$$\langle \xi, (X_{s*}^{-1}[A_j, A_i])_x \rangle = 0, \quad \text{for } 1 \le i, j \le r \text{ and } s < \sigma.$$

By repeating the above calculation with $[A_j, A_i]$ instead of A_i we get

$$\langle \xi, (X_{s*}^{-1}[A_j, [A_j, A_i]])_x \rangle = 0, \quad \text{for } 1 \le i, j \le r \text{ and } s < \sigma.$$

This allows us to cancel the bounded variation part in Eq. (8.6) which gives, in addition,

$$\langle \xi, (X_{s*}^{-1}[A_0, A_i])_x \rangle = 0, \quad \text{for } 1 \le i \le r \text{ and } s < \sigma.$$

By iteration, we see that if $A_{[I]}$ is any of the brackets appearing in (H_1), i.e.,

$$A_{[I]} \in \mathrm{Lie}\big(A_1, \dots, A_r,\ [A_i, A_j]_{0 \le i,j \le r},\ [A_i, [A_j, A_k]]_{0 \le i,j,k \le r}, \dots\big),$$

then

$$\langle \xi, (X_{s*}^{-1} A_{[I]})_x \rangle = 0, \quad s < \sigma.$$

In particular, by taking $s = 0$, we find that

$$\langle \xi, (A_{[I]})_x \rangle = 0.$$

But, since according to (H_1),

$$\mathrm{Lie}\big(A_1, \ldots, A_r,\ [A_i, A_j]_{0\le i,j\le r},\ [A_i,[A_j,A_k]]_{0\le i,j,k\le r}, \ldots\big)(x) = T_xM,$$

we conclude $\xi = 0$. □

In the sequel, we want to sketch the proof that, for given $x \in M$ and $t > 0$,

$$P_t(x, dy) = p_t(x, y)\,\mathrm{vol}(dy), \tag{8.7}$$

where $p_t(x,\cdot) \in C^\infty(M)$. This is the essential part in the stochastic proof of Corollary 8.6(b). To this end, we have to show that $\mu = P_t(x, dy)$ as a distribution is sufficiently smooth. This means that we have to find estimates for the distributional derivatives of μ.

Lemma 8.11. *Let μ be a probability measure on a manifold M ($\dim M = n$) such that*

$$|\langle f, D^{(\alpha)}\mu\rangle| \le C_\alpha \|f\|_\infty, \quad \text{for all } \alpha \in \mathbb{N}^n \text{ and } f \in C_c^\infty(M).$$

Then $\mu(dy) = \rho(y)\mathrm{vol}(dy)$ with $\rho \in L^1(dy) \cap C^\infty(M)$.

Hence to achieve (8.7) for the measure $\mu(dy) = P_t(x, dy)$, we have to show that

$$\Big|\mathbb{E}\big[(D^{(\alpha)} f)(X_t(x))\big]\Big| \le C_\alpha \|f\|_\infty, \quad \forall \alpha \in \mathbb{N}^n.$$

8.2 Girsanov's theorem.

In the sequel we shall use a basic fact from stochastic analysis. This is a special case of Girsanov's theorem [52] which specifies how to remove a drift by change of measure.

Theorem 8.12 (Girsanov). *Let B be a standard Brownian motion on $\mathbb{R}^r$ and let u_t be a continuous adapted process taking values in $\mathbb{R}^r$ as well, such that*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t |u_s|^2 ds\right)\right] < \infty.$$

Consider the Brownian motion with drift $\hat{B}$ defined as

$$d\hat{B}_t := dB_t + u_t\, dt.$$

Then, if B is a Brownian motion on $\mathbb{R}^r$ with respect to $\mathbb{P}$, then $\hat{B}$ is a Brownian motion on $\mathbb{R}^r$ with respect to $\hat{\mathbb{P}}$, where the new probability measure $\hat{\mathbb{P}}$ is given by

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}}\bigg|_{\mathscr{F}_t} = \exp\left(-\int_0^t u_s dB_s - \frac{1}{2}\int_0^t |u_s|^2 ds\right).$$

Hence, defining

$$G_t := \exp\left(-\int_0^t u_s dB_s - \frac{1}{2}\int_0^t |u_s|^2 ds\right),$$

we have

$$d\hat{\mathbb{P}} = G_t\, d\mathbb{P} \quad \text{on } \mathscr{F}_t.$$

In particular, for any measurable functional F on path space, we conclude that

$$\mathbb{E}_{\mathbb{P}}[F(B_\bullet)] = \mathbb{E}_{\hat{\mathbb{P}}}[F(\hat{B}_\bullet)]. \tag{8.8}$$

Eq. (8.8) specifies how a perturbation of a standard Brownian motion by an additive drift can be compensated via a change of measure.

8.3 Elementary stochastic calculus of variations. We fix a point $x \in M$ and consider $u_s = \lambda\, a_s$, where $\lambda \in T_x^*M$ and where a_s is a continuous adapted process taking values in

$$T_x M \otimes (\mathbb{R}^r)^* \equiv T_x M \otimes \mathbb{R}^r$$

such that

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t |\lambda\, a_s|^2\, ds\right)\right] < \infty,$$

for all λ in a small neighborhood U about 0.

In the SDE (8.4) defining the stochastic flow X, we add a drift to the driving Brownian motion B,

$$dB_t^\lambda := dB_t + \lambda a_t\, dt,$$

and compensate this perturbation by changing the measure from $\mathbb{P}$ to $\mathbb{P}^\lambda$,

$$\mathbb{P}^\lambda|\mathscr{F}_t = G_t^\lambda \cdot \mathbb{P}|\mathscr{F}_t,$$

where

$$G_t^\lambda = \exp\left(-\int_0^t \lambda\, a_s dB_s - \frac{1}{2}\int_0^t |\lambda\, a_s|^2 ds\right).$$

We denote by $X_\bullet^\lambda(x)$ the solution to SDE (8.4) when driven by B_t^λ instead of B_t.

By Girsanov's theorem, we may conclude that

$$\mathbb{E}\left[f\left(X_t^\lambda(x)\right) g(B_\bullet^\lambda)\, G_t^\lambda\right] \quad \text{is independent of } \lambda. \tag{8.9}$$

Here f is a smooth function on M and g is a functional of $B_\bullet^\lambda|[0,t]$ such that $g(B_\bullet^\lambda)$ is differentiable in λ. The explicit form of g will be determined later.

We may assume that supp (f) lies in a chart of M; then we write $(D_i f)(x) := (df)_x\, e_i$. Also, since $x \in M$ is fixed, we identify T_xM with $\mathbb{R}^n$.

From Eq. (8.9) we know that

$$\frac{\partial}{\partial \lambda_k}\bigg|_{\lambda=0} \mathbb{E}\left[f\left(X_t^\lambda(x)\right) g(B_\bullet^\lambda)\, G_t^\lambda\right] = 0,$$

which gives

$$\mathbb{E}\left[\sum_i (D_i f)\big(X_t(x)\big) \left(\frac{\partial}{\partial \lambda_k}\bigg|_{\lambda=0} X_t^\lambda(x)\right)^i g(B_\bullet)\right] = -\mathbb{E}\left[f\big(X_t(x)\big) \frac{\partial}{\partial \lambda_k}\bigg|_{\lambda=0} \Big(g(B_\bullet^\lambda)\, G_t^\lambda\Big)\right].$$

We write

$$\left(\frac{\partial}{\partial \lambda_k}\bigg|_{\lambda=0} X_t^\lambda(x)\right)^i = \big(\partial X_t(x)\big)_{ik}.$$

It is easily checked that

$$\partial X_t(x) \equiv \frac{\partial}{\partial \lambda}\bigg|_{\lambda=0} X_t^\lambda(x) = X_{t*} \int_0^t (X_{t*}^{-1} A)_x\, a_s\, ds \in \operatorname{Hom}(T_xM, T_xM).$$

Thus, if we take

$$a_s := (X_{s*}^{-1} A)_x^* \colon\ T_x^*M \to \mathbb{R}^r,$$

then

$$\partial X_t(x) = X_{t*} C_t(x).$$

Finally, taking

$$g(B_\bullet^\lambda) := \Big(C_t^\lambda(x)^{-1}\, (X_{t*}^\lambda)^{-1}\Big)_{kj}\, \gamma(B_\bullet^\lambda),$$

where $\gamma(B_\bullet^\lambda)$ is specified later, and then summing over k, we obtain

$$\mathbb{E}\left[(D_j f)(X_t(x))\, \gamma(B_\bullet)\right] = -\mathbb{E}\Bigg[f\big(X_t(x)\big) \underbrace{\sum_k \frac{\partial}{\partial \lambda_k}\bigg|_{\lambda=0} \Big(C_t^\lambda(x)^{-1} (X_{t*}^\lambda)^{-1}\Big)_{kj}\, \gamma(B_\bullet^\lambda)\, G_t^\lambda}_{=:\, \mathscr{H}_j(\gamma)}\Bigg].$$

By iteration, this shows that

$$\mathbb{E}\Big[(D_i D_j D_k \cdots f)\big(X_t(x)\big)\Big] = -\mathbb{E}\Big[f(X_t(x))\big(\cdots \mathscr{H}_k \mathscr{H}_j \mathscr{H}_i(\mathbb{1}_M)\big)\Big]. \tag{8.10}$$

From Eq. (8.10) we get the crucial estimate

$$\left|\mathbb{E}\Big[(D_iD_jD_k\cdots f)\big(X_t(x)\big)\Big]\right| \le \|f\|_\infty \times \|\cdots \mathscr{H}_k\mathscr{H}_j\mathscr{H}_i(\mathbb{1}_M)\|_1, \tag{8.11}$$

where $\mathbb{1}_M$ denotes the function on M which is identically equal to 1. Hence, to conclude, it remains to show that

$$\|\cdots \mathscr{H}_k\mathscr{H}_j\mathscr{H}_i(\mathbb{1}_M)\|_1 < \infty \tag{8.12}$$

for arbitrary indices $1 \le i, j, k, \ldots \le n$.

The terms appearing in the norm in (8.12) can easily be worked out explicitly by using formulas like

$$\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0} C_t^\lambda(x)^{-1} = -C_t(x)^{-1}\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0} C_t^\lambda C_t^\lambda(x)^{-1}, \tag{8.13}$$

$$\frac{\partial}{\partial\lambda_k}\bigg|_{\lambda=0} G_t^\lambda = -\left(\int_0^t (X_{s*}^{-1}A)_x\, dB_s\right)_k. \tag{8.14}$$

Note that apart from $C_t^{-1}(x)$, only polynomial expressions of quantities appear, which lie in each L^p-space $(1 \le p < \infty)$.

To conclude the proof of Theorem 8.8, the integrability condition (8.5) still needs to be verified. This requires some nontrivial technical estimates; see [39] for a detailed exposition, as well as the simplifications due to [49]. A unified treatment of these issues can be found in [50].

9 Future prospects

Given a sub-Riemannian structure on a differentiable manifold M we have discussed the problem of defining a canonical sub-Laplacian $L = \Delta_{\mathcal{H}}$ on M, either as $L = \mathrm{trace}_{\mathcal{H}}\nabla df$ by choosing a metric partial connection on $\mathcal{H}$, or by endowing M with a smooth volume measure and defining L as the divergence of the horizontal gradient. Such sub-Laplacians have a representation in Hörmander form as

$$L = A_0 + \frac{1}{2}\sum_{i=1}^r A_i^2$$

with vector fields $A_0, A_1, \ldots, A_r \in \Gamma(\mathcal{H}) \subset \Gamma(TM)$. Under the assumption that Hörmander's bracket-generating condition (H_1) is satisfied, the existence of a smooth heat kernel $p_t(x, y)$ in $C^\infty(]0,\infty[\times M \times M)$ is guaranteed,

$$P_t(x, dy) := \mathbb{P}\{X_t(x) \in dy\} = p_t(x, y)\,\mathrm{vol}(dy),$$

and probabilistic methods can be applied to investigate the asymptotics of $p_t(x, y)$ for small and large times. Heat kernel asymptotic expansion is well studied in Riemannian and sub-Riemannian geometry. The classical results of Ben Arous, Léandre, and others [12, 42, 41] include such asymptotic expansion for the cases of diagonal $p_t(x, x)$ and off-diagonal and off cut-locus $p_t(x, y)$; the on cut-locus case $p_t(x, y)$ is understood only up to the leading order [4]. For the application of Malliavin calculus in the study of heat kernel expansions, see [58].

In terms of the Γ-operator

$$\Gamma(f, g) = \frac{1}{2}\left(L(fg) - fLg - gLf\right), \quad f, g \in C^\infty(M), \tag{9.1}$$

the *Carnot–Carathéodory distance* on M is defined as

$$\mathsf{d}_{\mathrm{CC}}(x, y) := \sup\left\{|f(x) - f(y)|\colon\ f \in C_c^\infty(M),\ \Gamma(f, f) \leq 1\right\}. \tag{9.2}$$

Under the *strong Hörmander condition*,

$$\mathrm{Lie}(A_1, \ldots, A_r)(x) = T_x M, \quad x \in M,$$

the Carnot–Carathéodory distance is finite and (9.2) defines a metric structure on M.

As in Riemannian geometry, it is natural to investigate the radial process

$$R_t := \mathsf{d}_{\mathrm{CC}}(x_0, X_t(x)) \tag{9.3}$$

for large times [27]. On a Riemannian manifold, by means of classical Laplacian comparison theorems, the speed of the radial process can be controlled by lower (Ricci) curvature bounds. Defining curvature in sub-Riemannian geometry however is an intriguing problem [1]. Up to now, for instance, no direct probabilistic proof for nonexplosion in finite time of sub-Riemannian diffusion by controlling the radial process (9.3) under sub-Riemannian curvature bounds is known [23].

During recent years, several results have appeared linking sub-Riemannian geometric invariants to properties of diffusions of corresponding second-order operators and their heat semigroup; see [6, 7, 21, 22]. These so-called curvature-dimension inequalities are based on a generalization of the Γ_2-calculus for sub-Riemannian manifolds introduced by Baudoin and Garofalo [8].

Connections between the probabilistic behavior of subelliptic diffusions and analytic properties of the corresponding heat semigroups, most directly expressed in functional inequalities, have attracted a lot of attention [17, 48, 43, 3]. For instance, denoting by $P_t f$ the (minimal) heat semigroup generated by

$$L = A_0 + \frac{1}{2}\sum_{i=1}^r A_i^2,$$

acting on bounded functions $f \in C^\infty(M)$, one seeks to find a constant C such that

$$|\nabla^{\mathrm{hor}} P_t f|^2 \le C P_t |\nabla^{\mathrm{hor}} f|^2 \tag{9.4}$$

holds pointwise for any $t > 0$; see [17, 48]. Note that the squared norm of the horizontal gradient $\nabla^{\mathrm{hor}} f$ is given by

$$|\nabla^{\mathrm{hor}} f|^2 = \sum_{i=1}^{r} (A_i f)^2.$$

Conversely, functional inequalities of the type in (9.4) can be used to deduce nonexplosion of the underlying diffusion [9, 23].

Bibliography

[1] Andrei Agrachev, Davide Barilari, and Luca Rizzi, *Curvature: a variational approach*, To appear in Memoirs American Mathematical Society.

[2] Andrei A. Agrachev and Yuri L. Sachkov, *Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences, vol. 87, Springer, Berlin, 2004, Control Theory and Optimization, II. MR 2062547. Zbl 1062.93001.

[3] Dominique Bakry, Fabrice Baudoin, Michel Bonnefont, and Djalil Chafaï, *On gradient bounds for the heat kernel on the Heisenberg group*, J. Funct. Anal. **255** (2008), no. 8, 1905–1938. MR 2462581. Zbl 1156.58009.

[4] Davide Barilari, Ugo Boscain, and Robert W. Neel, *Small-time heat kernel asymptotics at the sub-Riemannian cut locus*, J. Differential Geom. **92** (2012), no. 3, 373–416. MR 3005058. Zbl 1270.53066.

[5] Fabrice Baudoin, *An introduction to the geometry of stochastic flows*, Imperial College Press, London, 2004. MR 2154760. Zbl 1085.60002.

[6] Fabrice Baudoin and Michel Bonnefont, *Log-Sobolev inequalities for subelliptic operators satisfying a generalized curvature dimension inequality*, J. Funct. Anal. **262** (2012), no. 6, 2646–2676. MR 2885961. Zbl 1254.58004.

[7] Fabrice Baudoin, Michel Bonnefont, and Nicola Garofalo, *A sub-Riemannian curvature-dimension inequality, volume doubling property and the Poincaré inequality*, Math. Ann. **358** (2014), no. 3-4, 833–860. MR 3175142. Zbl 1287.53025.

[8] Fabrice Baudoin and Nicola Garofalo, *Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries*, To appear in Journal of the EMS.

[9] Fabrice Baudoin and Jing Wang, *Curvature dimension inequalities and subelliptic heat kernel gradient bounds on contact manifolds*, Potential Anal. **40** (2014), no. 2, 163–193. MR 3152160. Zbl 1321.53034.

[10] Denis Bell, *The Malliavin calculus and hypoelliptic differential operators*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **18** (2015), no. 1, 1550001, 24. MR 3324717. Zbl 1310.60069.

[11] Denis R. Bell and Salah Eldin A. Mohammed, *An extension of Hörmander's theorem for infinitely degenerate second-order operators*, Duke Math. J. **78** (1995), no. 3, 453–475. MR 1334203. Zbl 0840.60053.

[12] Gérard Ben Arous, *Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus*, Ann. Sci. École Norm. Sup. (4) **21** (1988), no. 3, 307–331. MR 974408. Zbl 0699.35047.

[13] Gérard Ben Arous, Shigeo Kusuoka, and Daniel W. Stroock, *The Poisson kernel for certain degenerate elliptic operators*, J. Funct. Anal. **56** (1984), no. 2, 171–209. MR 738578. Zbl 0556.35036.

[14] Jean-Michel Bismut, *Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions*, Z. Wahrsch. Verw. Gebiete **56** (1981), no. 4, 469–505. MR 621660. Zbl 0445.60049.

[15] Ugo Boscain, Robert Neel, and Luca Rizzi, *Intrinsic random walks and sub-Laplacians in sub-Riemannian geometry*, ArXiv e-prints (2015).

[16] Roger W. Brockett, *Lie algebras and Lie groups in control theory*, Geom. Methods Syst. Theory, Proc. NATO advanced Study Inst., London, 43–82 (1973), 1973. Zbl 0305.93003.

[17] Bruce K. Driver and Tai Melcher, *Hypoelliptic heat kernel inequalities on the Heisenberg group*, J. Funct. Anal. **221** (2005), no. 2, 340–365. MR 2124868. Zbl 1071.22005.

[18] David Elworthy, *Geometric aspects of diffusions on manifolds*, École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87, Lecture Notes in Math., vol. 1362, Springer, Berlin, 1988, 277–425. MR 983375. Zbl 0658.58040.

[19] Michel Emery, *Stochastic calculus in manifolds*, Universitext, Springer, Berlin, 1989. With an appendix by P.-A. Meyer. MR 1030543. Zbl 0697.60060.

[20] Elisha Falbel, Claudio Gorodski, and Michel Rumin, *Holonomy of sub-Riemannian manifolds*, Internat. J. Math. **8** (1997), no. 3, 317–344. MR 1454476. Zbl 0881.53025.

[21] Erlend Grong and Anton Thalmaier, *Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations: Part I*, Math. Z. **282** (2016), no. 1-2, 99–130. Zbl 06538579.

[22] ———, *Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations, Part II*, Math. Z. **282** (2016), no. 1-2, 131–164. Zbl 06538580.

[23] ———, *Stochastic completeness and gradient representations for sub-Riemannian manifolds*, Preprint arXiv:1605.00785.

[24] Wolfgang Hackenbroch and Anton Thalmaier, *Stochastische Analysis*, Mathematische Leitfäden. [Mathematical Textbooks], B. G. Teubner, Stuttgart, 1994, Eine Einführung in die Theorie der stetigen Semimartingale. [An introduction to the theory of continuous semimartingales]. MR 1312827. Zbl 0815.60046.

[25] Martin Hairer, *On Malliavin's proof of Hörmander's theorem*, Bull. Sci. Math. **135** (2011), no. 6-7, 650–666. MR 2838095. Zbl 1242.60085.

[26] Lars Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171. MR 0222474. Zbl 0156.10701.

[27] Elton P. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics, vol. 38, American Mathematical Society, Providence, RI, 2002. MR 1882015. Zbl 0994.58019.

[28] Kanji Ichihara and Hiroshi Kunita, *A classification of the second order degenerate elliptic operators and its probabilistic characterization*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **30** (1974), 235–254. MR 0381007. Zbl 0326.60097.

[29] ______, *Supplements and corrections to the paper: "A classification of the second order degenerate elliptic operators and its probabilistic characterization" (Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **30** *(1974), 235–254)*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **39** (1977), no. 1, 81–84. MR 0488328. Zbl 0382.60069.

[30] Nobuyuki Ikeda and Shinzo Watanabe, *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland, Amsterdam; Kodansha, Tokyo, 1989. MR 1011252. Zbl 0684.60040.

[31] Velimir Jurdjevic, *Geometric control theory*, Cambridge Studies in Advanced Mathematics, vol. 52, Cambridge University Press, Cambridge, 1997. MR 1425878. Zbl 0940.93005.

[32] Shoshichi Kobayashi and Katsumi Nomizu, *Foundations of differential geometry. Vols I* and *II*, Wiley Classics Library, John Wiley & Sons, New York, 1996, Reprint of the 1963/1969 original, A Wiley-Interscience Publication. MR 1393940, MR 1393941.

[33] Hiroshi Kunita, *Supports of diffusion processes and controllability problems*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), Wiley, New York-Chichester-Brisbane, 1978, pp. 163–185. MR 536011. Zbl 0409.60063.

[34] ______, *On the decomposition of solutions of stochastic differential equations*, Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980), Lecture Notes in Math., vol. 851, Springer, Berlin-New York, 1981, pp. 213–255. MR 620992. Zbl 0474.60046.

[35] ______, *Stochastic differential equations and stochastic flows of diffeomorphisms*, École d'été de probabilités de Saint-Flour, XII—1982, Lecture Notes in Math., vol. 1097, Springer, Berlin, 1984, pp. 143–303. MR 876080. Zbl 0554.60066.

[36] ______, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990. MR 1070361. Zbl 0743.60052.

[37] Shigeo Kusuoka, *Malliavin calculus revisited*, J. Math. Sci. Univ. Tokyo **10** (2003), no. 2, 261–277. MR 1987133. Zbl 1031.60048.

[38] Shigeo Kusuoka and Daniel Stroock, *Applications of the Malliavin calculus. I*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 271–306. MR 780762. Zbl 0546.60056.

[39] ______, *Applications of the Malliavin calculus. II*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **32** (1985), no. 1, 1–76. MR 783181. Zbl 0568.60059.

[40] ______, *Applications of the Malliavin calculus. III*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **34** (1987), no. 2, 391–442. MR 914028. Zbl 0633.60078.

[41] Rémi Léandre, *Majoration en temps petit de la densité d'une diffusion dégénérée*, Probab. Theory Related Fields **74** (1987), no. 2, 289–294. MR 871256. Zbl 0587.60073.

[42] ______, *Minoration en temps petit de la densité d'une diffusion dégénérée*, J. Funct. Anal. **74** (1987), no. 2, 399–414. MR 904825. Zbl 0637.58034.

[43] Hong-Quan Li, *Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg*, J. Funct. Anal. **236** (2006), no. 2, 369–394. MR 2240167. Zbl 1106.22009.

[44] Paul Malliavin, *C^k-hypoellipticity with degeneracy*, Stochastic analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978), Academic Press, New York-London, 1978, pp. 199–214. MR 517243. Zbl 0449.58022.

[45] ______, *C^k-hypoellipticity with degeneracy. II*, Stochastic analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978), Academic Press, New York-London, 1978, pp. 327–340. MR 517250. Zbl 0449.58023.

[46] ______, *Stochastic calculus of variation and hypoelliptic operators*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), Wiley, New York-Chichester-Brisbane, 1978, pp. 195–263. MR 536013. Zbl 0411.60060.

[47] ______, *Stochastic analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 313, Springer, Berlin, 1997. MR 1450093. Zbl 0878.60001.

[48] Tai Melcher, *Hypoelliptic heat kernel inequalities on Lie groups*, Stochastic Process. Appl. **118** (2008), no. 3, 368–388. MR 2389050. Zbl 1138.22008.

[49] James Norris, *Simplified Malliavin calculus*, Séminaire de Probabilités, XX, 1984/85, Lecture Notes in Math., vol. 1204, Springer, Berlin, 1986, pp. 101–130. MR 942019. Zbl 0609.60066.

[50] David Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer, Berlin, 2006. MR 2200233. Zbl 1099.60003.

[51] Philip E. Protter, *Stochastic integration and differential equations*, Stochastic Modelling and Applied Probability, vol. 21, Springer, Berlin, 2005, second edition. Version 2.1, corrected third printing. MR 2273672. Zbl 1041.60005.

[52] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third edition, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer, Berlin, 1999. MR 1725357. Zbl 0917.60006.

[53] Ichiro Shigekawa, *On stochastic horizontal lifts*, Z. Wahrsch. Verw. Gebiete **59** (1982), no. 2, 211–221. MR 650613. Zbl 0487.60056.

[54] Daniel W. Stroock and S. R. S. Varadhan, *On degenerate elliptic-parabolic operators of second order and their associated diffusions*, Comm. Pure Appl. Math. **25** (1972), 651–713. MR 0387812. Zbl 0344.35041.

[55] ———, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory, Univ. California Press, Berkeley, Calif., 1972, pp. 333–359. MR 0400425. Zbl 0255.60056.

[56] Héctor J. Sussmann and Velimir Jurdjevic, *Controllability of nonlinear systems*, J. Differential Equations **12** (1972), 95–116. MR 0338882. Zbl 0242.49040.

[57] Philippe Tondeur, *Foliations on Riemannian manifolds*, Universitext, Springer, New York, 1988. MR 934020. Zbl 0643.53024.

[58] Shinzo Watanabe, *Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels*, Ann. Probab. **15** (1987), no. 1, 1–39. MR 877589. Zbl 0633.60077.

[59] David Williams, *Probability with martingales*, Cambridge Mathematical Textbooks, Cambridge University Press, Cambridge, 1991. MR 1155402. Zbl 0722.60001.