Technische Universität Dresden
Fachrichtung Mathematik

Institut für Mathematische Stochastik

The Bismut-Elworthy-Li Formula and Gradient Estimates for Stochastic Differential Equations

Diplomarbeit
zur Erlangung des ersten akademischen Grades

Diplommathematiker

vorgelegt von

Name: Baumgarth Vorname: Robert

geboren am: 03.10.1987 in: Löbau

Tag der Einreichung: 11.11.2015

Betreuer: Prof. Dr. rer. nat. René L. Schilling

Contents

Index of notation

Unless otherwise stated, **functions** are maps whose codomain is $\mathbb{R}^n$ whereas the term **map(ping)** normally refers to a map between arbitrary manifolds. Binary operations between functions such as $f \pm g$, $f \cdot g$, $f \wedge g$, $f \vee g$, comparisons $f \leqslant g$, $f < g$ or limits $f_j \xrightarrow{j\to\infty} f$, $\lim_j f_j$, $\liminf_j f_j$, $\sup_j f_j$ are understood pointwise. “Positive” and “negative” means “$\geqslant 0$” and “$\leqslant 0$”, respectively.

Analysis and measure theory

$\mathbb{N}$ $[\mathbb{N}_0]$	natural numbers [incl. 0]
$\mathbb{R}$	real numbers
$\inf\emptyset$	$\inf\emptyset = +\infty$
$a \wedge b, a \vee b$	minimum and maximum
$f(t-)$	left limit, $\lim_{s\uparrow t} f(s)$
càdlàg	right continuous with left limits, 59
$\lvert x\rvert$	Euclidian norm in $\mathbb{R}^n$ $\lvert x\rvert^2 = x_1^2 + \cdots + x_n^2$
$\lVert\cdot\rVert_\infty$	uniform norm
$\mathbb{1}_A$	$\mathbb{1}_A(x) := \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$
δ_x	Dirac measure at x, $\delta_x(A) = \mathbb{1}_A(x)$
pr_E	projection on E

Sets

$\hat{M}$	$\hat{M} := M \cup \{\infty\}$, 26
E^*	dual space of E
$F \subset_{\text{fin}} E$	F is a finite subset of E
$U \Subset A$	U is an open subset of A

Differential geometry

M	smooth manifold, 4
(U, φ)	(local) chart, 3
TM $[T^*M]$	[co]tangent bundle, 12
T_pM $[T_p^*M]$	[co]tangent space at $p \in M$, 6
$\pi : E \to M$	fibre or vector bundle, 8
$\pi : P \to M$	principcal G-bundle, 38
$\Gamma^r(E)$	C^r-sections of fibre bundle, 9
∂_i	$\left(\frac{\partial}{\partial x_i}\big\vert_p\right)$ basis for T_pM, 7
$\mathrm{d}x^i$	$\left(\mathrm{d}x^i\big\vert_p\right)$ basis for T_p^*M, 14
$g = \langle\cdot,\cdot\rangle_g$	Riemannian metric, 15
$\overline{g} = \langle\cdot,\cdot\rangle_{\mathbb{R}^n}$	Euclidean Riemannian metric, 16
$\sharp, \flat$	musical isomorphisms, 16
∇	covariant derivative, 19
$\nabla_h X$	directional derivative, 24
grad	gradient, 16
div	divergence, 24
tr	trace, contraction, 24
Δ	Laplace operator, $_{i\leqslant n}\partial_i^2$, 24

Δ_M	Laplace-Beltrami operator on manifold M, 24
T	torsion tensor, 22
$\Omega^n(M)$	differential n-forms on M, 14
$\mathscr{O}(M)$	orthonormal frame bundle, 39

Spaces of sets

$\mathscr{B}(E)$	space of Borel-measurable functions $f : E \to \mathbb{R}$
$\mathscr{B}_b(E)$	space of bounded, Borel-measurable functions $f : E \to \mathbb{R}$
$C(E)$	space of continuous functions $f : E \to \mathbb{R}$
$C_b(E)$	space of bounded, continuous functions $f : E \to \mathbb{R}$
$C_c^\infty(E)$	space of smooth functions $f : E \to \mathbb{R}$ with compact support
BC^r	space of C^r-functions with first r derivatives bounded
$\mathscr{S}(E)$	space of all continuous semimartingales on E
$\mathscr{M}(E)$	space of all local martingales on E
$\mathscr{A}(E)$ [$\mathscr{A}_0(E)$]	space of all continuous finite variation processes [starting at zero] on E
$\mathrm{Hom}(E, F)$	space of all homomorphisms from E to F

Probability theory

$\sim$	distributed as
$\overset{\mathrm{m}}{=}$	modulo local martingales, 32
a.s.	almost sure(ly)
a.e.	almost every(where)
iid	independent and identically distributed
$(\Omega, \mathscr{F}, \mathbb{P})$	(underlying) probability space
$\mathbb{P}, \mathbb{E} = \mathbb{E}_{\mathbb{P}}$	probability, expectation with respect to $\mathbb{P}$
$\mathbb{E}(\,\cdot \mid \mathscr{H})$	conditional expectation with respect to a σ-algebra $\mathscr{H}$

Stochastic processes and SDEs

$\mathscr{F}_t^X$	$\sigma(X_s : s \leqslant t)$ natural filtration
$\mathscr{F}_t = \mathscr{F}_{t+}$	$\mathscr{F}_{t+} := \bigcap_{r>t} \mathscr{F}_r$, $t \geqslant 0$, right-continuous filtration
τ	stopping time $\{\tau \leqslant t\} \in \mathscr{F}_t, t \geqslant 0$
X_t^τ	stopped process $X_{t \wedge \tau}$
ζ	lifetime, 26
$[X]$	square bracket $[X, X]$
$(P_t)_{t \geqslant 0}$	semigroup, $P_t f(x) := \mathbb{E} f(X_t(x))$, 53
A	generator of $(P_t)_{t \geqslant 0}$, 53
$B = (B_t)_{t \geqslant 0}$	Brownian motion (BM)
$\mathrm{BM}(E)$	BM on $E \subset \mathbb{R}^n$
$\mathrm{BM}(M, g)$	BM on (M, g), 34
$b(\mathrm{d}X, \mathrm{d}X)$	b-quadratic variation, 30
$L = (L_t)_{t \geqslant 0}$	Lévy process (LP), 59
(a, Σ, ν)	Lévy triplet
$\psi(\xi)$	Lévy exponent, 59
$\nu(\mathrm{d}y)$	Lévy measure, 59
$S = (S_t)_{t \geqslant 0}$	subordinator, 61
SDE	Stochastic differential equation
$\diamond$	Stratonovich circle, 26
BEL	Bismut-Elworthy-Li formula

Chapter 1

Introduction

Let M be a Riemannian manifold and $p_t(x,y)$ the minimal heat kernel on M. Bismut's formula is a probabilistic representation of the covariant derivative of $p_t(x,y)$

$$\nabla \log p_t(x,y) = \frac{1}{t}\mathbb{E}\left(\int_0^t (\mathbf{d}X_s)^* A(X_s) \mathrm{d}B_s \,\middle|\, X_t = y\right). \tag{1.1}$$

Here the righthand side is the expectation of a functional of Brownian motion X on M starting from x and conditioned to return to y at time T, the process B is the stochastic anti-development of X which is a Euclidean Brownian motion by definition. Using the heat semigroup which is also the transition semigroup of Brownian motion on M

$$P_t f(x) = \mathbb{E} f(X_t(x)) = \int_M p_t(x,y) f(y)\, \mathrm{d}y,$$

we get for all reasonable smooth functions f on M

$$\nabla_h P_t f(x) = \nabla_h \mathbb{E} f(X_t(x)) = \frac{1}{t}\mathbb{E}\left(f(X_t(x)) \int_0^t \nabla_h X_s(x)\, \mathrm{d}B_s\right). \tag{1.2}$$

Suppose $P_t f = e^{\Delta_M t/2} f = f$, Δ_M is the Laplace-Beltrami operator on M, then the left-hand side of (1.1) is simply the gradient of the function f. This potentially opens ways of studying gradient estimates of harmonic functions. Moreover, there have recently been developments in financial mathematics, where $\nabla_h \mathbb{E} f(X_t(x))$ is the rate of change (sensitivity) of the price of the derivative with respect to the initial prices of the underlying assets, cf. [Hsu07].

The aim of this work is to give a consistent introduction to Brownian motion on manifolds and typical problems arising. This leads to the classical Bismut-Elworthy-Li formula (BEL), cf. Theorem 4.6, for a large class of diffusion processes. Finally, we show a generalisation of the BEL for a nonlinear stochastic differential equation (SDE) driven by a subordinated Brownian motion.

Since this text is primarily written for probabilists with a reasonable background in basic Euclidean stochastic analysis, we introduce elementary results and definitions from differential geometry and agree upon notation in Chapter 2.

Chapter 3 introduces two constructions of Brownian motion on manifolds. To this end, we give a brief introduction to SDEs driven by a continuous semimartingale and to martingales on manifolds. This leads to two different concepts to define Brownian motion on M. The extrinsic approach qua a martingale problem and the intrinsic approach as a projection using the orthonormal frame bundle which is known as the Eells-Elworthy-Malliavin construction. Therefore additionally required differential geometry is reviewed on the spot.

Chapter 4 is concerned with the classical Bismut-Elworthy-Li formula. Elworthy and Li generalised in [EL94b] Bismut's formula for a much larger class of diffusion processes using simple martingale arguments. We give a proof in the case $M = \mathbb{R}^n$ and sketch the proof for an arbitrary compact Riemannian manifold as it depends heavily on the theory of stochastic flows.

Chapter 5 provides a brief overview of Lévy processes and subordination. In particular, we will consider subordinated Brownian motion as a special class of Lévy processes.

In the final Chapter 6, we show that the Bismut-Elworthy-Li formula also holds for nonlinear SDEs driven by a subordinated Brownian motion.

We emphasize that we do not attempt to state all results in their most general form since a complete discussion would go beyond the scope of a diploma thesis. However, we try to motivate abstract definitions and point out notable sources to deepen the understanding.

Lastly, I have the pleasant task of thanking the people who helped creating this thesis. My advisor René L. Schilling for his guidance and kindness in answering my questions. Franziska Kühn and Sebastian Blank for sacrificing some of their precious time to proof-read this thesis and pointing out quite a few misprints and minor errors. Without the patience and support of Martin this work would not have been possible.

Dresden, November 2015 Robert Baumgarth

Chapter 2

Elements of Differential Geometry

In this chapter we collect necessary preliminaries and agree upon the notation used, so most results will be presented without proof. The simplest manifolds are the *topological manifolds*, i.e. topological spaces which behave "locally like" the Euclidean space $\mathbb{R}^n$. But there is no reasonable way to define smoothness as a purely topological property, because it is not invariant under homeomorphisms. For example, a curve is called "smooth" if it has a tangent line that varies continuously from point to point. For a surface we would expect the same by replacing the tangent line trough a tangent surface (plane) etc. Therefore, as a well-known example, a circle, as a "smooth" curve, can be transformed homeomorphic to a square which is obviously not "smooth" in our sense. Hence, the key idea is to build a *smooth structure* on a topological manifold to get derivatives of reasonable meaning. The next step (especially for integration) is the concept of differential forms and smooth manifolds with boundary. Finally, we introduce a *Riemannian manifold*, that is, a smooth manifold equipped with an inner product. Solely by the so called *Riemannian metric*, we can describe fundamental geometric properties of the manifold like the curvature or parallel transport. The *parallel transport (or displacement)* will turn out to be the key ingredient to define Brownian motion on a manifold. In this work we restrict ourselves to smooth manifolds *without* boundary.

A profound introduction to smooth manifolds can be found in [Lee13] or [Jän01], fundamental theorems of Riemannian geometry and curvature are summarised in e.g. [Lee05]. A standard reference for the theory of fibre bundle and connections, in particular horizontal bundle and horizontal lifts, is [KN63]. Having a profound background of these topics (and being a German reader) you should take a look at [HT94] which provides a systematic treatment of these topics designed for stochastic analysis on manifolds.

2.1 Smooth Manifolds and Smooth Maps

Let M be a **topological manifold** of dimension n (**topological** n**-manifold**), that is, a topological Hausdorff space with second-countable basis for open sets, which is **locally Euclidean of dimension** n, that is, each point of M has a neighbourhood that is homeomorphic to an open subset of $\mathbb{R}^n$.[1]

2.1 Definition. Let $r \in \mathbb{N}_0 \cup \{\infty\}$ and I an index set. An n**-dimensional** C^r**-atlas** is a collection of pairs $(U_\alpha, \varphi_\alpha)_{\alpha \in I}$ on M such that for all $\alpha \in I$:

[1]The third property means, more specifically, that for each $p \in M$ we find a subset $U \Subset M$ containing p, a subset $\tilde{U} \Subset \mathbb{R}^n$, such that $\varphi : U \to \tilde{U}$ is a homeomorphism.

(i) $U_\alpha \Subset M$, that is, U_α is an open subset of M, and $\bigcup_{\alpha\in I} U_\alpha \supset M$, that is, U_α **cover** M.

(ii) For each $\alpha \in I$, the **(coordinate) chart** $\varphi_\alpha : U_\alpha \to \mathcal{O}_\alpha$ is a homeomorphism of the **chart domain (or coordinate neighbourhood)** U_α onto an $\mathcal{O}_\alpha \Subset \mathbb{R}^n$.

(iii) For all $\alpha, \beta \in I$ with $U_\alpha \cap U_\beta \neq \emptyset$, $\varphi_\alpha(U_\alpha \cap U_\beta)$ is open in $\mathbb{R}^n$ and the **transition maps**

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta) \in C^r.$$

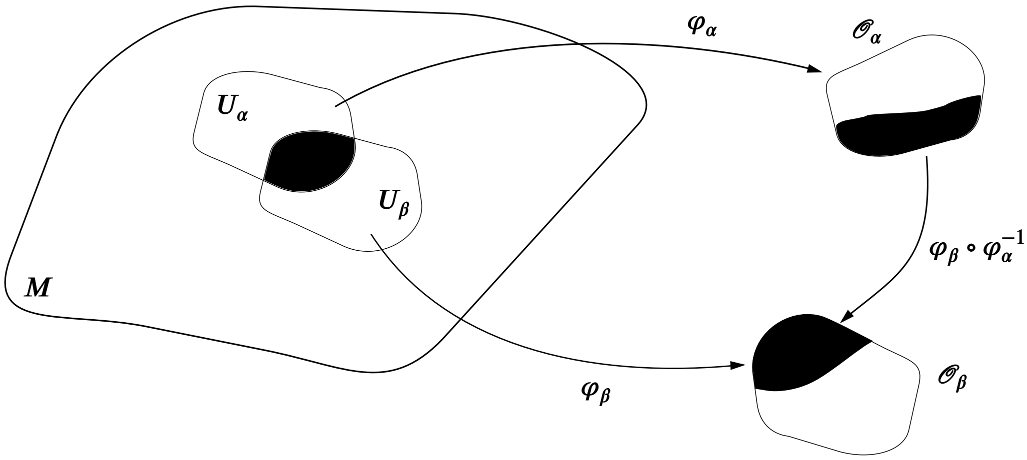

Figure 2.1: Charts on manifolds

To be more precise, we define an atlas for M to be a collection of charts whose domains cover M. Let $(U_\alpha, \varphi_\alpha)_{\alpha\in I}$ be an n-dimensional C^r-atlas. We always consider M as a topological space equipped with the topology

$$\left\{U \subset M \mid \varphi_\alpha(U_\alpha \cap U) \Subset \mathbb{R}^n \text{ for all } \alpha \in I\right\}.$$

In particular, the φ_α are homeomorphisms with respect to this topology for all $\alpha \in I$.

2.2 Definition. (a) Two n-dimensional C^r-atlases are **equivalent** if their union is again a C^r-atlas. An n-dimensional **differentiable structure** D on M is a maximal n-dimensional differentiable C^∞-atlas.

(b) An n-dimensional **smooth manifold** (n-**manifold**) is a pair $(M, \mathbf{D})$ consisting of topological manifold M and an n-dimensional differentiable structure **D** on M. Since we deal exclusively with smooth manifolds, we usually suppress the adjective *smooth*.

Every subset $U \Subset \mathbb{R}^n$ admits an n-manifold if we choose the only chart $\mathrm{id} : U \to U$, called **standard structure**. Unless we explicitly specify otherwise, we always use this smooth structure on $\mathbb{R}^n$.

Remember that, throughout this thesis we make a slight distinction between the terms *function* and *map*. We generally reserve the term **function** for a map whose codomain is $\mathbb{R}^n$ whereas the term **map(ping)** normally refers to a map between arbitrary manifolds.

2.3 Definition. A function $f : M \to \mathbb{R}^k$, $k \in \mathbb{N}$, is **smooth** at $p \in M$ if for some (hence every) chart (U, φ) around p, the function $f \circ \varphi^{-1}$ is smooth in a neighbourhood of $\varphi(p)$. The family of all smooth functions on M is denoted by $C^\infty(M)$.

Thus, if we want to study the local behaviour of a function near a point $p \in M$, the key idea is to choose a chart around p and "pull the map f down" from the manifold to an open subset where we can use the usual differentiability conditions.

2.4 Definition. Let M and N be manifolds. A map $F : M \to N$ is said to be **smooth**, denoted $F \in C^\infty$, if for all charts $\varphi : U \to \mathcal{O}$, $\psi : \tilde{U} \to \tilde{\mathcal{O}}$ ($\mathcal{O}, \tilde{\mathcal{O}} \Subset \mathbb{R}^n$) with $F(U) \subset \tilde{\mathcal{O}}$, the composition function $\psi \circ F \circ \varphi^{-1}$ is smooth in the usual sense.

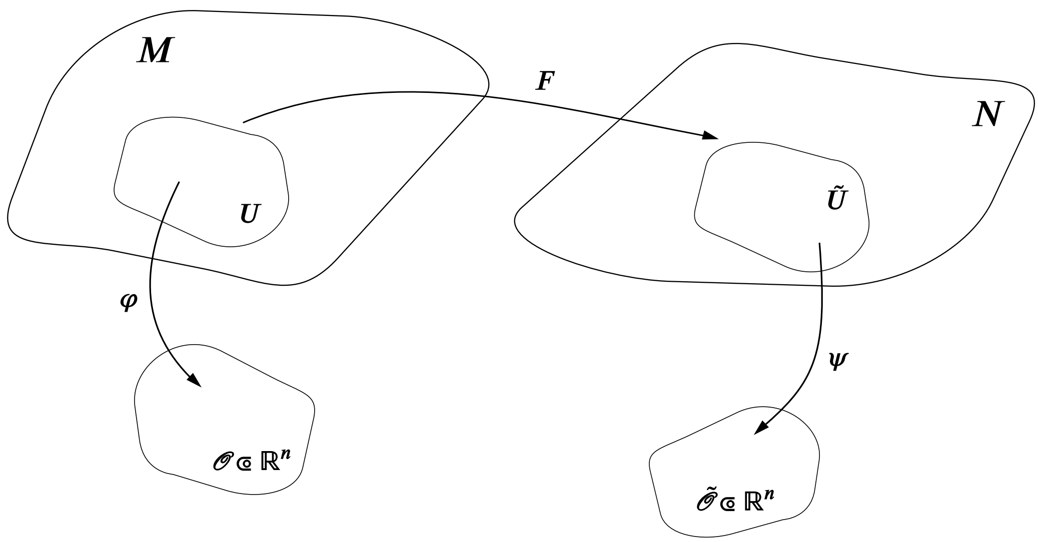

Figure 2.2: Smooth map between manifolds.

2.5 Definition. A **diffeomorphism** is a smooth bijection between manifolds such that its inverse map is also smooth.

2.6 Remark (Einstein summation convention). From now on, we use the handsome **Einstein summation convention**. This is an important notational convention that is commonly used in manifold theory, since we often have to deal with vectors and covectors and the inevitably excrescence of summation signs: If an index appears twice, once as a subscript and once as a superscript, we skip the summation symbol, e.g. we write

$$v^i \partial_i, \quad \omega_i \mathbf{d}x^i, \quad \frac{\partial x^i}{\partial x^{\bar{i}}} \frac{\partial x^j}{\partial x^{\bar{j}}} F_{ij},$$

instead of

$$\sum_{i=1}^{n} v^i \partial_i, \quad \sum_{i=1}^{n} \omega_i \mathbf{d}x^i, \quad \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial x^i}{\partial x^{\bar{i}}}\frac{\partial x^j}{\partial x^{\bar{j}}} F_{ij}.$$

Therefore, we will make the distinction and write (contravariant) vectors $e_1, ..., e_n \in V$ always with a subscript and the corresponding dual frame's covectors $\varepsilon^1, ..., \varepsilon^n \in V^*$ with a superscript.

2.2 The Tangent Space and Differential on Manifolds

One of the central ideas of calculus is that the derivative represents the "best linear approximation" of a map near a given point: For a function $f : \mathbb{R}^n \to \mathbb{R}^k$, the map $\mathbf{d}f_x : \mathbb{R}^n \to \mathbb{R}^k$ is characterised by $f(x+a) = f(x) + \mathbf{d}f_x \cdot a + \psi(a)$, where $\psi(a)/\|a\| \to 0$ as $a \to 0$, and given by the Jacobian matrix. Thus, we could consider the downstairs map $\mathbf{d}(\varphi \circ f \circ \varphi^{-1})$, but this definition obviously depends on the choice of charts. Thinking of the classical derivative of a function of one variable, we would think of a tangent line, in two variables of surface or tangent plane. Therefore, the idea is to approximate the manifold locally at p and $f(p)$ first, by the so called tangent space T_pM to a manifold at a point $p \in M$, and then define the differential as a linear map between them.

2.7 Definition. (a) Two real-valued functions, each defined and smooth in some neighbourhood of a point $p \in M$, are **equivalent** if they agree in a neighbourhood of p. The equivalence classes are called the **germs of smooth functions on M at p** and the set of these germs is denoted by $\mathscr{G}_p(M)$.

(b) Let M be a manifold, $p \in M$. A **tangent vector** to M at p is a **derivation** of the ring of germs $\mathscr{G}_p(M)$, that is, an $\mathbb{R}$-linear map

$$v : \mathscr{G}_p \to \mathbb{R}$$

that satisfies the product rule for all $f, g \in \mathscr{G}_p(M)$

$$v(f \cdot g) = v(f) \cdot g(p) + f(p) \cdot v(g). \tag{2.1}$$

We call the vector space of these derivations the **tangent space** to M at p, denoted by $T_p(M)$.

2.8 Remark. For any point $p \in M$ a **tangent vector** at p can be considered as an equivalence class of C^∞ curves that pass through 0: For all $\varepsilon > 0$

$$\gamma : (-\varepsilon, \varepsilon) \to M, \quad \gamma(0) = p,$$

such that under the equivalence relation $\gamma \sim \tilde{\gamma}$ for some (hence every) chart (U, φ) around p we have $(\varphi \circ \gamma)^\bullet(0) = (\varphi \circ \tilde{\gamma})^\bullet(0)$. The set of such (equivalence classes of) tangent vectors is called **(geometrically defined) tangent space** and also be denoted by $T_p(M)$. Since both spaces can be identified naturally, we either consider $v \in T_pM$ as a derivation or as a geometrically defined tangent vector, cf. [Lee13, Chapter 3].

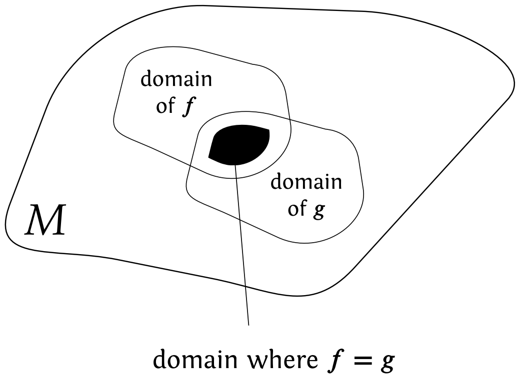

Figure 2.3: For $f \sim g$, it is not necessary that they agree on their common domain, only on a small neighbourhood of p.

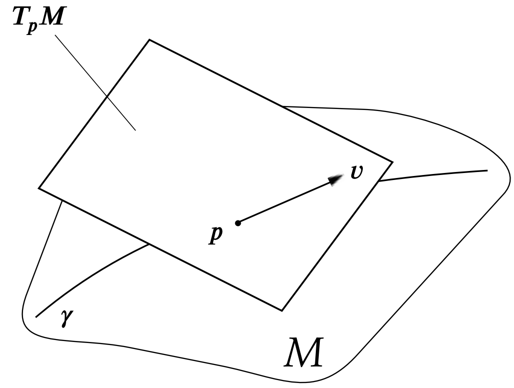

Figure 2.4: For every tangent vector $v \in T_pM$ we find a curve γ in M with $\gamma(0) = p$ and $\dot\gamma(0) = v$.

2.9 Definition. Let M and N be smooth manifolds and $F : M \to N$ a smooth map. For each $p \in M$, we define the **differential of F at p** as the linear map

$$\begin{aligned} \mathrm{d}F_p : T_pM &\to T_{F(p)}N \\ \mathrm{d}F_p(v)(f) &:= v(f \circ F), \end{aligned}$$

for $v \in T_pM$, $f \in C^\infty(M)$.

Because the differential "pushes" tangent vectors forward from the domain manifold to the codomain, it is also called **push-forward**, denoted $(F_*)_p := \mathrm{d}F_p$. We will use both notations interchangeably.

Let $F : M \to N$, $G : N \to L$ be smooth maps between manifolds. Then we have the chain rule $\mathrm{d}(G \circ F)_p = \mathrm{d}G_{F(p)} \circ \mathrm{d}F_p$.

2.10 Remark (Notation). Let $\left(u^1, ..., u^n\right)$ the standard coordinates in $\mathbb{R}^n$. Let $p \in U$, where (U, φ) is a chart with local coordinates $x^1, ..., x^n$, i.e. $\varphi = \left(x^1, ..., x^n\right)$. Then

$$\mathrm{d}\varphi_p : T_pM \xrightarrow{\sim} T_{\varphi(p)}\mathbb{R}^n \cong \mathbb{R}^n$$

provides an isomorphism between vector spaces over $\mathbb{R}$ and hence $\dim_{\mathbb{R}} T_pM = n$. Then the ith vector of the basis for T_pM, given in coordinates $\left(x^i\right)$, will be denoted by

$$\partial_i\big|_p := \frac{\partial}{\partial x^i}\bigg|_p \in T_pM,$$

where $\partial_i\big|_p := (\mathrm{d}\varphi_p)^{-1}(e_i) = \mathrm{d}(\varphi^{-1})_{\varphi(x)}(e_i)$ for $i = 1, ..., n$. We will sometimes write $\partial_i\big|_p$ instead of $\partial_i(p)$ to clarify notation, e.g. if ∂_i already has a subscript like above. Therefore, always mind that a local coordinate constitutes a *map* from U to $\mathbb{R}^n$ by $x^i(p) := u^i(\varphi(p))$. So, for any smooth function f, $\partial_i\big|_p$ is a derivation

$$f \mapsto \frac{\partial f}{\partial x^i}\bigg|_p := \frac{\partial\left(f \circ \varphi^{-1}\right)}{\partial u^i}\bigg|_{\varphi(p)}.$$

2.11 Definition. A smooth map $F : M \to N$ between manifolds is called an **immersion** if $\mathrm{d}F_p : T_pM \to T_{F(p)}N$ is injective for every $p \in M$. The map F is an **embedding** if F is an immersion and a homeomorphism onto its image $F(M) \subset N$ (in the subspace topology).

2.12 Definition. An **(embedded or regular) submanifold of** M is a subset $M_0 \subset M$ that is a manifold in the subspace topology, endowed with a smooth structure with respect to which the inclusion map $\iota : M_0 \hookrightarrow M$ is a smooth embedding.

2.13 Remark. This definition is equivalent to say that a subset $M_0 \subset M$ is a k-dimensional submanifold of M, if for every $p \in M_0$ there is a chart (U, φ) around $p \in M$ such that

$$\varphi(U \cap M_0) = \varphi(U) \cap (\mathbb{R}^k \times \{0\}) = \left\{p \in \varphi(U) : x^{k+1} = \dots = x^n = 0\right\}.$$

We will call $\operatorname{codim}(M_0) := n - k$ the **codimension of** M_0.

2.14 Example. (i) The sphere S^n is a smooth submanifold of $\mathbb{R}^{n+1}$.

(ii) For any smooth map $F : M \to N$ with $\dim M = m$, the graph

$$\operatorname{gr}(F) := \{(x, y) \in M \times N : y = F(x)\}$$

is a smooth m-dimensional submanifold of $M \times N$.

2.3 Fibre bundle and Vector Bundle

A fibre bundle is a space that is locally a product space, but globally may have a different topological structure. We restrict ourselves to fibre bundles between manifolds so that the total space E is manifold instead of a topological space. Thus, the trivialisations become diffeomorphisms (not homeomorphisms).

2.15 Definition. Let E, M and F be manifolds. A smooth map $\pi : E \to M$ or, more precisely, the tuple (E, M, π, F) is called **fibre bundle with (typical) fibre** F, if for every point in $p \in M$ there exists an open neighbourhood U of p in M and a diffeomorphism

$\varphi : \pi^{-1}(U) \xrightarrow{\sim} U \times F$ on U such that the diagram

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\varphi} & U \times \mathsf{F} \\ {\scriptstyle \pi}\downarrow & \swarrow {\scriptstyle \mathrm{pr}_U} & \\ U & & \end{array}$$

commutes. The pair (U, φ) is called **bundle chart** or **local trivialisation** of the fibre. A collection $(U_\alpha, \pi_\alpha)_{\alpha \in I}$ of bundle charts with $M = \bigcup_{\alpha \in I} U_\alpha$ is a **bundle atlas for** E. Morerover, M is called **base space**, E **total space**, π the **projection**, F **(typical) fibre** and $E_p := \pi^{-1}\{p\}$ **fibre above** $p \in M$. We also use the common notations $E \to M$ or E/M.

If $E = M \times \mathsf{F}$ and $\pi = \mathrm{pr}_M$, then E is a fibre bundle (of F) over M, hence it only has a global bundle chart. Any such fibre bundle is called a **trivial bundle**.

Next, we introduce a special class of fibre bundles, called vector bundle, whose fibres are vector spaces. Important examples include the tangent bundle TM and the cotangent bundle T^*M of a manifold (cf. Section 2.5 below).

2.16 Definition. A **rank** m **real smooth vector bundle with fibre** V is a tuple (E, M, π, V), such that (E, B, π, V) is a smooth fibre bundle such that the fibre V is a real vector space of dimension m and the following conditions hold:

(i) For every $p \in M$, the fibre $\pi^{-1}(p)$ is an m-dimensional (real) vector space.

(ii) For every trivialisation, $\varphi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times V$, for every $p \in U_\alpha$, the restriction of φ_α to the fibre $\pi^{-1}(p)$ is a linear isomorphism, $\pi^{-1}(p) \to V$.

2.17 Definition. Let $\pi : E \to M$ be an n-dimensional vector bundle and $k \leqslant n$. A subset $E_0 \subset E$ is called k-dimensional **subbundle of** E if for every point $p \in M$ there exists on each fibre a linear bundle chart $\varphi : E/U \to U \times \mathbb{R}^n$ for E with $\varphi(E_0/U) = U \times (\mathbb{R}^k \times \{0\})$. Then $\pi\big|_{E_0} : E_0 \to M$ itself is a k-dimensional vector bundle over M.

2.18 Definition. Let E, E' be vector bundles over M. A smooth map $\varphi : E \to E'$ is called **vector bundle homomorphism** if it is a linear map on each fibre, i.e. $\pi' \circ \varphi = \pi$,

$$\begin{array}{ccc} E & \xrightarrow{\quad\varphi\quad} & E' \\ & {\scriptstyle \pi}\searrow \quad \swarrow {\scriptstyle \pi'} & \\ & M & \end{array}$$

2.19 Definition. Let $\pi : E \to M$ be a vector bundle. A continuous map $\sigma : M \to E$ is a **section of** E if it is a continuous right-inverse for π, i.e. a continuous map $\sigma : M \to E$ such that $\pi \circ \sigma = \mathrm{id}_M$.[2] The family of all k-times differentiable sections of M is denoted by $\Gamma^k(E)$ and $\Gamma := \Gamma^0$ by convention.

[2] Revised version: corrected misprint.

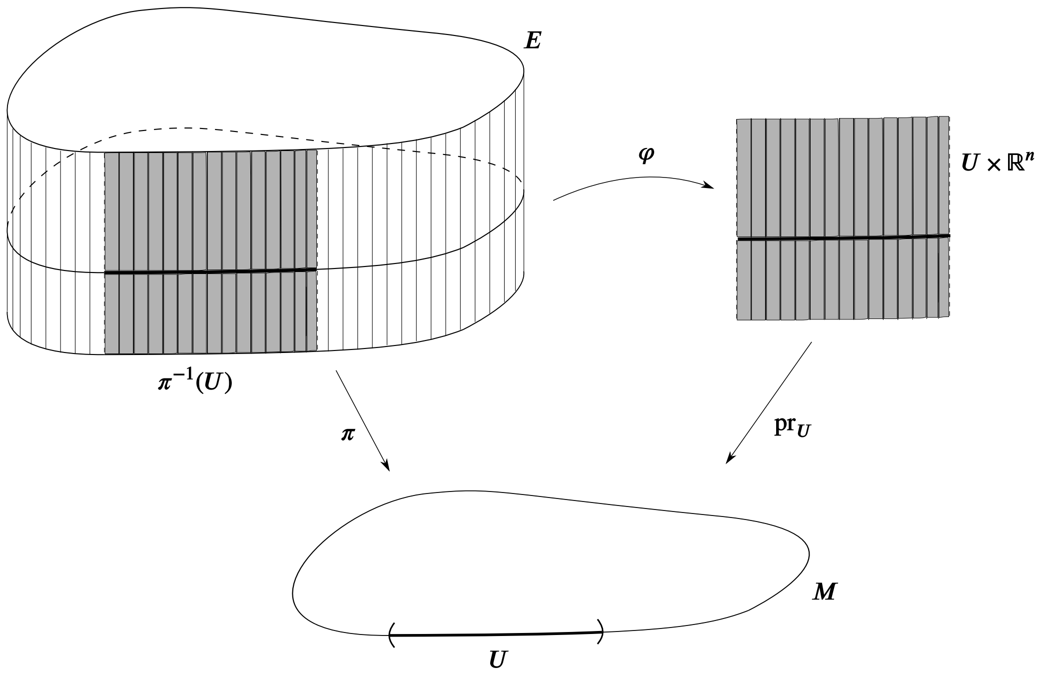

Figure 2.5: A local trivialisation of a vector bundle

2.20 Definition. Let $\pi : E \to M$ be an n-dimensional vector bundle and $p_0 \in M$. A **local frame for E at p_0** is a collection of sections $E_1, \ldots, E_n \in \Gamma(E/U)$ with an open neighbourhood U of p_0 such that for every $p \in U$ the ordered tuple $(E_1\big|_p, \ldots, E_n\big|_p)$ forms a real basis of E_p. For any $A \in \Gamma(E)$ there exist $a^i \in C^\infty(U)$ such that $A\big|_U = a^i E_i$.

2.21 Definition. Let $F : M \to N$ be a smooth map between manifolds and $\pi : E \to N$ a locally trivial fibre bundle. Then

$$F^*E := \bigsqcup_{p \in M} E_{F(p)} \to M$$

with the canoncial projection is also a locally trivial fibre bundle. The bundle charts (U, φ) for E provide "induced" pre-bundle charts $(F^{-1}(U), F^*\varphi)$ for

$$F^*E \equiv \{(p, e) \in M \times E : F(p) = \pi(e)\}$$

on each fibre, namely

$$F^*\varphi : F^*E\big/F^{-1}(U) \to F^{-1}(U) \times \mathsf{F},$$
$$F^*\varphi\big|_{(F^*E)_p} \equiv \varphi\big|_{E_{F(p)}} \text{ for } p \in F^{-1}(U).$$

This locally trivial fibre bundle with basis M is called the **pullback fibre by F** or **induced fibre of E by F**.

2.22 Example. Let $F : M \to N$ be a smooth map between manifolds and $\pi : E \to N$ a vector bundle. Then $F^*E \to M$ is also a vector bundle, the **pullback of E**

by F. For a bundle homomorphism $\varphi : E \to E'$ over N there is a bundle homomorphism $F^*\varphi : F^*E \to F^*E'$ over M defined on each fibre, namely $F^*\varphi|_{(F^*E)_p} = \varphi|_{E_{F(p)}}$.

2.23 Definition. Let $F : M \to N$ be a smooth map between manifolds, E a vector bundle over N. Any element of the $C^\infty(M)$ module

$$\Gamma(F^*E) = \{A : M \to E \mid A \text{ is smooth with } \pi \circ A = F\}$$

is called a **section along** F. In particular, the elements of the $C^\infty(M)$-module $\Gamma(F^*TN)$ are **vector fields along** F. For any interval $I \subset \mathbb{R}$ and smooth curve $\gamma : I \to M$, we have

$$\Gamma(\gamma^*E) \equiv \left\{\sigma : I \to E \mid \sigma \text{ is smooth with } \sigma(t) \in E_{\gamma(t)} \text{ for every } t \in I\right\}.$$

In particular, the vector field $\dot\gamma \in \Gamma(\gamma^*TN), \dot\gamma_t := \dot\gamma(t)$ is called **tangential vector field along** γ.

2.4 Tensors and Tensor Fields

Tensors provide a unified language to talk about multilinear maps. Thus, they generalise the concepts of vectors and covectors, i.e. elements of the dual space of a given vector space. Simple examples include covectors, inner products and determinants. Since manifold theory tries to interpret linear approximations of calculus in a coordinate independent way, tensors are a key ingredient to establish the theory of differential forms.

Let V be real-valued finite-dimensional vector space with basis $e_1, ..., e_n$ and denote by V^* the corresponding dual space, i.e. the vector space of all linear functionals on V, with basis $\varepsilon^1, ..., \varepsilon^n$. Recall that a **tensor of type** (r, s), i.e. **covariant** r**-tensor on** V and **contravariant** s**-tensor on** V, is defined to be an element of the vector space

$$T^{(r,s)}(V) := (V^*)^{\otimes r} \otimes V^{\otimes s},$$

where $\otimes$ denotes the abstract tensor product. Usually, one can think of a multilinear map

$$\underbrace{V^* \times \cdots \times V^*}_{r \text{ copies}} \times \underbrace{V \times \cdots \times V}_{s \text{ copies}} \to \mathbb{R}.$$

In particular, a 0-tensor is a real number and a covariant 1-tensors is a covector.

The **bundle of mixed tensors of type** (r, s) is defined by

$$T^{(r,s)}TM := \bigsqcup_{p \in M} (T_p^*M)^{\otimes r} \otimes (T_pM)^{\otimes s}.$$

A **tensor field** is a smooth section of the tensor bundle.

This means in particular, any contravariant 1-tensor field is a vector field and any covariant 1-tensor field is covector field. Since a 0-tensor is a real number, a 0-tensor field can be interpreted as a continuous real-valued function.

In smooth local coordinates $\left(x^i\right)$ any (r, s)-tensor can be written locally in the form

$$T = T^{i_1,...,i_s}_{j_1,...,j_r}\, \partial_{i_1} \otimes \ldots \otimes \partial_{i_s} \otimes \mathbf{d}x^{j_1} \otimes \ldots \otimes \mathbf{d}x^{j_r},$$

for $i_1, ..., i_s, j_1, ..., j_r = 1, ..., n$.

2.5 Vector Fields and Differential Forms

We define the set of all tangent vectors at all points of a manifold, the so called **tangent bundle**, denoted TM, to be the disjoint union of the tangent spaces at all points of M

$$TM = \bigsqcup_{p\in M} T_pM.$$

Similarly, we define the **cotangent bundle** by $T^*M = \bigsqcup_{p\in M} T_p^*M$.

Let $F : M \to N$ be a smooth map between manifolds. The differential induces a vector bundle homomorphism $\mathbf{d}F : TM \to F^*TN$ over M from the tangent bundle TM to the tangent bundle TN, also denoted $\mathbf{d}F$ or F_*, by the following commutative diagram

$$\begin{array}{ccc} TM & \xrightarrow{\mathbf{d}F} & TN \\ {\scriptstyle \mathrm{pr}_M}\downarrow & & \downarrow{\scriptstyle \mathrm{pr}_N} \\ M & \xrightarrow[F]{} & N \end{array}$$

and thus defined pointwise on each fibre by the differential $\mathbf{d}F_p : T_pM \to T_{F(p)}N$.

2.24 Definition. A **vector field on** M is a section of the tangent bundle $\pi : TM \to M$. More precisely, that is a continuous map $X : M \to TM$, denoted $p \mapsto X_p$, with

$$\pi \circ X = \mathrm{id}_M \iff X_p \in T_pM \quad (p \in M).$$

Thus, $\Gamma(TM)$ denotes the set of all vector fields on M, $\Gamma^\infty(TM)$ the set of all smooth vector fields on M, i.e. Xf is smooth for every $f \in C^\infty(M)$. Multiplication with real-valued functions is defined pointwise: for $f \in C^\infty(M)$, $X \in \Gamma^\infty(TM)$, define $fX : TM \to M$ by $(fX)_p := f(p)X_p$. Note that every vector field $M \to TM$ can be identified with a derivation as a linear map $C^\infty(M) \to C^\infty(M)$.

2.25 Definition. Let V be a real vector space. An **alternating n-form** ω on V is a multilinear map $\omega : V^n \to \mathbb{R}$ with the additional property that interchanging two of the variables switches the sign: For $i < j$

$$\omega(v_1, ..., v_i, v_j, ..., v_n) = -\omega(v_1, ..., v_j, v_i, ..., v_n),$$ [3]

or equivalently, for any permutation $\rho : \{1, ..., n\} \to \{1, ..., n\}$,

$$\omega(v_1, ..., v_n) = \mathrm{sgn}(\rho)\omega(v_{\rho(1)}, ..., v_{\rho(n)}).$$ [4]

The vector space of alternating n-forms on V will be denoted by Λ^nV with the convention $\Lambda^0V := \mathbb{R}$. Thus, the alternating 0-forms are the real numbers and $\Lambda^1V = V^*$ the dual space of V.

[3] Revised version: Slight reformulation for better readability.
[4] Revised version: Slight reformulation for better readability.

Next, we introduce differential forms, i.e. we explore the theory of alternating tensor fields on a manifold which are the key ingredient to define integration on manifolds and generalise classical vector calculus operations like gradient, divergence and curl.

The subspace of all alternating covariant n-tensors on V is denoted by $\Lambda^n(V^*) \subset T^{(n,0)}(V)$. Define a projection, called **alternation**,

$$\begin{aligned} \mathrm{Alt} : T^{(n,0)}(V) &\to \Lambda^n(V^*) \\ (\mathrm{Alt}\,\alpha)(v_1, ..., v_n) &\mapsto \frac{1}{n!} \sum_{\rho \in S_n} (\operatorname{sgn} \rho)\ \alpha\left(v_{\rho(1)}, ..., v_{\rho(n)}\right), \end{aligned}$$

where S_n denotes the symmetric group of n elements. Next, we define the **elementary alternating tensor** $\varepsilon^I \in \Lambda^n(V^*)$ by

$$\varepsilon^I(v_1, ..., v_n) = \begin{pmatrix} \varepsilon^{i_1}(v_1) & \dots & \varepsilon^{i_1}(v_n) \\ \vdots & \ddots & \vdots \\ \varepsilon^{i_n}(v_1) & \dots & \varepsilon^{i_n}(v_n) \end{pmatrix} = \det \begin{pmatrix} v_1^{i_1} & \dots & v_n^{i_1} \\ \vdots & \ddots & \vdots \\ v_n^{i_1} & \dots & v_n^{i_n} \end{pmatrix}.$$

Now, let $\dim V = n$ and $\left(\varepsilon^I\right)$ be a basis for V^*. Then, it is easy to see that a basis for $\Lambda^n(V^*)$ is given by

$$\mathscr{E} = \left\{\varepsilon^I : I \in \mathbb{N}^n \text{ an increasing multiindex}\right\},$$

and therefore $\dim \Lambda^n(V^*) = \binom{n}{k}$. In particular, if $k > n$, then $\dim \Lambda^n(V^*) = 0$ and $\dim \Lambda^n(V^*) = 1$.

2.26 Definition. For any $\omega \in \Lambda^n(V^*)$, $\eta \in \Lambda^m(V^*)$, we define the **wedge product** or **exterior product** to be the $(n+m)$-covector[5]

$$\omega \wedge \eta := \frac{(n+m)!}{n!\, m!} \mathrm{Alt}(\omega \otimes \eta).$$

The wedge product is bilinear, associative and anticommutative. Moreover, for any multiindex $I = (i_1, ..., i_n) \in \mathbb{N}^n$,

$$\varepsilon^{i_1} \wedge ... \wedge \varepsilon^{i_n} = \varepsilon^I,$$

so we will use both elements interchangeably. For any other multiindex $J \in \mathbb{N}^m$ with $J = (j_1, ..., j_m)$ we have $\varepsilon^I \wedge \varepsilon^J = \varepsilon^{IJ}$, where the composition IJ is defined as $IJ = (i_1, ..., i_n, j_1, ..., j_m)$. For any $\omega^1, ..., \omega^n \in V^*$, $v_1, ..., v_n \in V$

$$\omega^1 \wedge ... \wedge \omega^n(v_1, ..., v_n) = \det \omega^j(v_i).$$

2.27 Example. The dual basis for $\left(\mathbb{R}^3\right)^*$ is given by (e^1, e^2, e^3). Since $e^j(e_i) = \delta_i^j$, we have

$$\begin{aligned} e^{13}(v, w) &= \begin{vmatrix} e^1(v) & e^1(w) \\ e^3(v) & e^3(w) \end{vmatrix} = v^1 w^3 - w^1 v^3, \\ e^{123}(v, w, u) &= \det(v, w, u). \end{aligned}$$

[5]Note that the coefficient is chosen only for computational reasons.

Like for tensor fields, we define in a natural way a smooth subbundle and vector bundle of rank $\binom{n}{k}$ over M as the subset of $T^{(n,0)}TM$ of alternating tensors

$$\Lambda^n T^* M := \bigsqcup_{p \in M} \Lambda^n(T_p^* M).$$

2.28 Definition. A section of $\Lambda^n T^* M$ is said to be a **(differential)** n**-form**, i.e. a (continuous) tensor field, whose value at each point is an alternating tensor. The vector space of all smooth n-forms is denoted by $\Omega^n(M) := \Gamma^\infty(\Lambda^n T^* M)$.

Since $\Lambda^0 T_p M = \mathbb{R}$, we have $\Omega^0(M) = C^\infty(M)$. The wedge product is extended pointwise: $(\omega \wedge \eta)_p := \omega_p \wedge \eta_p$ and, since $\Omega^0(M) = C^\infty(M)$, we use the shorthand $f \wedge \eta =: f\eta$. Thus, for any smooth coordinate chart, ω can be written locally as

$$\omega = \sum_I \omega_I \, \mathbf{d}x^I = \sum_I \omega_I \, \mathbf{d}x^{i_1} \wedge \ldots \wedge \mathbf{d}x^{i_n} \quad (I \text{ increasing index}),$$

where the coefficients ω^I are continuous functions defined on the coordinate domain and $\mathbf{d}x^I$ is a shorthand for $\mathbf{d}x^{i_1} \wedge \ldots \wedge \mathbf{d}x^{i_n}$.

2.29 Theorem ([Lee13, Proposition 14.23]). *On a manifold M, there exists a unique operator $\mathbf{d} : \Omega^n(M) \to \Omega^{n+1}(M)$ for all $n \in \mathbb{N}_0$, the* **exterior derivative***, with the following properties.*

(i) $\mathbf{d}$ is linear over $\mathbb{R}$.

(ii) If $\omega \in \Omega^r(M)$, $\eta \in \Omega^1(M)$, then

$$\mathbf{d}(\omega \wedge \eta) = \mathbf{d}\omega \wedge \eta + (-1)^r \omega \wedge \mathbf{d}\eta.$$

(iii) $\mathbf{d} \circ \mathbf{d} = 0$.

(iv) For any $f \in \Omega^0(M)$ with differential $\mathbf{d}f$ of f, we have

$$\mathbf{d}f(X) = Xf.$$

In any smooth chart $\mathbf{d}$ is given by

$$\mathbf{d}\left(\sum_I \omega_I \, \mathbf{d}x^I\right) := \sum_I \mathbf{d}\omega_I \wedge \mathbf{d}x^I \quad (I \text{ increasing index}), \tag{2.2}$$

where $\mathbf{d}\omega_I$ is the differential of the continuous coefficient functions ω_I defined on the coordinate domain.

Thus, $\mathbf{d}f$ is an element of the dual space $T_p^* M$ of $T_p M$ and in this sense, we have $\mathbf{d}f_p(v) = v(f)$, for $v \in T_p M$, the **differential of** f. This means in particular, for any chart $\varphi = (x^1, \ldots, x^n)$ on U, we can take the differentials $\mathbf{d}x^i \in \Omega^1(U)$ of the coordinate function x^j themselves. More precisely,

$$\mathbf{d}x^i(\partial_j) = \partial_j x^i = \delta^i_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

In other words, the coordinate covector field is none other than the differential $\mathbf{d}x^i$. This justifies the notation $\mathbf{d}x^i$. Moreover, we have shown the following lemma.

2.30 Lemma. *At each point $p \in U$, the basis T_p^*M dual to $\left(\partial_1, ..., \partial_n\right)$ is $\left(\mathbf{d}x^1, ..., \mathbf{d}x^n\right)$, where $\mathbf{d}x^1, ..., \mathbf{d}x^n \in \Omega^1(U)$ are the differentials of the coordinate functions $x^i : U \to \mathbb{R}$ of a chart.*

If we choose $\omega \in \Omega^1(M)$, then, using the anticommutativity of differential forms,

$$\begin{aligned}\mathbf{d}\left(\omega_j\, \mathbf{d}x^j\right) &= \sum_{i,j} \frac{\partial \omega_j}{\partial x^i}\, \mathbf{d}x^i \wedge \mathbf{d}x^j \\ &= \sum_{i<j} \frac{\partial \omega_j}{\partial x^i}\, \mathbf{d}x^i \wedge \mathbf{d}x^j + \sum_{i>j} \frac{\partial \omega_i}{\partial x^j}\, \mathbf{d}x^j \wedge \mathbf{d}x^i \\ &= \sum_{i<j} \left(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j}\right) \mathbf{d}x^i \wedge \mathbf{d}x^j,\end{aligned}$$

which reduces for any function $f \in \Omega^0(M)$ to

$$\mathbf{d}f = \frac{\partial f}{\partial x^i}\mathbf{d}x^i,$$

i.e. the usual differential of f.

2.31 Definition. Let $F : M \to N$ a smooth map between manifolds, E a vector bundle over N and $n \in \mathbb{N}_0$. For any multilinear form $\alpha \in \Gamma(T^*N^{\otimes n} \otimes E)$ with values in E we define the **pullback form (induced by F)** by

$$(F^*\alpha)_p(v_1, ..., v_n) := \alpha_{F(p)}(\mathbf{d}F_p(v_1), ..., \mathbf{d}F_p(v_n)), \quad v_i \in T_pM, p \in M.$$

Then, $F^*\alpha \in \Gamma(T^*M^{\otimes n} \times F^*E)$ is a multilinear form with values in F^*E. In particular, for every $X \in \Gamma(E)$ we get $F^*X \in \Gamma(F^*E)$ with $(F^*X)_p = X_{F(p)}$, for $p \in M$.

If $E_1, ..., E_n \in \Gamma(E/U)$ is a local frame for E, then $F^*E_1, ..., F^*E_n \in \Gamma(F^*E/F^{-1}(U))$ is a local frame for F^*E. For any section $Y \in \Gamma(F^*E)$ there are uniquely determined functions $b^1, ..., b^n \in C^\infty(F^{-1}(U))$ with $Y\big|_{F^*(U)} = b^i F^*E_i$.

2.6 Riemannian Manifolds and the Musical Isomorphisms

2.32 Definition. A **Riemannian metric** $g = \langle \cdot, \cdot \rangle_g$ on M is a smooth symmetric covariant 2-tensor field on M that is positive definite at each point. A **Riemannian manifold** is a pair (M, g), where M is a smooth manifold and g is a Riemannian metric on M.

In any smooth local coordinates $\left(x^i\right)$, a Riemannian metric can be written as

$$g = g_{ij}\mathbf{d}x^i \otimes \mathbf{d}x^j,$$

where (g_{ij}) is a symmetric positive definite matrix of smooth functions. Using the **symmetric product** of two 1-forms $\omega\eta := \frac{1}{2}(\omega \otimes \eta + \eta \otimes \omega)$, this can be reduced to

$$g = g_{ij}\mathbf{d}x^i\mathbf{d}x^j$$

2.33 Example. The simplest example of a Riemannian manifold is the $\mathbb{R}^n$ with its **Euclidean metric** $\overline{g}$, which is defined as the inner product on each tangent space $T_p\mathbb{R}^n \cong \mathbb{R}^n$ for all $p \in M$. In standard coordinates, we get

$$\overline{g} = \delta_{ij}\mathbf{d}x^i\mathbf{d}x^j = \sum_i \mathbf{d}x^i\mathbf{d}x^i = \sum_i \left(\mathbf{d}x^i\right)^2. \tag{2.3}$$

Thus, the matrix in these coordinates is just $\overline{g} = \delta_{ij}$. Applied to vectors $v, w \in T_pM$, this yields

$$\overline{g}_p(v, w) = \delta_{ij}v^iw^j = \sum_{i=1}^{n} v^iw^i = v \cdot w.$$

Consequently, $\overline{g}$ is the 2-tensor field whose value at each point is the Euclidean dot product. To this end, we denote it by $\overline{g} = \langle \cdot, \cdot \rangle_{\mathbb{R}^n}$.

A Riemannian metric determines an inner product on each tangent space T_pM, which is typically written $\langle X, Y\rangle_g := g(X, Y)$ for $X, Y \in T_pM$. By Riesz' representation theorem, g provides a natural isomorphism between tangent and cotangent bundle given by $X \mapsto \langle X, \cdot\rangle$,

$$TM \underset{\sharp}{\overset{\flat}{\rightleftarrows}} T^*M.$$

More precisely, we define the **flat operator** $\flat : TM \to T^*M$, $X^\flat(Y) := g(X, Y)$. In coordinates, we have

$$X^\flat = \left\langle X^i\partial_i, \cdot\right\rangle = g(X^i\partial_i, \cdot) = g_{ij}X^i\mathbf{d}x^j = X_j\mathbf{d}x^j, \quad \text{where } X_j := g_{ij}X^i.$$

Since g is invertible, we define the **sharp operator** analogously by $\sharp : T^*M \to TM$, $\omega^\sharp(\eta) = g(\omega, \eta)$, in local coordinates this becomes

$$\omega^\sharp = \left\langle \omega_i\mathbf{d}x^i, \cdot\right\rangle = g(\omega_i\mathbf{d}x^i, \cdot) = g^{ij}\omega_j\partial_j = \omega^i\partial_i, \quad \text{where } \omega^i := g^{ij}\omega_j,$$

and g^{ij} are the components of the inverse of the metric tensor, so that $g_{ij}g^{jk} = \delta_i^k$.

The most important $\sharp$-operation is the generalisation of the gradient as a vector field on manifolds. For all smooth functions f on M, we define a vector field called the **gradient of f** by

$$\nabla f := \operatorname{grad} f := (\mathbf{d}f)^\sharp.$$

Hence, ∇f is the unique vector field that satisfies[6]

$$\langle \nabla f, X \rangle = Xf, \qquad X \in \Gamma^\infty(TM), \tag{2.4}$$

and this means

$$\langle \nabla f, \cdot \rangle = \mathbf{d} f, \tag{2.5}$$

the gradient is the dual of the differential $\mathbf{d}f$. In smooth coordinates ∇f has the expression

$$\nabla f = g^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j}, \tag{2.6}$$

which shows that ∇f is smooth. In particular, on $\mathbb{R}^n$ with Euclidean metric this reduces to the well-known formula from classical calculus

$$\nabla f = \delta^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j} = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^i}.$$

Now, we add an orientation.

2.34 Definition. A Riemannian manifold is called **orientable** if there is a smooth n-form which is nonzero at any point.

2.35 Proposition ([Lee13, Proposition 15.29]). *Suppose (M, g) is an oriented Riemannian n-manifold ($n \in \mathbb{N}$). Then there exists a unique $\mathrm{vol}_g \in \Omega^n(M)$, called the* **Riemannian volume form**, *that satisfies*

$$\mathrm{vol}_g(E_1, \ldots, E_n) = 1,$$

for some (hence every) local oriented orthonormal frame (E_i) for M.

Frequently, it is useful to have an expression for the Riemannian volume form in coordinates.

2.36 Proposition ([Lee13, Proposition 15.31]). *Let (M, g) be an oriented Riemannian n-manifold ($n \in \mathbb{N}$). In any oriented smooth coordinates x^i, the Riemannian volume form is given by*

$$\mathrm{vol}_g = \mathrm{G}\, \mathbf{d}x^1 \wedge \ldots \wedge \mathbf{d}x^n,$$

where $\mathrm{G} := \sqrt{\left|\det(g_{ij})\right|}$ *is the $n \times n$ matrix of the components in these coordinates.*

2.7 Connections and Laplace-Beltrami Operator

Next, we want to study local invariants on Riemannian manifolds. The key idea of connections is to generalise the directional derivative of a vector field in an coordinate-invariant

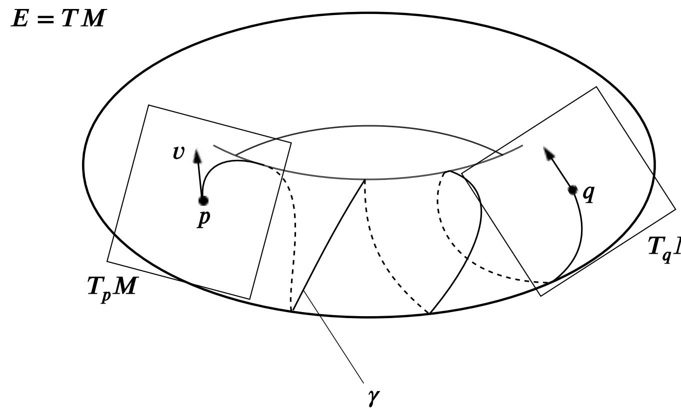

Figure 2.6: Parallel displacement on a manifold

way. Note that we focus on linear connection, i.e. connections in the tangent bundle. In geometrical terms, given a vector bundle $\pi : E \to M$ over M, say TM, each fibre is a tangent space at a point $p \in M$. Let $\gamma : [0,1] \to M$ be a smooth curve with $\gamma(0) = p$ and $\gamma(1) = q$. We are looking for a natural way of transporting vectors with their local geometry (length, angles) from $v \in E_p$ along γ to E_q. This requires an additional structure, called *linear connection*, which allows us to identify the fibres in a natural way, more precisely, to translate elements of one fibre from E along a curve to another. There are different, but equivalent, ways to define linear connections in a vector bundle: as parallel displacement, as covariant derivative and as horizontal splitting of TE.

First, we introduce the intuitive definition of parallel transport.

2.37 Definition. Let $\pi : E \to M$ be a vector bundle over M. A **parallel transport (translation) in E** assigns to every smooth path γ from p to q in M a linear isomorphism $L_\gamma : E_p \to E_q$ with the following properties:

(i) (Invariance under Reparametrisation) If $\gamma : [a,b] \to M$ is a smooth curve and $\sigma : [a',b'] \to [a,b]$ smooth with $\sigma(a') = a$ and $\sigma(b') = b$, then $L_{\gamma\circ\sigma} = L_\gamma$.
(Transitivity) For $\gamma : [a,b] \to M$ and $a \leqslant c \leqslant b$ we have $L_\gamma\big|_{[a,b]} = L_\gamma\big|_{[c,b]} \circ L_\gamma\big|_{[a,c]}$.
(Pullback) $L_{\gamma^-} = L_\gamma^{-1}$ for $\gamma^- : [a,b] \to M, t \mapsto \gamma(a+b-t)$.

(ii) (Dependence on Parameters) If γ depends smoothly on a parameter, so does L_γ.

(iii) (First Order Axiom) Let $\gamma : [a,b] \to M$ be a smooth curve and $\sigma \in \Gamma(\gamma^* E)$. We define the **covariant derivative $\nabla_D\sigma \in \Gamma(\gamma^* E)$ of σ to L along γ** to be

$$(\nabla_D\sigma)(t) := \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} L_{\gamma|_{[t,t+\varepsilon]}}\sigma(t+\varepsilon) \in E_{\gamma(t)}.$$

The First Order Axiom means that for $X \in \Gamma(E)$ and $v \in T_pM$ the **covariant deriva-**

[6] Indeed, the mapping $X \mapsto Xf$ is a linear functional on T_pM and by Riesz representation theorem, there exists a unique vector $h \in T_pM$ such that this linear functional is given by $\langle h, X\rangle_g$. Hence, we set $\nabla f(x) = h$.

tive $\nabla_v X$ of X in the direction v

$$\nabla_v X := \nabla_D (X \circ \gamma)(0) \in E_p,$$

is well-defined, for every smooth curve $\gamma : [-\varepsilon, \varepsilon] \to M$ with $\gamma(0) = 0$ and $\dot\gamma(p) = v$.

Next, we turn our attention to the abstract definition of a covariant derivative on a vector bundle E.

2.38 Definition. Let E be a vector bundle over M. A **covariant derivative on E** is an $\mathbb{R}$-linear map

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

satisfying, for $X \in \Gamma(E)$, $f \in C^\infty(M)$,

$$\nabla(fX) = \mathbf{d}f \otimes X + f\nabla X.$$

A section $X \in \Gamma(E)$ is **parallel** if $\nabla X = 0$. Since $\Gamma(T^*M \otimes E) \cong \mathrm{Hom}_{C^\infty(M)}(\Gamma(TM), \Gamma(E))$ we can also consider the covariant derivative ∇ on E as an $\mathbb{R}$-bilinear map

$$\begin{aligned} \Gamma(TM) \times \Gamma(E) &\to \Gamma(E) \\ (X, Y) &\mapsto \nabla_X Y := (\nabla Y)X. \end{aligned}$$

Hence,

(i) $X \mapsto \nabla_X Y$ is **linear over $C^\infty(M)$**, that is, for all $f, g \in C^\infty(M)$,

$$\nabla_{fX_1 + gX_2} Y = f\nabla_{X_1} Y + g\nabla_{X_2} Y,$$

(ii) and $\nabla_X Y$ satisfies a derivation-type product rule, for all $f \in C^\infty(M)$,

$$\nabla_X(fY) = f\nabla_X Y + (Xf)Y.$$

∇ is called **del** and $\nabla_X Y$ the **covariant derivative of Y in the direction of X**.

2.39 Remark. (i) Let $Y \in \Gamma(TM)$, $X \in \Gamma(E)$ and $p \in M$. Then $(\nabla_Y X)_p$ depends only on $Y_p \in T_pM$. Therefore, for $v \in T_pM$ and $Y \in \Gamma(TM)$ with $Y_p = v$, also $\nabla_v X := (\nabla_Y X)_p \in E_p$ is well-defined and called **covariant derivative of X in the direction v**.

(ii) For $v \in T_pM$, $X \in \Gamma(E)$, $\nabla_v X$ only depends on the germ of X in p.

2.40 Notation. Let ∇ be a covariant derivative on an n-dimensional vector bundle E over M, $d = \dim M$ and $e_1, \ldots, e_n \in \Gamma(E/U)$ a local frame for E. If (φ, U) is a chart for M, then $\partial_1, \ldots, \partial_d \in \Gamma(TM/U)$ is a local frame for TM. By Remark 2.39, the section $\nabla_{\partial_i} e_j \in \Gamma(E/U)$ is well-defined. The uniquely determined functions $\Gamma_{ij}^k \in C^\infty(U)$ with $\nabla_{\partial_i} e_j = \Gamma^k e_k$ are called **Christoffel symbols** of the covariant derivative ∇ (with respect to (φ, U) and $e_1, \ldots, e_n \in \Gamma(E/U)$). Moreover, they characterise the covariant derivative of ∇ in E/U.

2.41 Definition. Let ∇ be a covariant derivative on a vector bundle E over M and $\gamma : I \to M$ a smooth curve.

(i) (Covariant derivative of sections along curves) For sections $X \in \Gamma(\gamma^* E)$ along γ we call $\nabla_D X \in \Gamma(\gamma^* E)$ the **covariant derivative of X along γ**. Herein $D = \frac{\mathrm{d}}{\mathrm{d}t}$ is the canonical vector field on I.

(ii) (Parallel sections along curves) A section $X \in \Gamma(\gamma^* E)$ along γ is called **parallel along γ (with respect to ∇)**, if $\nabla_D X = 0$. $\Gamma_{\mathrm{par}}(\gamma^* E)$ denotes the real vector subspace of $\Gamma(\gamma^* E)$, the family of all parallel sections along γ.

2.42 Theorem ([HT94, Satz 7.83]). *Let ∇ be a covariant derivative on a vector bundle E over M and $\gamma : I \to M$ a smooth curve, $t_0 \in I$ and $e \in E_{\gamma(t_0)}$. Then there is a unique parallel section $X \in \Gamma_{\mathrm{par}}(\gamma^* E)$ along γ with $X_{t_0} = e$.*

Proof. Without loss of generality we can assume a global chart (h, M) for M and global frame $e_1, ..., e_n \in \Gamma(E)$ for E. Then there are uniquely determined coefficients $b^i \in \mathbb{R}$ with $e = b^i e_i$ and $\Gamma_{ij}^k \in C^\infty(M)$ (cf. Remark 2.40). Since $X \in \Gamma_{\mathrm{par}}(\gamma^* E)$, for fixed $t_0 \in I$ with $\gamma(t_0) \in M$, we have $X = X^i \gamma^* e_i$ locally at t_0. Thus, locally for t around t_0, using $\dot\alpha(t) = \dot\alpha^i(t)(\partial_i)_{\alpha(t)}$,

$$\begin{aligned}(\nabla_D X)(t) &= \dot X^j(t)(e_j)_{\gamma(t)} + X^j(t)\nabla_{\dot\gamma(t)} e_j \\ &= \dot X^j(t)(e_j)_{\gamma(t)} + X^j(t)\dot\gamma^i(t)\nabla_{(\partial_i)_{\gamma(t)}} e_j,\end{aligned}$$

i.e. locally around t_0

$$\nabla_D X = \Big(D(X^k) + X^j D(\gamma^i)(\Gamma_{ij}^k \circ \gamma)\Big)\gamma^* e_k.$$

Define $c_{jk} := -\dot\gamma^i(\Gamma_{ij}^k \circ \gamma) \in C^\infty(I)$, then from above $\nabla_D X = 0, X_{t_0} = e$ is equivalent to solve the system of linear differential equations

$$\dot X^k = -c_{kj} X^j, \quad X^k(t_0) = b^k, \quad k = 1, ..., n.$$

Mind that the uniquely determined solution of the system is defined on the whole interval I. ■

2.43 Definition. Let ∇ be a covariant derivative on a vector bundle E over M and $\gamma : I \to M$ a smooth curve. Define for $s, t \in I$ an isomorphism,

$$\begin{aligned}//_{s,t} &: E_{\gamma(s)} \to E_{\gamma(t)} \\ //_{s,t}(e) &:= X_t,\end{aligned}$$

where $X \in \Gamma_{\mathrm{par}}(\gamma^* E)$ with $X_s = e$, the **parallel displacement $//_{s,t}$ from $E_{\gamma(s)}$ to $E_{\gamma(t)}$ along γ**.

2.44 Remark. (i) We write $//_t := //_{0,t}$ for short.

(ii) It holds $/\!/_{s,t}^{-1} = /\!/_{t,s}$ and $/\!/_{t,t} = \mathrm{id}_{E_{\gamma(t)}}$. Every basis $(e_1, ..., e_n)$ for $E_{\gamma(s)}$ can be extended to a global frame $\bar{e}_1, ..., \bar{e}_n \in \Gamma(\gamma^* E)$ for $\gamma^* E$ by setting $\bar{e}_{i,t} := /\!/_{s,t}(e_i)$.

(iii) The parallel displacement $/\!/_{s,t}$ provides a parallel transport in the sense of definition 2.37 and constitutes the underlying covariant derivative: Let $X \in \Gamma(E), v \in T_pM$ and $\gamma : I \to M$ a smooth curve with $\dot{\gamma}(0) = v$, then

$$\nabla_v X = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \left(/\!/_t^{-1} X_{\gamma(t)}\right) \in E_p. \tag{2.7}$$

Proof. Let (e_i) and $(\bar{e}_i)$ be as in (i) and $a^i \in C^\infty(I)$ with $\gamma^* X = a^i \bar{e}_i$. Thus, we get $/\!/_t^{-1}(\gamma^* X)_t = a^i(t) e_i$ and

$$\nabla_v X = \nabla_D(\gamma^* X)_0 = \dot{\gamma}^i(0) e_i + \gamma^i(0)(\nabla_D \bar{e}_i)_0 = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \left(/\!/_t^{-1} X_{\gamma(t)}\right),$$

and the result follows. ∎

Thus, the parallel displacement in E and the covariant derivative on E define the same structure on M. We need a third equivalent concept.

2.45 Definition. Let $\pi : E \to M$ be a vector bundle over a manifold M. A subbundle $H \subset TE$ is called **horizontal bundle** or **horizontal splitting of TE** if

(i) $TE = H \oplus \pi^* E$, i.e. $T_e E = H_e \oplus E_{\pi(e)}$ for $e \in E$,

(ii) $\rho_* H_e = H_{s.e}$ for $\rho_s : E \to E, e \mapsto se$ and $s \in \mathbb{R} \setminus \{0\}$.

2.46 Remark. Let $p = \pi(e)$. The projection $\pi : E \to M$ is a **submersion**, i.e. $(\mathbf{d}\pi)_e : T_e E \to T_e M$ is surjective, with $\ker(\mathbf{d}\pi)_e = T_e(\pi^{-1}p) = T_e(E_p) \cong E_p \subset T_e E$. Then $h = (\mathbf{d}\pi|_H)^{-1} : \pi^* TM \xrightarrow{\sim} H \subset TE$ is smooth and called **horizontal lift**.

2.47 Notation. Let $TE = H \oplus V$ with $V := \pi^* E$ be a horizontal bundle of TE and $w \in T_e E$ with $e \in E$ and $p = \pi(e) \in M$. We call w **horizontal**, if $w \in H_e$ and **vertical** (in the sense "tangential at the submanifold E_p to E") if $w \in V_e \equiv E_p$. Even though there is naturally a vertical subspace V_e to every fibre $T_e E$ of TE, there is no canonical choice of a horizontal subspace H_e. A horizontal splitting is just the choice of an $\mathbb{R} \setminus \{0\}$-invariant complement space H_e to V_e (smooth in e) for every $e \in E$.

2.48 Theorem ([HT94, Satz 7.88]). *The Horizontal bundle of TE and covariant derivatives on E determine identical structures. More precisely, for every vector bundle $\pi : E \to M$ over M holds:*

(i) Every covariant derivative ∇ on E provides naturally a horizontal splitting H of TE, namely for $e \in E$ with $p := \pi(e)$ by

$$H_e := \left\{ X_* v : v \in T_pM, X \in \Gamma(E) \text{ with } X_p = e \text{ and } \nabla_v X = 0 \right\} \subset T_e E.$$

(ii) Conversely, every horizontal splitting H of TE provides a covariant derivative ∇ on E, namely:

(a) If $X \in \Gamma(E)$, then $\nabla X \in \Gamma(T^*M \otimes E)$ is determined by the following vector bundle morphism over M:

$$TM \xrightarrow{X_*} X^*TE \equiv X^*H \oplus X^*V \xrightarrow{\mathrm{pr}_V} X^*V = X^*\pi^*E = E.$$

(b) If $\sigma \in \Gamma(\gamma^*E)$ for a smooth curve γ on M, then $\nabla_D\sigma \in \Gamma(\gamma^*E)$ is determined as follows: $(\nabla_D\sigma)(t_0)$ is the image of $(\partial_t)_{t_0}$ under

$$\mathbb{R} = T\mathbb{R} \xrightarrow{\sigma_*} \sigma^*TE \equiv \sigma^*H \oplus \sigma^*V \xrightarrow{\mathrm{pr}_V} \sigma^*V = \sigma^*\pi^*E = E,$$

i.e. $(\partial_t)_{t_0}$ is mapped on $\dot\sigma(t_0)$ and then projected on the vertical component.

By Theorem 2.48, a section $X \in \Gamma(E)$ is parallel, i.e. $\nabla_v X = 0$ for all $v \in TM$, if and only if X_*v is horizontal for every $v \in TM$. Furthermore, a section $\sigma \in \Gamma(\gamma^*E)$ along a curve $\gamma : I \to M$ is parallel, i.e. $\nabla_D\sigma = 0$, if and only if $\dot\sigma(t) \in H_{\sigma(t)}$ for all $t \in I$. As a consequence of Theorem 2.42 we also get the following theorem.

2.49 Theorem ([HT94, Satz 7.89]). *Let $\pi : E \to M$ be a vector bundle over M and H a horizontal bundle of TE. Let $\gamma : I \to M$ be a smooth curve and $e \in E_{\gamma(t_0)}$ for some $t_0 \in I$. Then there is a unique horizontal lift of γ to a horizontal curve $u : I \to E$ with $u(t_0) = e$, i.e. $\pi \circ u = \gamma$ and $\dot u(t) \in H_{u(t)}$ for $t \in I$, $u(t_0) = e$.*

2.50 Definition. Let $\pi : E \to M$ be a vector bundle over a manifold M. A **linear connection on E** is a covariant derivative on E (equivalently a parallel displacement in E or a horizontal bundle of TE). A linear connection on TM is often called a **linear connection on M**.

2.51 Definition. (a) The **torsion tensor** T of a connection ∇ is defined by

$$\begin{aligned} T : \Gamma(TM)\times\Gamma(TM) &\to \Gamma(TM) \\ T(X,Y) &= \nabla_X Y - \nabla_Y X - [X,Y], \end{aligned}$$

where $[X,Y] := XY - YX$ denotes the Lie bracket. T can be identified as a $(1,2)$-tensor $T \in \Gamma(T^*M^{\otimes 2} \otimes TM)$. A linear connection is said to be **symmetric** or **torsion-free** if its torsion vanishes identically, i.e. if

$$\nabla_X Y - \nabla_Y X \equiv [X,Y]. \tag{2.8}$$

(b) Let (M,g) be a Riemannian manifold. A linear connection is said to be **compatible with** g if it satisfies the following product rule for all $X,Y,Z \in \Gamma(TM)$

$$X\langle Y,Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle. \tag{2.9}$$

A trivial example of a connection on $\mathbb{R}^n$ is the Euclidean connection, defined by

$$\overline{\nabla}_X(Y^j\partial_j) := (XY^j)\partial_j, \tag{2.10}$$

i.e. $\overline{\nabla}_X Y$ is just the vector field whose components are the ordinary directional derivatives of the components of Y in the direction X. Moreover, $\overline{\nabla}$ has the pleasant property (2.9), which can be easily verified by computing in terms of the standard basis. The next theorem shows that on any Riemannian manifold there is naturally a unique connection satisfying (2.9) and (2.8). This motivates the compatibility definition.

2.52 Theorem (Levi-Civita). *Let M be a Riemannian manifold. Then there exists a unique linear connection ∇ on M, called* **Levi-Civita connection***, that is compatible with g and symmetric.*

2.53 Definition. Let (M, g) and (N, h) be Riemannian manifolds. A smooth diffeomorphism $F : M \to N$ is said to be an **isometry** if $g = F^*h$. More precisely, if

$$g_p(u, v) = (F^*h)_p(u, v) \overset{\text{def}}{=} h_{F(p)}\left((\mathbf{d}F)_p u, (\mathbf{d}F)_p v\right)$$

for every $u, v \in T_pM$.

2.54 Definition. Let M be a manifold equipped with a linear connection ∇. For any $\omega \in \Omega^1(M)$ and $A \in \Gamma(TM)$ we define $\nabla_A\omega \in \Omega^1(M)$ by

$$\nabla\omega(A, B) \equiv \nabla_A\omega(B) := A(\omega B) - \omega(\nabla_A B), \quad B \in \Gamma(TM). \tag{2.11}$$

$\nabla_A\omega$ is well-defined since the righthand side only depends C^∞-linear of B. In particular, if $\omega = \mathbf{d}f$ with $f \in C^\infty(M)$, we call

$$\begin{aligned} &\nabla\mathbf{d}f \in \Gamma(T^*M \otimes T^*M) \\ &\nabla_A\mathbf{d}f(B) \equiv \nabla\mathbf{d}f(A, B) = ABf - (\nabla_A B)f \end{aligned}$$

the **second fundamental form** or **Hessian** of f.

2.55 Remark. We have $\nabla f \equiv \mathbf{d}f$ and $\nabla_Y f = \nabla f(Y) = \mathbf{d}f(Y) = Yf$ if f is a smooth function. Using the declared properties of the covariant derivative we find that $\nabla_X(\nabla f(Y)) = \nabla\nabla f(X, Y) + \nabla_{\nabla_X Y} f$. Combining both gives

$$\nabla\nabla f(X, Y) = X(Yf) - (\nabla_X Y)f. \tag{2.12}$$

On a Riemannian manifold $(M, \langle\cdot,\cdot\rangle)$ this is characterised by the musical isomorphisms (cf. Section 2.6) $\langle \operatorname{grad} f, Y\rangle = \left\langle (\mathbf{d}f)^\sharp, Y\right\rangle = Yf$. By the compatibility of the metric (2.9) and (2.12), we get $\nabla\mathbf{d}f(X, Y) = \langle\nabla_X \operatorname{grad} f, Y\rangle$.

2.56 Definition. Let (M, g) be a Riemannian manifold and ∇ the Levi-Civita connection on M. For $f \in C^\infty(M)$ the **Laplace-Beltrami operator** Δ_M is defined to be

$$\Delta_M f := \operatorname{tr} \nabla \mathbf{d} f \in C^\infty(M).$$

More precisely, this means for any orthonormal basis $E_1, ..., E_n$ for T_pM we have

$$\Delta_M f(p) = \sum_{i=1}^{n} \nabla \mathbf{d} f(E_i, E_i).$$

2.57 Remark. The divergence of a vector field X is defined to be the contraction of the $(1,1)$-tensor ∇X, i.e. $\operatorname{div} X := \operatorname{tr} \nabla X = \langle \nabla_{E^i} X, E_i \rangle$, any orthonormal basis $E_1, ..., E_n$ for T_pM. Then, we can define the Laplace-Beltrami operator in the more familiar form $\Delta_M := \operatorname{div} \circ \operatorname{grad}$. Moreover, setting $X := \operatorname{grad} f$ in the definition of $\operatorname{div} X$ this is obviously equivalent to $\Delta_M f \equiv \operatorname{tr} \nabla^2 f = \sum_{i=1}^n \nabla^2 f(X_i, X_i)$. The trace can be understood as follows: Consider f as a $(0,2)$-tensor, transform it into at $(1,1)$-tensor via the metric by applying the $\sharp$-operator. Then a $(1,1)$-tensor can be interpreted as an endomorphism of T_pM of which the trace can be taken. Thus, Δ_M is a nondegenerate second order elliptic operator. In local coordinates $\operatorname{div} X$ is given by

$$\operatorname{div} X = \frac{1}{\mathrm{G}} \partial_i \left(\mathrm{G} X^i \right). \tag{2.13}$$

Combining this with (2.6), we get the representation for Δ_M

$$\Delta_M f = \frac{1}{\mathrm{G}} \partial_i \left(\mathrm{G} g^{ij} \partial_j f \right). \tag{2.14}$$

2.58 Example. Let $M = \mathbb{R}^n$ with the Euclidean metric $\overline{g}$ as in (2.3) and ∇ the Levi-Civita connection. We recover the usual Euclidean definitions in the following way.

The local coordinates are the constant functions $(x \mapsto e_i) \in C^\infty(\mathbb{R}^n)$ for $i = 1, ..., n$ and $\overline{g}(v, w) = \langle v, w \rangle_{\mathbb{R}^n} = v \cdot w$, for $v, w \in T_p\mathbb{R}^n \cong \mathbb{R}^n$, is the usual dot product. Using local coordinates it is easy to verify that

$$\nabla_{\partial_i} \partial_j = 0, \tag{2.15}$$

hence

$$\nabla_v w \overset{\text{def}}{=} v\, w^i \partial_i = v\, w^i e_i = \left(v(w^1), ..., v(w^n) \right). \tag{2.16}$$

In the literature, this naturally given connection ∇ on $\mathbb{R}^n$ is often denoted D, i.e. $\nabla_v A = D_v A$. By (2.16), $\nabla_v A = D_v A = \langle \nabla A, v \rangle = \langle \operatorname{grad} A, v \rangle$ is the directional derivative of $A \in C^\infty(\mathbb{R}^n; \mathbb{R}^n)$ in the direction of $v \in \mathbb{R}^n$. Moreover,

$$\nabla \mathbf{d} f(\partial_i, \partial_j) \overset{\text{def}}{=} \partial_i \partial_j f - \nabla_{\partial_i} \partial_j f \overset{(2.15)}{=} \partial_i \partial_j f.$$

Thus, we regain the usual Euclidean Laplacian

$$\Delta_M f(x) \overset{\text{def}}{=} \operatorname{tr} \nabla \mathbf{d} f = \sum_{i=1}^{n} \nabla \mathbf{d} f(\partial_i, \partial_i) = \sum_{i=1}^{n} \partial_i \partial_i f = \sum_{i=1}^{n} \partial_i^2 f.$$

Note that by definition $\left|\overline{g}\right| = 1$. So, we could also have used the local representation (2.14).

Chapter 3

SDEs and Brownian Motion on Manifolds

In this chapter we introduce the notion of SDEs on manifolds. The main problem arising is that in general there is no possibility to transport the process qua charts on the manifold, since Itô's formula is obviously not invariant under coordinate transformations. But we will see that there is a possibility to define Brownian motion as a martingale problem. The Laplace operator will be replaced by the Laplace-Beltrami operator which depends on the Riemannian metric chosen. Thus, we require an additional structure in form of the parallel displacement. Therefore, controlled by the Riemannian metric, Brownian motion will be a local object by definition. However, its stochastic behaviour determines global aspects of the topology and geometry of the manifold.

The main sources for this chapter are [HT94], [Elw13], [Hsu02] and [ÉM89]. [HT94, Chapter 7] provides a systematic treatment of the modern differential geometry necessary to understand the notion of stochastic analysis on manifolds. Section 3.3 is also based on the original work by Elworthy [Elw82] and a lecture given at St. Flour [Elw13]. [Hsu02] treats the subject with less generality and does not require an extensive background in differential geometry. A standard reference for the necessary differential geometry is [KN63].

In this chapter let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbb{P})$ denote a standard filtrated probability space. Assume that the filtration satisfies the **usual hypotheses**, i.e. it is a right-continuous complete filtration. A stochastic process $(X_t)_{t\geqslant 0}$ is **adapted** if X_t is an $\mathscr{F}_t$-measurable random variable for each time $t \geqslant 0$. Let $\mathscr{S}(E)$ be the family of all continuous semimartingales on a set E,

$$\mathscr{S} = \mathscr{M} \oplus \mathscr{A}_0, \tag{3.1}$$

where $\mathscr{M}$ is the family of all continuous local martingales and $\mathscr{A}_0$ the family of all continuous finite variation processes starting at zero (a.s.). Such a unique (canonical) decomposition always exists, cf. [Pro05, p. 131] or [WR90, p. 358]. We somtimes suppress the adjective *continuous*.

Note that we use the different but common notations for the sample space Ω and the differential n-forms $\Omega^n(M)$ (cf. Definition 2.28).

3.1 SDEs and Martingales on Manifolds

3.1 Definition. Let $A \in \Gamma(TM)$ and let $A^2(f) := A(A(f))$ for $f \in C^\infty(M)$. A map $\mathbf{A} : C^\infty(M) \to C^\infty(M)$ is called a **Hörmander type partial differential operator (PDO)**

if there exist vector fields $A_0, A_1, \dots, A_n$ on M such that $\mathbf{A}$ is of the form

$$\mathbf{A} = A_0 + \sum_{i=1}^{n} A_i^2.$$

In particular, if $M = \mathbb{R}^n$ and $A_0 = 0$, $A_i := \partial_i$ $(i = 1, \dots, n)$, then $\mathbf{A} = \sum_{i=1}^{n} A_i^2 = \Delta$ is the usual Euclidean Laplace operator.

3.2 Definition. Let M be a manifold, $\mathbf{A}$ a Hörmander type PDO on M and $x \in M$. An adapted continuous process X with valued in M and with $X_0 = x$ is said to be a **flow process to A** (with starting point x) if for every test function $f \in C_c^\infty(M)$ on M the process

$$N(f)_t := f(X_t) - f(x) - \int_0^t \mathbf{A} f(X_r) \mathrm{d}r, \quad t \geqslant 0,$$

is a martingale.

A flow process may only have a finite lifetime ζ. Then ζ is a predictable stopping time and X defined on $[0, \zeta)$ such that on $\{\zeta < \infty\}$ holds: $X_t \xrightarrow{\text{a.s.}} \infty$ in the one-point compactification $\hat{M} := M \cup \{\infty\}$ of M at $t \nearrow \zeta$. In this case there exists a continuous extension $(X_t)_{t \geqslant 0}$ with values in $\hat{M}$ by setting $X_t(\omega) := \infty$ for $t \geqslant \zeta(\omega)$ and $f(\infty) := 0$ by definition, $f \in C_c^\infty(M)$.

3.3 Definition. The pair (A, Z) is called a **stochastic differential equation on a manifold** M (SDE on M) if

(i) Z is a continuous semimartingale with values in a finite dimensional real vector space E,

(ii) $A : M \times E \to TM$ is a vector bundle homomorphism on M.

We will denote the SDE (A, Z) as $\mathrm{d}X = A(X) \diamond \mathrm{d}Z$. Herein, "$\diamond$" denotes the Stratonovich circle in order to distinguish it from the usual composition of maps, denoted "$\circ$".

More precisely, condition (ii) constitutes the following commutative diagram

$$\begin{array}{ccc} (x,e) & \longmapsto & A(x)e \\ M \times E & \xrightarrow{\;A\;} & TM \\ \Big\downarrow{\scriptstyle \mathrm{pr}_M} & & \Big\downarrow{\scriptstyle \pi} \\ M & \xrightarrow[\mathrm{id}_M]{} & M \end{array}$$

and for every $x \in M$ the map $A(x) : E \to T_x M$ is linear on each fibre; in particular $A(\cdot)e \in \Gamma(TM)$ for $e \in E$. The semimartingale $Z = (Z_t)_{t \geqslant 0}$ is defined on a standard filtrated probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geqslant 0}, \mathbb{P})$ and we can write $Z = Z^i e_i$, where $(e_i)_{1 \leqslant i \leqslant r}$ is any basis for E and Z^i real semimartingales. We will always assume that Z has an infinite lifetime. This can be assured through a time shift if necessary, cf. [HT94, p. 364].

3.4 Definition. Let (A, Z) be an SDE and $\Xi : \Omega \to M$ an $\mathscr{F}_0$-measurable random variable. A **solution of the stochastic differential equation**

$$\mathrm{d}X = A(X) \diamond \mathrm{d}Z \tag{3.2}$$

with **initial condition** $X_0 = \Xi$ is a continuous adapted process $(X_t)_{t<\zeta}$ with values in M such that for every test function $f \in C_c^\infty(M)$ the composed process $f(X)$ is a real semimartingale and satisfies the integral equation

$$f(X_\tau) = f(\Xi) + \int_0^\tau (\mathbf{d}f)_X A(X) \diamond \mathrm{d}Z, \quad \mathbb{P} - \text{a.s.}, \tag{3.3}$$

for every stopping time τ with $0 \leqslant \tau < \zeta$. A solution of (3.2) with maximal lifetime is called **maximal solution** of the SDE (3.2); the SDE is **nonexplosive**. In this case (if necessary after passing over to the extension of f on $\hat{M}$ and X on $[0, \infty) \times \Omega$) up to indistinguishability

$$f(X_t) = f(X_0) + \int_0^t (\mathbf{d}f)_X A(X) \diamond \mathrm{d}Z, \quad t \geqslant 0. \tag{3.4}$$

Maximal lifetime of the continuous M-valued process X means that

$$\{\zeta < \infty\} \subset \left\{ \lim_{t \nearrow \zeta} X_t = \infty \text{ in } \hat{M} \right\}, \quad \mathbb{P} - \text{a.s.} \tag{3.5}$$

A solution of (3.2) is a semimartingale on M by definition (in the sense of L. Schwartz): Every adapted M-valued process X is a **semimartingale on** M if for every $f \in C_c^\infty(M)$ the composition $f(X)$ is a real-valued semimartingale. Mind that for the maximal lifetime of X the semimartingale $f(X)$ is well-defined on the hole line $[0, \infty)$. Moreover, the compositions $f(X)$ with smooth functions f are real semimartingales but, in general, only defined up to the lifetime of X.

Before we move on, let us get more familiar with the different admissible notations. For every $x \in M$ the composition

$$E \xrightarrow{A(x)} T_x M \xrightarrow{\mathbf{d}f_x} \mathbb{R}$$

is linear by definition. Therefore, if we write the semimartingale Z with a fixed basis $(e_i)_{1 \leqslant i \leqslant r}$ for E as $Z = \sum_1^r Z^i e_i$, we get

$$(\mathbf{d}f)_X A(X) \diamond \mathrm{d}Z \equiv \sum_{i=1}^r (\mathbf{d}f)_X A(X) e_i \diamond \mathrm{d}Z^i.$$

The bundle homomorphism A is naturally determined through the vector fields $A_i := A(\cdot)e_i$ for $i = 1, \dots, r$. Thus, we can symbolically write (3.2) as

$$\mathrm{d}X = \sum_{i=1}^r A_i(X) \diamond \mathrm{d}Z^i, \tag{3.6}$$

which should be read, for every test function $f \in C_c^\infty(M)$, as

$$\mathrm{d}f(X) = \sum_{i=1}^{r} (\mathbf{d}f)_X A_i(X) \diamond \mathrm{d}Z^i.$$

But $(\mathbf{d}f)_x A_i(x) = (A_i f)(x)$ so that the above equation is equal to

$$\mathrm{d}f(X) = \sum_{i=1}^{r} (A_i f)(X) \diamond \mathrm{d}Z^i, \quad f \in C_c^\infty(M).$$

Conversely, for a fixed basis $(e_i)_{1\leqslant i\leqslant r}$ for E and arbitrary vector fields $A_1, ..., A_r \in \Gamma(TM)$ on M there is a unique bundle homomorphism $A \in \Gamma(\mathrm{Hom}(M \times E, TM))$ with $A_i := A(\cdot)e_i$. Thus, the equations (3.2) and (3.6) are equivalent. Consequently, without loss of generality we can assume $E = \mathbb{R}^n$.

3.5 Example. Let $E = \mathbb{R}^{n+1}$ and $Z = (t, Z^1, ..., Z^n)$, where Z^i are real semimartingales. Let $A_0, A_1, ..., A_n \in \Gamma(TM)$ be given. Then (3.6) reads

$$\mathrm{d}X = A_0(X)\mathrm{d}t + \sum_{i=1}^{n} A_i(X) \diamond \mathrm{d}Z^i. \tag{3.7}$$

Thus, the composition $f(X)$ is a real semimartingale, for every $f \in C_c^\infty(M)$, and

$$\begin{aligned} \mathrm{d}f(X) &= (A_0 f)(X)\mathrm{d}t + \sum_{i=1}^{n} (A_i f)(X) \diamond \mathrm{d}Z^i \\ &= (A_0 f)(X)\mathrm{d}t + \sum_{i=1}^{n} (A_i f)(X)\mathrm{d}Z^i + \frac{1}{2}\sum_{i=1}^{n} \mathrm{d}\left((A_i f)(X)\right)\mathrm{d}Z^i, \end{aligned}$$

by the well-known conversion from Stratonovich to Itô differentials. Since

$$\mathrm{d}(A_i f(X)) = (A_0 A_i f)(X)\mathrm{d}t + \sum_{j=1}^{n} (A_j A_i f)(X) \diamond \mathrm{d}Z^j,$$

we get $\mathrm{d}\left((A_i f)(X)\right)\mathrm{d}Z^i = \sum_{j=1}^{n} (A_j A_i f)(X)\mathrm{d}Z^j\mathrm{d}Z^i$,[1] i.e.

$$\mathrm{d}f(X) = (A_0 f)(X)\mathrm{d}t + \frac{1}{2}\sum_{i,j=1}^{n} (A_i A_j f)(X)\mathrm{d}[Z^i, Z^j] + \sum_{i=1}^{n} (A_i f)(X)\mathrm{d}Z^i.$$

In particular, if we set $Z = (t, B^1, ..., B^n)$ where B is an n-dimensional Brownian motion we get, for every $f \in C_c^\infty(M)$,

$$\mathrm{d}f(X) = (A_0 f)(X)\mathrm{d}t + \frac{1}{2}\sum_{i=1}^{n} (A_i^2 f)(X)\mathrm{d}t + \sum_{i=1}^{n} (A_i f)(X)\mathrm{d}B^i,$$

where we used $\mathrm{d}B^i\mathrm{d}B^j = \delta_{ij}\mathrm{d}t$. But this means for $\mathbf{A} = A_0 + \frac{1}{2}\sum_1^n A_i^2$

$$\mathrm{d}f(X) - (\mathbf{A}f)(X)\mathrm{d}t = \mathrm{d}(\text{martingale})$$

[1] Revised version: corrected misprint.

3.6 Corollary. *Every maximal solution of the SDE*

$$\begin{aligned} \mathrm{d}X &= A_0(X)\mathrm{d}t + A_i(X) \diamond \mathrm{d}B^i \\ X_0 &= x \in M \end{aligned} \tag{3.8}$$

is a flow process X staring in x with generator $\mathbf{A} = A_0 + \frac{1}{2}\sum_{i=1}^n A_i^2$.

3.7 Theorem (Existence and Uniqueness, [HT94, Satz 7.43]). *Let (A, Z) be an SDE on M and Ξ an $\mathscr{F}_0$-measurable random variable. Then there exists a unique maximal solution X of (3.2) with lifetime $\zeta > 0$ $\mathbb{P}$-a.s. and initial condition $X_0 = \Xi$. Uniqueness holds in the following sense: For any other solution $(Y_t)_{t<\tau}$ of (3.2) with the same initial condition, it holds $(X_t)_{t<\tau} = Y$ $\mathbb{P}$-a.s. for every $\tau \leqslant \zeta$.*

3.8 Corollary. *Let $\mathrm{d}X = A(X)\diamond\mathrm{d}Z$ be the SDE (3.2) on M and $N \subset M$ a closed submanifold with $A(x)e \in T_xN$ for every $x \in N$ and $e \in E$. Then every solution of (3.2) starting on N stays up to its lifetime $\mathbb{P}$-a.s. on N.*

The proof of Theorem 3.7 can be found in [Elw82] or [HT94]. It is based on the famous

3.9 Theorem (Whitney's Embedding Theorem, [Lee13, Theorem 6.15]). *Every smooth n-manifold (with or without boundary) admits a proper smooth embedding into $\mathbb{R}^{2n+1}$ considered as a closed submanifold.*

The idea is very simple. Taking such a (Whitney) embedding ι, we can identify M with its image $M \overset{\iota}{\hookrightarrow} \iota(M) \subset \mathbb{R}^{2n+1}$, so that it is a submanifold of $\mathbb{R}^{2n+1}$. In particular, for every $x \in M$, the tangent space is a subspace of $\mathbb{R}^{2n+1}$ by

$$T_xM \overset{\mathrm{d}\iota|_x}{\hookrightarrow} T_x\mathbb{R}^{2n+1} = \mathbb{R}^{2n+1}.$$

Using a C^∞ partition of unity to extend A to a map

$$\overline{A} : \mathbb{R}^{2n+1} \times \mathbb{R}^k \to \mathbb{R}^{2n+1} \times \mathbb{R}^{2n+1},$$

we see that if X is a solution of (3.2) on M with $X_0 = \Xi$, then $\overline{X} := \iota \circ X$ is a solution to the new SDE $\mathrm{d}X = \overline{A}(X) \diamond \mathrm{d}Z$ on $\mathbb{R}^{2n+1}$ with $\overline{X}_0 = \iota \circ \Xi$. Therefore, also uniqueness follows. The main problem is to show that $\{t < \zeta\} \subset \left\{X_t \in M\right\}$ holds for every solution $\overline{X}$. This approach is often called *extrinsic*, since it relies on embedding the manifold in the ambient Euclidean space by a proper extension.

A solution X of (3.2) on M is, by definition, an M-valued semimartingale in the sense that all compositions $f(X)$ with $f \in C^\infty(M)$ are continuous real semimartingales on $[0, \zeta)$ (with ζ the lifetime of X). The converse is also true.

3.10 Theorem (M-valued semimartingales as solutions of SDEs, [HT94, Satz 7.47]). *Every semimartingale on a manifold can be written as the solution of (3.2).*

Finally, we generalise the quadratic variation for a semimartingale on M. In the case $M = \mathbb{R}^n$ with the Euclidean metric $\overline{g}$ (cf. (2.3)) this reduces to the classical concept of quadratic variation of a semimartingale. On a manifold, there exists nothing like multiplication, so the idea is to replace it by an arbitrary "bilinear form", i.e. a twice covariant tensor.

3.11 Proposition ([HT94, Satz 7.57]). *Let $X \in \mathcal{S}(M)$. Then there exists a unique linear map $\Gamma(T^*M \otimes T^*M) \to \mathcal{A}$, denoted by $b \mapsto \int b(\mathrm{d}X, \mathrm{d}X)$, such that, for all $f, g \in C^\infty(M)$,*

$$\mathbf{d}f \otimes \mathbf{d}g \mapsto [f(X), g(X)] \tag{3.9}$$

$$f \cdot b \mapsto \int f(X) b(\mathrm{d}X, \mathrm{d}X). \tag{3.10}$$

By definition, $b(\mathrm{d}X, \mathrm{d}X) := \mathrm{d} \int b(\mathrm{d}X, \mathrm{d}X)$.

3.12 Definition. The process $\int b(\mathrm{d}X, \mathrm{d}X)$ is said to be the **integral of b along X** or **b-quadratic variation of X**. Its value at time t will be denoted $\int_0^t b(\mathrm{d}X_s, \mathrm{d}X_s)$ instead of $\left(\int b(\mathrm{d}X, \mathrm{d}X)\right)_t$.

Proof. Every $b \in \Gamma(T^*M \otimes T^*M)$ can be written uniquely as $b = \sum_{i,j=1}^l b_{ij} \mathbf{d}\varphi^i \otimes \mathbf{d}\varphi^j$ for finitely many functions $b_{ij} \in C^\infty(M)$ and $\varphi^1, \ldots, \varphi^l \in C^\infty(M)$ (cf. Lemma A.3 (ii)). Set

$$\int b(\mathrm{d}X, \mathrm{d}X) := \sum_{i,j} \int b_{ij}(X)\ \mathrm{d}[\varphi^i(X), \varphi^j(X)], \tag{3.11}$$

then uniqueness follows immediately using (3.9) and (3.10). To show existence we verify that (3.11) is well-defined, i.e.

$$b = \sum_{\mathrm{fin}} u_k \mathbf{d}f^k \otimes \mathbf{d}g^k = 0 \quad \Longrightarrow \quad \sum_k u_k(X) \mathrm{d}[f^k(X), g^k(X)] = 0.$$

We write $\tilde{u}_k = u_k(\varphi)$, $f^k = \tilde{f}^k(\varphi)$ and $g^k = \tilde{g}^k(\varphi)$ with suitable extensions $\tilde{u}_k, \tilde{f}^v$, $\tilde{g}^k \in C^\infty(\mathbb{R}^l)$ (cf. Lemma A.3 (i)). Set $\tilde{X} = \varphi(X)$, then we have

$$\begin{aligned}
\sum_k u_k(X)\ \mathrm{d}[f^k(X), g^k(X)] &= \sum_k \tilde{u}_k(X)\ \mathrm{d}[\tilde{f}^k(\tilde{X}), \tilde{g}^k(\tilde{X})] \\
&= \sum_{i,j} \sum_k \tilde{u}^k(\tilde{X}) \partial_i \tilde{f}^k(\tilde{X}) \partial_j \tilde{g}^k(\tilde{X})\ \mathrm{d}[\tilde{X}^i, \tilde{X}^j] \\
&= \sum_{i,j} \underbrace{\left(\sum_k u_k \mathbf{d}f^k \otimes \mathbf{d}g^k \right)}_{=0} \left(\frac{\partial}{\partial \varphi_i}\Big|_X, \frac{\partial}{\partial \varphi_j}\Big|_X \right)\ \mathrm{d}[\tilde{X}^i, \tilde{X}^j] \\
&= 0.
\end{aligned}$$

∎

3.13 Remark (Properties of $b(\mathrm{d}X, \mathrm{d}X)$, cf. [ÉM89, (3.12)]). (i) The quadratic variation $\int b(\mathrm{d}X, \mathrm{d}X)$ depends only on the symmetric part of b. In particular, if b is anti-symmetric, then $\int b(\mathrm{d}X, \mathrm{d}X) = 0$.

(ii) The quadratic variation $\int b(\mathrm{d}X, \mathrm{d}X)$ commutes with changes of probability measure and continuous changes of time: If $\mathbb{Q} \ll \mathbb{P}$, then the integral $\mathbb{P}\text{-}\int b(\mathrm{d}X, \mathrm{d}X)$ is a version of $\mathbb{Q}\text{-}\int b(\mathrm{d}X, \mathrm{d}X)$. Since this is true for classical stochastic integrals and quadratic variations, it is clear from (3.9) and (3.10). If $(\tau_t)_{t\geqslant 0}$ is a continuous change of time (cf. [Pro05, p.247]), then the M-valued semimartingale $Y_t = Y_{\tau_t}$ (for the filtration $\mathscr{F}_{\tau_t}$) has the quadratic variation

$$\int_0^t b(\mathrm{d}Y, \mathrm{d}Y) = \int_{\tau_0}^{\tau_t} b(\mathrm{d}X, \mathrm{d}X).$$

In particular, $\int b(\mathrm{d}X^\tau, \mathrm{d}X^\tau) = \left(\int b(\mathrm{d}X, \mathrm{d}X)\right)^\tau$ for every finite stopping time τ.

3.14 Proposition ([HT94, Satz 7.61]). *Let $F : M \to N$ be a smooth map between manifolds and $b \in \Gamma(T^*N \otimes T^*N)$. For any $X \in \mathcal{S}(M)$, we have*

$$\int F^* b(\mathrm{d}X, \mathrm{d}X) = \int b\left(\mathrm{d}F(X), \mathrm{d}F(X)\right). \tag{3.12}$$

Proof. The lefthand side of (3.12) admits the defining properties of the b-quadratic variation for $F(X)$. ■

3.15 Proposition ([HT94, Satz 7.62]). *Let $X \in \mathcal{S}(M)$. Then there exists a unique linear map $\Omega^1(M) \to \mathcal{S}$, $\omega \mapsto \int \omega(\diamond \mathrm{d}X) =: \int_X \omega$ such that, for all $f \in C^\infty(M)$,*

$$\mathbf{d}f \mapsto f(X) - f(X_0) \tag{3.13}$$

$$f \cdot \omega \mapsto \int f(X) \diamond \omega(\diamond \mathrm{d}X). \tag{3.14}$$

By definition, $f(X) \diamond \omega(\diamond \mathrm{d}X) := f(X) \diamond \mathrm{d}\left(\int_X \omega\right)$. The commutation properties in Corollary 3.13 (ii) apply equally to $\int_X \omega$.

3.16 Definition. We call $\int_X \omega$ the **Stratonovich integral of the (differential) form ω along X.**

Proposition 3.15 is proved completely analogously to Proposition 3.11 using Lemma A.3 (iii), so we will not present it here.

3.17 Example. Let X be a smooth deterministic M-valued curve, i.e. $X_t = x(t)$, then

$$\int_X \omega = \int \omega(\dot{x}(t))\mathrm{d}t, \quad \omega \in \Omega^1(M).$$

3.18 Proposition (Pullback formula for Stratonovich integrals of differential forms, cf. [HT94, Satz 7.66]). *Let $F : M \to N$ be a smooth map, $\omega \in \Omega^1(M)$ and $X \in \mathcal{S}(M)$. Then*

$$\int_X F^* \omega = \int_{F(X)} \omega. \tag{3.15}$$

Proof. The lefthand side of (3.15) admits the defining properties of the Stratonovich integral of ω along $F(X)$. ■

The next formula is used to calculate the quadration variation of two forms. We will use it in the proof of Theorem 3.54.

3.19 Remark. Let $\omega, \eta \in \Gamma(T^*M)$, then $\omega \otimes \eta \in \Gamma(T^*M \otimes T^*M)$ and

$$\left[\int_X \omega, \int_X \eta\right] = \int (\omega \otimes \eta)(\mathrm{d}X, \mathrm{d}X). \tag{3.16}$$

The next result describes a pullback formula for the b-quadratic variation, cf. Definition 2.31. It will be used in the proof of Theorem 3.54.

3.20 Remark (cf. [HT94, Satz 7.72], [HT94, Satz 7.74]). (i) Let $X \in \mathcal{S}(M)$. In (3.11), we can replace the multiplier $f \in C^\infty(M)$ by a continuous adapted $\mathbb{R}$-valued process Y. Thus, there is a unique linear map $B \mapsto \int B(\mathrm{d}X, \mathrm{d}X)$ from the real vector space of continuous adapted $(T^*M \otimes T^*M)$-valued processes B to $\mathcal{A}$ such that $B_t \in T^*_{X_t}M \otimes T^*_{X_t}M$, $t \geqslant 0$, and

$$\begin{aligned} b(X) &\mapsto \int b(\mathrm{d}X, \mathrm{d}X) \\ Y \cdot B &\mapsto \int Y B(\mathrm{d}X, \mathrm{d}X), \text{ where } \int Y B(\mathrm{d}X, \mathrm{d}X) := \int Y \mathrm{d} \int B(\mathrm{d}X, \mathrm{d}X). \end{aligned}$$

This is proved in the same way as Proposition 3.11 but using part (iv) of Lemma A.3.

(ii) Let $X \in \mathcal{S}(M)$ and $F : M \to N$ be a smooth function. For any continuous adapted $(T^*N \otimes T^*N)$-valued process B with $B_t \in T^*_{F(X)}N \otimes T^*_{F(X)}N$, we get the pull back formula

$$\int F^*B(\mathrm{d}X, \mathrm{d}X) = \int B(\mathrm{d}F(X), \mathrm{d}F(X)). \tag{3.17}$$

Now we are able to define the notion of martingales on a manifold.

3.21 Definition. Let M be a manifold equipped with a linear connection ∇ and X be an M-valued semimartingale. Then X is said to be a **(∇)-martingale** if, for every $f \in C^\infty(M)$,

$$\mathbf{d}f(X) \overset{\mathrm{m}}{=} \frac{1}{2} \nabla \mathbf{d}f(\mathrm{d}X, \mathrm{d}X). \tag{3.18}$$

3.22 Remark. (i) In Definition 3.21, the notation $\overset{\mathrm{m}}{=}$ denotes **equality modulo continuous local martingales**, i.e. $X \overset{\mathrm{m}}{=} Y$ means that $X - Y$ is a continuous local martingale. For example, this gives another way of stating the uniqueness in the Doob-Meyer decomposition (3.1) by saying that two continuous processes with finite variation X and Y such that $X_0 = Y_0$ and $X \overset{\mathrm{m}}{=} Y$ are equal.

(ii) Since the notion of martingales is invariant under time changes, cf. Corollary 3.13 (ii), we can always assume an infinite lifetime through a time shift.

(iii) Note that martingales on M, by definition, have a.s. continuous sample paths.

(iv) Because $\nabla \mathbf{d} f(\mathrm{d}X, \mathrm{d}X)$ in (3.18) depends only on the symmetric part of $\nabla \mathbf{d} f$, we restrict ourselves to torsion-free linear connections, whereas the linear connection can be symmetrised as well.

3.23 Example. Let $M = \mathbb{R}^n$ equipped with the canonical linear connection, cf. Example 2.58. The family of all ∇-martingales is the family $\mathscr{M}$ of all continuous local martingales. This is obvious, since, by Itô's formula, a continuous $\mathbb{R}^n$-semimartingale is a local martingale if and only if, for every $f \in C^\infty(\mathbb{R}^n)$, it holds

$$\mathbf{d} f(X) - \frac{1}{2} \sum_{i,j} \partial_i \partial_j f(X) \, \mathrm{d}[X^i, X^j] \in \mathrm{d}\mathscr{M}.$$

But this is equivalent to $\mathbf{d} f(X) \stackrel{\mathrm{m}}{=} \frac{1}{2} \nabla \mathbf{d} f(\mathrm{d}X, \mathrm{d}X)$.

3.24 Remark (Martingales as solutions of SDEs). Let M be equipped with torsion-free linear connection. Let X be a solution of the SDE

$$\mathrm{d}X = A_0(X)\mathrm{d}t + A(X) \diamond \mathrm{d}Z,$$

where $A_0 \in \Gamma(TM)$, $A \in \Gamma(\mathrm{Hom}(M \times \mathbb{R}^n, TM))$ and Z a continuous $\mathbb{R}^n$-valued semimartingale. Set $A_i := A(\cdot)e_i$, $i = 1, \dots, n$, such that for every $f \in C^\infty(M)$

$$\mathrm{d} f(X) = (A_0 f)(X)\mathrm{d}t + \sum_{i=1}^n (A_i f)(X)\mathrm{d}Z^i + \frac{1}{2} \sum_{i,j=1}^n (A_i A_j f)(X) \, \mathrm{d}[Z^i, Z^j].$$

On the one hand $\nabla \mathbf{d} f(A_i, A_j) \stackrel{\mathrm{def}}{=} A_i A_j f - (\nabla_{A_i} A_j) f$ and on the other hand

$$\nabla \mathbf{d} f(\mathrm{d}X, \mathrm{d}X) = \sum_{i=1}^n \nabla \mathbf{d} f(A_i, A_j)(X) \, \mathrm{d}[Z^i, Z^j],$$

we find $\mathrm{d} f(X) - \frac{1}{2} \nabla \mathbf{d} f(\mathrm{d}X, \mathrm{d}X)$ as

$$(A_0 f)(X)\mathrm{d}t + \sum_{i=1}^n (A_i f)(X)\mathrm{d}Z^i + \frac{1}{2} \sum_{i,j=1}^n (\nabla_{A_i} A_j f)(X) \, \mathrm{d}[Z^i, Z^j].$$

In particular, if Z is a $\mathrm{BM}(\mathbb{R}^n)$, then X is a ∇-martingale if

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

3.2 Extrinsic Definition

We are now in the position to define Brownian motion as solution to a martingale problem on a Riemannian manifold.

3.25 Definition. Let (M, g) be a Riemannian manifold and X an adapted M-valued process with maximal lifetime ζ. The process X is a **Brownian motion on** (M, g) if, for every $f \in C^\infty(M)$, the real process

$$f(X) - \frac{1}{2}\int \Delta_M f(X)\mathrm{d}t$$

is a local martingale (with lifetime ζ). The family of all Brownian motions on (M, g) will be denoted by $\mathrm{BM}(M, g)$.

Next, we generalise Lévy's characterisation for M-valued Brownian motions.

3.26 Definition. Let X be a semimartingale with values on a Riemannian manifold $(M, g = \langle\cdot,\cdot\rangle)$. Then

$$[X, X] := \int g(\mathrm{d}X, \mathrm{d}X) = \int \langle \mathrm{d}X, \mathrm{d}X\rangle$$

is the **Riemannian quadratic variation** of X.

3.27 Theorem ([ÉM89, (5.18)], [HT94, Satz 7.116]). *Let X be a semimartingale with values on a Riemannian manifold (M, g) and maximal lifetime. Then the following are equivalent.*

(i) X is a Brownian motion.

(ii) X is a ∇-martingale with $[f(X), f(X)] = \int \|\mathrm{grad}\, f(X)\|^2\, \mathrm{d}t$ for every $f \in C^\infty(M)$.

In this case, for all $b \in \Gamma(T^*M \otimes T^*M)$

$$\int b(\mathrm{d}X, \mathrm{d}X) = \int \mathrm{tr}\, b(X)\mathrm{d}t. \tag{3.19}$$

In particular, the Riemannian quadratic variation of a Brownian motion X is given by

$$\int_0^t g(\mathrm{d}X, \mathrm{d}X) = (\dim M)t.$$

Proof. Let $X \in \mathcal{S}(M)$ with $[f(X), f(X)] = \int \|\mathrm{grad}\, f(X)\|^2\, \mathrm{d}t$ for every $f \in C^\infty(M)$. By polarisation, for every $f_1, f_2 \in C^\infty(M)$, also

$$[f_1(X), f_2(X)] = \int \langle \mathrm{grad}\, f_1(X), \mathrm{grad}\, f_2(X)\rangle_{T_X M}\, \mathrm{d}t.$$

Since for any orthonormal basis (E_i) of $T_X M$,

$$\begin{aligned}
\mathrm{tr}\, \mathbf{d}f_1 \otimes \mathbf{d}f_2 &\overset{\mathrm{def}}{=} \sum_i (\mathbf{d}f_1 \otimes \mathbf{d}f_2)(E_i, E_i) \\
&= \sum_i (\mathbf{d}f_1)(E_i)(\mathbf{d}f_2)(E_i) \\
&\overset{(2.5)}{=} \sum_i \langle \mathrm{grad}\, f_1, E_i\rangle \langle \mathrm{grad}\, f_2, E_i\rangle = \langle \mathrm{grad}\, f_1, \mathrm{grad}\, f_2\rangle,
\end{aligned}$$

we get

$$[f_1(X), f_2(X)] = \int (\mathbf{d}f_1 \otimes \mathbf{d}f_2)(\mathrm{d}X, \mathrm{d}X) = \int \mathrm{tr}(\mathbf{d}f_1 \otimes \mathbf{d}f_2)(X)\mathrm{d}t.$$

By uniqueness in Theorem 3.11 this means

$$\int b(\mathrm{d}X, \mathrm{d}X) = \int \mathrm{tr}\, b(X)\mathrm{d}t, \tag{3.20}$$

for an arbitrary $b \in \Gamma(T^*M \otimes T^*M)$.

(ii) $\Longrightarrow$ (i) If we set $b = \nabla\mathbf{d}f$ in (3.20), it follows that

$$\mathrm{d}f(X) \overset{\mathrm{m}}{=} \frac{1}{2}\nabla\mathbf{d}f(\mathrm{d}X, \mathrm{d}X) \overset{(3.20)}{=} \frac{1}{2}\Delta_M f(X)\mathrm{d}t,$$

i.e. X is a $\mathrm{BM}(M, g)$.

(i) $\Longrightarrow$ (ii) Let X a $\mathrm{BM}(M, g)$ and $f \in C^\infty(M)$. Then $\nabla\mathbf{d}f^2 = 2(f\nabla\mathbf{d}f + \mathbf{d}f \otimes \mathbf{d}f)$ implies that $\Delta(f^2) = 2f\Delta f + 2\,\|\mathrm{grad}\, f\|^2$, also

$$\mathrm{d}f^2(X) \overset{\mathrm{m}}{=} \frac{1}{2}\Delta_M f^2(X)\mathrm{d}t = (f\Delta_M f)(X)\mathrm{d}t + \|\mathrm{grad}\, f\|^2 \circ X\,\mathrm{d}t.$$

On the other hand, by Itô's formula,

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

Now it follows from the uniqueness of the Doob-Meyer decomposition that

$$[f(X), f(X)] = \int \|\mathrm{grad}\, f(X)\|^2\, \mathrm{d}t,$$

thus using (3.20) again, we conclude

$$\nabla\mathbf{d}f(\mathrm{d}X, \mathrm{d}X) \overset{(3.20)}{=} \mathrm{tr}\,\nabla\mathbf{d}f(X)\mathrm{d}t \overset{\mathrm{def}}{=} \Delta_M f(X)\mathrm{d}t.$$

Therefore, X is a ∇-martingale. ■

3.28 Example. By Theorem 3.27 it is clear that a Brownian motion in the sense of Definition 3.25 reduces to the usual definition of an $\mathbb{R}^n$-valued Brownian motion if we choose the $(\mathbb{R}^n, \overline{g})$, i.e. the Euclidean metric (2.3).

Finally, we show that an M-valued Brownian motion can be constructed as the solution of a suitable SDE.

3.29 Proposition ([HT94, Satz 7.118]). *Let (M, g) be a Riemannian manifold with the Levi-Civita connection ∇ on M and*

$$\mathrm{d}X = A(X) \diamond \mathrm{d}B + A_0(X)\mathrm{d}t, \tag{3.21}$$

where $A_0 \in \Gamma(TM)$, $A \in \Gamma(\mathrm{Hom}(M \times \mathbb{R}^n, TM))$ and B a $\mathrm{BM}(\mathbb{R}^n)$. Then the solution of (3.21) is a $\mathrm{BM}(M, g)$ if the following conditions are satisfied.

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

(ii) $A(x) \in \mathrm{Hom}(\mathbb{R}^n, T_x M)$ *is the projection on* $T_x M$ *for* $x \in M$*, i.e.* $A(x)A(x)^* = \mathrm{id}_{T_x M}$*, where* $A(x)^*$ *is the adjoint to* $A(x)$*.*

Proof. Let X be a solution of (3.21) such that both conditions are satisfied. By Remark 3.24, condition (i) assures that X is a ∇-martingale. Moreover, for every $f \in C^\infty(M)$,

$$\mathrm{d}f(X) \overset{\mathrm{m}}{=} \frac{1}{2}\sum_{i=1}^{n} \nabla \mathbf{d} f(A_i, A_i)(X)\mathrm{d}t.$$

Therefore, it is sufficient to show that $\sum_i \nabla \mathbf{d} f(A_i, A_i) = \Delta_M f$. For every $x \in M$, fix an orthonormal basis $(a_1, \dots, a_n)$ for $T_x M$. Then by definition

$$\begin{aligned}\Delta_M f(x) = \mathrm{tr}(\nabla \mathbf{d} f)_x &= \sum_i (\nabla \mathbf{d} f)_x(a_i, a_i) \\ &= \sum_i (\nabla \mathbf{d} f)_x(A(x)A(x)^* a_i, A(x)A(x)^* a_i).\end{aligned}$$

Now, extending $(A(x)^* a_1, \dots, A(x)^* a_n)$ to an orthonormal basis $(\tilde{e}_1, \dots, \tilde{e}_n)$ for $\mathbb{R}^n$ and denoting the standard basis for $\mathbb{R}^n$ by $(e_1, \dots, e_n)$, we get

$$\begin{aligned}\Delta_M f(x) &= \sum_{i=1}^{n} \nabla \mathbf{d} f_x(A(x)\tilde{e}_i, A(x)\tilde{e}_i) \\ &= \sum_{i=1}^{n} \nabla \mathbf{d} f_x(A(x)e_i, A(x)e_i) = \sum_{i=1}^{n} \nabla \mathbf{d} f_x(A_i(x), A_i(x)),\end{aligned}$$

and the result follows. ■

The conditions (i) and (ii) in Proposition 3.29 can always be achieved for n large enough: Let $M \hookrightarrow \mathbb{R}^n$ a Whitney embedding on the manifold M such that $T_x M \hookrightarrow \mathbb{R}^n$ is a vector subspace for every $x \in M$. Define $A \in \Gamma(\mathrm{Hom}(M \times \mathbb{R}^n, TM))$ on each fibre as orthogonal projection $A(x) : \mathbb{R}^n \to T_x M$ on $T_x M$ and define $A_0 \in \Gamma(TM)$ by $A_0 = -\frac{1}{2}\nabla_{A_i} A_i$, then every solution to (3.21) provides a Brownian motion on (M, g).

However, there is no canonical choice of these coefficients A_0 and A on an arbitrary Riemannian manifold. But there is a canonical SDE on the orthonormal frame bundle $\mathcal{O}(M)$ over M, whose solution (projected by $\pi : \mathcal{O}(M) \to M$) defines a Brownian motion on (M, g). This is the Eells-Elworthy-Malliavin construction which is presented below.

3.3 The Eells-Elworthy-Malliavin Construction or Intrinsic Definition

In the previous section, we have shown that Brownian motion can be defined as the solution of a suitable SDE on a Riemannian manifold M. However, there is no Hörmander type representation of the Laplace-Beltrami operator Δ_M if the manifold is not parallelizable, i.e. the tangent bundle is not trivial (cf. [Lee13, Corollary 10.20]). But it holds the fundamental relation (cf. Theorem 3.53 below)

$$\Delta_{\mathcal{O}(M)}\pi^* = \pi^*\Delta_M, \tag{3.22}$$

i.e. there exists a *lifted* version of the Laplace-Beltrami operator, called *horizontal Laplacian*, on the orthonormal frame bundle $\mathcal{O}(M) \to M$ over M. Each element $u \in \mathcal{O}(M)$ is an isometry $u : \mathbb{R}^n \to T_{\pi(u)}M$. The set of tangent vectors of horizontal curves passing through a fixed point $u \in \mathcal{O}(M)$ is the horizontal splitting $H_u\mathcal{O}(M)$ (cf. 3.34, 3.37) with

$$T_u\mathcal{O}(M) = H_u\mathcal{O}(M) \oplus V_u O(M),$$

and n well-defined unique horizontal vectors $L_i(u) \in H_u\mathcal{O}(M)$ whose projection is the ith unit vector ue_i of the orthonormal frame, i.e. $\pi_* L_i(u) = ue_i$, where (e_i) is the canonical basis for $\mathbb{R}^n$. Using this relation, it is due to Malliavin, Eells and Elworthy that there always exists a lifted Brownian motion as solution of the globally defined SDE

$$\mathrm{d}U_t = L_i(U_t) \diamond \mathrm{d}B_t^i,$$

where B is an n-dimensional Brownian motion. A solution is a diffusion generated by $\Delta_{\mathcal{O}(M)}$. By Itô's formula for $\tilde{f} \in C^\infty(\mathcal{O}(M))$

$$\mathrm{d}\tilde{f}(U_t) = L_i\tilde{f}(U_t)\mathrm{d}B_t^i + \frac{1}{2}\Delta_{\mathcal{O}(M)}\tilde{f}(U_t)\mathrm{d}t.$$

Applying this to the lift $\tilde{f} := f \circ \pi$ we get, using (3.22),

$$\mathrm{d}f(X_t) = L_i f(X_t)\mathrm{d}B_t^i + \frac{1}{2}\Delta_M f(X_t)\mathrm{d}t,$$

where $X_t = \pi(U_t)$ is the projection of the lifted Brownian motion U_t on the manifold M. It follows that X_t is a Brownian motion on M starting from $X_0 = \pi U_0$. Therefore, the key idea was to solve conversely the SDE on the orthonormal frame bundle $\mathcal{O}(M)$ and project the solution back down to M by $\pi : \mathcal{O}(M) \to M$, cf. [Elw82], [Mal78].

In geometrical terms, the idea is to "roll" our manifold M by means of the (stochastic) parallel displacement along the paths of an $\mathbb{R}^n$-valued Brownian motion ("rolling without slipping"), known as *stochastic development*. Starting in $p \in M$, the resulting Brownian motion X on M can be thought of as footprints left behind by the paths of the Euclidean Brownian motion B in the tangent space $T_pM \cong \mathbb{R}^n$ if M is rolled along the paths of B, cf. Figure 3.3. The procedure is known as *Cartan development* in the deterministic case. We will see that it can be adopted to work with a suitable Stratonovich SDE.

To this end, we define another special class of fibre bundles, called principal bundles. As a vector bundle is a fibre bundle whose typical fibre is a finite-dimensional vector space, a principal bundle is a fibre bundle whose typical fibre is a Lie group. An important example will be the orthonormal frame bundle $\mathcal{O}(M)$ on a Riemannian manifold M with structure group $G = \mathrm{O}(n)$. Therefore, we briefly sum up the necessary background on Lie groups and define principal bundles. Furthermore, we generalise the main geometrical concepts, like connections and parallel displacement. As it would be beyond the scope of this work we will not provide every notion in detail.

A **Lie group** G is a group which is also a smooth manifold such that the group operation $G \times G \to G, (g,h) \mapsto gh^{-1}$ is a smooth map. It is traditional to denote the identity element by the symbol e (for German *Einselement*).

Let G be a Lie group and M be a manifold. A G**-left action on** M is a smooth map

$$M \times G \to M, \quad (p,g) \mapsto g.p =: L_g(p),$$

that satisifes

(a) $g_1.(g_2.p) = (g_1 g_2).p$ for all $g_1, g_2 \in G$ and $p \in M$,

(b) $e.p = p$ for all $p \in M$.[2]

Thus, a G-left action is a group homomorphism from G into the group of smooth diffeomorphisms on M such that the map $M \times G \to M$ is smooth. A G-left action is said to be **effective**, if $g \mapsto L_g$ is injective and **free** if $g.p = p$ for $p \in M$ implies $g = e$ (*transitivity* is the existence and *freedom* is the uniqueness). A typical example is the natural left action, more precisely, left translation of G itself. A G**-right action** is defined analogously. Mind that, if $M \times G \to M, (p,g) \mapsto p.g$ is a G-right action on M, then $M \times G \to M, (p,g) \mapsto pg^{-1}$ is a G-left action on M. For $g \in G$, $\mathrm{ad}\, g$ is the inner automorphism of G defined by $(\mathrm{ad}\, g)h := R_{g^{-1}} L_g h = ghg^{-1}$ for every $h \in G$, the **adjoint action of** G.

3.30 Example. (i) The **general linear group** $\mathrm{GL}(n,\mathbb{R})$ of invertible $n \times n$ matrices with real entries is a Lie group.

(ii) The group $\mathrm{O}(n)$ of all $n \times n$ orthogonal matrices is a Lie group.

3.31 Definition. A **principal** G**-bundle over** M is a fibre bundle $\pi : P \to M$ together with a continuous right action $R_g : P \times G \to P$ of G on P such that

(i) R_g preserves the fibres of P, i.e. if $u \in P_x$ then $ug \in P_x$ for all $g \in G$ and $x \in M$, and acts freely and transitively on them;

(ii) there is an open cover $(U_\alpha, \varphi_\alpha)_{\alpha \in I}$ of M with local trivialisations

$$\varphi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times G$$

such that, under φ_α, the action of G on $\pi^{-1}(U_\alpha)$ is from the right, i.e. if $\varphi_\alpha(u) = (\pi(u), p)$), then $\varphi_\alpha(R_g(u)) = (\pi(u), p.g)$.

[2] Revised version: corrected misprint.

3.32 Example. For any manifold M and any Lie group G, the trivial fibre bundle $\pi : P(= M \times G) \to M$ is a principal G-bundle, with the right action R given by $R_g(p,h) = (p, h.g)$, for all $(p,h) \in P = M \times G$ and $g \in G$.

3.33 Example. If $\pi : E \to M$ is equipped with a Riemannian metric, the bundle charts are linear on each fibre and provide an $\mathcal{O}(n)$ atlas, i.e. a restriction of the structure group $\mathrm{GL}(n,\mathbb{R})$ to $\mathrm{O}(n)$.

An important example is the **frame bundle** $P = \mathrm{GL}(M) \to M$ over M consisting of all frames $u \in P_x$ at some point $x \in M$, i.e. linear isomorphisms $u : \mathbb{R}^n \to T_xM$ which we can identify with the $\mathbb{R}$-basis for T_xM

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

where $(e_1, \dots, e_n)$ is the standard basis for $\mathbb{R}^n$. The group $G = \mathrm{GL}(n,\mathbb{R})$ operates on $\mathrm{GL}(M)$ from the right

$$u.g : \mathbb{R}^n \xrightarrow{g} \mathbb{R}^n \xrightarrow{u} T_xM,$$

where $g = (g_{ij}) \in G$. Then $ug \in \mathrm{GL}(M)$ by $(ug)_j = \sum_i g_{ij}u_i$.

Restricting the structure group $\mathrm{GL}(n,\mathbb{R})$ to $\mathrm{O}(n)$, the frames $u \in P_x$ at a point $x \in M$ become isometries. We call $P = \mathcal{O}(M) \to M$ consisting of all frames $u \in P_x$ at a point $x \in M$, i.e. linear isometries $u : \mathbb{R}^n \to T_xM$, the **orthonormal frame bundle over** M.

To identify fibres (intrinsically) we need an additional structure, namely a generalised form of a connection (covariant derivative, parallel displacement). But we will see that this can be generalised on principal $\mathrm{GL}(n,\mathbb{R})$-bundles.

3.34 Definition. Let $\pi : P \to M$ a principal bundle over M with structure group G. A G-**connection** in a principal bundle is a smoothly-varying assignment to each point $u \in P$ of a subspace H_P of T_uP such that

(i) $T_uP = V_u \oplus H_u$ for all $u \in P$, i.e. $TP = V \oplus H := \ker \mathbf{d}\pi \oplus h(\pi^*TM)$,

(ii) $H_{R_g(u)} = (\mathbf{d}R_g)H_u$ for all $g \in G, u \in P$ (G-**invariance**).

We call H_u the **horizontal space** at u and

$$V_u \overset{\text{def}}{=} \ker \mathbf{d}\pi \equiv \left\{ v \in T_uP : (\mathbf{d}\pi)v = 0 \right\}$$

the **vertical space** at u. The bundle isomorphism

$$h : \pi^*TM \xrightarrow{\sim} H \hookrightarrow TP \tag{3.23}$$

is called **horizontal lift** of the G-connection, i.e. on each fibre $h_u : T_{\pi(u)}M \xrightarrow{\sim} H_u$ for $u \in P$.

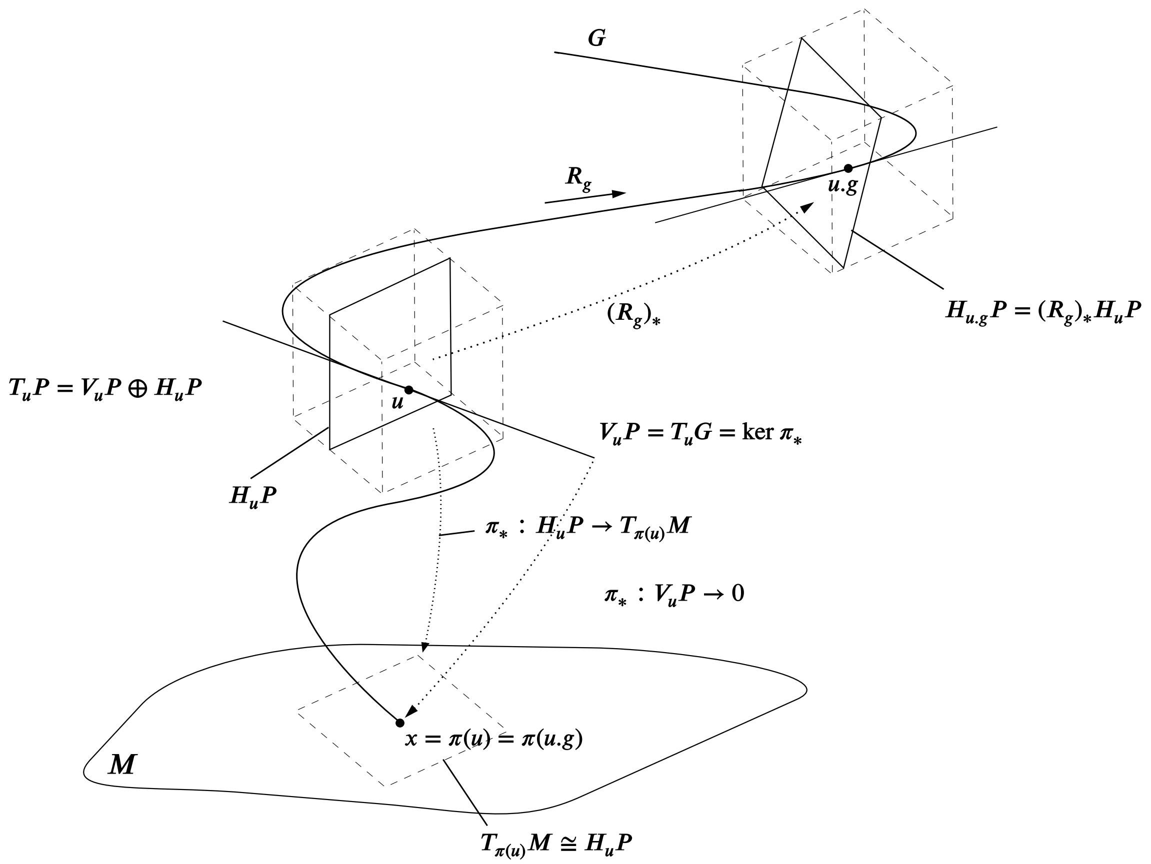

Figure 3.1: G-connection of a principal G-bundle

Mind that, for every $u \in P$, the vertical space V_u is naturally given and G-invariant. But in general, there is no canonical choice of a complementary space H_u. However, a G-connection in a principal bundle splits every vector field $X \in \Gamma(TP)$ in a horizontal and a vertical part: $X = X^{\mathrm{hor}} + X^{\mathrm{vert}}$.

3.35 Definition. Let $\pi : P \to M$ a principal G-bundle over M. Every $u \in P$ defines an embedding $I_u : G \hookrightarrow P, g \mapsto u.g$. The differential ι_u of I_u at the Einselement $e \in \mathfrak{g}$,

$$\iota_u := (\mathrm{d}I_u)_e : T_eG \longrightarrow T_uP, \quad A \mapsto \tilde{A}(u), \tag{3.24}$$

provides an identification $\kappa_u : \mathfrak{g} \xrightarrow{\sim} V_u$ from the Lie algebra $\mathfrak{g} := T_eG$ of G with vertical fibre V_u at u. The vertical vector field $\tilde{A} \in \Gamma(TP)$, defined by (3.24), is the **standard vertical vector field on P associated to $A \in \mathfrak{g}$**.

3.36 Definition. Let $\pi : P \to M$ a principal G-bundle over M with a G-connection. By

$$\tilde{\omega}_u(X_u) := \kappa_u^{-1}(X^{\mathrm{vert}})_u, \quad X \in \Gamma(TP), \tag{3.25}$$

we define a $\mathfrak{g}$-valued 1-form $\tilde{\omega} \in \Gamma(T^*P \otimes \mathfrak{g})$ on P, called **connection form** of the G-connection.

The connection form is by definition **horizontal**, i.e. $\tilde\omega(X) = 0$ if and only if X is a horizontal vector field on P.

A G-connection in P is determined uniquely by its connection form $\tilde\omega$: for all $u \in P$, $\tilde\omega_u : T_uP \to \mathfrak{g}$ is a linear map with $\ker\tilde\omega_u = H_u$. Conversely, every equivariant differential form, i.e. $\tilde\omega((\mathrm{d}R_g)X) = \mathrm{ad}(g^{-1})\tilde\omega(X)$ for $X \in \Gamma(TP)$,

$$\tilde\omega \in \Gamma(T^*P \otimes \mathfrak{g}) \text{ with } \tilde\omega(\tilde A) = A \text{ for } A \in \mathfrak{g}$$

defines a G-connection in P whose connection form is given by $\tilde\omega$.

3.37 Remark. In the language of differential geometry a G-connection in P is a smooth splitting of the exact sequence (3.26) of vector bundles over P. This will help us to understand the different notions and the generalisation of the covariant derivative below. We may think of it as a way to visualise the definitions from above. We can describe the G-connection

$$\begin{array}{ccccccccc} & & P\times\mathfrak{g} & & V\oplus H & & & & \\ & & \| & & \| & & & & \\ 0 & \longrightarrow & \ker \mathrm{d}\pi & \xrightarrow{\iota} & TP & \xrightarrow{\mathrm{d}\pi} & \pi^*TM & \longrightarrow & 0 \\ & & & \overset{\tilde\omega}{\dashleftarrow} & & \overset{h}{\dashleftarrow} & & & \end{array} \tag{3.26}$$

(i) as horizontal lift h: $\mathrm{d}\pi \circ h = \mathrm{id}$,

(ii) as differential form $\tilde\omega \in \Gamma(T^*P \otimes \mathfrak{g})$ with $\tilde\omega \circ \iota = \mathrm{id}$,

(iii) or as horizontal subbundle H: $V \oplus H = TP$.

On each fibre (3.26) reads

$$\begin{array}{ccccccccc} & & & & V\oplus H & & & & \\ & & & & \| & & & & \\ 0 & \longrightarrow & \mathfrak{g} & \xrightarrow{\iota_u} & TP & \xrightarrow{\mathrm{d}\pi|_u} & T_{\pi(u)}M & \longrightarrow & 0 \\ & & & \overset{\tilde\omega_u}{\dashleftarrow} & & \overset{h_u}{\dashleftarrow} & & & \end{array} \tag{3.27}$$

(i) as $(\mathrm{d}R_g)h = h$ for every $g \in G$,

(ii) as $\tilde\omega((\mathrm{d}R_g)X) = \mathrm{ad}(g^{-1})\tilde\omega(X)$ for $X \in \Gamma(TP)$,

(iii) or as $(\mathrm{d}R_g)H_u = H_{u.g}$ for every $u \in P$.

There is a one-to-one correspondence between connections in principal bundles and linear connections in the associated vector bundles (cf. [HT94, p. 7.129]). Therefore, we also have a one-to-one correspondence between the Levi-Civita connection in TM and the $\mathrm{O}(n)$-connections in the orthonormal frame bundle $\mathcal{O}(M)$. To this end, in the following we call $\mathrm{O}(n)$-connections in the orthonormal frame bundle $\mathcal{O}(M)$ also **Levi-Civita connections on M.**

3.38 Theorem ([HT94, Satz 7.130]). *Let $\pi : P \to M$ be a principal G-bundle over M with a G-connection. Let $x : I \to M$, $t \mapsto x(t)$ be a smooth curve and $t_0 \in I$. Then there exists for every $u_0 \in P$ with $\pi(u_0) = x(t_0)$ a unique horizontal curve $u : I \to P$ with $u(t_0) = u_0$ above $t \mapsto x(t)$, i.e. $\pi \circ u(t) = x(t)$ and $\dot{u}(t) \in H_{u(t)}$ for every $t \in I$.*

3.39 Corollary. *Let P be the principal G-bundle over a manifold M. Every G-connection in P naturally defines a parallel displacement on P along smooth curves $t \mapsto x(t)$ in M, namely for $t_0, t_1 \in I$ as*

$$\overline{//_{t_0,t_1}} : P_{x(t_0)} \xrightarrow{\sim} P_{x(t_1)}, \quad u_0 \mapsto u(t_1), \tag{3.28}$$

where $t \mapsto u(t)$ is the uniquely determined horizontal lift of $t \mapsto x(t)$ on P with $u(t_0) = u_0$.

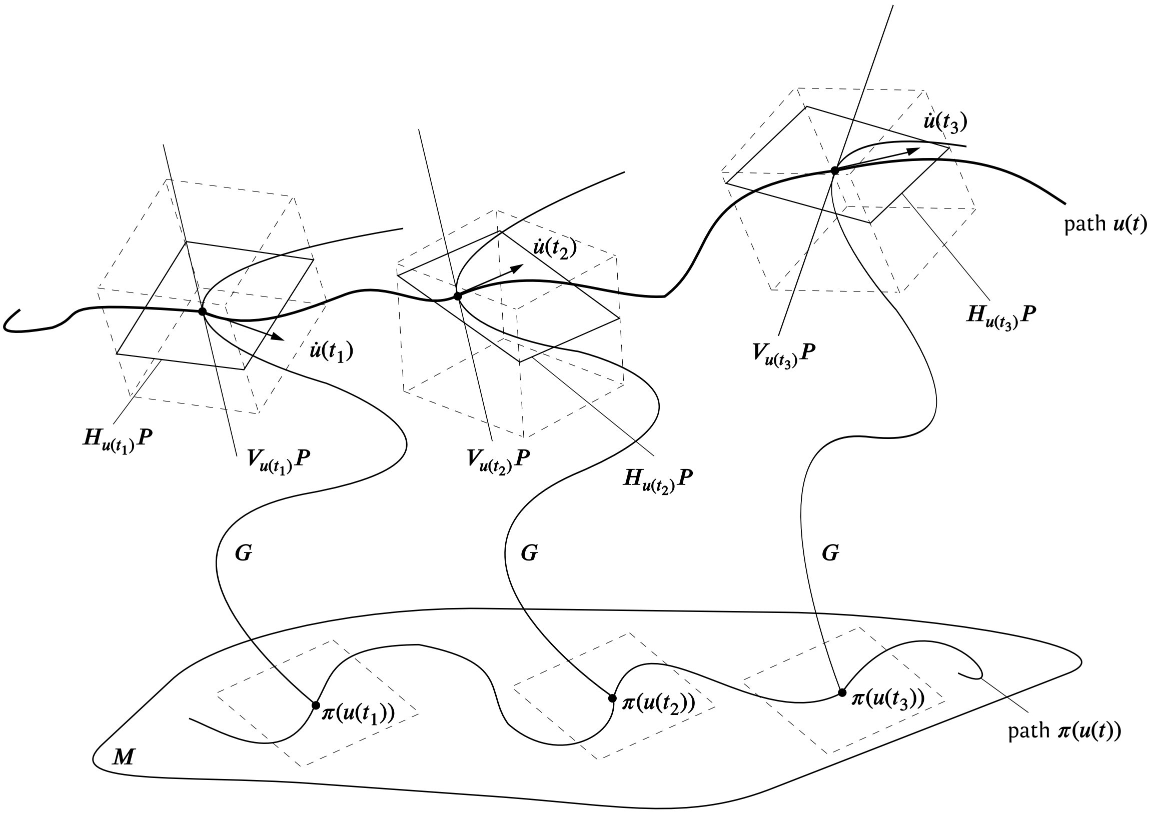

Figure 3.2: Horizontal lift $u(t)$ through principal G-bundle

For the rest of this chapter we restrict ourselves to the principal G-bundle $P = \mathcal{O}(M)$ over a Riemannian manifold M with $G = \mathrm{O}(n)$. The associated Lie algebra is given by the algebra of matrices $\mathfrak{g} = \{A \in \mathrm{M}(n \times n; \mathbb{R}) : A \text{ skew-symmetric}\}$. Fix a G-connection in P with

$$\tilde{\omega} \in \Gamma(T^*P \otimes \mathfrak{g}), \quad \tilde{\omega}_u(X_u) = \kappa_u^{-1}(X^{\mathrm{vert}})_u, \quad u \in P \text{ and } X \in \Gamma(TP), \tag{3.29}$$

the $\mathfrak{g}$-valued connection form (cf. Definition 3.36) and

$$\vartheta \in \Gamma(T^*P \otimes \mathbb{R}^n), \quad \vartheta_u(X_u) := u^{-1}(\mathrm{d}\pi X_u), \quad u \in P \text{ with } X \in \Gamma(TP), \tag{3.30}$$

the so called **canonical** 1-**form** on the principal bundle $\pi : P \to M$. We read every $u \in P = \mathcal{O}(M)$ as isomorphism $u : \mathbb{R}^n \xrightarrow{\sim} T_{\pi(u)}M$. Note that the definition of a connection form depends on the G-connection, but not the canonical 1-form ϑ.

3.40 Theorem ([HT94, Satz 7.131]). *The orthonormal frame bundle $P = \mathcal{O}(M)$ as a Riemannian manifold is parallelizable, i.e. the tangent bundle $T\mathcal{O}(M) \to \mathcal{O}(M)$ is a trivial bundle.*

Proof. Let $P = \mathcal{O}(M)$. Choose a G-connection in P with $G = \mathrm{O}(n)$ and split $TP = V \oplus H$. A canonical trivialisation for TP is given as follows: the vertical subbundle V is trivialised by the standard vertical vector fields $\tilde{A}$ to A, where A passes through a basis for $\mathfrak{g}$. The horizontal subbundle H is trivialised by the **standard horizontal vector fields** $L_1, ..., L_n$ in $\Gamma(TP)$, defined by $L_i(u) := h_u(ue_i)$. Then, for every $u \in P$,

$$\left(\tilde{A}(u), L_i(u) : A \in \text{basis for } \mathfrak{g},\ i = 1, ..., n\right)$$

is a basis for $T_uP = V_u \oplus H_u$, since $\mathfrak{g} \xrightarrow{\sim} V_u$, $A \mapsto \tilde{A}(u)$ and $h_u : T_{\pi(u)}M \xrightarrow{\sim} H_u$ are isomorphisms. ■

The standard vertical and horizontal vector fields are obviously given by

$$\vartheta(\tilde{A}) = 0 \text{ and } \vartheta(L_i) = e_i \text{ or } \tilde{\omega}(\tilde{A}) = 0 \text{ and } \tilde{\omega}(L_i) = 0.$$

3.41 Definition. The second order differential operator

$$\Delta_{\mathcal{O}(M)} := \sum_{i=1}^{n} L_i^2$$

is called **horizontal Laplacian** on $\mathcal{O}(M)$.

3.42 Remark (Notation). Let $\pi : P \to M$ be a principal G-bundle over M with a G-connection. For $X \in \Gamma(TM)$ we denote by $\overline{X} \in \Gamma(TP)$ the corresponding horizontal lifts with $\overline{X}_u = h_u(X_{\pi(u)})$ for $u \in P$.

For $g \in G = \mathrm{O}(n)$, a function $f : P \to \mathbb{R}^n$ is **equivariant** if $f(u.g) = g^{-1} f(u)$. The key idea for the next Theorem will be the relation

$$\vartheta_u(\overline{Y}_u) \equiv u^{-1} Y_{\pi(u)} =: f(u) \tag{3.31}$$

for $Y \in \Gamma(TM)$. Unrevealing the definitions of the canonical 1-form in (3.30) this is equivalent to

$$u^{-1} Y_{\pi(u)} = u^{-1} \left(\mathrm{d}\pi(h_u(Y_{\pi(u)}))\right),$$

so (3.31) follows by diagram (3.27) and $\mathrm{d}\pi \circ h = \mathrm{id}$.

3.43 Theorem ([HT94, Satz 7.132]). *Let M be a Riemannian manifold and $P = \mathcal{O}(M)$ with a G-connection in P. Then, for all vector fields $X, Y \in \Gamma(TM)$, the covariant derivative $\nabla_X Y \in \Gamma(TM)$ with respect to the linear connection on TM is given by*

$$(\nabla_X Y)_x = u\left(\overline{X}_u \vartheta(\overline{Y})\right) \qquad \text{for some } u \in P \text{ with } \pi(u) = x.$$

Equivalently, using the notation from above,

$$(\nabla_Y X)_x = u(\overline{X}_u f). \tag{3.32}$$

Proof. We choose a curve γ on M with $\gamma(0) = x$, $\dot\gamma(0) = X_x$. Then, by (2.7),

$$(\nabla_X Y)_x = \lim_{\varepsilon \to 0+} \frac{//_{\varepsilon,0} Y_{\gamma(\varepsilon)} - Y_{\gamma(0)}}{\varepsilon}, \tag{3.33}$$

where $//_{\varepsilon,0}$ is the parallel displacement from $T_{\gamma(\varepsilon)}M$ to $T_{\gamma(0)}M$ along γ. Since we can interpret $u \in \mathcal{O}(M)$ as isometry $u : \mathbb{R}^n \xrightarrow{\sim} T_{\pi(u)}M$ and let $t \mapsto u(t)$ a horizontal lift of $t \mapsto \gamma(t)$ on P with $u(0) = u$, then by Corollary 3.39

$$//_{0,\varepsilon} : T_{\gamma(0)}M \xrightarrow{\sim} T_{\gamma(\varepsilon)}M, \quad v \mapsto u(\varepsilon)u(0)^{-1}v,$$

and thus $//_{\varepsilon,0} \equiv u(0)u(\varepsilon)^{-1}$. Then

$$\begin{aligned}
(\nabla_X Y)_{\pi(u)} &= \lim_{\varepsilon \to 0+} \frac{u(0)u(\varepsilon)^{-1} Y_{\pi(u(\varepsilon))} - Y_{\pi(u(0))}}{\varepsilon} \\
&= \lim_{\varepsilon \to 0+} \frac{u\left(u(\varepsilon)^{-1} Y_{\pi(u(\varepsilon))} - u(0)^{-1} Y_{\pi(u(0))}\right)}{\varepsilon} \\
&= u\left(\lim_{\varepsilon \to 0+} \frac{f(u(\varepsilon)) - f(u(0))}{\varepsilon}\right) = u\left(\overline{X}_u(f)\right),
\end{aligned}$$

since $\dot u(t) = h_{u(t)}\dot\gamma(t)$ implies $\dot u(0) = h_u(X_u) = \overline{X}_u$. ∎

Analogously to equation (3.32) we can describe covariant derivatives of differential forms. The next theorem is proved in the same way as Theorem 3.43.

3.44 Theorem ([HT94, Satz 7.133]). *Let M be a Riemannian manifold and $P = \mathcal{O}(M)$ with a G-connection in P. Let $\omega \in \Omega^1(M)$ a form on M, which we write as an equivariant function $F : P \to \mathbb{R}^n$,[3] namely $F = (F^1, \ldots, F^n)$ with $F^i(u) := \omega_{\pi(u)}(ue_i)$, $i = 1, \ldots, n$. Then, for every $X \in \Gamma(TM)$ with horizontal lift $\overline{X} \in \Gamma(TP)$ and $u \in P$,*

$$(\nabla_X \omega)_{\pi(u)}(ue_i) = \overline{X}_u F^i, \quad i = 1, \ldots, n. \tag{3.34}$$

3.45 Definition. Let P be a principal G-bundle over a manifold M and $\tilde\omega \in \Gamma(T^*P \otimes \mathfrak{g})$ a connection form of a G-connection in P. For every P-valued semimartingale U the

[3] An equivariant function $F : P \to \mathbb{R}^n$, i.e. $F(u.g) = g^* F(u)$ with g^* dual linear map to $g \in \mathrm{O}(n)$.

Stratonovich integral $\int_U \tilde{\omega}$ provides a semimartingale with values in the Lie algebra $\mathfrak{g}$, namely component-by-component with respect for a basis $\mathfrak{g}$ as $\int_U \tilde{\omega} := \left(\int_U \tilde{\omega}^1, \ldots, \int_U \tilde{\omega}^n\right)$ if $\tilde{\omega} = \left(\tilde{\omega}^1, \ldots, \tilde{\omega}^n\right)$. We call U **horizontal** if $\int_U \tilde{\omega} = 0$ a.s. If X is an M-valued semimartingale, then we call a P-valued semimartingale U **horizontal lift of** X if U is horizontal and $\pi \circ U = X$ a.s.

Obviously, the concept of horizontal lifts of semimartingales generalises the concept of horizontal lifts of M-valued smooth curves (cf. Theorem 3.38) according to which a curve $t \mapsto u(t)$ above $t \mapsto x(t)$ is horizontal, i.e. $\pi \circ u = x$ and $\tilde{\omega}(\dot{u}) = 0$ (cf. Example 3.17). We prove the existence below.

For the rest of this chapter we concentrate on Riemannian n-manifolds M with the Levi-Civita connection. Recall that the corresponding splitting is given by $TP = V \oplus H$ where $V_u = \ker \vartheta_u$ and $H_u = \ker \tilde{\omega}_u$ for $u \in P$.

3.46 Definition. Let X be an M-valued semimartingale and U its horizontal lift with values in $P = \mathcal{O}(M)$. The $\mathbb{R}^n$-valued semimartingale

$$Z = \int_U \vartheta = \int \vartheta(\diamond \mathrm{d}U)$$

is the $\mathbb{R}^n$**-representation of** X (with initial basis U_0). In particular, with respect to the standard basis $\mathbb{R}^n$ we get $Z = (Z^1, \ldots, Z^n)$ with $Z^i = \int_U \vartheta^i$.

3.47 Theorem ([HT94, Satz 7.137]). *Let X be an M-valued semimartingale, U its horizontal lift on $P = \mathcal{O}(M)$ and Z an $\mathbb{R}^n$-representation of X.*

(i) $\int_U \sigma = \int \sigma(U) L_i(U) \diamond \mathrm{d}Z^i$ *for every differential form* $\sigma \in \Gamma(T^*P)$.

(ii) $\int_X \omega = \int \omega(X) U e_i \diamond \mathrm{d}Z^i$ *for every differential form* $\omega \in \Gamma(T^*M)$.

In particular, we get $\mathrm{d}f(U) = L_i f(U) \diamond \mathrm{d}Z^i$ *for every* $f \in C^\infty(P)$ *and we write*

$$\mathrm{d}U = L_i(U) \diamond \mathrm{d}Z^i. \tag{3.35}$$

Moreover, we get $\mathrm{d}f(X) = (U e_i)(f) \diamond \mathrm{d}Z^i$*, for every* $f \in C^\infty(M)$ *and we write*

$$\mathrm{d}X = U \diamond \mathrm{d}Z. \tag{3.36}$$

Proof. Equation (3.35) and (3.36) are a direct implication of (i) and (ii) if the set $\sigma = \mathbf{d}f$ for $f \in C^\infty(P)$ and $\omega = \mathbf{d}f$ for $f \in C^\infty(M)$.

(i) By Theorem 3.15, it is sufficient to show the righthand side satisfies the defining properties of $\int_U \sigma$. For every $f \in C^\infty(P)$ we have to show

$$\mathrm{d}f(U) = \mathbf{d}f(U) L_i(U) \diamond \mathrm{d}Z^i = (L_i f)(U) \diamond \mathrm{d}Z^i$$

or equivalently,

$$f(U) - f(U_0) = \int_U \sigma, \quad \text{with } \sigma \in \Gamma(T^*P), \sigma_u := L_i f(u)\vartheta^i_u. \tag{3.37}$$

For every $A \in T_u P$, we have

$$\begin{aligned} L_i f(u)\vartheta^i_u(A) &= (\mathbf{d}f)_u L_i(u)\vartheta^i_u(A) \\ &\overset{\text{Def}}{\underset{L_i,\vartheta_u}{=}} (\mathbf{d}f)_u h_u(ue_i)\left(u^{-1}(\mathbf{d}\pi)_u A\right)^i \\ &= (\mathbf{d}f)_u h_u\left(uu^{-1}(\mathbf{d}\pi)_u A\right) \\ &= (\mathbf{d}f)_u h_u((\mathbf{d}\pi)_u A) = \left((\mathbf{d}f)_u \circ \mathrm{pr}_{H_u}\right)(A). \end{aligned}$$

Thus, $L_i f(u)\vartheta^i_u = (\mathbf{d}f)_u \circ \mathrm{pr}_{H_u}$. However, using the Definitions 3.35 and 3.36, we get

$$(\mathbf{d}f \circ \mathrm{pr}_V)_u = (\mathbf{d}f)_u \kappa_u \tilde{\omega}_u = (\mathbf{d}f \circ I_u)_e \tilde{\omega}_u.$$

Since U is horizontal by assumption, we have $\int_U \mathbf{d}f \circ \mathrm{pr}_V = 0$ and therefore

$$f(U) - f(U_0) = \int_U \mathbf{d}f = \int_U \mathbf{d}f \circ \mathrm{pr}_H + \int_U \mathbf{d}f \circ \mathrm{pr}_V = \int_U \sigma$$

and (3.37) follows. The second defining property of the Stratonovich integral is evident.

(ii) It is sufficient to show that, for every $f \in C^\infty(M)$,

$$\mathrm{d}f(X) = (\mathbf{d}f)(X)Ue_i \diamond \mathrm{d}Z^i = Ue_i(f) \diamond \mathrm{d}Z^i.$$

Using (i) and the relation $(\mathbf{d}\pi)_u L_i(u) = ue_i$,

$$\begin{aligned} \mathrm{d}f(\pi \circ U) &= \mathbf{d}(f \circ \pi)(U)L_i(U) \diamond \mathrm{d}Z^i \\ &= (\mathbf{d}f) \circ \pi(U)\mathbf{d}\pi(U)L_i(U) \diamond \mathrm{d}Z^i = (\mathbf{d}f)(X)Ue_i \diamond \mathrm{d}Z^i, \end{aligned}$$

and the result follows. ■

3.48 Theorem ([HT94, Satz 7.138]). *Let X be an M-valued semimartingale, U its horizontal lift on $P = \mathcal{O}(M)$ and X an $\mathbb{R}^n$-representation of X.*

(i) $\int a(\mathrm{d}U, \mathrm{d}U) = \int a(U)(L_i(U), L_j(U))\mathrm{d}[Z^i, Z^j]$ *for every* $a \in \Gamma(T^*P \otimes T^*P)$.

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

The next, fundamental theorem shows the existence of horizontal lifts to M-valued semimartingales.

3.49 Theorem ([HT94, Satz 7.141]). *Let P be a principal G-bundle over a manifold M with G-connection. Let x_0 be an M-valued random variable and u_0 a P-valued random variable above x_0, i.e. $\pi \circ u_0 = x_0$ a.s. Then, for every M-valued semimartingale X with $X_0 = x_0$ there is a unique horizontal lift U on P with $U_0 = u_0$ a.s.*

Sketch of the Proof. By Theorem 3.10, every semimartingale can be written as solution of a Stratonovich SDE of the form

$$\mathrm{d}X = A_i(X) \diamond \mathrm{d}Z^i, \quad X_0 = x_0, \tag{3.38}$$

where Z is an $\mathbb{R}^n$-valued semimartingale. If $\overline{A}_i \in \Gamma(TP)$ is the horizontal lift of $A_i \in \Gamma(TM)$, i.e. $\overline{A}_i(u) = h_u(A_i(\pi u))$ for $u \in P$, then a good candidate of the desired horizontal lift is given by the *horizontal lifted SDE* on P

$$\mathrm{d}U = \overline{A}_i(U) \diamond \mathrm{d}Z^i, \quad U_0 = u_0. \tag{3.39}$$

We have $\mathbf{d}\pi(U) = (\mathbf{d}\pi)_U \overline{A}_i(U) \diamond \mathrm{d}Z^i = A_i(\pi \circ U) \diamond \mathrm{d}Z^i$ with $\pi \circ U_0 = u_0$. As (3.38) has a unique solution, we get $\pi \circ U = X$. However, $\int_U \tilde{\omega} = \int \tilde{\omega}(U)\overline{A}_i(U) \diamond \mathrm{d}Z^i = 0$.

It remains to show uniqueness and that U has the same lifetime as X. ■

Given a Riemannian manifold M with Levi-Civita connection. Above M we have the orthonormal frame bundle $P = \mathscr{O}(M)$ with induced $\mathrm{O}(n)$-connection. Then, we have shown so far:

(1) Let u_0 be a P-valued random variable and $x_0 = \pi \circ u_0$. If X is an M-valued semimartingale with $X_0 = x_0$, then there is a unique P-valued horizontal lift U of X with $U_0 = u_0$ (cf. Theorem 3.49). By Definition 3.46, this defines an $\mathbb{R}^n$-representation Z of X (with initial frame u_0) as $Z = \int_U \vartheta$. Then, every of the processes X, U, Z is determined by the other two as follows.

   (a) Z determines U as solution of the SDE $\mathrm{d}U = L_i(U) \diamond \mathrm{d}Z^i$ with $U_0 = u_0$,

   (b) U determines X by $X = \pi \circ U$,

   (c) X determines Z as $Z = \int_U \vartheta$, where U is the uniquely determined horizontal lift of X with $U_0 = u_0$.

   Mind that this procedure depends only trivially on the choice of u_0 above x_0. If we choose an $\mathscr{F}_0$-measurable P-valued random variable $\tilde{u}_0$ with $\pi \circ \tilde{u}_0 = x_0$ a.s. instead of u_0, then $u_0 = \tilde{u}_0.g_0$ for any $\mathscr{F}_0$-measurable random variable g_0 with values in the Lie group G of the orthogonal $n \times n$-matrices and U changes to $\tilde{U} = U.g_0$. Since $R_g^* \vartheta = g^{-1}\vartheta$, for $g \in G$, Z transforms into

$$\tilde{Z} = \int_{\tilde{U}} \vartheta = \int_U R_{g_0} \vartheta = \int_U g_0^{-1} \vartheta = g_0^{-1} Z. \tag{3.40}$$

(2) Usually, one starts vice versa. Choose a continuous $\mathbb{R}^n$-valued semimartingale Z with $Z_0 = 0$ and fix an $\mathscr{F}_0$-measurable random variable $u_0 : \Omega \to P$ as initial value. Define

U on P as the maximal solution to

$$\mathrm{d}U = L_i(U) \diamond \mathrm{d}Z^i, \quad U_0 = u_0,$$

and set $X := \pi \circ U$ as the projection from U on M with initial value $X_0 = \pi \circ u_0$. Thus, we obtain a horizontal process U over X with $U_0 = u_0$. In particular,

$$\mathrm{d}X = Ue_i \diamond \mathrm{d}Z^i = U \diamond \mathrm{d}Z.$$

Since $L_i(U) \diamond \mathrm{d}Z^i = h_U(Ue_i) \diamond \mathrm{d}Z^i$,

$$\mathrm{d}U = h_U(\diamond \mathrm{d}X).$$

Hence, we regain the original process up to the lifetime of U by the $\mathbb{R}^n$-representation $Z = \int_U \vartheta$ of X. We say X is the **stochastic development** of Z.

Fix $u \in P$, read as an isometry $u : \mathbb{R}^n \xrightarrow{\sim} T_xM$ with $x = \pi(u)$, then Z is equivalent to a T_xM-valued semimartingale $\tilde{Z} = uZ$. Stochastic development provides a one-to-one identification between the continuous semimartingales $\tilde{Z} := \tilde{Z}^1(ue_1) + \cdots + \tilde{Z}^n(ue_n)$ with $\tilde{Z}_0 = 0$ on the tangent space T_xM and the semimartingales X on the manifold M with $X_0 = x$, where $\tilde{Z} \mapsto X = \pi \circ U$ and U is the solution of the SDE

$$\mathrm{d}U = L_i(U)u^{-1} \diamond \mathrm{d}\tilde{Z}^i, \quad U_0 = u.$$

3.50 Definition. Let M be a Riemannian manifold equipped with the Levi-Civita connection. Let X be a semimartingale on M and U a horizontal lift of X on $\mathcal{O}(M)$. For every $0 \leqslant s \leqslant t$, we define $//_{s,t} := U_t \circ U_s^{-1}$ by

$$\begin{array}{ccc} T_{X_s}M & \xrightarrow{\quad\sim\quad} & T_{X_t}M \\ {\scriptstyle U_s}\nwarrow & & \nearrow{\scriptstyle U_t} \\ & \mathbb{R}^n & \end{array}$$

and $//_{t,s} := //_{s,t}^{-1}$. We call the isometries $//_t := //_{0,t} : T_{X_0}M \to T_{X_t}M$ **stochastic parallel displacement along** X.

The stochastic parallel displacement $//_t$ extends naturally from TM to $T^{(r,s)}M$. For $\omega \in \Omega^1(M)$ and $A \in \Gamma(TM)$ we have $(//_t\omega_{X_0})(A_{X_t}) = \omega_{X_0}(//^{-1}A_{X_t})$.

Finally, we want to illustrate the procedure of stochastic development geometrically. Let $P = \mathcal{O}(M)$ above a Riemannian manifold M. Fix $u \in \mathcal{O}(M)$, as always read as an isometry $u : \mathbb{R}^n \xrightarrow{\sim} T_{\pi(u)}M$, and set $x := \pi(u)$.

We may think of X as the trace which constitutes by the contact (or touch) of M with T_XM on M, more precisely, by the contact of the paths of Z by the virtue of the identification $U^{-1} : T_XM \xrightarrow{\sim} \mathbb{R}^n$ on M if we roll M along $t \mapsto Z_t$ ("rolling without slipping"). Mind that, in general, the trajectories of Z are not differentiable so that a path-wise interpretation makes no sense.

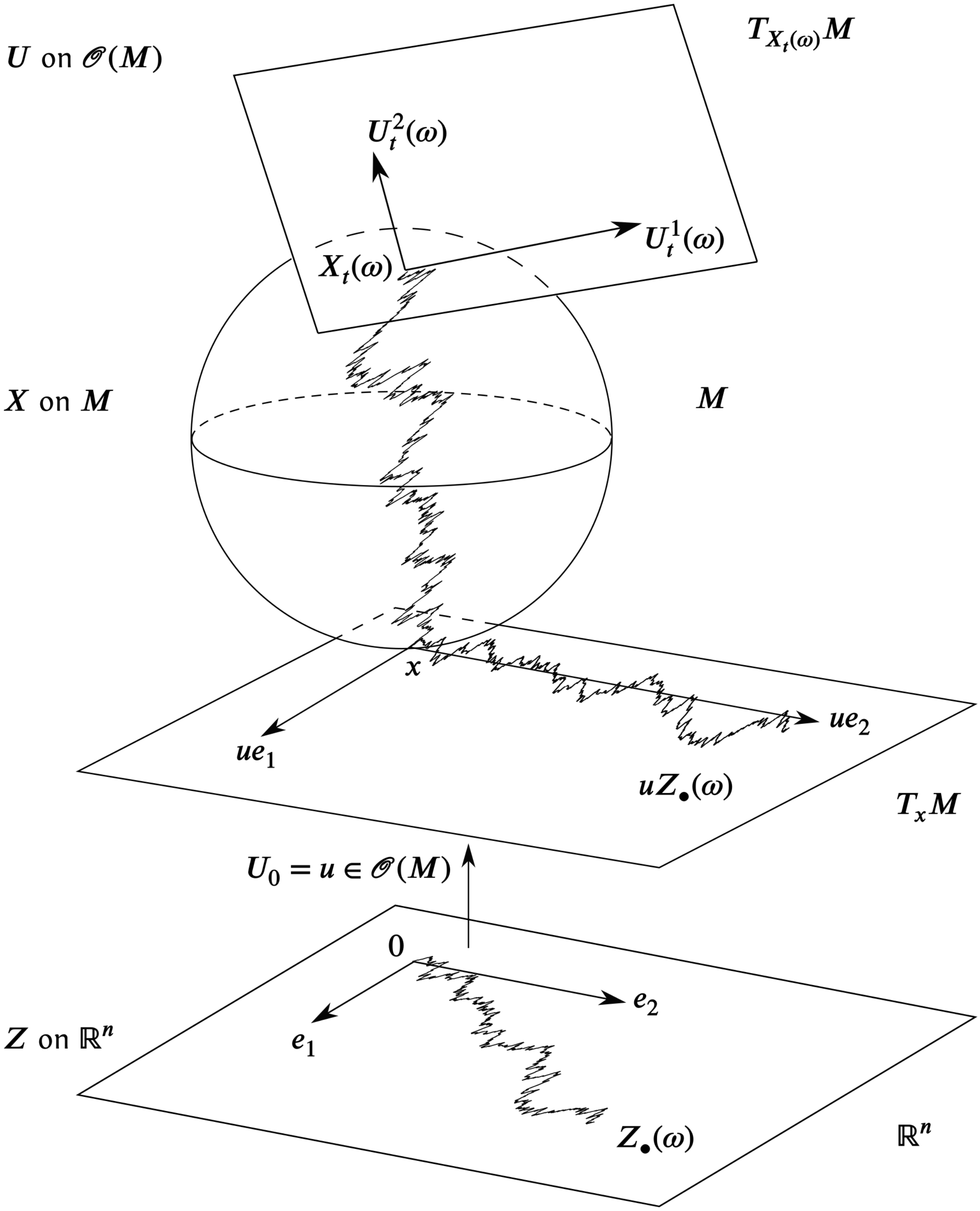

Figure 3.3: Stochastic Development

The next technical lemma will be used in the proof of Theorem 3.54.

3.51 Lemma ([HT94, Satz 7.142]). *Let $(V_k)_{k\in\mathbb{N}}$ be a countable open cover of M and X a continuous adapted M-valued process. Then there exists an increasing sequence $(\tau_n)_{n\in\mathbb{N}_0}$ of stopping times with $\tau_0 = 0$ and $\sup_n \tau_n = \infty$ such that the process X takes values in every of the* *intervals*

$$[\tau_n, \tau_{n+1}] \cap \left([0,\infty) \times \left\{\tau_n < \tau_{n+1}\right\}\right).$$

Proof. We choose a refinement $(W_k)_{k\in\mathbb{N}}$ of $(V_k)_{k\in\mathbb{N}}$ such that for every $k \in \mathbb{N}$ holds $\overline{W}_k \subsetneqq V_{n(k)}$. We construct a sequence $(\tau_n^k)_{0\leqslant k\leqslant n}$ of stopping times that satisfies (maybe new numbered) the assumption. Let $\tau_0^0 := 0$. If τ_n^k is constructed up to n, then set

$$\tau_{n+1}^0 := \tau_n^n, \text{ and } \tau_{n+1}^k := \inf\left\{t \geqslant \tau_{n+1}^{k-1} \,:\, X_t \notin W_k\right\} \text{ for } k = 1, ..., n+1.$$

It remains to show that $\sup_{n\geqslant 0}\sup_{k\leqslant n}\tau_n^k = \infty$. If $\omega \in \Omega$ with $t_0 := \sup_{n\geqslant 0}\sup_{k\leqslant n}\tau_n^k < \infty$, then $X_{t_0}(\omega) \in W_l$ for some l and by continuity $X_t(\omega) \in W_l$ for every $t \in [t_0-\varepsilon, t_0+\varepsilon]$ for $\varepsilon > 0$ small enough. By definition of t_0 there is an $n_0 \in \mathbb{N}, n_0 \geqslant l$ with $\tau_{n_0}^0(\omega) > t_0 - \varepsilon$. But from above we should have $\tau_{n_0}^l(\omega) \geqslant t_0 + \varepsilon$, which is a contradiction. ■

3.52 Theorem (Geometric Itô formula, [HT94, Satz 7.145]). *Let M be a Riemannian manifold equipped with the Levi-Civita connection ∇. Let X be an M-valued semimartingale, U its horizontal lift on $\mathcal{O}(M)$ and $Z = \int_U \vartheta$ the $\mathbb{R}^n$-representation of X. Then, for every $f \in C^\infty(M)$,*

$$\mathrm{d}f(X) = \mathbf{d}f(X)(Ue_i)\mathrm{d}Z^i + \frac{1}{2}\nabla\mathbf{d}f(X)(Ue_i, Ue_i)\mathrm{d}[Z^i, Z^j]. \tag{3.41}$$

We write

$$\mathrm{d}f(X) = \mathbf{d}f(U\,\mathrm{d}Z) + \frac{1}{2}\nabla\mathbf{d}f(\mathrm{d}X, \mathrm{d}X). \tag{3.42}$$

Proof. From $\mathrm{d}U = L_i(U) \diamond \mathrm{d}Z^i$ follows that

$$\begin{aligned}\mathrm{d}f(X) &= \mathrm{d}f(\pi \circ U) = L_i(f \circ \pi)(U) \diamond \mathrm{d}Z^i \\ &= L_i(f\circ\pi)(U)\mathrm{d}Z^i + \frac{1}{2}L_iL_j(f\circ\pi)(U)\mathrm{d}[Z^i, Z^j].\end{aligned}$$

By the definition of the differential and the chain rule,

$$L_i(f\circ\pi)(u) = \mathbf{d}(f\circ\pi)_u L_i(u) = (\mathbf{d}f)_{\pi(u)}(\mathbf{d}\pi)_u h_u(ue_i) = (\mathbf{d}f)_{\pi(u)}(ue_i),$$

we have $L_i(f\circ\pi)(u) = F^i(u)$, where $F : \mathcal{O}(M) \to \mathbb{R}^n$ is the corresponding equivariant function to $\mathbf{d}f$ defined by $F^i(u) = (\mathbf{d}f)_{\pi(u)}(ue_i)$ (cf. Theorem 3.44). By (3.34),

$$L_iL_j(f\circ\pi)(u) = (L_iF^j)(u) = (h_u(ue_i))F^j \overset{(3.34)}{=} \left(\nabla_{ue_i}\mathbf{d}f\right)_{\pi(u)}(ue_j) = \nabla\mathbf{d}f(ue_i, ue_j)$$

and the result follows. ■

From the last line in the proof of Theorem 3.52 we receive the following relation.

3.53 Theorem. *Let M be a Riemannian manifold equipped with Levi-Civita connection. Then*

$$\Delta_{\mathcal{O}(M)}\pi^* = \pi^*\Delta_M. \tag{3.43}$$

Thus, the horizontal Laplace operator $\Delta_{\mathcal{O}(M)}$ is the lift of the Laplace-Beltrami operator Δ_M to the orthonormal frame bundle $\mathcal{O}(M)$. More precisely, this reads, for every $f \in C^\infty(M)$ and $\tilde{f} := f \circ \pi$ its lift to $\mathcal{O}(M)$,

$$\Delta_{\mathcal{O}(M)} \tilde{f}(u) = \Delta_M f(x) \quad \text{for any } u \in \mathcal{O}(M) \text{ with } x = \pi(u).$$

Proof. For $u \in \mathcal{O}(M)$, we have

$$\sum_{i=1}^n L_i^2 (f \circ \pi)(u) = \sum_{i=1}^n \nabla \mathbf{d} f(ue_i, ue_i) = (\operatorname{tr} \nabla \mathbf{d} f)\pi(u) = (\Delta_M f) \circ \pi(u). \qquad \blacksquare$$

The procedure of stochastic development of $\mathbb{R}^n$-valued semimartingales allows us to construct for every family of $\mathbb{R}^n$-valued semimartingales a corresponding M-valued semimartingale. Next, we show that local martingales on $\mathbb{R}^n$ correspond to ∇-martingales on M through stochastic development. In particular, on a Riemannian manifold correspond $\mathrm{BM}(\mathbb{R}^n)$ and $\mathrm{BM}(M, g)$ qua stochastic development.

3.54 Theorem ([HT94, Satz 7.147]). *Let M be a Riemannian manifold equipped with the Levi-Civita connection ∇. Let X be an M-valued semimartingale and U_0 an $\mathcal{O}(M)$-valued $\mathscr{F}_0$-measurable random variable with $\pi \circ U_0 = X_0$ a.s. and $Z = \int_U \vartheta$ the $\mathbb{R}^n$-representation of X with initial frame U_0.*

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

(ii) Since a Riemannian metric is given on M, U_0 takes values in $\mathcal{O}(M)$. Then X is a $\mathrm{BM}(M, g)$ if Z is a $\mathrm{BM}(\mathbb{R}^n)$. More precisely, an $\mathbb{R}^n$-valued Brownian motion stopped at ζ, where ζ is the lifetime of X.

Proof. (i) By Definition 3.21, X is a ∇-martingale if, for every $f \in C^\infty(M)$,

$$\mathrm{d} f(X) - \frac{1}{2} \nabla \mathbf{d} f(\mathrm{d}X, \mathrm{d}X) \overset{\mathrm{m}}{=} 0.$$

The Geometric Itô formula 3.52 implies that $\mathbf{d} f(X)(Ue_i) \diamond \mathrm{d}Z^i \overset{\mathrm{m}}{=} 0$ for every $f \in C^\infty(M)$. Using Lemma 3.51 this is equivalent to Z being a local martingale.

(ii) By Definition 3.25, X is a $\mathrm{BM}(M, g)$ if, for every $f \in C^\infty(M)$,

$$\mathrm{d} f(X) - \frac{1}{2} \Delta_M f(X) \mathrm{d}t \overset{\mathrm{m}}{=} 0.$$

If Z is a $\mathrm{BM}(\mathbb{R}^n)$, then X must be a $\mathrm{BM}(M, g)$ by (3.41). Conversely, if X is a $\mathrm{BM}(M, g)$, then by Lévy's characterisation of M-valued Browian motions 3.27 X is a ∇-martingale and Z is also a local martingale by part (i) of the proof. By definition, $Z^i = \int_U \vartheta^i$ with $\vartheta_u^i = \langle \mathbf{d}\pi(\cdot), ue_i \rangle = \pi^* \langle \cdot, ue_i \rangle$ such that the quadratic variation of Z can be calculated using (3.16) and (3.17):

$$\begin{aligned}\mathrm{d}[Z^i, Z^j] &= \left[\int_U \vartheta^i, \int_U \vartheta^j\right] = \left(\vartheta^i \otimes \vartheta^j\right)(\mathrm{d}U, \mathrm{d}U)\\ &= \pi^*\left(\langle\cdot, Ue_i\rangle \otimes \left\langle\cdot, Ue_j\right\rangle\right)(\mathrm{d}U, \mathrm{d}U)\\ &= \left(\langle\cdot, Ue_i\rangle \otimes \left\langle\cdot, Ue_j\right\rangle\right)(\mathrm{d}X, \mathrm{d}X)\\ &= \mathrm{tr}\left(\langle\cdot, Ue_i\rangle \otimes \left\langle\cdot, Ue_j\right\rangle\right)(X)\mathrm{d}t = \delta_{ij}\mathrm{d}t.\end{aligned}$$

Lévy's characterisation of $\mathrm{BM}(\mathbb{R}^n)$ yields that Z is a Brownian motion. ■

Theorem 3.54 provides a natural procedure to construct a Brownian motion on Riemannian manifolds. More precisely, we construct a $\mathrm{BM}(M, g)$ with initial point $x \in M$ as follows: Let B a n-dimensional Brownian motion on a standard filtered probability space. Choose $u \in \mathscr{O}(M)$ with $\pi(u) = x$ and solve the SDE

$$\mathrm{d}U = L_i(U) \diamond \mathrm{d}B^i, \quad U_0 = 0.$$

By Theorem 3.54, we get $X = \pi \circ U$ is a $\mathrm{BM}(M, g)$ with initial point $X_0 = x$. Mind that this construction depends on the choice of $u \in \pi^{-1}\{x\}$ and the choice of the driving Euclidean Brownian motion B. For every $g \in \mathscr{O}(M)$ with B, gB is a $\mathrm{BM}(\mathbb{R}^n)$. Furthermore, if X is constructed by ug and B is constructed by u and uB, then they are equal up to indistinguishability. i.e. different choices of $u \in \pi^{-1}\{x\}$ just causes a change in the underlying Brownian motion B. In particular, the distribution of X is independent of these choices.

Finally, we note that all results go through in the same way for general manifolds if we replace the bundle of all orthonormal frames $\mathscr{O}(M)$ by the bundle of all linear frames $\mathrm{GL}(M)$ with structure group $\mathrm{GL}(n, \mathbb{R})$ and require the linear connection ∇ to be torsion-free (cf. Definition 2.51 (a)).

Chapter 4

The Bismut-Elworthy-Li Formula

In this chapter, we prove the Bismut-Elworthy-Li formula on $\mathbb{R}^n$ and sketch Elworthy and Li's original proof. Our main sources are [EL94a], the original paper due to Elworthy and Li [EL94b] and a previous work of Elworthy [Elw92].

In 1984, Bismut proved a remarkable connection (cf. [Bis84, Theorem 2.14]) for a smooth, compact and connected Riemannian n-manifold M with $p_t(x, y)$ the (smooth) heat kernel with respect to the Laplace-Beltrami operator on M.

Bismut's Formula. *For every* $x_0, y_0 \in M$, $t > 0$,

$$\frac{\operatorname{grad}_{x_0} p_t(x_0, y_0)}{p_t(x_0, y_0)} = \frac{1}{t} \mathbb{E}^{\mathbb{P}^t_{x_0, y_0}} \int_0^t \tilde{E}'_s \delta \beta_s.$$

Herein $\mathbb{P}^t_{x,y}$ denotes the Brownian bridge measure, i.e. conditioned Brownian paths on M to start in a point x and run into y with time t. β is a Brownian motion and $\tilde{E}'$ a certain semimartingale both living on the tangent space $T_x M$ at x. Bismut's proof is based on Malliavin calculus and calculus of variations. In [Elw92], Elworthy used rather elementary martingale based arguments to prove the result. Soon after, Elworthy and Li provided a systematical treatment for a large class of diffusion processes in [EL94b]. We will illustrate their basic ideas in rather modern notation.

Recall that a (operator) **semigroup** $(P_t)_{t \geqslant 0}$ on a Banach space $(L, \|\cdot\|)$ is a family of linear operators $P_t : L \to L$ satisfying

$$P_{t+s} = P_t P_s, \quad P_0 = \mathrm{id}, \quad \text{for all } s, t \geqslant 0.$$

$(P_t)_{t \geqslant 0}$ is **(of class)** C_0 or a **strongly continuous semigroup** if $\lim_{t \to 0} \left\| P_t f - f \right\| = 0$ for all $f \in L$, and a **contraction** on $(L, \|\cdot\|)$ if $\left\| P_t f \right\| \leqslant \|f\|$ for all $f \in L$. $(P_t)_{t \geqslant 0}$ **preserves positivity**, if $P_t f \geqslant 0$, for $f \geqslant 0$, and is **minimal** if for any other positivity preserving contractive semigroup $(Q_t)_{t \geqslant 0}$, we have $P_t f \leqslant Q_t f$ a.e. for every $f \in L$.

4.1 Definition. The **(infinitesimal) generator A** of a C_0-contraction semigroup $(P_t)_{t \geqslant 0}$ on $(L, \|\cdot\|)$ is the operator

$$\mathbf{A} f := \lim_{t \to 0} \frac{P_t f - f}{t}, \tag{4.1}$$

where the domain $\mathscr{D}(\mathbf{A})$ of $\mathbf{A}$ is the family of all f where the limit (4.1) exists.

Note that $\mathbf{A}$ is densely defined and closed in L. Moreover, the generator uniquely determines the semigroup. By Taylor's formula, (4.1) can alternatively be written as

$$P_t f = f + t\mathbf{A} f + o(t),$$

so that the generator $\mathbf{A}$ gives the first-order approximation to P_t for small t. For some basic properties of the transition semigroup and the generator see e.g. [SP14, Chapter 7].

4.2 Definition. The **minimal heat semigroup** $P_t : \mathscr{B}_b(M) \to \mathscr{B}_b(M), t \geqslant 0$, is a minimal positivity preserving C_0-contraction semigroup with generator $\mathbf{A} = \frac{1}{2}\Delta_M + A_0$.

The minimal heat semigroup $(P_t)_{t\geqslant 0}$ admits a smooth heat kernel $p_t : M \times M \to \mathbb{R}$, $t > 0$, with respect to the Riemannian volume element $\mathrm{vol}_g = \mathrm{d}y$, i.e.

$$P_t f(x) = \int_M p_t(x, y) f(y)\, \mathrm{d}y, \quad f \in \mathscr{B}_b(M), \tag{4.2}$$

and $u(t, x) := P_t f(x)$ is the unique solution of the heat equation

$$\partial_t u = \frac{1}{2}\Delta_M u, \quad u(0, \cdot) = f.$$

For a concise introduction, especially on the existence of (4.2), we refer the reader to [Cha84, p. 187 ff.] and [Gri09, Chapter 7]. A more probabilistic approach can be found in [Hsu02, Chapter 4].

Let M be a smooth manifold. Consider the nondegenerate and nonexplosive SDE

$$\mathrm{d}X = A(X) \diamond \mathrm{d}B + A_0(X)\mathrm{d}t \tag{4.3}$$

on M, where $A_0 : \mathbb{R}^m \to T_x M$, $A \in \Gamma(TM)$ and B is an $\mathbb{R}^m$-valued Brownian motion on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F})_{t\geqslant 0}, \mathbb{P})$ and the filtration $(\mathscr{F}_t)_{t\geqslant 0}$ satisfies the usual hypotheses (cf. Chapter 3). The infinitesimal generator (second order elliptic operator) $\mathbf{A}$ can be written as $\mathbf{A} = \frac{1}{2}\Delta_M + A_0$, where Δ_M denotes the Laplace-Beltrami operator for the induced Riemannian metric on M and A_0 is a smooth vector field on M, cf. Corollary 3.6. The SDE is **nondegenerate**, if A is nondegenerate, i.e. $A(x) : \mathbb{R}^m \to T_x M$ is surjective for each $x \in M$. However by Proposition 3.29 (ii), there exists the (pointwise) adjoint $A^*(x)$ as a right inverse of $A(x)$ and $A(x)$ is the projection (which is surjective). We denote by BC^r the space of C^r-functions with their first r derivatives bounded (using the Riemannian metric on the manifold).

4.1 Bismut-Elworthy-Li Formula on $\mathbb{R}^n$

First, we consider the case $M = \mathbb{R}^n$. We can take the Itô form of (4.3),

$$\mathrm{d}X_t = A(X_t)\mathrm{d}B_t + A_0(X_t)\mathrm{d}t, \tag{4.4}$$

where $A : \mathbb{R}^n \to \mathbb{R}^{m\times n}$ and $A_0 : \mathbb{R}^n \to \mathbb{R}^n$ are smooth.

The proofs of Theorem 4.3 and Theorem 4.6 partly depend on the theory of stochastic flows. We only give the key ideas. Details can be found in [Elw92, I, §4] and [Elw82, pp. VII, VIII]. The problem is to find nice versions of the map $(t, x, \omega) \mapsto F_t(x, \omega) \equiv X_t(x, \omega) \in M$ ($\omega \in \Omega$) which assigns to $x \in M$ the solution of (4.4) starting in $X_0 = x$. By the usual Whitney embedding argument we can apply the Kolmogorov-Chentsov Theorem A.1 to show that there is a version of F such that:

(i) For all $x \in M$, $(F_t(x,\cdot))_{t\geqslant 0}$ solves (4.3) with $F_0(x,\cdot) = x$.

(ii) Each map $[0,\infty) \times M \to M$, $(t,x) \mapsto F_t(x,\cdot)$ is continuous.

The idea of stochastic flows was heavily advanced by Kunita, cf. [Kun97]. We have the following existence and uniqueness result, even for noncompact manifolds (cf. [Elw82, VIII, Theorem 2B]). Suppose A and A_0 are in C^r, for $r \geqslant 2$. Then there is a (partially) defined flow $(F_t(\cdot), \zeta(\cdot))$ such that, for each $x \in M$, $(F_t(x), \zeta(x))$ is a maximal solution to (4.4) with lifetime $\zeta(x)$ and if

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

then there is a set Ω_0 with probability 1 such that, for all $\omega \in \Omega_0$, the solution flow has the following properties:

(i) For each $t \geqslant 0$, $\zeta(\cdot,\omega)$ is lower semicontinuous on M and therefore $M_t(\omega)$ an open subset of M.

(ii) $X_t(\cdot,\omega)$ is a diffeomorphism from $M_t(\omega)$ onto an open subset of M.

(iii) For each $t \geqslant 0$ the map $s \mapsto X_s(\cdot,\omega)$ from $[0,t]$ into $C^\infty(M_t(\omega), M)$ (endowed with its C^∞ topology) is continuous.

Note that (4.4) is **complete**, i.e. $\zeta(x,\omega) = \infty$ a.s. for each $x \in M$ so that the set $\{(x,\omega) \in M \times \Omega : t < \zeta(x,\omega)\}$ has full Leb $\otimes\, \mathbb{P}$ measure. Moreover, there is a **derivative process** $v_t := DF_t(x_0)(v_0) := \mathbf{d}X_t\big|_x(v_0) \overset{\text{def}}{=} \nabla_{v_0}X_t(x)$ to $X_t(x)$ starting from v_0, i.e. a derivative (in probability) of F_t at x_0 in the direction of some given $v_0 \in \mathbb{R}^n$ (cf. [Elw82, VII, Theorem 8E]). It can be shown that v_t is given as the solution of the formally differentiated SDE (4.4) (up to a modification).

For any 1-form $\Theta : \mathbb{R}^n \to \mathrm{Lin}(\mathbb{R}^n;\mathbb{R})$ we define a heat semigroup[1]

$$\hat{P}_t(\Theta) : \mathbb{R}^n \to \mathrm{Lin}(\mathbb{R}^n;\mathbb{R})$$
$$\left(\hat{P}_t(\Theta)\right)_x(v_0) := \mathbb{E}\Theta_{X_t(x)}(v_t),$$

whenever the righthand side exists. Then

$$\mathbf{d}(P_tf)_x(v_0) = \left(\hat{P}_t(\mathbf{d}f)\right)_x(v_0). \tag{4.5}$$

Unrevealing the definitions, this is equivalent to

$$\nabla_{v_0}\mathbb{E}f(X_t(x)) = \mathbb{E}\nabla_{v_0}f(X_t(x)).$$

So the idea is that formal differentiation under the expectation holds true. Then we get the following result.

4.3 Theorem (Bismut-Elworthy-Li Formula on $\mathbb{R}^n$, [EL94b, Theorem 2.1]). *Let (4.4) be non-explosive, so there exists a right inverse map $A^*(x)$ to $A(x)$ for each $x \in \mathbb{R}^n$, smooth in x. Let*

[1] Revised version: Slight reformulation for better readability.

$f : \mathbb{R}^n \to \mathbb{R} \in \mathrm{BC}^1$ with $\hat{P}_t(\mathbf{d}f) = \mathbf{d}(P_t f)$ a.s. for $t \geqslant 0$. Then for almost all $X_0 = x \in \mathbb{R}^n$ and $t > 0$,

$$\mathbf{d}(P_t f)_x(v_0) = \frac{1}{t}\mathbb{E}\left(f(X_t(x)) \int_0^t \langle \mathbf{d}X_s(x)(A^* v_s), \mathrm{d}B_s \rangle_{\mathbb{R}^m} \right), \tag{4.6}$$

$v_0 \in \mathbb{R}^n$, provided $\int_0^t \langle \mathbf{d}X_s(A^* v_s), \mathrm{d}B_s \rangle_{\mathbb{R}^m}$, $t \geqslant 0$, is a martingale.

Proof. Let $T > 0$. Recall that for a C^2-function $f = f(t,x) : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}$ Itô's formula is of the form

$$f(t, X_t) - f(s, X_s) = \int_s^t \nabla_x f(r, X_r)\mathrm{d}B_r + \int_s^t \left(\frac{1}{2}\Delta_x f(r, B_r) + \frac{\partial}{\partial t} f(r, B_r) \right) \mathrm{d}r, \tag{4.7}$$

where Δ_x denotes the Euclidean Laplacian with respect to x. Applying Itô's formula to $(t,x) \mapsto P_{T-t}f(x)$, $0 \leqslant t \leqslant T$, we get

$$P_{T-t}f(X_t) = P_T f(x) + \int_0^t \mathbf{d}(P_{T-s}f)_{X_s}(A(X_s)\mathrm{d}B_s),$$

for $0 \leqslant t < T$. The parabolic regularity and specific form of the semigroup ensures that the last term[2] in (4.7) vanishes. Taking the limit as $t \to T$, we have

$$f(X_T) = P_T f(x) + \int_0^T \mathbf{d}(P_{T-s}f)_{X_s}(A(X_s)\mathrm{d}B_s). \tag{4.8}$$

Multiplying (4.8) through by our martingale and then taking expectations using the fact that f is bounded and $A(x)A^*(x) = \mathrm{id}$, we obtain

$$\begin{aligned}
\mathbb{E}\left[f(X_T) \int_0^T \langle \mathbf{d}X_s(A^* v_s), \mathrm{d}B_s \rangle \right] &\overset{\text{A.4}}{=} \mathbb{E} \int_0^T \mathbf{d}(P_{T-s}f)_{X_s}(v_s)\mathrm{d}s \\
&\overset{(4.5)}{=} \mathbb{E} \int_0^T \left(\hat{P}_{T-s}(\mathbf{d}f) \right)_{X_s}(v_s)\mathrm{d}s \\
&\overset{(*)}{=} \int_0^T \left(\hat{P}_s \hat{P}_{T-s}(\mathbf{d}f) \right)_x (v_0)\mathrm{d}s \\
&= \int_0^T \left(\hat{P}_T(\mathbf{d}f) \right)_x (v_0)\mathrm{d}s = T \cdot \hat{P}_T(\mathbf{d}f)_x(v_0),
\end{aligned}$$

using the semigroup property of $\hat{P}_t$. Choosing $\Theta = \hat{P}_{T-s}(\mathbf{d}f)$ in (4.5) justifies $(*)$. ■

4.4 Remark. The proof is based on the equality (4.5). This can always be assured provided the map $x \mapsto \mathbb{E}\left|\nabla_{.} X_t(x)\right|$ is continuous (cf. [SP14, 15 and 16]). The same holds for the generalisation (4.11) in the next section.

[2]The heat equation makes everything work: The inversion in the time index of the semigroup forces us to switch from the heat equation $\partial_t - \mathbf{A} = 0$ to the space-time-harmonic context $\partial_t + \mathbf{A} = 0$.

4.5 Remark (Notation). By Example 2.33, the Euclidean metric $\overline{g} = \langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ on $\mathbb{R}^n$ is just the scalar product. Using that

$$\mathbf{d}(P_t f)_x(v_0) \overset{(2.5)}{\equiv} \langle \nabla P_t f(x), v_0 \rangle_{\mathbb{R}^m} = \nabla_{v_0} P_t f(x) \underset{P_t}{\overset{\text{def}}{=}} \nabla_{v_0} \mathbb{E} f(X_t(x)),$$

we can rewrite (4.6) as

$$\nabla_{v_0} \mathbb{E} f(X_t(x)) = \frac{1}{t} \mathbb{E} \left[f(X_t(x)) \int_0^t A^* \nabla_{v_s} X_s(x) \, \mathrm{d}B_s \right].$$

4.2 Bismut-Elworthy-Li Formula on M

The proof of the manifold case (cf. Theorem 4.6 below) requires also good knowledge of covariant derivative formulas for stochastic flows. We only give the idea.

We consider the Stratonovich SDE (4.3) for a general Riemannian manifold M. We continue to assume nonexplosion. By Proposition 3.29 (ii), $A(x) : \mathbb{R}^m \to T_x M$ is surjective and let A_0 be a smooth vector field on M. Write $A_i(x) := A(x) e_i$ for an orthonormal basis $(e_1, \dots, e_m)$ for $\mathbb{R}^m$. Thus, (4.3) becomes

$$\mathrm{d}X = A_i(X_t) \diamond \mathrm{d}B^i + A_0(X_t) \mathrm{d}t. \tag{4.9}$$

Here $(B_t^i)_{t \geqslant 0}$ are independent $\mathrm{BM}(\mathbb{R}^1)$ for $i = 1, \dots, m$. Recall that, using the Levi-Civita connection on M (cf. Remark 3.24), we have

$$A_0(x) = -\frac{1}{2} \sum_{i=1}^m \nabla_{A_i} A_i(x).$$

The formally differentiated version of SDE (4.9) can be expressed as a covariant derivative equation

$$DV_t = \nabla_{V_t} A \diamond \mathrm{d}B_t + \nabla_{V_t} A_0 \mathrm{d}t, \tag{4.10}$$

where

$$DV_t := /\!/_s \mathrm{d} \left(/\!/_s^{-1} V_s(x) \right),$$

is the covariant Itô differential of the stochastic parallel transport $/\!/_s : T_{X_0} M \to T_{X_s} M$ along the paths of $(X_t)_{t \geqslant 0}$ (cf. Definition 3.50). Now, $V_t = \mathbf{d}X_t|_{x_0} V_0 = \nabla_{V_0} X_t(x_0)$, the derivative at x_0 in the direction V_0, for $V_0 \in T_{X_0} M$ defined as a map $TM \to TM$.

In this situation the previous proofs go through in the same way, using covariant differentiation to replace the usual differentiation and the adjoint $X(x)^* : T_x M \to \mathbb{R}^m$ as right inverse for $X(x)$.

4.6 Theorem (Bismut-Elworthy-Li Formula on M, [EL94b, Theorem 3.1]). *Let M be a complete Riemannian manifold and $\mathsf{A} = \frac{1}{2}\Delta_M + A_0$. Assume (4.9) is complete. Let $f : M \to \mathbb{R}$ in BC^1 with $\hat{P}_t(\mathbf{d}f) = \mathbf{d}(P_t f)$ a.e. for $t \geqslant 0$. Then for almost all $X_0 \in M$,*

$$\mathbf{d}P_t f(V_0) = \frac{1}{t} \mathbb{E} \left(f(X_t) \int_0^t \langle V_s, A(X_s) \mathrm{d}B_s \rangle_{X_s} \right), \quad V_0 \in T_{X_0} M, \tag{4.11}$$

provided $\int_0^t \langle V_s, A(X_s) \mathrm{d}B_s \rangle_{X_s}$ is a martingale.

Finally, we get the following Bismut-type formula as an easy consequence of Theorem 4.6.

4.7 Corollary ([EL94b, Corollary 3.2]). *Suppose that* $\hat{P}_t(\mathbf{d}f) = \mathbf{d}(P_t f)$ *for all* $t > 0$, $f \in C_c^\infty(M)$ *and* $p_t : M \times M \to \mathbb{R}$, $t > 0$, *be the heat kernel (with respect to the Riemannian volume form). Then, for* $t > 0$,

$$\nabla \log p_t(X_0, y) = \frac{1}{t}\mathbb{E}\left(\int_0^t (\mathbf{d}X_s)^* A(X_s)\mathrm{d}B_s \,\middle|\, X_t = y\right) \tag{4.12}$$

for almost all $y \in M$, *provided* $\int_0^t \langle V_s, A(X_s)\mathrm{d}B_s\rangle$ *is a martingale.*

Note that in Corollary 4.7 $(\mathbf{d}X_s)^*$ denotes the *adjoint* of $\mathbf{d}X_s$, not the pullback.

Proof. Let $f \in C_c^\infty(M)$. By the smoothness of p_t for $t > 0$, we can differentiate (4.2) to obtain

$$\mathbf{d}(P_t f)(V_0) = \int_M \langle \nabla p_t(\cdot, y), V_0\rangle_{X_0} f(y)\,\mathrm{d}y.$$

On the other hand, we may rewrite (4.11) as

$$\mathbf{d}(P_t f)(V_0) = \int_M p_t(x_0, y) f(y)\,\mathbb{E}\left(\frac{1}{t}\int_0^t \langle \mathbf{d}X_s(V_0), A(X_s)\,\mathrm{d}B_s\rangle \,\middle|\, X_t = y\right)\mathrm{d}y.$$

Comparing the last two equations, we get

$$\nabla p_t(\cdot, y)(X_0) = p_t(X_0, y_0)\mathbb{E}\left(\frac{1}{t}\int_0^t (\mathbf{d}X_s)^*(A\,\mathrm{d}B_s) \,\middle|\, X_t = y\right),$$

since $\langle \mathbf{d}X_s(V_0), A(X_s)\mathrm{d}B_s\rangle = \langle \mathbf{d}X_s^*(A\,\mathrm{d}B_s), V_0\rangle$. ■

Chapter 5

Lévy Processes and Subordination

Let us recall some basic definitions and properties of Lévy processes. We follow essentially [App09] and [Sat13]. Furthermore, we introduce the concept of subordination. To this end we will use some well-known results and connections to Bernstein functions which simplifies the theory tremendously. Our main source for Bernstein functions is [SSV10]. Finally, we investigate subordinate Brownian motions which form a large class of symmetric Lévy processes, yet much more tractable than general symmetric Lévy processes, cf. [KSV11]. Subordinate Brownian motions are used in mathematical finance, as the subordinator can be thought of as the "operational time" or "intrinsic time". Most results will be presented without proof.

In this chapter, $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbb{P})$ denotes a standard filtrated probability space and the filtration satisfies the usual hypotheses (cf. Chapter 3).

5.1 Definition. An adapted process $L = (L_t)_{t\geqslant 0}$ with $L_0 = 0$ a.s. is a **(n-dimensional) Lévy process** if

(L1) L has **increments independent of the past**, that is, $L_t - L_s \perp\!\!\!\perp \mathscr{F}_s^L$, $0 \leqslant s < t$, where $\mathscr{F}_s^L := \sigma(L_r : r \leqslant s)$ denotes the natural filtration;

(L2) L has **stationary increments**, that is, $L_t - L_s \sim L_{t-s}$, $0 \leqslant s < t$;

(L3) $t \mapsto L_t(\omega)$ is càdlàg[1] for all $\omega \in \Omega$.

By the Lévy-Khintchine formula, the characteristic function of L_t is given by

$$\mathbb{E}e^{i\xi\cdot L_t} = e^{-t\psi(\xi)} \quad (\xi \in \mathbb{R}^n, t \geqslant 0) \tag{5.1}$$

where, for every $\xi \in \mathbb{R}^n$,

$$\psi(\xi) = -ia\cdot\xi + \frac{1}{2}\xi\cdot\Sigma\xi + \int_{\mathbb{R}^n\setminus\{0\}} \left(1 - e^{iy\cdot\xi} + iy\cdot\xi\,\mathbb{1}_{\{|y|\leqslant 1\}}\right)\nu(\mathrm{d}y)$$

denotes the **Lévy exponent** of L. The Lévy process is completely determined by the so called **Lévy triplet** (a, Σ, ν): the **drift** $a \in \mathbb{R}^n$, the **diffusion coefficient** $\Sigma \in \mathbb{R}^{n\times n}$ symmetric nonnegative-definite and the **Lévy measure** $\nu(\mathrm{d}y)$ on $\mathbb{R}^n \setminus \{0\}$ which satisfies $\int_{\mathbb{R}^n\setminus\{0\}} \left(|y|^2 \wedge 1\right)\nu(\mathrm{d}y) < \infty$.

5.2 Definition. An $\mathbb{R}^n$-valued random variable X with law μ_X is **infinitely divisble** if, for all $n \in \mathbb{N}$, there exist iid random variables $X_1, \ldots, X_n$ such that $X \sim X_1 + \ldots + X_n$.

[1] A càdlàg (French *continue à droite, limite à gauche*) function is a function defined on the real numbers (or a subset of them) that is everywhere right-continuous and has left limits everywhere.

Denote by $\mu^{\star n}$ the n-fold convolution of a probability measure μ with ifself, that is, $\mu^n = \mu \star \cdots \star \mu$. Then it is easy to see that a distribution μ_X on $\mathbb{R}^n$ is infinitely divisible if and only if, for each $n \in \mathbb{N}$, there exists $\mu^{1/n}$ on $\mathbb{R}^n$ such that $\mu_X = \left[\mu^{1/n}\right]^{\star n}$. Simple examples for infinitely divisible random variables are Gaussian random variables and Poisson random variables.

5.3 Proposition. *If L is a Lévy process, then L_t is infinitely divisible, for each $t \geqslant 0$.*

Proof. For each $n \in \mathbb{N}$, we can write L_t as the telescoping series

$$L_t = \sum_{i=1}^{n} \left(L_{t\frac{i}{n}} - L_{t\frac{i-1}{n}} \right).$$

Thus, using (L1) and (L2), the increments $L_{t\frac{i}{n}} - L_{t\frac{i-1}{n}}$ are iid and the result follows. ■

An $\mathbb{R}^n$-valued random variable X is called **α-stable** if there exist real-valued sequences $(a_k)_{k\in\mathbb{N}} \subset (0,\infty)$ and $(b_k)_{k\in\mathbb{N}}$ such that

$$X_1 + \ldots + X_k \sim a_k X + b_k, \tag{5.2}$$

where $X_1, \ldots, X_k$ are independent copies of X. X is said to be **strictly stable** if $b_k = 0$. It can be shown that the only possible choice (up to a multiplicative constant) for a_k in (5.2) is $k^{1/\alpha}$, where $0 < \alpha \leqslant 2$ (cf. [App09, p. 34]). The parameter α plays a key role in the investigation of stable random variables and is called the **index of stability**. It follows immediately from (5.2) that all stable random variables are infinitely divisible. The definition extends naturally to processes in the following way: A process $X = (X_t)_{t\geqslant 0}$ is **α-stable** if all its finite-dimensional distributions are α-stable. For example, every Gaussian process is a 2-stable process.

Let $0 < \alpha \leqslant 2$. If μ is a strictly α-stable distribution on $\mathbb{R}$, then (cf. [Sat13, Theorem 14.19]) the Fourier transform $\hat{\mu}$ of μ is

$$\hat{\mu}(\xi) = \exp\left(-c_1 \left|\xi\right|^\alpha e^{-i(\pi/2)\vartheta\alpha \operatorname{sgn}\xi}\right), \tag{5.3}$$

where $c_1 > 0$ and $\vartheta \in \mathbb{R}$ with $|\vartheta| \leqslant \frac{2-\alpha}{\alpha} \wedge 1$. The parameters c_1 and ϑ are uniquely determined by μ. Conversely, for any c_1 and ϑ, there is a strictly α-stable distribution μ satisfying (5.3). In particular, in the case $\alpha = 2$, ϑ must be 0 (cf. [Sat13, Remark 14.10]).

Furthermore, recall that an $\mathbb{R}^n$-valued process $(X_t)_{t\geqslant 0}$ is **rotation invariant** if $X_t \sim UX_t$ for every orthogonal matrix U, $t \geqslant 0$. Note that this is equivalent to say that $\hat{\mu}(\xi) = \mathbb{E}e^{i\xi X_t}$ is a function only of $|\xi|$. In particular, if $n = 1$, we get $X_t \sim -X_t$, $t \geqslant 0$, i.e. X is **symmetric**. Symmetric stable random variables are of special interest to us, since symmetry displays self-similarity. Moreover, the characteristic function has a simple explicit form (cf. [Sat13, Theorem 14.14]): An $\mathbb{R}^n$-valued random variable X is rotation invariant and α-stable with $0 < \alpha \leqslant 2$ if and only if

$$\mathbb{E}e^{i\xi\cdot X} = e^{-c|\xi|^\alpha}, \quad c > 0. \tag{5.4}$$

Note that when $\alpha \neq 2$, the Lévy measure takes the form (cf. [Sat13, pp.77-80])

$$\nu(B) = \int_0^\infty \int_{S^{n-1}} \mathbb{1}_B(r\vartheta)\, \mathrm{d}S(\vartheta)\, \frac{\mathrm{d}r}{r^{1+\alpha}}, \quad B \in \mathscr{B}(\mathbb{R}^n).$$

If it is rotation invariant this simplifies to

$$\nu(\mathrm{d}r) = c\frac{\mathrm{d}r}{|r|^{n+\alpha}}, \quad c > 0.$$

5.4 Definition. An $\mathbb{R}$-valued Lévy process $(X_t)_{t\geqslant 0}$ is said to be **increasing** if $t \mapsto X_t(\omega)$ is increasing a.s. An increasing Lévy process $S = (S_t)_{t\geqslant 0}$ taking values in $[0,\infty)$ (with $S_0 = 0$) is called a **subordinator**.

A smooth function $g : (0,\infty) \to \mathbb{R}$ is called a **Bernstein function** if $(-1)^n g^{(n)} \leqslant 0$ for every positive integer n. From the definition it is easy to see that $\lambda \mapsto \lambda^\alpha$ is a Bernstein function if and only if $0 \leqslant \alpha \leqslant 1$. Every Bernstein function admits a representation (cf. [SSV10, Theorem 3.2])

$$g(\lambda) = a + b\lambda + \int_{(0,\infty)} \left(1 - e^{-\lambda t}\right) \nu(\mathrm{d}t), \tag{5.5}$$

where $a, b \geqslant 0$ and ν is a measure on $(0,\infty)$ satisfying $\int_{(0,\infty)}(1 \wedge t)\nu(\mathrm{d}t) < \infty$, also called Lévy measure. In particular, the (Lévy) triplet (a, b, ν) determines g uniquely and vice versa. Moreover (cf. [SSV10, Theorem 5.2]), for a Bernstein function g, there exists a unique convolution semigroup of sub-probability measures $(\nu_t)_{t\geqslant 0}$ on $[0,\infty)$ such that the Laplace transform $\mathscr{L}$ of ν_t is given by $\mathscr{L}\nu_t = e^{-tg}$, for all $t \geqslant 0$. The converse is also true: A nonnegative function g on $(0,\infty)$ is the Laplace exponent of a subordinator if and only if it is a Bernstein function with $g(0+) = 0$. Consequently, a subordinator S is completely characterised by its Laplace exponent g via

$$\mathbb{E}e^{-\lambda S_t} = e^{-tg(\lambda)}, \quad \lambda > 0 \text{ with } g(\lambda) = -\psi(i\lambda). \tag{5.6}$$

Let L be a Lévy process and S an independent subordinator. Then we find

$$\begin{aligned}
\mathbb{E}e^{i\xi\cdot L_{S_t}} &\overset{L \perp\!\!\!\perp S}{=} \int \mathbb{E}e^{i\xi\cdot L_s}\mathbb{P}\left(S_t \in \mathrm{d}s\right) \\
&\overset{(5.1)}{=} \int e^{-s\psi(\xi)}\mathbb{P}\left(S_t \in \mathrm{d}s\right) \\
&= \mathbb{E}e^{-\psi(\xi)S_t} = e^{-t\, g\circ\psi(\xi)}.
\end{aligned}$$

Written more compactly this reads $\psi_X = g_S \circ \psi_L$.

In particular, choosing $L = B$ an n-dimensional Brownian motion and S an α-stable subordinator for $0 < \alpha < 1$, then it is clear that $\psi(\xi) = |\xi|^2$ and $g(\lambda) = \lambda^\alpha$ force $a = b = 0$ and $\nu(\mathrm{d}t) = c_\alpha \frac{\mathrm{d}t}{t^{1+\alpha}}$ in (5.5), respectively. This follows by a well-known trick of establishing a double integral and then changing the order of integration.

$$\begin{aligned}\int_0^\infty \left(1-e^{-\lambda t}\right)\frac{\mathrm{d}t}{t^{1+\alpha}} &= -\int_0^\infty \left(\int_0^t \lambda e^{-\lambda s}\,\mathrm{d}s\right) t^{-1-\alpha}\,\mathrm{d}t\\ &= -\int_0^\infty \left(\int_s^\infty t^{-1-\alpha}\,\mathrm{d}t\right)\lambda e^{-\lambda s}\,\mathrm{d}s\\ &= \frac{\lambda}{\alpha}\int_0^\infty e^{-\lambda s}s^{-\alpha}\,\mathrm{d}s\\ &= \frac{\lambda^\alpha}{\alpha}\int_0^\infty e^{-t}t^{-\alpha}\,\mathrm{d}t = \frac{\lambda^\alpha}{\alpha}\,\Gamma(1-\alpha),\end{aligned}$$

wherein Γ denotes the Gamma function. Thus we get, for $\lambda > 0$ and $0 < \alpha < 1$,

$$\lambda^\alpha = \frac{\alpha}{\Gamma(1-\alpha)}\int_0^\infty \left(1-e^{-\lambda t}\right)\frac{\mathrm{d}t}{t^{1+\alpha}}.$$

The subordinator S corresponding to $g(\lambda) = \lambda^{\alpha/2}$, for $0 < \alpha < 2$, is called an **$\alpha/2$-stable subordinator**. In particular, by (5.6), it follows that

$$\mathbb{E}e^{i\xi S_t} = e^{t|\xi|^{\alpha/2}}.$$

5.5 Definition. Let $B = (B_t)_{t\geqslant 0}$ be an n-dimensional Brownian motion and let $S = (S_t)_{t\geqslant 0}$ be an independent subordinator. The process $X = (X_t)_{t\geqslant 0}$ defined by $X_t := B_{S_t}$, $t \geqslant 0$, is called a **subordinated Brownian motion**.

For any subordinated Brownian motion $X = B_S$ it follows from the above considerations that $\psi_{B\circ S}(\xi) = |\xi|^\alpha$, i.e. $(B_{S_t})_{t\geqslant 0}$ is a rotation invariant α-stable process.

Lastly, we investigate generators of Lévy processes, especially α-stable subordinators. For any Lévy process, the **transition semigroup** $P_t f$ on $C_b(\mathbb{R}^n)$ is defined by

$$P_t f(x) := \mathbb{E}^x f(L_t) := \mathbb{E} f(L_t + x), \quad f \in C_b(\mathbb{R}^n). \tag{5.7}$$

Then, the semigroup properties are not difficult to see by $L_0 = 0$ a.s. and (L1). Let $\mu = \mathbb{P}_{L_1}$ be the corresponding infinitely divisible distribution on $\mathbb{R}^n$, i.e. $\mu^t = \mathbb{P}_{L_t}$. Then the spatially homogenous transition function $p_t(x, \mathrm{d}y)$ is

$$p_t(x, B) = \mu^t(B - x), \quad t \geqslant 0, x \in \mathbb{R}^n, B \in \mathscr{B}(\mathbb{R}^n).$$

Thus (cf. [Sat13, Theorem 10.5]), every Lévy process is a Markov process with transition function $p_t(x, B)$ and vice versa. By the Lévy-Khintchine formula we can state an explicit formula for the generator.

5.6 Theorem ([Sat13, Theorem 31.5]). *The transition semigroup $(P_t)_{t\geqslant 0}$ of a Lévy process $(L_t)_{t\geqslant 0}$ on $\mathbb{R}^n$ defined in (5.7) is a strongly continuous semigroup on $C_b(\mathbb{R}^n)$ with norm $\|P_t\| = 1$. Moreover, $C_c^\infty \subset C_b^2 \subset \mathfrak{D}(\mathrm{A})$ and the generator is given by*

$$\begin{aligned}\mathrm{A}f(x) = a(x)\cdot\nabla f(x) + \operatorname{tr}(\Sigma\cdot D^2 f(x)) +\\ \int_{\mathbb{R}^n\setminus\{0\}} \left(f(x+y) - f(x) - y\cdot\nabla f(x)\mathbb{1}_{\{|y|\leqslant 1\}}\right)\nu(\mathrm{d}y),\end{aligned} \tag{5.8}$$

for $f \in C_b^2$*, where* (a, Σ, ν) *is the Lévy triplet of* $(L_t)_{t \geqslant 0}$.

Note that the pointwise limit (sometimes called *weak generator*) coincides with the uniform limit as in (4.1) only since $C_c^\infty \subset \mathscr{D}(\mathbf{A})$ (cf. [SP14, Theorem 7.22] or [Sat13, Theorem 31.7]).

5.7 Example. (i) If $X = B$ is a Brownian motion on $\mathbb{R}^n$, then X has the Lévy triplet $(0, I, 0)$. So, we have $\mathbf{A} = \frac{1}{2}\Delta$ on $C_b^2(\mathbb{R}^n)$, where Δ is the usual Euclidean Laplacian.

(ii) Let $X = B_S$ be a subordinated Brownian motion on $\mathbb{R}^n$. The corresponding Bernstein function is given by $g(\lambda) = \lambda^{\alpha/2}$ for $0 < \alpha < 2$ and the Lévy measure by $\nu(\mathrm{d}y) = c_\alpha \frac{\mathrm{d}y}{|y|^{n+\alpha}}$, respectively. Thus, the generator $\mathbf{A}$ is[2]

$$\mathbf{A}f(x) = \mathrm{p.v.} \int_{\mathbb{R}^n} (f(x+y) - f(x)) \frac{\mathrm{d}y}{|y|^{n+\alpha}}. \tag{5.9}$$

[2]Revised version: corrected misprint.

Chapter 6

BEL for SDEs Driven by α-stable Processes

In the final chapter, following [Zha13] we generalise the classical Bismut-Elworthy-Li formula on $\mathbb{R}^n$ for *nonlinear* SDEs driven by subordinated Brownian motion. We will use a deterministic time change method due to [Léa88] and [Kus10] which they originally used to study Malliavin calculus. We present their main ideas below.

Let us consider the following nonlinear SDE in $\mathbb{R}^n$ driven by a subordinated Brownian motion B_{S_t}

$$\begin{aligned} \mathrm{d}X_t(x) &= b_t(X_t(x))\mathrm{d}t + \sigma \cdot \mathrm{d}B_{S_t}, \\ X_0(x) &= x, \end{aligned} \tag{6.1}$$

where $\sigma \in \mathbb{R}^{n\times n}$ is invertible and $b : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$.

By Definition 5.5, we know that a subordinated Brownian motion $\left(B_{S_t}\right)_{t\geqslant 0}$ is a symmetric α-stable process and, using (5.9), the generator of the solution $\left(X_t(x)\right)_{t\geqslant 0, x\in\mathbb{R}^n}$ is given by[1]

$$\mathsf{A}_t f(x) := b_t(x)\cdot\nabla f(x) + \mathrm{p.v.}\int_{\mathbb{R}^n} (f(x+\sigma y) - f(x))\frac{\mathrm{d}y}{|y|^{n+\alpha}},$$

If $(B_t)_{t\geqslant 0}$ is a canonical n-dimensional Brownian motion, then (6.1) reduces to an SDE in $\mathbb{R}^n$,

$$\begin{aligned} \mathrm{d}X_t(x) &= b_t(X_t(x))\mathrm{d}t + \sigma \cdot \mathrm{d}B_t, \\ X_0(x) &= x. \end{aligned} \tag{6.2}$$

We have shown in Chapter 4 that the Bismut-Elworthy-Li formula holds: for every $f \in C_b^1(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$

$$\nabla_h \mathbb{E} f(X_t(x)) = \frac{1}{t}\mathbb{E}\left(f(X_t(x))\int_0^t \sigma^{-1}\cdot\nabla_h X_s(x)\mathrm{d}B_s\right). \tag{6.3}$$

Moreover, we assume that $\nabla_h X_t(x)$ satisfies the following linear equation[2]

$$\nabla_h X_t(x) = h + \int_0^t \nabla b_s(X_s(x))\cdot\nabla_h X_s(x)\mathrm{d}s. \tag{6.4}$$

Now, our aim is to establish a Bismut-Elworthy-Li formula for $\nabla_h \mathbb{E} f(X_t(x))$ for (6.1), i.e. an SDE driven by a *subordinated* Brownian motion.

[1] Revised version: corrected misprint.

[2] In [Zha13] is stated that $\nabla_h X_t(x)$ has to satisfy (6.4) so that the Bismut-Elworthy-Li formula (6.3) holds true. Following Elworthy and Li's original paper in Chapter 4 we strongly doubt this assertion.

6.1 Theorem ([Zha13, Theorem 1.1]). *Let $b \in \mathrm{BC}^1$. Then for any function $f \in C_b^1(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$, we have*

$$\nabla_h \mathbb{E} f(X_t(x)) = \mathbb{E}\left(\frac{1}{S_t} f(X_t(x)) \int_0^t \sigma^{-1} \cdot \nabla_h X_s(x) \mathrm{d}B_{S_s}\right), \tag{6.5}$$

where $\nabla_h X_s(x)$ is determined by (6.4). In particular, for any $\alpha \in (0,2)$ and $p \in (1,\infty]$, there exists a constant $C = C(\alpha, p) > 0$ such that for all $t > 0$,

$$\left|\nabla_h \mathbb{E} f(X_t(x))\right| \leqslant C \left\|\sigma^{-1}\right\| \mathrm{e}^{\|\nabla b\|_\infty t}\, t^{-\frac{1}{\alpha}} \left(\mathbb{E}\left|f(X_t(x))\right|^p\right)^{1/p}, \tag{6.6}$$

where $\left\|\sigma^{-1}\right\| := \sup_{|x|=1}\left|\sigma^{-1}x\right|$ and $|\cdot|$ denotes the Euclidian norm.

6.2 Remark. From (6.4), using Grönwall's inequality A.2, we see that $s \mapsto \nabla_h X_s(x)$ is a bounded, and as solution a continuous $\sigma\left(B_{S_r} : r \leqslant s\right)$-adapted process. Thus, the stochastic integral in (6.5) makes sense.

We now introduce the main idea of the deterministic time change method (cf. [Léa88], [Kus10]). Let $\mathbb{W}$ be the classical Wiener space, i.e. the space of all continuous functions from $[0,\infty)$ to $\mathbb{R}^n$ vanishing at starting point 0, which is endowed with the locally uniform convergence topology and the Wiener measure $\mu_{\mathbb{W}}$ so that the coordinate process

$$B_t(w) = w_t$$

is a canonical n-dimensional Brownian motion. Let $\mathbb{S}$ be the space of all increasing càdlàg functions from $(0,\infty)$ to $(0,\infty)$ with $\lim_{s\downarrow 0} \ell_s = 0$, which is endowed with the Skorohod metric and the probability measure $\mu_{\mathbb{S}}$ so that the coordinate process

$$S_t(\ell) := \ell_t$$

is an $\alpha/2$-stable subordinator S_t for $\alpha \in (0,2)$ (cf. [Sat13]). Consider the following product probability space

$$(\Omega, \mathscr{F}, \mathbb{P}) := (\mathbb{W} \times \mathbb{S}, \mathscr{B}(\mathbb{W}) \times \mathscr{B}(\mathbb{S}), \mu_{\mathbb{W}} \times \mu_{\mathbb{S}}),$$

and define

$$L_t(w,\ell) := w_{\ell_t}.$$

Then $\left(L_t\right)_{t\geqslant 0}$ is an α-stable process on $(\Omega, \mathscr{F}, \mathbb{P})$. We shall use the following two natural filtrations associated to the Lévy process L_t and the Brownian motion B_t:

$$\mathscr{F}_t := \sigma\left(L_s(w,\ell) : s \leqslant t\right), \qquad \mathscr{F}_t^{\mathbb{W}} := \sigma\left(B_s(w) : s \leqslant t\right).$$

In particular, we can regard the solution $X_t(x)$ of SDE (6.1) as an $\mathscr{F}_t$-adapted functional on Ω and

$$\mathbb{E} f(X_t(x)) = \int_{\mathbb{S}}\int_{\mathbb{W}} f(X_t(x; w_{\ell_\cdot}))\mu_{\mathbb{W}}(\mathrm{d}w)\mu_{\mathbb{S}}(\mathrm{d}\ell).$$

For $\ell \in \mathbb{S}$, let $X_t^\ell(x)$ solve the SDE

$$\mathrm{d}X_t^\ell(x) = b_t(X_t^\ell(x))\mathrm{d}t + \sigma \cdot \mathrm{d}B_{\ell_t}, \quad X_0^\ell(x) = x. \tag{6.7}$$

Now, our task is to establish a formula for $\nabla_h \mathbb{E}^{\mu_\mathbb{W}} f(X_t^\ell(x))$. This is not obvious since $t \mapsto B_{\ell_t}$ is not continuous and the classical Bismut-Elworthy-Li formula cannot be used directly. In Section 6.1, we prove a formula for $\nabla_h \mathbb{E}^{\mu_\mathbb{W}} f(X_t^\ell(x))$ by suitable approximation of Steklov's average for ℓ_t and in Section 6.2 we prove Theorem 6.1.

6.1 Derivative formula of SDEs under deterministic time change

In this section, we fix $\ell \in \mathbb{S}$ and consider SDE (6.7). Unless otherwise stated, all expectations are taken on the Wiener space $(\mathbb{W}, \mathscr{B}(\mathbb{W}), \mu_\mathbb{W})$. Note that $t \mapsto B_{\ell_t}$ is a Gaussian process with mean zero and independent increments. In particular, B_{ℓ_t} is a càdlàg $\mathscr{F}_{\ell_t}^{\mathbb{W}}$-martingale.

Thus if $b \in \mathrm{BC}^1$, it is well-known that for each $x \in \mathbb{R}^n$, SDE (6.7) admits a unique càdlàg $\mathscr{F}_{\ell_t}^{\mathbb{W}}$-adapted solution $X_t^\ell(x)$ (cf. [Pro05, p. 255, Theorem 6]).

The main aim of this section is to establish the following formula:

6.3 Theorem ([Zha13, Theorem 2.1]). *Let $b \in \mathrm{BC}^1$. Then for any function $f \in C_b^1(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$, we have*

$$\nabla_h \mathbb{E} f(X_t^\ell(x)) = \mathbb{E}\left(\frac{1}{\ell_t} f(X_t^\ell(x)) \int_0^t \sigma^{-1} \cdot \nabla_h X_s^\ell(x) \mathrm{d}B_{\ell_s}\right), \tag{6.8}$$

where $\nabla_h X_s^\ell(x)$ is determined by the linear equation

$$\nabla_h X_t^\ell(x) = h + \int_0^t \nabla b_s(X_s^\ell(x)) \cdot \nabla_h X_s^\ell(x) \mathrm{d}s. \tag{6.9}$$

To prove this formula, we use a time change argument to transform the SDE (6.7) into an SDE driven by a canonical Brownian motion and then use the classical Bismut-Elworthy-Li formula (6.3). To this end, for $\varepsilon \in (0,1)$, we define

$$\ell_t^\varepsilon := \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \ell_s \mathrm{d}s + \varepsilon t = \int_0^1 \ell_{\varepsilon s + t} \mathrm{d}s + \varepsilon t. \tag{6.10}$$

Since $t \mapsto \ell_t$ is increasing and right continuous, it follows that for each $t \geqslant 0$,

$$\ell_t^\varepsilon \downarrow \ell_t \quad \text{as} \quad \varepsilon \downarrow 0. \tag{6.11}$$

Moreover, $t \mapsto \ell_t^\varepsilon$ is absolutely continuous and strictly increasing. Let γ^ε be the inverse function of ℓ^ε, i.e.

$$\ell_{\gamma_t^\varepsilon}^\varepsilon = t, \ t \geqslant \ell_0^\varepsilon \quad \text{and} \quad \gamma_{\ell_t^\varepsilon}^\varepsilon = t, \ t \geqslant 0.$$

By definition, γ_t^ε is also absolutely continuous on $[\ell_0^\varepsilon, \infty)$. Define

$$Y_t^{\ell^\varepsilon}(x) := X_{\gamma_t^\varepsilon}^{\ell^\varepsilon}(x), \quad t \geqslant \ell_0^\varepsilon.$$

By (6.7) and the change of variables, we see that for $t \geqslant \ell_0^\varepsilon$,

$$\begin{aligned} Y_t^{\ell^\varepsilon}(x) &= x + \int_0^{\gamma_t^\varepsilon} b_s(X_s^{\ell^\varepsilon}(x))\mathrm{d}s + \sigma \cdot B_t \\ &= x + \int_{\ell_0^\varepsilon}^t b_{\gamma_s^\varepsilon}(Y_s^{\ell^\varepsilon}(x))\dot{\gamma}_s^\varepsilon \mathrm{d}s + \sigma \cdot B_t. \end{aligned}$$

Hence, we can use the classical Bismut-Elworthy-Li formula (6.3) to derive that

$$\nabla_h \mathbb{E} f(Y_t^{\ell^\varepsilon}(x)) = \frac{1}{t}\mathbb{E}\left(f(Y_t^{\ell^\varepsilon}(x)) \int_0^t \sigma^{-1} \cdot \nabla_h Y_s^{\ell^\varepsilon}(x)\mathrm{d}B_s \right), \ t \geqslant \ell_0^\varepsilon,$$

where $\nabla_h Y_t^{\ell^\varepsilon}(x)$ satisfies

$$\nabla_h Y_t^{\ell^\varepsilon}(x) = h + \int_{\ell_0^\varepsilon}^t \nabla b_{\gamma_s^\varepsilon}(Y_s^{\ell^\varepsilon}(x)) \nabla_h Y_s^{\ell^\varepsilon}(x) \dot{\gamma}_s^\varepsilon \mathrm{d}s.$$

Clearly, for each $t \geqslant 0$,

$$Y_{\ell_t^\varepsilon}^{\ell^\varepsilon}(x) = X_t^{\ell^\varepsilon}(x), \quad \nabla_h Y_{\ell_t^\varepsilon}^{\ell^\varepsilon}(x) = \nabla_h X_t^{\ell^\varepsilon}(x),$$

and therefore

$$\begin{aligned} \nabla_h \mathbb{E} f(X_t^{\ell^\varepsilon}(x)) &= \frac{1}{\ell_t^\varepsilon}\mathbb{E}\left(f(X_t^{\ell^\varepsilon}(x)) \int_0^{\ell_t^\varepsilon} \sigma^{-1} \cdot \nabla_h Y_s^{\ell^\varepsilon}(x)\mathrm{d}B_s \right) \\ &= \frac{1}{\ell_t^\varepsilon}\mathbb{E}\left(f(X_t^{\ell^\varepsilon}(x)) \int_0^t \sigma^{-1} \cdot \nabla_h X_s^{\ell^\varepsilon}(x)\mathrm{d}B_{\ell_s^\varepsilon} \right). \end{aligned} \tag{6.12}$$

Now, we want to let $\varepsilon \to 0$. We need several lemmas. The following lemma is a typical application of Grönwall's inequality A.2.

6.4 Lemma ([Zha13, Lemma 2.2]). *For any $p \geqslant 1$ and $t \geqslant 0$, we have*

$$\lim_{\varepsilon \downarrow 0} \mathbb{E}\left(\sup_{x \in \mathbb{R}^n} |X_t^{\ell^\varepsilon}(x) - X_t^\ell(x)|^p \right) = 0, \tag{6.13}$$

and for any $x, h \in \mathbb{R}^n$,

$$\lim_{\varepsilon \downarrow 0} \mathbb{E}\left(\sup_{0 \leqslant s \leqslant t} |\nabla_h X_s^{\ell^\varepsilon}(x) - \nabla_h X_s^\ell(x)|^p \right) = 0, \tag{6.14}$$

where

$$\nabla_h X_t^{\ell^\varepsilon}(x) = h + \int_0^t \nabla b_s(X_s^{\ell^\varepsilon}(x)) \cdot \nabla_h X_s^{\ell^\varepsilon}(x)\mathrm{d}s. \tag{6.15}$$

Proof. For simplicity of notation, we drop the variable "x" below. From (6.7), we have

$$|X_t^{\ell^\varepsilon} - X_t^\ell| \leqslant \|\nabla b\|_\infty \int_0^t |X_s^{\ell^\varepsilon} - X_s^\ell| \mathrm{d}s + |\sigma \cdot B_{\ell_t^\varepsilon} - \sigma \cdot B_{\ell_t}|.$$

By Grönwall's inequality A.2, we get

$$|X_t^{\ell^\varepsilon} - X_t^\ell| \leqslant \mathrm{e}^{\|\nabla b\|_\infty t} |\sigma \cdot B_{\ell_t^\varepsilon} - \sigma \cdot B_{\ell_t}|,$$

which then gives (6.13) as $\ell_t^\varepsilon \xrightarrow{\varepsilon \downarrow 0} \ell_t$. As for (6.14), by (6.15) and the linear equation (6.9) we have

$$\begin{aligned} \left|\nabla_h X_t^{\ell^\varepsilon} - \nabla_h X_t^\ell\right| \leqslant & \int_0^t \left|\nabla b_s(X_s^{\ell^\varepsilon})\right| \cdot \left|\nabla_h X_s^{\ell^\varepsilon} - \nabla_h X_s^\ell\right| \mathrm{d}s \\ & + \int_0^t \left|\nabla b_s(X_s^{\ell^\varepsilon}) - \nabla b_s(X_s^\ell)\right| \cdot \left|\nabla_h X_s^\ell\right| \mathrm{d}s, \end{aligned}$$

which yields by Grönwall's inequality A.2 that

$$|\nabla_h X_t^{\ell^\varepsilon} - \nabla_h X_t^\ell| \leqslant \left(1 + \mathrm{e}^{\|\nabla b\|_\infty t} \ \|\nabla b\|_\infty t\right) \int_0^t |\nabla b_s(X_s^{\ell^\varepsilon}) - \nabla b_s(X_s^\ell)| \cdot | \nabla_h X_s^\ell| \mathrm{d}s.$$

From the definition in (6.9), using the elementary estimate $(a+b)^p \leqslant 2^p(a^p + b^p)$ and the Hölder inequality, we get

$$\begin{aligned} \mathbb{E}\left|\nabla_h X_s^\ell\right|^p & \leqslant 2^p \, |h|^p + \mathbb{E}\left|\int_0^s \nabla b_r(X_r^\ell) \cdot \nabla_h X_r^\ell \ \mathrm{d}r\right|^p \\ & \leqslant 2^p \, |h|^p + s^{p-1} \, \|\nabla b\|_\infty \int_0^s \mathbb{E}\left|\nabla_h X_r^\ell\right|^p \ \mathrm{d}r. \end{aligned}$$

Thus, again by Grönwall's inequality A.2, it is easy to see that for any $p \geqslant 1$,

$$\sup_{0 \leqslant s \leqslant t} \mathbb{E}|\nabla_h X_s^\ell|^p \leqslant C.$$

Limit (6.14) now follows by the dominated convergence theorem, (6.13) and the continuity of $x \mapsto \nabla b_s(x)$. ∎

We also need the following lemma, where the second part is due to Kusuoka [Kus10, Lemma 2.3]. The proof can be found in [Zha13, Lemma 2.3].

6.5 Lemma ([Zha13, Theorem 2.3]). *(i) Assume that ξ_t is a bounded continuous and $\mathscr{F}_{\ell_t}^{\mathbb{W}}$-adapted $\mathbb{R}^n$-valued process. For each $p, T > 0$, we have*

$$\lim_{\varepsilon \downarrow 0} \mathbb{E}\left|\int_0^T \xi_s \mathrm{d}B_{\ell_s^\varepsilon} - \int_0^T \xi_s \mathrm{d}B_{\ell_s}\right|^p = 0, \tag{6.16}$$

where ℓ_s^ε is defined by (6.10).

(ii) Assume that ξ_t is a left continuous and $\mathscr{F}^{\mathbb{W}}_{\ell_t}$-adapted $\mathbb{R}^n$-valued process and satisfies, for some $p > 0$,

$$\mathbb{E}\left(\int_0^T |\xi_s|^2 \mathrm{d}\ell_s\right)^{p/2} < +\infty, \quad \textit{for all } T \geqslant 0. \tag{6.17}$$

Then there exists a constant $C_p > 0$ such that for all $T \geqslant 0$,

$$\mathbb{E}\left(\sup_{t\leqslant T}\left|\int_0^t \xi_s \mathrm{d}B_{\ell_s}\right|^p\right) \leqslant C_p \mathbb{E}\left(\int_0^T |\xi_s|^2 \mathrm{d}\ell_s\right)^{p/2}. \tag{6.18}$$

6.6 Remark. By Burkholder's inequality (cf. [SP14, Theorem 18.19]), we have, for $p \geqslant 1$,

$$\mathbb{E}\left(\sup_{t\leqslant T}\left|\int_0^t \xi_s \mathrm{d}B_{\ell_s}\right|^p\right) \asymp \mathbb{E}\left(\int_0^T |\xi_s|^2 \mathrm{d}[B_\ell]_s\right)^{p/2}, \tag{6.19}$$

where $\asymp$ means that both sides are comparable by multiplying a constant. Except for the case of $p = 2$, it is not known whether the righthand side of (6.18) and (6.19) are comparable.

The quadratic variation $[B_\ell]_t$ of $(B_{\ell_t})_{t\geqslant 0}$ given by (cf. Appendix B)

$$[B_\ell]_t = \ell_t - \sum_{0<s\leqslant t} \Delta\ell_s + \sum_{0<s\leqslant t} \left|\Delta B_{\ell_s}\right|^2.$$

6.7 Lemma ([Zha13, Lemma 2.5]). *For all $t \geqslant 0$, we have*

$$\lim_{\varepsilon\downarrow 0} \mathbb{E}\left(\int_0^t \sigma^{-1}\cdot\nabla_h X_s^{\ell^\varepsilon}(x)\mathrm{d}B_{\ell_s^\varepsilon}\right) = \mathbb{E}\left(\int_0^t \sigma^{-1}\cdot\nabla_h X_s^{\ell}(x)\mathrm{d}B_{\ell_s}\right). \tag{6.20}$$

Proof. For simplicity of notation, we drop the variable "x" below. Since $\ell_s^\varepsilon \geqslant \ell_s$ by definition, $X_s^{\ell^\varepsilon}$ and X_s^ℓ are $\mathscr{F}_{\ell_s^\varepsilon}$-adapted so that the stochastic integrals in (6.22) make sense. To establish (6.20), it is sufficient to prove the following two limits.

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\left|\int_0^t \left(\sigma^{-1}\cdot\nabla_h X_s^{\ell^\varepsilon} - \sigma^{-1}\cdot\nabla_h X_s^\ell\right)\mathrm{d}B_{\ell_s^\varepsilon}\right|^2 = 0 \tag{6.21}$$

and

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\left|\int_0^t \sigma^{-1}\cdot\nabla_h X_s^\ell \mathrm{d}B_{\ell_s^\varepsilon} - \int_0^t \sigma^{-1}\cdot\nabla_h X_s^\ell \mathrm{d}B_{\ell_s}\right|^2 = 0. \tag{6.22}$$

For (6.21), by the Itô isometry and (6.14), we have

$$\begin{aligned}
&\lim_{\varepsilon\downarrow 0}\mathbb{E}\left|\int_0^t \sigma^{-1}\cdot\left(\nabla_h X_s^{\ell^\varepsilon} - \nabla_h X_s^\ell\right)\mathrm{d}B_{\ell_s^\varepsilon}\right|^2 \\
&\quad = \lim_{\varepsilon\downarrow 0}\mathbb{E}\int_0^t \left|\sigma^{-1}\cdot\left(\nabla_h X_s^{\ell^\varepsilon} - \nabla_h X_s^\ell\right)\right|^2 \mathrm{d}\ell_s^\varepsilon \\
&\quad \leqslant \left\|\sigma^{-1}\right\|^2 \lim_{\varepsilon\downarrow 0}\mathbb{E}\left(\sup_{0\leqslant s\leqslant t}\left|\nabla_h X_s^{\ell^\varepsilon} - \nabla_h X_s^\ell\right|^2\right)\ell_t \overset{(6.14)}{=} 0.
\end{aligned}$$

Limit (6.22) follows from Lemma 6.5. ∎

Proof (of Theorem 6.3). As $\ell_t^\varepsilon \xrightarrow{\varepsilon\downarrow 0} \ell_t$, the righthand side of (6.12)

$$\frac{1}{\ell_t^\varepsilon}\mathbb{E}\left(f(X_t^{\ell^\varepsilon}(x))\int_0^t \sigma^{-1}\cdot\nabla_h X_s^{\ell^\varepsilon}(x)\mathrm{d}B_{\ell_s^\varepsilon}\right) \xrightarrow[(6.20)]{(6.13)} \mathbb{E}\left(\frac{1}{\ell_t}f(X_t^{\ell}(x))\int_0^t \sigma^{-1}\cdot\nabla_h X_s^{\ell}(x)\mathrm{d}B_{\ell_s}\right)$$

converges to the one of (6.8) as $\varepsilon\downarrow 0$. On the other hand, by (6.13) and (6.14), we have

$$\begin{aligned}\nabla_h\mathbb{E}f(X_t^{\ell^\varepsilon}(x)) &= \mathbb{E}\left(\nabla f(X_t^{\ell^\varepsilon}(x))\cdot\nabla_h X_t^{\ell^\varepsilon}(x)\right)\\ &\xrightarrow{\varepsilon\downarrow 0}\mathbb{E}\left(\nabla f(X_t^{\ell}(x))\cdot\nabla_h X_t^{\ell}(x)\right) = \nabla_h\mathbb{E}f(X_t^{\ell}(x)).\end{aligned}$$ ■

6.2 Proof of Theorem 6.1

The following lemma follows by the monotone class theorem [Pro05, p. 7, Theorem 8].

6.8 Lemma. *Let $t\geqslant 0$ and $A\in\mathscr{F}_t$. For any $\ell\in\mathbb{S}$, we have*

$$\left\{w\in\mathbb{W} : w_\ell\in A\right\}\in\mathscr{F}^{\mathbb{W}}_{\ell_t}.$$

The first inequality below is due to Giné and Marcus [G+83, Theorem 3.5].

6.9 Theorem ([Zha13, Theorem 3.2]). *Let ξ_t be a left continuous $\mathscr{F}_t$-adapted $\mathbb{R}^n$-valued process that satisfies, for all $T>0$,*

$$\int_0^T \mathbb{E}\left|\xi_r\right|^\alpha \mathrm{d}r < +\infty.$$

Then there exists a constant $C=C(\alpha)>0$ such that for all $\delta>0$ and $T>0$,

$$\mathbb{P}\left(\sup_{t\leqslant T}\left|\int_0^t \xi_r\mathrm{d}B_{S_r}\right|\geqslant\delta\right)\leqslant C\delta^{-\alpha}\int_0^T\mathbb{E}\left|\xi_r\right|^\alpha\mathrm{d}r.$$

In particular, for any $p\in(0,\alpha)$ and some $C=C(\alpha,p)>0$,

$$\mathbb{E}\left(\sup_{t\leqslant T}\left|\int_0^t \xi_r\mathrm{d}B_{S_r}\right|^p\right)\leqslant C\left(\int_0^T\mathbb{E}\left|\xi_r\right|^\alpha\mathrm{d}r\right)^{p/\alpha}. \tag{6.23}$$

Proof. We only proof (6.23). Define

$$\eta := \sup_{t\leqslant T}\left|\int_0^t \xi_r\mathrm{d}B_{S_r}\right|.$$

Then for any $\kappa\geqslant 0$,

$$\begin{aligned}\mathbb{E}\eta^p &= p\int_0^\infty \delta^{p-1}\mathbb{P}(\eta>\delta)\mathrm{d}\delta = p\left(\int_0^\kappa+\int_\kappa^\infty\right)\delta^{p-1}\mathbb{P}(\eta>\delta)\mathrm{d}\delta\\ &\leqslant p\int_0^\kappa\delta^{p-1}\mathrm{d}\delta + pC\int_0^T\mathbb{E}|\xi_r|^\alpha\mathrm{d}r\int_\kappa^\infty\delta^{p-\alpha-1}\mathrm{d}\delta\\ &\leqslant \kappa^p+\frac{Cp}{p-\alpha}\kappa^{p-\alpha}\int_0^T\mathbb{E}|\xi_r|^\alpha\mathrm{d}r.\end{aligned}$$

Setting $\kappa=\left(\int_0^T\mathbb{E}|\xi_r|^\alpha\mathrm{d}r\right)^{1/\alpha}$, we obtain (6.23). ■

Below we prove a substitution formula for stochastic integrals with respect to the subordinated Brownian motion B_{S_t}.

6.10 Proposition ([Zha13, Proposition 3.3]). *Assume that $\xi_t(B_S)$ is a bounded and left continuous $\mathscr{F}_t$-adapted $\mathbb{R}^n$-valued process. Then for any $T \geqslant 0$, we have*

$$\int_0^T \xi_r(B_S)\mathrm{d}B_{S_r} = \int_0^T \xi_r(B_\ell)\mathrm{d}B_{\ell_r}\Big|_{\ell=S} \quad \mathbb{P}-a.s. \tag{6.24}$$

Moreover, for any nonnegative random variable g on $\mathbb{S}$ and $p > 0$, we have

$$\mathbb{E}\left(g(S)\sup_{t\leqslant T}\left|\int_0^t \xi_r(B_S)\mathrm{d}B_{S_r}\right|^p\right) \leqslant C_p\int_{\mathbb{S}} g(\ell)\mathbb{E}^{\mu_{\mathbb{W}}}\left(\int_0^T |\xi_r(B_\ell)|^2\mathrm{d}\ell_r\right)^{p/2}\mu_{\mathbb{S}}(\mathrm{d}\ell). \tag{6.25}$$

Proof. Without loss of generality, we assume $T = 1$. For given $n \in \mathbb{N}$, set $t_k := k/n$, $k = 0, 1, \dots, n$ and define

$$\xi_r^n(B_S) := \xi_0(B_S)1_{\{0\}}(r) + \sum_{i=0}^{n-1}\xi_{t_i}(B_S)1_{(t_i,t_{i+1}]}(r).$$

Then we have, by definition,

$$\begin{aligned}\int_0^1 \xi_r^n(B_S)\mathrm{d}B_{S_r} &= \sum_{i=0}^{n-1}\xi_{t_i}^n(B_S)(B_{S_{t_{i+1}}} - B_{S_{t_i}})\\ &= \sum_{i=0}^{n-1}\xi_{t_i}^n(B_\ell)(B_{\ell_{t_{i+1}}} - B_{\ell_{t_i}})\Big|_{\ell=S} = \int_0^1 \xi_r^n(B_\ell)\mathrm{d}B_{\ell_r}\Big|_{\ell=S},\end{aligned}$$

and by the left continuity of $r \mapsto \xi_r(B_S)$,

$$\lim_{n\to\infty}|\xi_r^n(B_S) - \xi_r(B_S)| = 0, \quad r \geqslant 0. \tag{6.26}$$

By (6.23), we have for $p \in (0, \alpha)$

$$\mathbb{E}\left|\int_0^1 (\xi_r^n(B_S) - \xi_r(B_S))\mathrm{d}B_{S_r}\right|^p \leqslant C\int_0^1 \mathbb{E}|\xi_r^n(B_S) - \xi_r(B_S)|^p\mathrm{d}r \to 0.$$

On the other hand, by Lemma 6.8, for each $\ell \in \mathbb{S}$ and $r \geqslant 0$, $\xi_r(B_\ell)$ is $\mathscr{F}_{\ell_r}^{\mathbb{W}}$-adapted. Thus, by (6.18), we have

$$\begin{aligned}\mathbb{E}\left|\int_0^1 (\xi_r^n(B_\ell) - \xi_r(B_\ell))\mathrm{d}B_{\ell_r}\Big|_{\ell=S}\right|^p &= \int_{\mathbb{S}}\mathbb{E}^{\mu_{\mathbb{W}}}\left|\int_0^1 (\xi_r^n(B_\ell) - \xi_r(B_\ell))\mathrm{d}B_{\ell_r}\right|^p\mu_{\mathbb{S}}(\mathrm{d}\ell)\\ &\leqslant C\int_{\mathbb{S}}\mathbb{E}^{\mu_{\mathbb{W}}}\left(\int_0^1 |\xi_r^n(B_\ell) - \xi_r(B_\ell)|^2\mathrm{d}\ell_r\right)^{p/2}\mu_{\mathbb{S}}(\mathrm{d}\ell)\\ &= C\mathbb{E}\left(\int_0^1 |\xi_r^n(B_S) - \xi_r(B_S)|^2\mathrm{d}S_r\right)^{p/2},\end{aligned}$$

which converges to zero by (6.26) and the dominated convergence theorem. Combining the above calculations, we obtain (6.24). Lastly, (6.25) is an easy consequence of (6.24) and the estimate (6.18). ■

We are now in a position to prove Theorem 6.1.

Proof (of Theorem 6.1). Combining (6.8) in Theorem 6.3 with (6.24), we have

$$\begin{aligned}\nabla_h \mathbb{E} f(X_t(x)) &\overset{(6.24)}{=} \nabla_h \mathbb{E} f(X_t^\ell(x))\big|_{\ell=S} \\ &\overset{(6.8)}{=} \mathbb{E}\left(\frac{1}{\ell_t} f(X_t^\ell(x)) \int_0^t \sigma^{-1}\cdot \nabla_h X_s^\ell(x)\mathrm{d}B_{\ell_s}\right)\Bigg|_{\ell=S} \\ &\overset{(6.24)}{=} \mathbb{E}\left(\frac{1}{S_t} f(X_t(x)) \int_0^t \sigma^{-1}\cdot \nabla_h X_s(x)\mathrm{d}B_{S_s}\right)\end{aligned}$$

and formula (6.5) follows. We now prove the gradient estimate (6.6). By Hölder's inequality, we have

$$\begin{aligned}|\nabla_h \mathbb{E} f(X_t(x))| &\leqslant \mathbb{E}\left(\frac{1}{S_t}\left|f(X_t(x)\int_0^t \sigma^{-1}\cdot \nabla_h X_s(x)\mathrm{d}B_{S_s}\right|\right) \\ &\leqslant \left(\mathbb{E}|f(X_t(x))|^p\right)^{1/p}\left(\mathbb{E}\left|\frac{1}{S_t}\int_0^t \sigma^{-1}\cdot \nabla_h X_s(x)\mathrm{d}B_{S_s}\right|^q\right)^{1/q},\end{aligned}$$

where $\frac{1}{p}+\frac{1}{q}=1$. By (6.4), using Grönwall's inequality A.2,

$$|\nabla_h X_t(x)| \leqslant |h| e^{\|\nabla b\|_\infty t}.$$

Thus by the estimate (6.25), we have

$$\begin{aligned}\mathbb{E}\left|\frac{1}{S_t}\int_0^t \sigma^{-1}\cdot \nabla_h X_s(x)\mathrm{d}B_{S_s}\right|^q &\leqslant C_q \int_{\mathbb{S}} \frac{1}{\ell_t^q}\mathbb{E}^{\mu_{\mathbb{W}}}\left(\left(\int_0^t |\sigma^{-1}\cdot \nabla_h X_s(x)|^2 \mathrm{d}\ell_s\right)^{q/2}\right)\mu_{\mathbb{S}}(\mathrm{d}\ell) \\ &\leqslant C_q \left\|\sigma^{-1}\right\|^q |h|^q e^{q\|\nabla b\|_\infty t}\int_{\mathbb{S}}\frac{\mu_{\mathbb{S}}(\mathrm{d}\ell)}{\ell_t^{q/2}}.\end{aligned}$$

The distributional density of an α-stable subordinator satisfies (cf. [BSS03, (14)])

$$\mathbb{P}\circ S_t^{-1}(\mathrm{d}s) \leqslant C t s^{-1-\frac{\alpha}{2}} e^{-ts^{-\frac{\alpha}{2}}}\mathrm{d}s.$$

Therefore, we have

$$\int_{\mathbb{S}}\frac{\mu_{\mathbb{S}}(\mathrm{d}\ell)}{\ell_t^{q/2}} = \mathbb{E}\left(\frac{1}{S_t^{q/2}}\right) \leqslant C\int_0^\infty t s^{-1-\frac{\alpha+q}{2}} e^{-ts^{-\frac{\alpha}{2}}}\mathrm{d}s = C t^{-\frac{q}{\alpha}}\int_0^\infty u^{q/\alpha}e^{-u}\mathrm{d}u,$$

where the last equality is due to the change of variable $u = ts^{-\frac{\alpha}{2}}$ and C only depends on α, q. Combining the above calculations, we obtain (6.6). ■

Chapter 7
Conclusion

The purpose of our work was to present the Brownian motion on manifolds and the Bismut-Elworthy-Li formula. BEL type formulae open ways to study gradient estimates.

The Eells-Elworthy-Malliavin construction in Section 3.3, depending heavily on differential geometry of principal bundles, requires further elaboration by the reader, although the main ideas should be clear. A major open question is to give the notion of Lévy processes on manifolds a reasonable meaning. First attempts have been made. For example, in [App01], Applebaum established a concept of *isotropic Lévy processes* on a Riemannian manifold following the idea to project from the frame bundle back down to the manifold.

We have only proved the classical BEL on a specific type of 1-forms, namely the differential of functions on manifolds. However, already in Elworthy and Li's original paper [EL94b] there are similar results for n-forms and the Hessian flow. Moreover, if the manifold is explosive, i.e. the manifold is noncompact, the lifetime ζ may be finite. Considerations of this type lead to the concepts of strong completeness [EL94a] in Elworthy and Li's point of view, but there are other approaches. Thalmaier overcomes this problem by generalizing the derivative process to a time derivative process taking values in T_xM in [Tha97]. In [TW98], Thalmaier and Wang obtained gradient estimates for harmonic functions on a regular open domain in M up to a constant based on the Ricci curvature. There are also extensions of Bismut's formula to vector bundles in [Hsu07] and [Hsu02].

As we pointed out in Chapter 2, we deal only with manifolds *without* boundary. Some considerations about smooth boundary problems can also be found in [Tha97] and [TW98].

Moreover, the classical BEL formula for diffusion processes has been proved to be a useful tool in various other aspects like strong Feller properties (cf. [Zha13, Section 4]) and functional inequalities. Therefore, it seems obvious to establish an analogous formula for jump-diffusion processes. In [CF06], Cass and Fritz extended the BEL formula to nondegenerate jump diffusions. Following [Zha13], we have shown that the BEL formula also holds for nonlinear SDEs driven by subordinated Brownian motion on $\mathbb{R}^n$. An open question is whether there is a generalisation for compact manifolds. An extension on $\mathbb{R}^n$ for the *multiplicative* case, namely $\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}B_{S_t}$, has recently been made in [WXZ13]. A major open topic is the extension of the BEL formula for more general types of Lévy processes.

Appendix A

Tools and Inequalities From Analysis & Stochastics

In this chapter, we gather some famous and useful result from stochastic and analysis. Proofs can be found e.g. in [SP14].

Typically, in order to show the continuity of a solution of a stochastic differential equation $x \mapsto X_t^x$, one uses the following famous result, known as the Kolmogorov-Chentsov theorem. Since we work with random fields, i.e. stochastic processes with a multi-dimensional index set, we state the general version.

A.1 Theorem (Kolmogorov 1934; Slutsky 1937; Chentsov 1956). *Denote by $(\xi(x))_{x\in\mathbb{R}^n}$ a stochastic process on $(\Omega, \mathscr{F}, \mathbb{P})$ with values in $\mathbb{R}^n$ and index set in $\mathbb{R}^n$. If*

$$\mathbb{E}\,|\xi(x) - \xi(y)|^\alpha \leqslant c\,|x-y|^{n+\beta} \qquad \text{for all } x, y \in \mathbb{R}^n, \tag{A.1}$$

holds for some constants $c > 0$ and $\alpha, \beta > 0$, then $(\xi(x))_{x\in\mathbb{R}^n}$ has a modification $\left(\tilde{\xi}(x)\right)_{x\in\mathbb{R}^n}$ with exclusively continuous sample paths.

The next theorem (from the theory of ordinary differential equations) is often used to show the uniqueness of the solution.

A.2 Theorem (Grönwall's inequality). *Let $u, a, b : [0,\infty) \to [0,\infty)$ be positive measurable functions satisfying the inequality*

$$u(t) \leqslant a(t) + \int_0^t b(s)u(s)\,\mathrm{d}s, \qquad \text{for each } t \geqslant 0.$$

Then

$$u(t) \leqslant a(t) + \int_0^t a(s)b(s)\,\exp\left(\int_0^t b(r)\,\mathrm{d}r\right)\,\mathrm{d}s, \qquad \text{for each } t \geqslant 0. \tag{A.2}$$

In particular, if $a(t) = a$ and $b(t) = b$ are constants, the estimate (A.2) reads

$$u(t) \leqslant ae^{bt}, \qquad \text{for each } t \geqslant 0. \tag{A.3}$$

The next Lemma is used in the proof of Proposition 3.11, Proposition 3.15 and Remark 3.20.

A.3 Lemma. *For any manifold M, there exist finitely many functions $\varphi^1, \ldots, \varphi^k \in C^\infty(M)$ such that:*

(i) For every $f \in C^\infty(M)$, there exists $\overline{f} \in C^\infty(\mathbb{R}^k)$ such that $f = \overline{f} \circ (\varphi^1, \ldots, \varphi^k)$.

(ii) *Every section* $b \in \Gamma(T^*M \otimes T^*M)$ *can be written uniquely as* $b = b_{ij}\mathbf{d}\varphi^i \otimes \mathbf{d}\varphi^j$, *where* $b_{ij} \in C^\infty(M)$.

(iii) *Every differential form* $\omega \in \Omega^1(M)$ *can be written uniquely as* $\omega = \omega_i \mathbf{d}\varphi^i$, *where* $\omega_i \in C^\infty(M)$.

(iv) *Let* $X \in \mathcal{S}(M)$. *Any continuous adapted* $(T^*M \otimes T^*M)$*-valued process* B *above* X, *i.e.* $B_t \in T^*_{X_t}M \otimes T^*_{X_t}M$ *for* $t \geqslant 0$, *admits a representation* $B = B_{ij}\mathbf{d}\varphi^i \otimes \mathbf{d}\varphi^j(X)$, *where* B_{ij} *are continuous adapted* $\mathbb{R}$*-valued processes.*

(v) *Let* $X \in \mathcal{S}(M)$. *Any continuous adapted* T^*M*-valued process* Ψ *above* X *admits a representation* $\Psi = \Psi_i\, \mathbf{d}\varphi^i(X)$, *where* Ψ_i *are continuous adapted* $\mathbb{R}$*-valued processes.*

Proof. We use the Whitney embedding $\iota : M \hookrightarrow \mathbb{R}^k$ so that M is a closed submanifold of $\mathbb{R}^k$. Then there exists a smooth partition of unity $(\chi_\lambda)_{\lambda\in\Lambda}$ on M, indexed by a set Λ, and a family $(I_\lambda)_\lambda$ of subsets $I_\lambda \subset \{1, ..., k\}$ such that: For every $\lambda \in \Lambda$ there is a chart $(\varphi^i)_{i\in I_\lambda}$ of M on an open neighbourhood of $\operatorname{supp} \chi_\alpha$ (cf. [Lee13, Chapter 2]).

(i) Define $f|_{\iota(M)} = \overline{f} \circ \iota$ and extend $\overline{f}$ locally around $M \cong \iota(M)$ such that $\overline{f} = 0$ on the normal space to T_pM. Last, multiply by a cut-off function which is identically 1 locally around $\iota(M)$ and vanishes on a neighbourhood large enough.

(ii) Let $\chi_\lambda b = b^\lambda_{ij}\mathbf{d}\varphi^i \otimes \mathbf{d}\varphi^j$ with $b^\lambda_{ij} \in C^\infty(M)$ such that $\operatorname{supp} b^\lambda_{ij} \subset \operatorname{supp} \chi_\lambda$ and $b^\lambda_{ij} := 0$ if $\{i, j\} \not\subset I_\lambda$. Then

$$b = b_{ij}\mathbf{d}\varphi^i \otimes \mathbf{d}\varphi^j \text{ where } b_{ij} := \sum_\lambda b^\lambda_{ij}.$$

(iii) This part is proved in the same way as assertion (ii).

(iv) As in (ii), we have $\chi_\lambda(X)B = B^\lambda_{ij}\mathbf{d}\varphi^i \otimes \mathbf{d}\varphi^j(X)$, where $B^\lambda_{ij} := \partial_i \otimes \partial_j(\chi_\lambda \circ X)$ are suitable continuous $\mathbb{R}$-valued processes with $\{i, j\} \subset I_\lambda$ and $B^\lambda_{ij} = 0$ if $\{i, j\} \not\subset I_\lambda$. By $B_{ij} = \sum_\lambda B^\lambda_{ij}$ the result follows.

(v) This part is proved in the same way as assertion (iv). ■

The following result is an easy consequence of Ito's isometry. It is used in the proof of the Bismut-Elworthy-Li formula 4.3.

A.4 Lemma. *Let* $(B_t)_{t\geqslant 0}$ *be an* $\mathbb{R}^n$*-valued Brownian motion and* $T > 0$. *Then for all processes* $f, g \in L^2([0, T] \times \Omega, \mathrm{Leb}|_{[0,T]} \otimes \mathbb{P})$ *which have a progressively measurable representative*

$$\mathbb{E}\left[\left(\int_0^t f(s)\,\mathrm{d}B_s\right) \cdot \left(\int_0^t g(s)\,\mathrm{d}B_s\right)\right] = \mathbb{E}\int_0^t f(s) \cdot g(s)\,\mathrm{d}s.$$

Proof. This follows by polarisation from Itô's isometry.

$$
\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t f(s)\,\mathrm{d}\boldsymbol{B}_s\right)\cdot\left(\int_0^t g(s)\,\mathrm{d}\boldsymbol{B}_s\right)\right]\\
&\quad=\frac{1}{4}\left[\mathbb{E}\left(\int_0^t (f(s)+g(s))\,\mathrm{d}\boldsymbol{B}_s\right)^2-\mathbb{E}\left(\int_0^t (f(s)-g(s))\,\mathrm{d}\boldsymbol{B}_s\right)^2\right]\\
&\quad=\frac{1}{4}\left(\mathbb{E}\int_0^t (f(s)+g(s))^2\,\mathrm{d}s-\mathbb{E}\int_0^t (f(s)-g(s))^2\,\mathrm{d}s\right)\\
&\quad=\mathbb{E}\int_0^t f(s)g(s)\,\mathrm{d}s.
\end{aligned}
$$

■

Appendix B

Quadratic Variation of B_{ℓ_t}

Let $(B_t)_{t\geqslant 0}$ a Brownian motion on $\mathbb{R}^n$ and $\ell \in \mathbb{S}$, where $\mathbb{S}$ is the space of all increasing càdlàg function from $(0,\infty)$ to $(0,\infty)$ with $\lim_{s\downarrow 0} = 0$. In Remark 6.6, we stated that then the quadratic variation of $(B_{\ell_t})_{t\geqslant 0}$ is given by

$$\left[B_\ell\right]_t = \ell_t - \sum_{0<s\leqslant t} \Delta \ell_s + \sum_{0<s\leqslant t} \left|\Delta B_{\ell_s}\right|^2 . \tag{B.1}$$

Recall that, if X, Y are semimartingales, where X^c, Y^c denotes their continuous martingale parts, we can then write the decomposition

$$[X,Y]_t = \langle X^c, X^c\rangle_t + \sum_{0<s\leqslant t} \Delta X_s \Delta Y_s.$$

where $\langle X^c, Y^c\rangle_t = [X,Y]_t^c$ denotes the continuous and $\sum_{0<s\leqslant t} \Delta X_s \Delta Y_s = [X,Y]_t^d$ the purely discontinuous part (cf. [Pro05, pp. 68-70]). The well-known fact that $\sum_{s\leqslant t} \Delta L_s^2 < \infty$ for any Lévy process $(L_t)_{t\geqslant 0}$ implies that $\sum_{s\leqslant t} \Delta B_{\ell_s}^2 < \infty$ for any $t \geqslant 0$. So heuristically, the decomposition (B.1) appears to be true: the square bracket composes in a purely continuous part which is well-known for Brownian motion and a purely discontinuous part. But, $s \mapsto \ell_s$ is càdlàg, so we have to subtract the jumps parts as well.

We have to calculate $[B_\ell]_t^c$. Denote by $J(\varepsilon)$ the jump times of ℓ up to time t with jump height exceeding ε

$$J(\varepsilon) := \{ s \in [0,t] \,:\, |\Delta \ell_s| \geqslant \varepsilon \} .$$

Since $s \mapsto \ell_s$ is càdlàg, J is a.s. finite for each $\varepsilon > 0$. Let $\Pi = \left\{0 = t_1 \leqslant \ldots \leqslant t_n = t\right\}$ be a partition of the interval $[0,t]$. We want to show that

$$[B_\ell]_t^c = \mathbb{P} - \lim_{|\Pi|\to 0} \sum_{j:[t_j,t_{j+1})\cap J(\varepsilon)=\varnothing} \left(B_{\ell(t_{j+1})} - B_{\ell(t_j)}\right)^2 = \ell_t - \sum_{0<s\leqslant t} \Delta \ell_s.$$

By the triangle inequality, we have

$$\mathbb{P}\left(\left|\sum_{j:[t_j,t_{j+1})\cap J(\varepsilon)=\varnothing} \left(B_{\ell(t_{j+1})} - B_{\ell(t_j)}\right)^2 - \left(\ell_t - \sum_{0<s\leqslant t} \Delta \ell_s\right)\right| > \delta\right) \leqslant I_1 + I_2,$$

where

$$I_1 := \mathbb{P}\left(\left| \sum_{j:[t_j,t_{j+1})\cap J(\varepsilon)=\varnothing} \left(B_{\ell(t_{j+1})} - B_{\ell(t_j)}\right)^2 - \sum_{j:[t_j,t_{j+1})\cap J(\varepsilon)=\varnothing} \left(\ell(t_{j+1}) - \ell(t_j)\right)\right| > \frac{\delta}{2}\right),$$

$$I_2 := \mathbb{P}\left(\left| \sum_{j:[t_j,t_{j+1})\cap J(\varepsilon)=\varnothing} \left(\ell(t_{j+1}) - \ell(t_j)\right) - \left(\ell_t - \sum_{0<s\leqslant t} \Delta \ell_s\right)\right| > \frac{\delta}{2}\right).$$

Using exactly the same argumentation as for the quadratic variation of a canonical Brownian motion, i.e. $\ell(t) \equiv t$, we find[1]

$$\mathbb{E}\left|\sum_{j:(t_{j-1},t_j]\cap J=\varnothing} (B_{\ell_{t_j}} - B_{\ell_{t_{j-1}}})^2 - \sum_{j:(t_{j-1},t_j]\cap J=\varnothing} (\ell_{t_j} - \ell_{t_{j-1}})\right|^2 \leqslant C \sup_{j:(t_{j-1},t_j]\cap J=\varnothing} (\ell_{t_{j-1}} - \ell_{t_j}),$$

for some constant $C > 0$ independent of Π and ε; the idea is basically to use the scaling property of Brownian motion, for more details cf. e.g. [SP14, Theorem 9.1]. Consequently,

$$\limsup_{|\Pi|\to 0} \mathbb{E}\left(\left|\sum_{j:(t_{j-1},t_j]\cap J=\varnothing} (B_{\ell_{t_j}} - B_{\ell_{t_{j-1}}})^2 - \sum_{j:(t_{j-1},t_j]\cap J=\varnothing} (\ell_{t_j} - \ell_{t_{j-1}})\right|^2\right) \leqslant C\varepsilon. \tag{B.2}$$

Applying Markov's inequality for $p = 2$, it follows just as in the continuous case that

$$\limsup_{|\Pi|\downarrow 0} I_1 \leqslant C\varepsilon \xrightarrow{\varepsilon \to 0} 0.$$

Again by the Markov inequality and triangle equality, we get

$$\begin{aligned}
&\mathbb{P}\left(\left|\sum_{j:[t_j,t_{j+1})\cap J(\varepsilon)=\varnothing} \left(\ell_{t_{j+1}} - \ell_{t_j}\right) - \left(\ell_t - \sum_{0<s\leqslant t} \Delta\ell_s\right)\right| > \frac{\delta}{2}\right) \\
&\quad \leqslant \frac{2}{\delta}\mathbb{E}\left|\sum_{j:[t_j,t_{j+1})\cap J(\varepsilon)=\varnothing} \left(\ell_{t_{j+1}} - \ell_{t_j}\right) - \left(\ell_t - \sum_{0<s\leqslant t,\Delta\ell_s\geqslant\varepsilon} \Delta\ell_s\right)\right| + \frac{2}{\delta}\mathbb{E}\sum_{\substack{0<s\leqslant t,\\ \Delta\ell_s<\varepsilon}} \Delta\ell_s. \\
&\quad \xrightarrow{|\Pi|\downarrow 0} \frac{2}{\delta}\sum_{\substack{0<s\leqslant t,\\ \Delta\ell_s<\varepsilon}} \Delta\ell_s \xrightarrow{\varepsilon\downarrow 0} 0.
\end{aligned}$$

[1] We would like to thank Franziska Kühn who helped elaborating the original argument.

Bibliography

[AE09] H. Amann and J. Escher. *Analysis III*. Basel; Berlin [u.a.]: Birkhäuser, 2009.

[App01] D. Applebaum. "Lévy processes in stochastic differential geometry". In: *Lévy Processes*. Springer, 2001, pp. 111–137.

[App09] D. Applebaum. *Lévy processes and stochastic calculus*. 2. ed. Cambridge [u.a.]: Cambridge University Press, 2009.

[Bis84] J.-M. Bismut. *Large deviations and the Malliavin calculus*. Vol. 45. Birkhauser, 1984.

[BSS03] K. Bogdan, Stósandrzej, and P. Sztonyk. "Harnack inequality for stable processes on d-sets". In: *Studia Math* 158.2 (2003), pp. 163–198.

[CF06] T. R. Cass and P. K. Friz. "The Bismut-Elworthy-Li formula for jump-diffusions and applications to Monte Carlo pricing in finance". In: *ArXiv Mathematics e-prints* (Apr. 2006). eprint: math/0604311v3.

[Cha84] I. Chavel. *Eigenvalues in Riemannian geometry*. Vol. 115. Academic press, 1984.

[EL94a] K. D. Elworthy and X.-M. Li. "Differentiation of heat semigroups and applications". In: *Probability Theory and Mathematics Stats (Proc. 6th Vilnius conferences)*. Ed. by B. G. et al. VSP/TEV. Vol. 6. In New Titles in Probability and Statistics Series. Utrecht and Vilnius., 1994, pp. 239–251.

[EL94b] K. D. Elworthy and X.-M. Li. "Formulae for the derivatives of heat semigroups". In: *Journal of Functional Analysis* 125 (1994), pp. 252–286.

[Elw06] K. D. Elworthy. *MA482 Stochastic Analysis*. Lecture Notes. Term 1, 2005–2006.

[Elw13] K. D. Elworthy. "Stochastic differential geometry at Saint-Flour Stochastic differential geometry at SaintFlour". In: *Probability at Saint-Flour*. Ed. by A. Ancona, K. D. [Elworthy, M. [Émery, and H. [Kunita. Probability at Saint-Flour. Springer, Heidelberg [u.a.]: Springer, 2013, pp. 113–262.

[Elw82] K. D. Elworthy. *Stochastic Differential Equations on Manifolds*. Vol. 70. Cambridge Lecture Notes Series. London; New York: Cambridge University Press, 1982.

[Elw92] K. D. Elworthy. *Stochastic flows on Riemannian manifolds*. Ed. by M. A. Pinsky and V. Wihstutz. Vol. 2. Progress in probability ; 27. Boston: Birkhäuser, 1992.

[ÉM89] M. Émery and P. A. Meyer. *Stochastic calculus in manifolds*. Berlin; Heidelberg; New York; London; Paris; Tokyo; Hong Kong: Springer, 1989.

[G+83] E. Gine, M. B. Marcus, et al. "The central limit theorem for stochastic integrals with respect to Lévy processes". In: *The Annals of Probability* 11.1 (1983), pp. 58–77.

[Gri09] A. Grigoryan. *Heat Kernel and Analysis on Manifolds*. Ed. by S.-T. Yau. Studies in Advanced Mathematics. AMS/IP, 2009.

[GW12] A. Guillin and F.-Y. Wang. "Degenerate Fokker-Planck equations: Bismut formula, gradient estimate and Harnack inequality". In: *Journal of Differential Equations* 253 (2012), pp. 20–40. DOI: 10.1016/j.jde.2012.03.014. arXiv: 1103.2817 [math.PR].

[Hsu02] E. P. Hsu. *Stochastic analysis on manifolds*. Providence, RI: American Mathematical Society, 2002.

[Hsu07] E. P. Hsu. "Heat equations on manifolds and Bismut's formula". In: *Stochastic Analysis and Partial Differential Equations, Contemporary Mathematics* 429 (2007), pp. 121–130.

[HT94] W. Hackenbroch and A. Thalmaier. *Stochastische Analysis: eine Einführung in die Theorie der stetigen Semimartingale*. Stuttgart: Teubner, 1994.

[Jac01] N. Jacob. *Pseudo differential operators and Markov processes*. Vol. 1. Dresden: Imperial College Press, 2001.

[Jän01] K. Jänich. *Vector analysis*. New York; Berlin; Heidelberg [u.a.]: Springer, 2001.

[KN63] S. Kobayashi and K. Nomizu. *Foundations of differential geometry*. Vol. 1. Interscience tracts in pure and applied mathematics ; 15. New York: Interscience Publ., 1963.

[KSV11] P. Kim, R. Song, and Z. Vondracek. "Potential Theory of Subordinate Brownian Motions Revisited". In: *ArXiv e-prints* (Feb. 2011). arXiv: 1102.1369 [math.PR].

[Küh10] W. Kühnel. *Differentialgeometrie: Kurven - Flächen - Mannigfaltigkeiten*. 5. Auflage. Vieweg+Teubner, 2010.

[Kun97] H. Kunita. *Stochastic flows and stochastic differential equations*. Vol. 24. Cambridge university press, 1997.

[Kus10] S. Kusuoka. "Malliavin calculus for stochastic differential equations driven by subordinated Brownian motions". In: *Kyoto J. Math.* 50.3 (2010), pp. 491–520. DOI: 10.1215/0023608X-2010-003.

[Léa88] R. Léandre. "Calcul des variations sur un brownien subordonné". In: *Séminaire de Probabilités XXII*. Springer, 1988, pp. 414–433.

[Lee05] J. M. Lee. *Riemannian manifolds: an introduction to curvature*. Springer, 2005.

[Lee13] J. M. Lee. *Introduction to Smooth Manifolds (Second Edition)*. Vol. 218. Graduate Texts in Mathematics. Springer, 2013.

[Mal02] P. Malliavin. *Stochastic analysis*. 1. Repr. Vol. 313. Berlin; Heidelberg [u.a.]: Springer, 2002.

[Mal78] P. Malliavin. *Géométrie différentielle stochastique*. Montreal, Presses de l'universite de Montreal, 1978.

[Øks07] B. Øksendal. *Stochastic Differential Equations. Sixth Editions*. Springer, 2007.

[Pla03] H. Plank. "Stochastic representation of the gradient and Hessian of diffusion semigroups on Riemannian manifolds". PhD thesis. 2003.

[Pro05] P. Protter. *Stochastic integration and differential equations*. 2nd edition. Berlin; Heidelberg; New York: Springer, 2005.

[PZ08] E. Priola and J. Zabczyk. "Structural properties of semilinear SPDEs driven by cylindrical stable processes". In: *ArXiv e-prints* (Oct. 2008). arXiv: 0810.5063 [math.AP].

[Rao04] M. M. Rao. *Real and stochastic analysis: new perspectives*. Birkhäuser, 2004.

[S+12] R. L. Schilling, P. Sztonyk, J. Wang, et al. "Coupling property and gradient estimates of Lévy processes via the symbol". In: *Bernoulli* 18.4 (2012), pp. 1128–1149.

[Sat13] K.-i. Sato. *Lévy processes and infinitely divisible distributions*. Rev. ed., corr. paperback ed. Cambridge [u.a.]: Cambridge University Press, 2013.

[Sch05] R. L. Schilling. *Measures, integrals and martingales*. Cambridge New York: Cambridge University Press, 2005.

[Shi04] I. Shigekawa. *Stochastic analysis*. Providence, RI: American Mathematical Society, 2004.

[SP14] R. L. Schilling and L. Partzsch. *Brownian motion: an introduction to stochastic processes*. 2nd revised edition. Berlin [u.a]: De Gruyter, 2014.

[SSV10] R. L. Schilling, R. Song, and Z. Vondracek. *Bernstein functions: theory and applications*. Berlin: de Gruyter, 2010.

[Tha97] A. Thalmaier. "On the differentiation of heat semigroups and Poisson integrals". In: *Stochastics: An International Journal of Probability and Stochastic Processes* 61.3-4 (1997), pp. 297–321.

[TW98] A. Thalmaier and F.-Y. Wang. "Gradient Estimates for Harmonic Functions on Regular Domains in Riemannian Manifolds". In: *Journal of Functional Analysis* 155.1 (1998), pp. 109–124. DOI: http://dx.doi.org/10.1006/jfan.1997.3220.

[Wan11a] F.-Y. Wang. "Gradient estimate for Ornstein-Uhlenbeck jump processes". In: *Stochastic Processes and their Applications* 121.3 (2011), pp. 466–478.

[Wan11b] F.-Y. Wang. "Derivative Formula and Harnack Inequality for Linear SDEs Driven by Lévy Processes". In: *ArXiv e-prints* (Apr. 2011). arXiv: 1104.5531 [math.PR].

[Wan14] F.-Y. Wang. *Analysis for diffusion processes on Riemannian manifolds*. New Jersey [u.a.]: World Scientific, 2014.

[Wil00] M. [Willem. *Development of mathematics 1950–2000*. Basel; Boston; Berlin: Birkhäuser, 2000.

[WR90] D. Williams and L. C. G. Rogers. *Diffusions, Markov processes and martingales - 2: Itô calculus*. Vol. 2. Chichester [u.a.]: Wiley, 1990.

[WW14] F.-Y. Wang and J. Wang. "Harnack inequalities for stochastic equations driven by Lévy noise". In: *Journal of Mathematical Analysis and Applications* 410.1 (2014), pp. 513–523.

[WXZ13] F.-Y. Wang, L. Xu, and X. Zhang. "Gradient estimates for SDEs Driven by Multiplicative Lévy Noise". In: *arXiv preprint arXiv:1301.4528* (2013).

[WZ11] F.-Y. Wang and X.-C. Zhang. "Derivative Formula and Applications for Degenerate Diffusion Semigroups". In: *ArXiv e-prints* (June 2011). arXiv: 1107.0096 [math.PR].

[Zha10] X. Zhang. "Stochastic flows and Bismut formulas for stochastic Hamiltonian systems". In: *Stochastic Processes and their Applications* 120.10 (2010), pp. 1929–1949.

[Zha12] X. Zhang. "Stochastic functional differential equations driven by Lévy processes and quasi-linear partial integro-differential equations". In: *The Annals of Applied Probability* 22.6 (2012), pp. 2505–2538.

[Zha13] X. Zhang. "Derivative formula and gradient estimates for SDEs driven by α-stable processes". In: *Stochastic Processes and their Applications* 123.4 (2013), pp. 1213–1228.

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die am heutigen Tag eingereichte Diplomarbeit zum Thema

"The Bismut-Elworthy-Li Formula and Gradient Estimates for Stochastic Differential Equations"

unter der Betreuung von

Prof. Dr. rer. nat. René Schilling

selbstständig erarbeitet, verfasst und Zitate kenntlich gemacht habe. Andere als die angegebenen Hilfsmittel wurden von mir nicht benutzt.

Datum | Unterschrift