MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Stochastic Analysis of Brownian Motion on Riemannian Manifolds: Characterizations and Constructions"

verfasst von / submitted by

Aymane Amine BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2024 / Vienna, 2024

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Mathematik
Betreut von / Supervisor:	Ass.-Prof. Dr. Julio Daniel Backhoff-Veraguas
Mitbetreut von / Co-Supervisor:	

Acknowledgements

Firstly, I would like to thank Julio Backhoff for entrusting me with the freedom to investigate a topic of my choice. His role as a supervisor has been pivotal, providing huge support and motivation that turned out to be necessary for the completion of this work.

Secondly, I would like to express my deepest gratitude to Professor Abdol-Reza Mansouri from Queens University in Kingston, Ontario, for his invaluable guidance, unwavering support, meticulously proof-checking of the thesis in the last stages, and insightful feedback throughout my entire research process. His expertise and tutoring in the subject of Stochastic Differential Geometry have been instrumental in shaping the trajectory of this thesis, and his expertise in different fields of mathematics has been a source of inspiration during my research.

Likewise, I am immensely thankful to Professor Elton P.Hsu from Northwestern University for taking the time to answer my questions and proof-check my thesis at the end, his constructive comments and thoughtful suggestions, have immensely enriched the quality of this thesis as well as helped me gain a deeper understanding of the material. Last but not least, gratitude extends to my brother, whose brilliant insights emerged as a cornerstone for the existence of this thesis.

My sincere appreciation extends to the University of Vienna as an establishment for being my home for the last couple of years. I would also like to thank my professors at the University of Vienna who were a source of inspiration and a reference point for me, to name a few: Günther Hörmann for his competence and his unlimited support, Michael Eichmair for his exceptional teaching style and his positive energy, Roland Donninger for his outstanding teaching skills, Bernhard Lamel for providing us a solid foundation in Analysis, and Nathanael Berestycki for his competency, empathy, and joy at teaching. I am grateful to have had you all as instructors. Occasional financial support for third-world students received from ÖH Sozialfonds is acknowledged with the highest level of gratitude.

I owe a debt of gratitude to my family, for their unwavering support and encouragement throughout this academic journey. Lastly, I want to emphasize, that the love of my lovely family as well as the kind professors I met has had a crucial effect throughout my years abroad.

I would like to express my heartfelt thanks to my old friends at the University of Vienna, the same goes for all the nice friends I made during my Erasmus in Trondheim, they played a significant role in overcoming all sorts of challenges and making my stay in Trondheim enjoyable, despite the cold-dark Nordic weather!

Lastly, the enriching experience gained from living and working alongside my studies, especially in the culturally rich city of Vienna, deserves acknowledgment. Therefore, I would like to take the occasion to confess my eternal love for Vienna, its streets, buildings, and cafes.

Abstract

This thesis explores a topic at the intersection of stochastic calculus and Riemannian geometry, namely, it investigates constructions of Brownian motion on non-flat geometries. To this end, we study extensions of results and techniques from stochastic calculus in the Euclidean space setting into their analogues on Riemannian manifolds and use these to define and study Brownian motion on compact Riemannian manifolds. The thesis is organized as follows: In the first chapter, we introduce some basic concepts and results related to compact Riemannian manifolds. Chapter 2 introduces the necessary concepts from stochastic calculus in $\mathbb{R}^d$. The third chapter is the main chapter and starts with an introduction to frame bundles, horizontal lifts, anti-developments, and their use in the context of stochastic differential equations on manifolds. We then study a fundamental construction of Brownian motion on a Riemannian manifold termed the Eells-Elworthy construction. Furthermore, we compare the description of Brownian motion by the method mentioned above with simpler descriptions in the case where the manifold is the sphere. We conclude the chapter by giving an equivalent characterization of Brownian motion on a manifold using the Laplace-Beltrami operator. We show that our previous constructions of Brownian motion on the sphere and torus fulfill it. Finally, we complete our study by examining the concept of Gamma martingales on a manifold and give an example of a Γ-martingale on the torus.

Kurzfassung

Das ist eine deutsche Kurzfassung meiner in Englisch verfassten Masterarbeit.:

In dieser Masterarbeit handelt es sich um ein Thema an der Schnittstelle von stochastischem Analysis und Riemannscher Geometrie, nämlich diese Arbeit untersucht Konstruktionen des Brownschen Bewegung auf nicht-flachen Geometrien. Zu diesem Zweck untersuchen wir Erweiterungen von Ergebnissen und Techniken aus dem Stochastischen Analysis im euklidischen setting in ihre Analoga auf Riemannschen Mannigfaltigkeiten und verwenden diese, um Brownsche Bewegung auf kompakten Riemannschen Mannigfaltigkeiten zu definieren und zu untersuchen. Die Arbeit ist wie folgt strukturiert:

Im ersten Kapitel führen wir die wesentlichen Konzepte und Ergebnisse im Zusammenhang mit kompakten Riemannschen Mannigfaltigkeiten ein. Kapitel 2 führt einige grundlegende Konzepte aus dem stochastischen Analysis in $\mathbb{R}^d$ ein. Das dritte Kapitel ist das Hauptkapitel und beginnt mit einer Einführung in Rahmenbündel, horizontale Lifts, Anti-abwicklungen und ihre Verwendung im Kontext stochastischer Differentialgleichungen auf Mannigfaltigkeiten. Anschließend untersuchen wir eine grundlegende Konstruktion der Brownschen Bewegung auf einer Riemannschen Mannigfaltigkeit, die als Eells-Elworthy-Konstruktion bezeichnet wird. Darüber hinaus vergleichen wir die Beschreibung der Brownschen Bewegung mit der oben genannten Methode mit einfacheren Beschreibungen im Fall, dass die Mannigfaltigkeit die Sphäre oder der Torus ist. Wir schließen das Kapitel ab, indem wir eine äquivalente Charakterisierung der Brownschen Bewegung auf einer Mannigfaltigkeit unter Verwendung des Laplace-Beltrami-Operators geben. Wir zeigen, dass unsere vorherigen Konstruktionen der Brownschen Bewegung auf der Sphäre und dem Torus dies erfüllen. Schließlich vervollständigen wir unsere Untersuchung, indem wir das Konzept der Gamma-Martingale auf einer Mannigfaltigkeit untersuchen und ein Beispiel für ein Γ-Martingal auf dem Torus schildern.

Contents

1. Notions of Riemannian Geometry

1.1. Preliminaries and notations

In this section, we collect the necessary tools and introduce the notation used to work with differentiable manifolds. Most of the results are standard and can be found in e.g. [Mil63], [GN12], [KS17] and [Lee00], so most of the results will be presented without proofs.

1.1.1. Differentiable Manifolds and differentiable maps

Recall that an n-dimensional topological manifold is a Hausdorff topological space with a countable basis of open sets such that each point possesses a neighborhood homeomorphic to an open subset of $\mathbb{R}^n$. Each pair (U, φ), where U is an open subset of $\mathbb{R}^n$ and $\varphi : U \to \varphi(U) \subset M$ is a homeomorphism of U to an open subset of M, is called a parametrization. $(\varphi(U), \varphi^{-1})$ is called a coordinate system or chart, and the set $\varphi(U) \subset M$ is called a coordinate neighborhood.

Definition 1.1.1. *An n-differentiable or smooth manifold M is a topological manifold of dimension n and a family of parametrizations $\varphi_\alpha : U_\alpha \to M$ defined on open sets $U_\alpha \subset \mathbb{R}^n$, such that:*

1. *the coordinate neighborhoods cover M i.e. $M = \bigcup_\alpha \varphi_\alpha(U_\alpha)$;*
2. *for each pair of indices α, β such that*

$$W := \varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta) \neq \varnothing,$$

   The overlap maps

$$\begin{aligned} \varphi_\beta^{-1} \circ \varphi_\alpha &: \varphi_\alpha^{-1}(W) \to \varphi_\beta^{-1}(W) \\ \varphi_\alpha^{-1} \circ \varphi_\beta &: \varphi_\beta^{-1}(W) \to \varphi_\alpha^{-1}(W) \end{aligned}$$

   are C^∞;

3. *the family $\mathcal{A} = (U_\alpha, \varphi_\alpha)$ is maximal with respect to (1) and (2), meaning that if $\varphi_0 : U_0 \mapsto M$ is a parametrization such that $\varphi_0^{-1} \circ \varphi$ and $\varphi^{-1} \circ \varphi_0$ are C^∞ for all φ in $\mathcal{A}$, then $((U_0, \varphi_0))$ is in $\mathcal{A}$*

Remark 1.1.1. *Any family $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}$ that satisfies (1) and (2) is called a C^∞-atlas for M. If $\mathcal{A}$ also satisfies (3) it is called a maximal atlas or a differentiable structure.*

Definition 1.1.2. *Let M and N be two differentiable manifolds of dimension m and n, respectively. A map $f : M \mapsto N$ is said to be differentiable or smooth at a point $p \in M$ if there exist parametrizations (U, φ) of M at p (i.e. $p \in \varphi(U)$) and (V, ψ) of N at $f(p)$, with $f(\varphi(U)) \subset \psi(V)$, such that the map*

$$\psi^{-1} \circ f \circ \varphi : U \subset \mathbb{R}^m \to \mathbb{R}^n$$

is smooth i.e., C^∞.
The map f is said to be differentiable on a subset of M if it is differentiable at every point of this set

Definition 1.1.3. *A bijection $f : M \to N$ is a diffeomorphism if it is C^∞ and its inverse is C^∞*

Remark 1.1.2. *(Einstein summation convention) A repeated dummy index that appears twice in the same term — once as a subscript and once as a superscript always indicates a tacit summation over the contextual range of that index. We illustrate this convention with the following example. The range of summation is $\{1, 2, \dots n\}$ and the following expression*

$$v^i \partial_i, \omega_i dx^i, g^{ij}(\omega.\omega_i)\omega_j$$

stands for

$$\sum_{i=1}^{n} v^i \partial_i, \sum_{i=1}^{n} \omega_i dx^i, \sum_{i,j=1}^{n} g^{ij}(\omega.\omega_i)\omega_j$$

1.1.2. Tangent space

From elementary vector calculus, we know that a vector $v \in \mathbb{R}^3$ is said to be tangent to a surface $S \subset \mathbb{R}^3$ at a point $p \in S$ if there exist a differentiable curve $c : (-\epsilon, \epsilon) \mapsto S \subset \mathbb{R}^3$ such that $c(0) = p$ and $\dot{c}(0) = v$. The set T_pS of all these vectors is a 2-dimensional vector space, called the tangent space to S at p, and can be identified with a two-dimensional subspace of $\mathbb{R}^3$. To generalize this to an n-dimensional abstract manifold we need to find a description of v which does not involve the ambient space $\mathbb{R}^3$. To do so, we notice that the components of v are

$$v^i = \frac{d(x^i \circ c)}{dt}(0),$$

where $x^i : \mathbb{R}^3 \mapsto \mathbb{R}$ is the i-th canonical coordinate function. If we ignore the ambient space, $x^i : S \to \mathbb{R}$ is just a differentiable function, and $v^i = v(x^i)$, where for any differentiable function $f : S \mapsto \mathbb{R}$, we define

$$v(f) := \frac{d(f \circ c)}{dt}(0)$$

This notation suggests that we look at v as an operator acting on smooth functions i.e. $v : C^\infty(S) \to \mathbb{R}$

Definition 1.1.4. *let $c : (-\epsilon, \epsilon) \mapsto M$ be a differentiable curve on a smooth manifold M. Consider the set $C^\infty(p)$ of all functions $f : M \to \mathbb{R}$ that are differentiable at $c(0) = p$. The tangent vector to the curve c at p is the operator $\dot{c}(0) : C^\infty(p) \to \mathbb{R}$ given by*

$$\dot{c}(0)(f) = \frac{d(f \circ c)}{dt}(0).$$

The tangent space at p is the space T_pM of all tangent vectors at p, moreover, T_pM has a canonical real vector space structure induced by the C^∞ structure of M.
Choosing a reparametrization $\varphi : U \subset \mathbb{R}^n \to M$ around p, the curve c is given in local coordinates by the curve in U

$$\hat{c}(t) := \left(\varphi^{-1} \circ c\right)(t) = \left(x^1(t), \ldots, x^n(t)\right),$$

and the operator $\dot{c}(0) = \sum_{i=1}^{n} \dot{x}^i(0) \left(\frac{\partial}{\partial x^i}\right)_p$, where $\left(\frac{\partial}{\partial x^i}\right)_p$ denotes the operator associated to the vector tangent to the curve c_i at p given in local coordinates by

$$\hat{c_i}(t) = \left(x^1, \ldots, x^{i-1}, x^i + t, x^{i+1}, \ldots, x^n\right), \text{ with } \left(x^1, \ldots, x^n\right) = \varphi^{-1}(p).$$

Then,

$$\begin{aligned}
\dot{c}(0)(f) &= \frac{d(f \circ c)}{dt}(0) \\
&= \frac{d}{dt}\left(\overbrace{(f \circ \varphi)}^{\hat{f}} \circ \overbrace{(\varphi^{-1} \circ c)}^{\hat{c}}\right)_{|t=0} \\
&= \frac{d}{dt}\left(\hat{f}\left(x^1(t), \ldots, x^n(t)\right)\right)_{|t=0} = \sum_{i=1}^{n} \frac{\partial \hat{f}}{\partial x^i}\left(\hat{c}(0)\right) \frac{dx^i}{dt}(0) = \\
&= \left(\sum_{i=1}^{n} \dot{x}^i(0) \left(\frac{\partial}{\partial x^i}\right)_{\varphi^{-1}(p)}\right)(f \circ \varphi).
\end{aligned}$$

For the last equality we used $\hat{c}(0) = \left(\varphi^{-1} \circ c\right)(0) = \varphi^{-1}(c(0)) = \varphi^{-1}(p)$.

Remark 1.1.3.

1. *The basis $\left\{\left(\frac{\partial}{\partial x^i}\right)_p\right\}_{i=1}^{n}$ determined by the chosen parameterization around p is called the associated basis to that parameterization.*

2. *The definition of the tangent space at p only uses functions that are differentiable on a neighborhood of p. Hence, if U is an open set of M containing p, the tangent space T_pU is naturally identified with T_pM*

Definition 1.1.5.

The tangent bundle is defined as the disjoint union of all tangent spaces

$$TM = \bigcup_{p\in M} T_pM = \{v \in T_pM \mid p \in M\}$$

it admits a differentiable structure naturally determined by M. Moreover, it is equipped with a natural projection map $\pi : TM \to M$ which takes $v \in T_pM$ to p. Furthermore, (TM, M, π) has a real vector bundle structure.

After defining tangent spaces, we can define the derivative at a point p of a differentiable map $f : M \to N$ between smooth manifolds. Clearly, we want this derivative to be a linear transformation

$$(df)_p : T_pM \to T_{f(p)}N$$

of the corresponding tangent spaces so as to agree with the usual derivative of f when M and N are Euclidean spaces, and to satisfy the chain rule.

Definition 1.1.6. *Let $f : M \to N$ be a differentiable map between smooth manifolds. For $p \in M$, the derivative of f at p is the map*

$$\begin{aligned}(df)_p : T_pM &\to T_{f(p)}N \\ v &\mapsto \overset{\cdot}{\overbrace{f \circ c}}(0)\end{aligned}$$

where $c : (-\epsilon, \epsilon) \to M$ is a curve satisfying $c(0) = p$ and $\dot{c}(0) = v$

Definition 1.1.7. *(Vector fields) A vector field on a smooth manifold M is a map that to each point $p \in M$ assigns a vector tangent to M at p:*

$$\begin{aligned}X : M &\to TM \\ p &\mapsto X(p) := X_p \in T_pM\end{aligned}$$

The vector field is said to be differentiable if this map is differentiable. The set of all differentiable vector fields on M is denoted by $\mathfrak{X}(M)$. Locally we have:

Proposition 1.1.1. *Let W be a coordinate neighborhood on M (that is, $W = \varphi(U)$ for some parametrization $\varphi : U \to M$), and let $x := \varphi^{-1} : W \to \mathbb{R}^n$ be the corresponding coordinate chart. Then a map $X : W \to TW$ is a differentiable vector field on W if and only if,*

$$X_p = a^1(p)\left(\frac{\partial}{\partial x^1}\right)_p + \cdots + a^n(p)\left(\frac{\partial}{\partial x^n}\right)_p$$

for some differentiable functions $a^i : W \to \mathbb{R}$ for $i = 1, \ldots, n$.

Definition 1.1.8. *A vector field defines a derivation on $C^\infty(TM)$ as follows:*

$$(Xf)_p = X_pf$$

The Lie bracket is the vector field corresponding to the derivation given by:

$$[X, Y]f = X(Yf) - Y(Xf),$$

for $f \in C^\infty(M)$

Definition 1.1.9. *Locally, the Lie bracket of two vector fields $X, Y \in \mathfrak{X}(M)$ is the vector field such that*

$$[X_\varphi, Y_\varphi] = X_\varphi Y_\varphi^j - Y_\varphi X_\varphi^j = \sum_{i=1}^{n} (X_\varphi^i \partial_i Y_\varphi^j - Y_\varphi^i \partial_i X_\varphi^j)$$

for every chart $\varphi : U \to W$ of M.

Definition 1.1.10. *A map $f : M \to N$ is an immersion if its tangent map*

$$(df)_p : T_pM \to T_{f(p)}N$$

is injective function at every point $p \in M$.

Definition 1.1.11. *A map $f : M \to N$ is an embedding if it is an immersion and a homomorphism onto its image*

Definition 1.1.12. *An embedded submanifold of M is a subset $S \subseteq M$ endowed with a smooth structure with respect to which the inclusion map $S \hookrightarrow M$ is a smooth embedding*

Proposition 1.1.2. *Let $M' \subset M$ be a submanifold (For a thorough definition of submanifolds see [Lee00] chapter 5. A prototype of this concept is for instance $\mathbb{S}^n$, the n-dimensional sphere, being a smooth submanifold of $\mathbb{R}^n$). Let $p \in M'$, as a subspace of T_pM, the tangent space of M' is given by:*

$$T_pM' = \{v \in T_pM | vf = 0 \textit{ whenever } f \in C^\infty(M) \textit{ and } f|_{M'} = 0\}$$

So a vector field $X \in \mathfrak{X}(M)$ is tangent to a submanifold M' if and only if, $X_p \in T_pM'$ for every $p \in M'$ and $Xf = 0$ for every $f \in C^\infty(M)$ that vanishes on M'.

1.1.3. Tensors and Tensor fields

Let V be an n-dimensional vector space. A k-tensor on V is a real multilinear function defined on the product $V \times \dots V$ of k copies of V. The set of k-tensors is itself a vector space and is usually denoted by $\mathcal{T}^k(V^*)$ where V^* is the dual vector space. T_p^*M is the dual of the tangent space T_pM at a point $p \in M$, usually called the cotangent space to M at p.

Example 1.1.1. *The space of 1-tensors $\mathcal{T}^1(V^*)$ is equal to V^*, the dual space of V that is the space of real-valued linear functions on V.*

We can also consider mixed (k, m)-tensors on V, that is, multilinear functions defined on the product $V \times \cdots \times V \times V^* \times \cdots \times V^*$ of k-copies of V and m-copies of V^*. A (k, m)-tensor is then k times covariant and m-times contravariant on V. The space of all (k, m)-tensors on V is denoted by $\mathcal{T}^{k,m}(V^*, V)$.

Definition 1.1.13. *A (k, m)-tensor field is a map that to each point $p \in M$ assigns a tensor $T \in \mathcal{T}^{k,m}(T_p^*M, T_pM)$*

Remark 1.1.4. *In smooth local coordinates (x^i) any (r,s)-tensor can be written locally in the form*

$$T = T^{i_1\cdots i_s}_{j_1\cdots j_r}\left(\frac{\partial}{\partial x^{i_1}}\right)_p \otimes \cdots \otimes \left(\frac{\partial}{\partial x^{i_s}}\right) \otimes (dx^{j_1})_p \otimes \cdots \otimes (dx^{j_r})_p$$

1.2. Riemannian Manifolds

In the classical setting of differential geometry of surfaces in $\mathbb{R}^n$ one finds that all geometric properties of the surface in question depend on the scalar product generated on the tangent spaces by the standard scalar product of the ambient Euclidean space. In this section we will generalize these notions to the setting of abstract manifolds and hence define the notion of Riemannian metric for this, we will also provide some basic examples along the way.

Definition 1.2.1. *A Riemannian metric g on a smooth manifold M is a symmetric positive definite smooth covariant 2-tensor field g. Thus, it is a smooth assignment of an inner product*

$$g_x : T_xM \times T_xM \to \mathbb{R}$$

to each of the tangent spaces T_xM of M. In other words, for each $x \in M, g = g_x$ is bilinear and satisfies

1. *$g(u,v) = g(v,u)$ for all $u, v \in T_xM$.*
2. *$g(u,u) \geq 0$ for all $u \in T_xM$.*
3. *$g(u,u) = 0$ if and only if $u = 0$.*

Furthermore, g is smooth in the sense that for any smooth vector fields X and Y, the function $x \mapsto g_x(X_x, Y_x)$ is smooth. In local coordinates $\{x^i\}$ on V, we can express the 2-tensor field g as

$$g|_V = g_{ij}dx^i \otimes dx^j$$

Locally, a metric can be described in terms of its coefficients in a local chart, defined by $g_{ij} = g(\partial_i, \partial_j)$. The smoothness of g is equivalent to the smoothness of all the coefficient functions g_{ij} in some chart. If (V, φ^{-1}) is a chart of M with coordinates $\varphi^{-1} = (x^1, \ldots, x^n)$ and natural basis vector fields $\partial_i = \dfrac{\partial}{\partial x^i}$ we write

$$g_{ij} = \langle \partial_i, \partial_j \rangle \quad (1 \leq i, j \leq n)$$

for the local components of g on V.

Example 1.2.1 (The standard inner product on Euclidean space)**.**

$$\langle v_p, w_p \rangle = v \cdot w = \sum_i v^i w^i$$

We will always consider $\mathbb{R}^n$ equipped with this Riemannian metric.

Example 1.2.2. *An important situation where a manifold can be given a Riemannian metric is when it is a submanifold of some Euclidean space $\mathbb{R}^n$. The inclusion into $\mathbb{R}^n$ gives a natural identification of tangent vectors to M with vectors in $\mathbb{R}^n$: Explicitly, if we write the inclusion as a map $\iota : M \hookrightarrow \mathbb{R}^n$, then we identify a vector $u \in TM$ with its image $d\iota(u)$ under the differential of the inclusion. The inner product on $\mathbb{R}^n$ can then be used to induce a Riemannian metric on M, by defining*

$$g(u, v) = \langle d\iota(u), d\iota(v)\rangle.$$

Having built the notion of a Riemannian metric we can now talk about lengths and distances. In fact, a Riemannian metric allows us to compute the length $\|v\| = \langle v, v\rangle^{\frac{1}{2}}$ of any vector $v \in TM$ (as well as the angle between two vectors with the same base point). Therefore we can measure the length of curves:

Definition 1.2.2. *If $(M, \langle\cdot,\cdot\rangle)$ is a Riemannian manifold and $c : [a, b] \to M$ is a differentiable curve, the length of c is defined by:*

$$l(c) = \int_a^b \|\dot{c}(t)\| dt.$$

Proposition 1.2.1. *"[And], Proposition 9.3.2" If M is a Riemannian manifold with metric g. Then M is a metric space with the distance function d defined as above. The metric topology agrees with the manifold topology.*

A Riemannian metric determines an inner product on each tangent space T_xM, it is written as $\langle v, w\rangle := g_x(v, w)$ for $v, w \in T_xM$. Now if we look at g as a map, g provides a natural isomorphism between the tangent space TM and its dual T^*M. In the following definition, we define this correspondence between TM and T^*M.

Definition 1.2.3. *In the literature, the map from $TM \to T^*M$ is usually named the **Flat operator** $\flat$. More precisely, given $X \in TM$ we define:*

$$X^\flat := g(X, .)$$

*Similarly, the map from $T^*M \to TM$ is usually named the **Sharp operator** $\sharp$. More precisely, given $\theta \in T^*M$ (A 1-form for instance), we define:*

$$\theta := g(\theta^\sharp, .)$$

Remark 1.2.1. *The most important $\sharp$-operation is the generalization of the gradient as a vector field on manifolds. For instance, for all $f \in C^\infty(M)$, we define a new vector field which we call the gradient of f as*

$$\vec{\nabla} f := \operatorname{grad} f := df^\sharp$$

This says the gradient of f is the unique, dual of the differential (1-form) df defined by the relation:

$$\langle \operatorname{grad} f, .\rangle = df(.)$$

In particular:

$$\langle \operatorname{grad} f, X\rangle = df(X) = X(f)$$

for all vector fields $X \in \mathfrak{X}(M)$.
The fact that it is unique can be justified as follows: The mapping $X \mapsto Xf$ is a linear functional on T_xM, there exists a unique vector $v \in T_xM$ such that this linear functional is given by $\langle v, X\rangle$. Thus, we set $\vec{\nabla} f|_x := v$
In smooth local coordinates, the gradient takes the expression:

$$\vec{\nabla} f = g^{ij}\frac{\partial f}{\partial x^i}\frac{\partial}{\partial x^j}$$

When $M = \mathbb{R}^n$ with Euclidean metric:

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

Here the gradient of f is thought of as a (Column) vector described in the basis $\left\{\frac{\partial}{\partial x^j}\right\}$.

Our main aim in the remainder of this chapter is to define the notion of geodesics, the Riemannian generalization of the notion of straight lines in a Euclidean space. To achieve this we first need to define affine connections, which provide a notion of parallelism of vectors along curves, and consequently of geodesics (curves whose tangent vector is parallel).

1.3. Covariant Differentiation

Definition 1.3.1. *Let M be a smooth manifold. An affine connection at a point $p \in M$ is a function that assigns to each tangent vector $X_p \in TM$ and to each vector field Y a new tangent vector*

$$\nabla_{X_p} Y \in T_pM$$

called the covariant derivative of Y in the direction X_p. This is required to be bi-linear as a function of X_p and Y. Furthermore, if

$$f : M \to \mathbb{R}$$

is a real-valued function, and if fY denotes the vector field

$$(fY)_q = f(q)Y_q$$

then the covariant derivative of fY in the direction X_p is required to satisfy the identity

$$\nabla_{X_p}(fY) = (X_pf)Y_p + f(p)\nabla_{X_p} Y$$

(As usual, X_pf denotes the directional derivative of f in the direction X_p).

Definition 1.3.2. *A* ***global affine*** *connection (or briefly a* ***connection****) on M is a function which assigns to each $p \in M$ an affine connection ∇_p at p, satisfying the following smoothness condition:*

1. *If X and Y are smooth vector fields on M then the vector field $\nabla_X Y$, defined by the identity*
$$(\nabla_X Y)_p = \nabla_{X_p} Y$$
must also be smooth.

2. *$\nabla_X Y$ is bilinear as a function of X and Y*

3. *$\nabla_{(fX)} Y = f(\nabla_X Y)$*

4. *$\nabla_X(fY) = (Xf)Y + f(\nabla_X Y)$.*
Conditions (1), (2), (3), (4) can be taken as the definition of a ***connection****.*

Remark 1.3.1. *In terms of local coordinates $u^1, \ldots, u^n$ defined on a coordinate neighborhood $U \subset M$, the connection ∇ is determined by n^3 smooth real valued functions Γ_{ij}^k on U (Also called the Christoffel symbols of the connection ∇), as follows. Let ∂_k denote the vector field $\dfrac{\partial}{\partial u^k}$ on U. Then any vector field X on U can be expressed uniquely as*

$$X = \sum_{k=1} a^k \partial_k \tag{1.1}$$

where the a^k are real valued functions on U. In particular, the vector field $\nabla_{\partial_i} \partial_j$ can be expressed as

$$\nabla_{\partial_i} \partial_j = \sum_{k=1} \Gamma_{ij}^k \partial_k \tag{1.2}$$

These functions Γ_{ij}^k determine the ***connection*** *completely on U. In fact, given vector fields $X = \sum a^i \partial_i$ and $Y = \sum b^j \partial_j$, one can expand $\nabla_X Y$ by the rules $(2), (3)$ in the following way:*

$$\nabla_X Y = \sum_i a^i \nabla_{\partial_i} \left[\sum_j b^j \partial_j \right]$$

Now using the rule (4), the right hand side becomes

$$\sum_{i,j} a^i \partial_i b^j \partial_j + a^i b^j \left[\nabla_{\partial_i} \partial_j \right]$$

and applying 1.2, we get

$$\sum_{i,j} a^i \partial_i b^j \partial_j + \sum_{i,j,k} a^i b^j \left[\Gamma_{ij}^k \partial_k \right]$$

Changing the index j to k in the first sum, we get the formula

$$\nabla_X Y = \sum_{k=1}^{n} \left(\sum_{i,j} a^i \partial_i b^j + a^i b^j \Gamma_{ij}^k \right) \partial_k$$

Conversely, given any smooth real-valued functions Γ_{ij}^k on U, one can define $\nabla_X Y$ by the above equation. The result clearly satisfies the conditions $(1), (2), (3), (4), (5)$.

Intuition 1.3.1. *To establish some intuition, let us consider the situation in $\mathbb{R}^n$. Let $X, Y \colon \mathbb{R}^n \to \mathbb{R}^n$ be vector fields. Let c be a smooth curve $c(t) := p + tX_p$, such that $c(0) = p$ and $\dot{c}(0) = X_p$, the standard definition of the directional derivative of the vector field X in the direction of the vector field Y at a point $p \in \mathbb{R}^n$ goes as follows*

$$(\nabla_X Y)_p := \lim_{t \to 0} \frac{Y(c(t)) - Y_p}{t}.$$

The result is a vector field on $\mathbb{R}^n$, and the operation ∇ defined above, clearly satisfies the conditions $(1), (2), (3), (4), (5)$, where $f \colon \mathbb{R}^n \to \mathbb{R}$ is a scalar function, and the function Yf "At a point p" is the directional derivative of f at p in the direction Y_p.

Now, if we try to proceed in the same manner on a general manifold M, we face an issue:

1. *We need to subtract $Y_p \in T_pM$ from $Y(c(t)) \in T_{c(t)}M$, while each tangent vector lies in a different vector space.*

Therefore, without extra structure, there is no way of identifying tangent spaces at different points of M in a canonical way. The definition of an affine connection supplies the manifold M with additional structure, in other words, an operation $\nabla \colon \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ which satisfies $(1), (2), (3), (4), (5)$ and mimic the behaviour of the familiar derivative as alluded to at the beginning. As a result, this operation enables us to differentiate vector fields along tangent vectors. Towards the end of this chapter, we will give some possible choices of affine connections as well as provide concrete examples by calculating their respective Christoffel symbols.

Using a **connection**, the covariant derivative of a vector field along a curve in M can be defined as below.

Definition 1.3.3. *A vector field defined along a differentiable curve $c : I \to M$ is a differentiable map $V : I \to TM$ such that $V(t) \in T_{c(t)}M$ for all $t \in I$, and let $\dot{c}(t)$ be the tangent vector. If V is a vector field defined along the differentiable curve $c : I \to M$ with $\dot{c} \neq 0$, its covariant derivative along c is the vector field defined along c given by*

$$\frac{DV}{dt}(t) := \nabla_{\dot{c}(t)} V = \nabla_{\dot{c}(t)} Y \tag{1.3}$$

for any vector fields $X, Y \in \mathfrak{X}(M)$ such that $X_{c(t)} = \dot{c}(t)$ and $Y_{c(s)} = V(s)$, with $s \in (t - \varepsilon, t + \varepsilon)$ for some $\varepsilon > 0$. Note that the definition of the covariant derivative makes sense only if $\dot{c}(t)$ were a vector field.

Remark 1.3.2. *The covariant derivative of a vector field with respect to (In the direction of) a tangent vector $\dot{c}(t)$ is well defined, however, we would like also to be able to take the covariant derivation of a vector field along a vector field, to do so we can extend vector fields along curves to vector fields on the ambient manifold as this can always be done, as shown in1.1. It follows that, the result $(\nabla_X Y)$ does not depend on the chosen extension X of the tangent vector $\dot{c}(t)$ along the curve c and the extension Y of the vector field V along the differentiable curve c.*

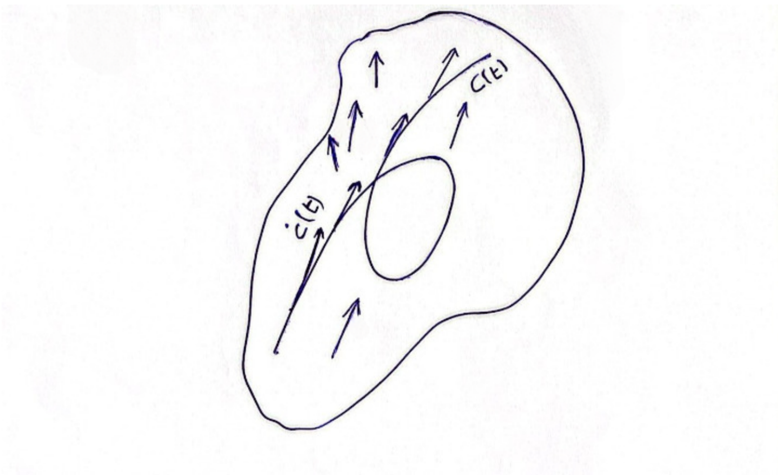

Figure 1.1.: Extendible vector fields

In local coordinates $x : W \to \mathbb{R}^n$ we have $x^i(t) := x^i(c(t))$. Let $X_i = \dfrac{\partial}{\partial x^i}$ and $V = \sum_{i=1}^{n} V^i X_i$, then:

$$\frac{DV}{dt} = \sum_{i=1}^{n} \frac{DV^i}{dt} X_i + V^i \frac{DX_i}{dt}$$

But we know that $\dfrac{DX_i}{dt} = \nabla_{\dot{c}(t)} X_i = \nabla_{\sum_j \frac{dx_j}{dt} X_j} X_i = \sum_j \dfrac{dx_j}{dt} \nabla_{X_j} X_i$. Now we use 1.2 to conclude

$$\frac{DV}{dt} = \sum_{i=1}^{n} \left(\dot{V}^i + \sum_{j,k=1}^{n} \Gamma^i_{jk}(c(t)) \dot{x}^j(t) V^k \right) X_i.$$

We can now define parallel displacement:

Definition 1.3.4. *A vector field V defined along γ is said to be parallel along γ if:*

$$\frac{DV}{dt} := \nabla_{\dot{\gamma}(t)} V = 0$$

for all $t \in I$. The curve γ is said to be a geodesic of the connection ∇ if $\dot{\gamma}$ is parallel along γ, i.e., if

$$\frac{D\dot{\gamma}}{dt} = 0$$

for all $t \in I$.

Remark 1.3.3. *In local coordinates $x : W \to \mathbb{R}^n$, the condition for V to be parallel along γ is written as*

$$\dot{V}^i + \sum_{j,k=1}^{n} \Gamma^i_{jk} \dot{x}^j V^k = 0 \quad (i = 1, \dots, n) \tag{1.4}$$

Picard-Lindelöf Theorem then says that given a curve $\gamma : I \to M$, a point $p \in \gamma(I)$ and a vector $v_p \in T_pM$, there exists a unique vector field $V : I \to TM$ parallel along γ such that $V(\gamma(0)) = v_p$, which is called the parallel transport of v_p along γ 1.2. Moreover, the

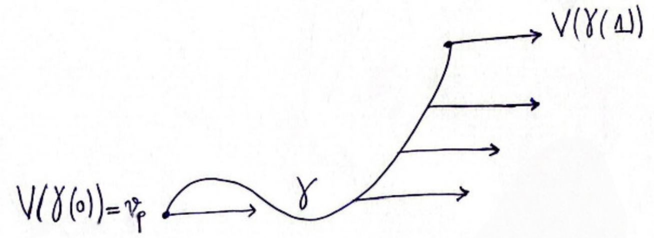

Figure 1.2.: The parallel transport of v_p along γ

geodesic equations are

$$\ddot{x}^i + \sum_{j,k=1}^{n} \Gamma^i_{jk} \dot{x}^j \dot{x}^k = 0 \quad (i = 1, \dots, n) \tag{1.5}$$

This is a system of second-order (nonlinear) ODE's for the coordinates of $\gamma(t)$. Therefore the Picard-Lindelöf Theorem implies that, given a point $p \in M$ and a vector $v_p \in T_pM$, there exists a unique geodesic $\gamma : I \to M$, defined on a maximal open interval I such that $0 \in I$, satisfying $c(0) = p$ and $\dot{c}(0) = v_p$.

Remark 1.3.4. *We saw that a choice of an affine connection is equivalent to the notion of parallel transport, one might consider an important class of connections, namely connections for submanifolds of Euclidean space. Let p and q be two sufficiently close points on $\mathbb{S}^2$, assume that $\mathbb{S}^2$ can be embedded in $\mathbb{R}^3$, consider the vector $v_1 \in T_p\mathbb{S}^2$, and let γ be a curve from p to q. A possible way to parallel displace v_1 along the curve can be done given a particular choice of a connection. Indeed, to displace the vector v_1 along the curve, we consider it as a vector in $\mathbb{R}^3$ and perform a usual translation in $\mathbb{R}^3$. Doing this we get a vector v_2 in $T_q\mathbb{S}^2$, however, the inhabitant of the sphere can only see that portion of this vector lying in the tangent plane! We overcome this issue by projecting v_2 down in $T_q\mathbb{S}^2$. As a result, we define our parallel displacement of v_1 along γ as $\mathrm{Proj}(v_2)$.*

We elaborate on this particular selection of connections in more depth by looking at the following examples.

Example 1.3.1. *Let X be a vector and Y be a vector field on the unit sphere (which can be thought of as a map $\mathbb{S}^2 \to \mathbb{R}^3$), the aforementioned prospective connection is defined as*

$$\nabla_X^{\mathbb{S}^2} Y := \mathrm{Proj}\,|_{T\mathbb{S}^2}(\nabla_X^{\mathbb{R}^3} Y) := \nabla_X^{\mathbb{R}^3} Y - N\langle \nabla_X^{\mathbb{R}^3} Y, N\rangle \tag{1.6}$$

Where $\mathrm{Proj}\,|_{T\mathbb{S}^2}$ is the projection onto the tangent sphere, and $\nabla_X^{\mathbb{R}^3} Y$ is the derivative of Y in the direction of the vector X and N is the unit normal vector field in $\mathbb{R}^3$ to $\mathbb{S}^2$.

Example 1.3.2. *Let $\mathbb{S}^2 \subset \mathbb{R}^3$ be a sphere with radius ρ and consider the usual coordinates $(\theta, \varphi) \in \mathbb{S}^2$ defined by the parametrization $r : (0, \pi) \times (0, 2\pi) \to \mathbb{R}^3$ given by:*

$$r(\theta, \varphi) = (R \sin\theta \cos\varphi, R \sin\theta \sin\varphi, R \cos\theta)$$

Hence

$$\frac{\partial r}{\partial \theta} = \begin{pmatrix} R\cos\theta\cos\varphi \\ R\cos\theta\sin\varphi \\ -R\sin\theta \end{pmatrix}, \frac{\partial r}{\partial \varphi} = \begin{pmatrix} -R\sin\theta\sin\varphi \\ R\sin\theta\cos\varphi \\ 0 \end{pmatrix}, \mathbf{N} = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}$$

where $\dfrac{\partial r}{\partial \theta}, \dfrac{\partial r}{\partial \varphi}$ are tangent vectors and $\mathbf{N}$ is normal unit vector. Let us now calculate explicitly the prospective connection we mentioned in 1.6 on the sphere. First we calculate $\nabla_{\partial_\theta}\partial_\theta$:

$$\nabla_{\partial_\theta}\partial_\theta = \left(\frac{\partial^2 r}{\partial \theta^2}\right)_{\|}$$

where the symbol $\|$ means the tangent part of a vector decomposition. In fact

$$\frac{\partial^2 r}{\partial \theta^2} = \begin{pmatrix} -R\sin\theta\cos\varphi \\ -R\sin\theta\sin\varphi \\ -R\cos\varphi \end{pmatrix} = -R\mathbf{N}$$

is proportional to the normal vector and hence its tangent part $\left(\frac{\partial^2 r}{\partial \theta^2}\right)_{\|}$ *is zero. By 1.2 we can write* $\nabla_{\partial_\theta}\partial_\theta = 0$ *and thus* $\Gamma^{\varphi}_{\theta\theta} = \Gamma^{\theta}_{\theta\theta} = 0$.

Secondly, we calculate $\nabla_{\partial_\theta}\partial_\varphi$ *and* $\nabla_{\partial_\varphi}\partial_\theta$*:*

$$\nabla_{\partial_\theta}\partial_\varphi = \nabla_{\partial_\varphi}\partial_\theta = \left(\frac{\partial}{\partial \varphi}\frac{\partial r}{\partial \theta}\right)_{\|} = \left(\frac{\partial}{\partial \theta}\frac{\partial r}{\partial \varphi}\right)_{\|}$$

On the other hand,

$$\left(\frac{\partial}{\partial \varphi}\frac{\partial r}{\partial \theta}\right) = \begin{pmatrix} -R\cos\theta\sin\varphi \\ R\cos\theta\cos\varphi \\ 0 \end{pmatrix} = \cot\theta\frac{\partial r}{\partial \varphi}$$

Doing a small calculation, we see that $< \left(\frac{\partial}{\partial \theta}\frac{\partial r}{\partial \varphi}\right), \mathbf{N} >= 0$. *Therefore,* $\nabla_{\partial_\theta}\partial_\varphi$ *has no orthogonal part in its vector decomposition and we get that* $\Gamma^{\theta}_{\theta\varphi} = \Gamma^{\theta}_{\varphi\theta} = 0$ *and* $\Gamma^{\varphi}_{\theta\varphi} = \Gamma^{\varphi}_{\varphi\theta} = \cot\theta$.

Lastly, we compute $\nabla_{\partial_\varphi}\partial_\varphi$*:*

$$\nabla_{\partial_\theta}\partial_\varphi = \left(\frac{\partial^2 r}{\partial \varphi^2}\right)_{\|}$$

On the other hand, we have:

$$\left(\frac{\partial^2 r}{\partial \varphi^2}\right) = \begin{pmatrix} -R\sin\theta\cos\varphi \\ -R\sin\theta\sin\varphi \\ 0 \end{pmatrix}$$

Now we project this vector onto the surface of the sphere using the normal unit vector:

$$\left(\frac{\partial^2 r}{\partial \varphi^2}\right) = \left(\frac{\partial^2 r}{\partial \varphi^2}\right)_{\|} + \mathbf{N} < \left(\frac{\partial^2 r}{\partial \varphi^2}\right), \mathbf{N} >$$

Solving for our desired tangent vector and plugging the expression of the corresponding vectors, we get:

$$\nabla_{\partial_\theta}\partial_\varphi = -\sin\theta\cos\theta\frac{\partial r}{\partial \theta}$$

Thus, $\Gamma^{\theta}_{\varphi\varphi} = -\sin\theta\cos\theta$ *and* $\Gamma^{\varphi}_{\varphi\varphi} = 0$.

It turns out that a Riemannian manifold admits a preferred connection satisfying:

1. We want our connection to be compatible with the metric, in other words, we expect a pair of vectors to keep its scalar product constant when parallel-transported along a curve.

1.3.1. Levi-Civita connection

Theorem 1.3.1. *On a Riemannian manifold $(M, \langle .,. \rangle)$, there exist a unique affine connection ∇ that satisfies: For every $X, Y, Z \in \mathfrak{X}(M)$*

1. $\nabla_Y Z - \nabla_Z Y = [Y, Z] \iff \Gamma^k_{ij} = \Gamma^k_{ji}$ *(Torsion free)*
2. $X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle$. *(Metric compatibility)*

This connection ∇ is called the Levi-Civita connection and the Koszul formula characterizes it:

$$2 \langle \nabla_X Y, Z \rangle = X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle - \langle X, [Y, Z]\rangle + \langle Z, [X, Y]\rangle$$

Choosing local coordinates $\left(x^1, \ldots, x^m\right)$ and the vector fields $X = \frac{\partial}{\partial x^j}$, $Y = \frac{\partial}{\partial x^k}$ and $Z = \frac{\partial}{\partial x^l}$, where $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$, we can substitute in Koszul's formula to get:

$$2 \left\langle \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l} \right\rangle = \frac{\partial}{\partial x^j} \cdot g_{kl} + \frac{\partial}{\partial x^k} \cdot g_{jl} - \frac{\partial}{\partial x^l} \cdot g_{jk}$$

$$\iff \left\langle \sum_{i=1}^{n} \Gamma^i_{jk} \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^l} \right\rangle = \frac{1}{2} \left(\frac{\partial g_{kl}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^l} \right)$$

$$\iff \sum_{i=1}^{n} g_{il} \Gamma^i_{jk} = \frac{1}{2} \left(\frac{\partial g_{kl}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^l} \right)$$

In a compact form:

$$\Gamma^i_{jk} = \frac{1}{2} \sum_{l=1}^{n} g^{il} \left(\frac{\partial g_{kl}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^l} \right), \text{Where } (g)^{il} = (g_{il})^{-1} \tag{1.7}$$

Let us now look at some examples to see some of the notions defined so far in action:

Example 1.3.3. *Let (r, ϕ, z) be the cylindrical coordinate on $\mathbb{R}^3$ defined by the parametrization $\psi : (r, \phi, z) \to (r \cos \phi, r \sin \phi, z)$. Note that*

$$(x, y, z) = (r \cos \phi, r \sin \phi, z)$$

is a chart on $\mathbb{R}^3 \setminus \{x \geq 0, y = 0\}$. We obtain thus the basis vectors:

1. $\partial_r = \frac{\partial \psi}{\partial r} = \cos \phi \partial_x + \sin \phi \partial_y = (\cos \phi, \sin \phi, 0)$
2. $\partial_\phi = \frac{\partial \psi}{\partial \phi} = -r \sin \phi \partial_x + r \cos \phi \partial_y = (-r \sin \phi, r \cos \phi, 0)$
3. $\partial_z = \partial_z$

We use these to compute the expression of the Riemannian metric induced on the cylinder by the Euclidean metric on $\mathbb{R}^3$*. Think of* $x^1 = r$ *,* $x^2 = \phi$*,* $x^3 = z$ *Then, The metric* g_{ij} *can be computed as follows:*

$$g_{11} = \langle \partial_r, \partial_r \rangle = \cos\phi^2 + \sin\phi^2 = 1$$

$$g_{22} = \langle \partial_\phi, \partial_\phi \rangle = r^2(\cos\phi^2 + \sin\phi^2) = r^2$$

$$g_{33} = \langle \partial_z, \partial_z \rangle = 1$$

The metric matrix has the form

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Its inverse

$$g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^{-2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Thus

$$g = g_{ij} dx^i \otimes dx^j = dr^2 + r^2 d\phi^2 + dz^2$$

Now we compute the Christoffel symbols for the Levi-Civita connection, using 1.7 where we omit the summation sign:

1. *Note that all* g^{1l} *are zero except* $g^{11} = 1$*:*

$$\Gamma^1_{22} = \frac{1}{2} g^{1l} \left(\frac{\partial g_{2l}}{\partial x^2} + \frac{\partial g_{2l}}{\partial x^2} - \frac{\partial g_{22}}{\partial x^l} \right) = \frac{1}{2} g^{1l} \left[g_{2l,2} + g_{2l,2} - g_{22,l} \right] = \frac{1}{2} g^{11} \left[g_{21,2} + g_{21,2} - g_{22,1} \right] = -r$$

2. *Note that all* g^{2l} *are zero except* $g^{22} = r^{-2}$*:*

$$\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{2} g^{2l} \left(\frac{\partial g_{2l}}{\partial x^1} + \frac{\partial g_{1l}}{\partial x^2} - \frac{\partial g_{21}}{\partial x^l} \right) = \frac{1}{2} g^{22} \left[g_{22,1} + g_{12,2} - g_{21,2} \right] = \frac{1}{r}$$

3. *All the remaining* Γ^i_{jk} *are zero.*

Since $\nabla_{\partial_i} \partial_j = \Gamma^k_{ij} \partial_k$*, we conclude:*

$$\begin{cases} \nabla_{\partial_\phi} \partial_\phi = -r\partial_r, \\ \nabla_{\partial_\phi} \partial_r = \nabla_{\partial_r} \partial_\phi = \dfrac{1}{r} \partial_\phi, \\ \nabla_{\partial_i} \partial_j = 0, \text{ otherwise} \end{cases}$$

Hence, it is easy to see that ∂_r*,* ∂_ϕ *are parallel in the z-direction. Indeed, they satisfy the equation for parallel transport 1.3.4:*

$$\begin{cases} \nabla_{\partial_z} \partial_\phi = 0 \\ \nabla_{\partial_z} \partial_r = 0 \end{cases}$$

Let us consider the example of the sphere again. We should expect to get the same expressions for the Christoffel symbols since the Levi-Civita connection is unique.

Example 1.3.4. *As before, consider the parametrization $r : (0, \pi) \times (0, 2\pi) \to \mathbb{R}^3$ given by:*

$$r(\theta, \varphi) = (R \sin\theta \cos\varphi, R \sin\theta \sin\varphi, R \cos\theta)$$

So,

1. $\partial_\theta = \dfrac{\partial r}{\partial \theta} = \cos\varphi \cos\theta \partial_x + \cos\theta \sin\varphi \partial_y - \sin\theta \partial_z = (\cos\varphi \cos\theta, \cos\theta \sin\varphi, -\sin\theta)$
2. $\partial_\varphi = \dfrac{\partial r}{\partial \varphi} = -\sin\theta \sin\varphi \partial_x + \sin\theta \cos\varphi \partial_y + 0\partial_z = (-\sin\theta \sin\varphi, \sin\theta \cos\varphi, 0)$

$$g_{\theta\theta} = \langle \partial_\theta, \partial_\theta \rangle = 1$$

Similarly,

$$g_{\varphi\varphi} = \langle \partial_\varphi, \partial_\varphi \rangle = \sin^2\theta$$
$$g_{\theta\varphi} = g_{\varphi\theta} = \langle \partial_\varphi, \partial_\theta \rangle = 0.$$

We conclude that the metric induced on $\mathbb{S}^2$ by the Euclidean metric of $\mathbb{R}^3$ is given by

$$g = d\theta \otimes d\theta + \sin^2\theta d\varphi \otimes d\varphi.$$

Now our metric matrix $g = (g_{ij})$ has the following expression:

$$(g_{ij}) = \begin{pmatrix} g_{\theta\theta} & g_{\theta\varphi} \\ g_{\varphi\theta} & g_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}$$

and hence it's inverse

$$(g^{ij}) = (g_{ij})^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{1}{\sin^2\theta} \end{pmatrix}$$

The Christoffel symbols can be easily computed from these matrices. For instance

$$\Gamma^{\theta}_{\varphi\varphi} = \frac{1}{2} \sum_{i=1}^{2} g^{\theta i} \left(\frac{\partial g_{\varphi i}}{\partial \varphi} + \frac{\partial g_{\varphi i}}{\partial \varphi} - \frac{\partial g_{\varphi\varphi}}{\partial x^i} \right)$$
$$= \frac{1}{2} g^{\theta\theta} \left(0 + 0 - \frac{\partial \left(\sin^2\theta\right)}{\partial \theta} \right) = -\sin\theta\cos\theta.$$

Only three of the eight Christoffel symbols are nonzero: the one computed above and

$$\Gamma^{\varphi}_{\theta\varphi} = \Gamma^{\varphi}_{\varphi\theta} = \cot\theta.$$

1.3.2. The Hessian and Laplace-Beltrami operator

Throughout this section, we assume that M is a Riemannian manifold endowed with a metric g and a Levi-Civita connection ∇ compatible with the metric g. One of our aims in this section is to define the Hessian $\nabla^2 f$ of $f \in C^\infty(M)$. It is know that, for a (k,m)-tensor field ω, the covariant derivative $\nabla\omega$ is then a $(k, m+1)$-tensor field. Following this observation, the Hessian should be the covariant of a $(0,1)$-tensor field ∇f "1-form also denoted df", therefore, the Hessian is a $(0,2)$-tensor field $\nabla^2 f$. Let us give a precise definition.

Definition 1.3.5. *For a general $\omega \in \mathcal{T}^{0,1}(T^*M)$ and $X, Y \in \mathfrak{X}(M)$, we define $\nabla\omega \in \mathcal{T}^{0,2}(T^*M, T^*M)$ as:*

$$\nabla\omega(X,Y) := \nabla_X \omega(Y) := X(Y) - (\nabla_X Y)$$

In particular, if $\omega = df$, where $f \in C^\infty(M)$, we define the Hessian as the map

$$(X,Y) \mapsto \nabla df(X,Y) := \nabla_X df(Y) := X(Yf) - (\nabla_X Y)f$$

Remark 1.3.5. *As we explained in 1.2.1 $\nabla f = df$, then it is not hard to see that $\nabla_Y f = \nabla f(Y) = df(Y) = Yf$, if $f \in C^\infty(M)$. Using the properties of covariant differentiation 1.3.2, we get:*

$$\begin{aligned} \nabla_X(\nabla f(Y)) &= \nabla_X(Yf) = Y(\nabla_X f) + (\nabla_X Y)f & (1.8)\\ &= \nabla_Y \nabla f(X) + (\nabla_X Y)f & (1.9)\\ &= X(Yf) = \nabla\nabla f(X,Y) + (\nabla_X Y)f & (1.10) \end{aligned}$$

Hence,

$$\begin{aligned} \nabla\nabla f(X,Y) &= X(Yf) - (\nabla_X Y)f & (1.11)\\ &= \langle \nabla_X \nabla f, Y\rangle = X(\langle \operatorname{grad} f, Y\rangle) - \langle \operatorname{grad} f, \nabla_X Y\rangle & (1.12) \end{aligned}$$

But using the metric compatibility 1.3.1, where $Y = \operatorname{grad} f$ and $Z = Y$ we get:

$$\langle \nabla_X \operatorname{grad} f, Y\rangle = \langle \nabla_X \nabla f, Y\rangle$$

Thus, we get:

$$\nabla\nabla f(X,Y) = \nabla df(X,Y) = \langle \nabla_X \operatorname{grad} f, Y\rangle$$

Definition 1.3.6. *Let M be a Riemannian manifold. For $f \in C^\infty(M)$ the Laplace-Beltrami operator Δ^M is defined to be:*

$$\Delta^M f := \operatorname{tr}(\nabla\nabla f) = \operatorname{tr}(\nabla df) \in C^\infty(M) \tag{1.13}$$

For an orthonormal basis $\{F_i\}_{i=1}^n$ of T_xM we have an explicit formula:

$$\Delta^M f := \operatorname{tr}(\nabla^2 f) = \sum_{i=1}^n \nabla^2 f(F_i, F_i)$$

Remark 1.3.6. *In a coordinate chart* $x = \{x^i\}$ *of* M *the Laplace-Beltrami operator is defined as follows:*

$$\Delta^M f = \frac{1}{\sqrt{G}} \frac{\partial}{\partial x^i} \left[\sqrt{G} g^{ij} \frac{\partial f}{\partial x^j} \right] \tag{1.14}$$

$$= g^{ij} \frac{\partial}{\partial x^i} \frac{\partial f}{\partial x^j} + b^i \frac{\partial f}{\partial x^i}, \tag{1.15}$$

where $b^i = \dfrac{1}{\sqrt{G}} \dfrac{\partial[\sqrt{G} g^{ij}]}{\partial x^i}$, g *is the metric matrix and* $G := \det[g]$. *We refer to section 3.1 in [Hsu02] for a thorough definition of the Laplace-Beltrami operator locally.*

2. Notions of Stochastic Analysis

In this chapter, we introduce some essential concepts from stochastic analysis. These will serve as a building block and a tool to build on when we formulate their counterparts in a non-Euclidean setting. All results in this chapter can be found in most standard references, e.g. [Oks13], [Shr04], [Prö21],[Eva13] and are stated, mostly, without proofs as these can be found in the aforementioned works.

2.1. Preliminaries

2.1.1. Conditional Expectation, Stopping Times and Continuous Martingales

In this section, we recall some basic notions from probability theory, namely, conditional expectations, stopping times, and martingales.

Definition 2.1.1. *Let Ω be a fixed non-empty set. A system $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ (i.e., the power set of Ω) of subsets is called a σ-algebra, if the following properties hold:*

- $\Omega \in \mathcal{A}$
- $A \in \mathcal{A} \implies A^{\mathsf{c}} \in \mathcal{A}$
- $A_n \in \mathcal{A}, n \geq 1 \implies \cup_{n\geq 1} A_n \in \mathcal{A}$

We then call $(\Omega, \mathcal{A})$ a measurable space.

Remark 2.1.1. *Let $\mathcal{E} \subseteq \mathcal{P}(\mathcal{X})$ be an arbitrary family of sets. Then there is a unique smallest σ-algebra $\sigma(\mathcal{E})$ containing $\mathcal{E}$. It is given by:*

$$\sigma(\mathcal{E}) = \cap_{\mathcal{A}\subseteq\mathcal{P}(\mathcal{X}),\mathcal{E}\subseteq\mathcal{A}} \mathcal{A},$$

where $\mathcal{A}$ is a σ-algebra.
In other words, we intersect all σ-algebras containing the collection of sets $\mathcal{E}$ to obtain the smallest σ-algebra containing $\mathcal{E}$

Definition 2.1.2. *The family $\mathcal{B}_{\mathbb{R}}:=\sigma(\mathcal{I}_{\mathbb{R}})$ is the called the Borel σ-algebra on $\mathbb{R}$, where $\mathcal{I}_{\mathbb{R}}$ is the collection of half-open intervals. It is the smallest σ-algebra generated by the family of half-open intervals. We call $A \in \mathcal{B}_{\mathbb{R}}$ Borel-(measurable) sets in $\mathbb{R}$.*

Definition 2.1.3. *A set function $\mu : \mathcal{A} \to [0, \infty]$ is called a measure on the measurable space $(\Omega, \mathcal{A})$ if it the following conditions hold:*

1. *Non-negativity: For all* $A \in \mathcal{A}$, $\mu(A) \geq 0$
2. $\mu(\emptyset) = 0$
3. σ*-additivity: For all countable collections* $\{A_i\}_{i=1}^{\infty}$ *of pairwise disjoint sets in* $\mathcal{A}$,

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

Definition 2.1.4. *The triple* $(\Omega, \mathcal{A}, \mu)$ *is called a measure space*

Definition 2.1.5. *Let* μ, ν *be two measures. Then,* ν *is said to be absolutely continuous with respect to* μ *(Shorthand* $\nu << \mu$*) if* $\mu(A) = 0 \implies \nu(A) = 0$, *for every* $A \in \mathcal{A}$

Definition 2.1.6. *A measure space* $(\Omega, \mathcal{A}, \mu)$ *with* $\mu(\Omega) = 1$ *is called a probability space, and* μ *a probability measure. We often write* $(\Omega, \mathcal{A}, \mathbb{P})$ *for the probability space.* $\mathbb{P}(A)$ *is called the probability of the event* $A \in \mathcal{A}$. *We usually use the notation* $\mathcal{F}$ *for* $\mathcal{A}$.

Definition 2.1.7. *Let* $(\Omega, \mathcal{F})$ *be a measurable space:*

1. *a function* $X : \Omega \to \mathbb{R}$ *is called measurable if:*

$$\forall B \in \mathcal{B}_{\mathbb{R}}, \quad X^{-1}(B) \in \mathcal{F}.$$

2. *For* $(\Omega, \mathcal{F}, \mathbb{P})$ *a probability space, a random variable* X *is a measurable function,* $X : \Omega \to \mathbb{R}$.
3. *The* σ*-algebra generated by the random variable* X *is defined as*

$$\sigma(X) = \sigma(\{X^{-1}(B), B \in \mathcal{B}_{\mathbb{R}}\})$$

Definition 2.1.8. *(*L^p*spaces) Let* $(\mathcal{X}, \mathcal{A}, \mu)$ *be a measurable space, let* f *is* $\mathcal{F}$*-measurable be the set of all real-valued measurable functions defined on* $\mathcal{X}$. *The space* $L^p(\mu)$ *comprises that subset of functions of* $\mathcal{F}$ *that have finite p-th moment (*$p \geq 1$*). In other words:*

$$L^p(\mu) = \left\{ f \in \mathcal{F} : \int |f(x)|^p d\mu(x) < \infty \right\}$$

When the measure μ *is the probability* $\mathbb{P}$, *we obtain the class of all random variables on the probability space* $(\mathcal{X}, \mathcal{A}, \mu)$ *that have finite p-th moments.*

Theorem 2.1.1. *(Radon-Nikodym Theorem) [Dur05] Let* μ *and* ν *be two measures on* $(\Omega, \mathcal{F})$. *If* $\nu << \mu$ *and* ν *is a finite measure, then, there is a unique* $\mathcal{F}$*-measurable function* $f \in L^1(\mu)$, *so that for all* $A \in \mathcal{F}$

$$\int_A f d\mu = \nu(A) \tag{2.1}$$

f *is usually denoted* $\dfrac{d\nu}{d\mu}$ *and called the Radon-Nikodym derivative.*

Definition 2.1.9 (Conditional Expectation)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{G}$ be a sub- σ-algebra of $\mathcal{F}$. Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, such that $X \in L^1(\mathbb{P})$ almost surely. Then $\mathbb{E}(X|\mathcal{G})$ is defined to be any random variable Y that satisfies:*

1. *Y is $\mathcal{G}$-measurable.*
2. *For every set $A \in \mathcal{G}$,*

$$\int_A Y d\mathbb{P} = \int_A X d\mathbb{P} \tag{2.2}$$

Note 2.1.1. *All the relations involving conditional expectation are meant to be almost surely.*

Remark 2.1.2. 1. *Existence of $\mathbb{E}(X|\mathcal{G})$: There is always a random variable Y satisfying the above properties (provided that $\mathbb{E}|X| < \infty$), i.e., conditional expectations always exist. In fact, the existence is guaranteed by the Radon-Nikodym theorem, taking $\mu = \mathbb{P}$ and $\nu(A) = \displaystyle\int_A X d\mathbb{P}$, for $A \in \mathcal{G}$. It turns out that, the Radon-Nikodym derivative $\dfrac{d\mu}{d\nu}$ is a version of $\mathbb{E}(X|\mathcal{G})$ "See Page 207 [Dur05]"*

2. *Uniqueness of $\mathbb{E}(X|\mathcal{G})$: There can be more than one random variable Y satisfying the above properties, but if Y' is another one, then $Y = Y'$ almost surely, i.e.:*

$$\mathbb{P}\left\{\omega \in \Omega; Y(\omega) = Y'(\omega)\right\} = 1. \tag{2.3}$$

3. *For random variables X, Y, we define:*

$$\mathbb{E}(X \mid Y) = \mathbb{E}(X \mid \sigma(Y)).$$

The Conditional Expectation satisfies the following properties:

Proposition 2.1.1.

- *If $X \in L^1(\mathbb{P})$ and is $\mathcal{G}$-measurable, then $\mathbb{E}(X|\mathcal{G}) = X$.*
- *If $X_1,\ X_2 \in L^1(\mathbb{P})$, and a, b are $\mathcal{G}$-measurable, bounded, then $aX_1 + bX_2 \in L^1(\mathbb{P})$ and*

$$\mathbb{E}(aX_1 + bX_2|\mathcal{G}) = a\mathbb{E}(X_1|\mathcal{G}) + b\mathbb{E}(X_2|\mathcal{G})$$

  *(**Linearity**).*

- *If $X \geq 0$ almost surely, then $\mathbb{E}(X|\mathcal{G}) \geq 0$ (**Positivity**).*
- *If $\phi : \mathbb{R} \to \mathbb{R}$ is convex and $X \in L^1(\mathbb{P})$ and $\phi(X) \in L^1(\mathbb{P})$, then:*

$$\mathbb{E}(\phi(X)|\mathcal{G}) \geq \phi(\mathbb{E}(X|\mathcal{G}))$$

  *(**Jensen´s Inequality for conditional expectation**).*

- *If $\mathcal{H}$ is a sub- σ-algebra of $\mathcal{G}$, then $\mathbb{E}\left(\mathbb{E}(X|\mathcal{G})|\mathcal{H}\right) = \mathbb{E}(X|\mathcal{H})$ **(Tower property)**.*
- *For every $X \in L^1(\mathbb{P})$,*
$$\mathbb{E}(X|\{\emptyset, \Omega\}) = \mathbb{E}(X)$$

Corollary 2.1.1. *It follows from the last two properties, that:*

$$\mathbb{E}\left(\mathbb{E}(X|\mathcal{G})\right) = \mathbb{E}(X) \tag{2.4}$$

for all $X \in L^1(\mathbb{P})$.

Definition 2.1.10.
- *A stochastic process indexed by $t \in [0, T]$ is a family of $\mathbb{R}^d$-valued random variables $(X_t)_{t\in[0,T]}$ is called a **stochastic process**.*
- *For $\omega \in \Omega$, the map $X(\omega) : [0, T] \to \mathbb{R}^d, t \mapsto X_t(\omega)$ is called a **sample path** of $(X_t)_{t\in[0,T]}$.*
- *We say $(X_t)_{t\in[0,T]}$ is **continuous** or **right-continuous** if its sample paths are continuous or (right-continuous) almost surely.*
- *A **càdlàg** process X is a stochastic process for which the paths*
$$X(\omega) : [0, T] \to \mathbb{R}^d, t \mapsto X_t(\omega)$$
are right-continuous with left limits everywhere, with probability one.

Definition 2.1.11.
- *A family of σ-algebras $(\mathcal{F}_t)_{t\in[0,T]} \subset \mathcal{F}$ is called **filtration** if $\mathcal{F}_s \subseteq \mathcal{F}_t$, for $s < t$.*

Definition 2.1.12. *Let $(\mathcal{F}_t)$ be a filtration, $(X_t)_{t\in[0,t]}$ is called $(\mathcal{F}_t)$- **adapted** if X_t is $\mathcal{F}_t$ measurable.*

Definition 2.1.13.
- *A random variable $\tau : \Omega \to [0, T] \cup \{\infty\}$ is called an $(\mathcal{F}_t)$-**stopping time** if:*
$$\{\tau \le t\} \subset \mathcal{F}_t, \forall t \in [0, T].$$

Example 2.1.1. *Let $t \ge 0$ be fixed. The properties of σ-algebras imply that $\left\{\tau > t - \frac{1}{n}\right\} = \left\{\tau \le t - \frac{1}{n}\right\}^c \in \mathcal{F}_{t-\frac{1}{n}}$ for all positive integers n. Since every set in $\mathcal{F}_{t-\frac{1}{n}}$ is also in $\mathcal{F}_t$, we conclude that $\left\{\tau > t - \frac{1}{n}\right\}$ is in $\mathcal{F}_t$ for every n, and hence*

$$\{\tau = t\} = \{\tau \le t\} \cap \left(\bigcap_{n=1}^{\infty} \left\{\tau > t - \frac{1}{n}\right\}\right)$$

is also in $\mathcal{F}_t$. In other words, a stopping time τ has the property that the decision to stop at time t must be based only on information available at time t.

Definition 2.1.14. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number, and let $\{\mathcal{F}_t, 0 \leq t \leq T\}$, be a filtration of sub- σ-algebras of $\mathcal{F}$. Consider an adapted stochastic process $X(t), 0 \leq t \leq T$; such that $\mathbb{E}[|X(t)|] < \infty$ for all t.*

- *X_t is a martingale, if*

$$\mathbb{E}[X(t) \mid \mathcal{F}_s] = X(s), \text{ for all } 0 \leq s \leq t \leq T$$

- *X_t is a sub-martingale, if*

$$\mathbb{E}[X(t) \mid \mathcal{F}_s] \geq X(s), \text{ for all } 0 \leq s \leq t \leq T$$

- *X_t is a super-martingale, if*

$$\mathbb{E}[X(t) \mid \mathcal{F}_s] \leq X(s), \text{ for all } 0 \leq s \leq t \leq T$$

Remark 2.1.3. *In the literature, the martingale property is referred to as fair game, indeed, let us denote X_t to be the asset of a player at time t, then the martingale relation $\mathbb{E}[X(u) \mid \mathcal{F}_t] = X(t)$ reflects the fact that the expected asset of the player at a later time u given the information he has at time t is exactly his currently asset.*

2.2. Brownian Motion

Definition 2.2.1. *A real-valued stochastic process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $B = (B_t)_{t \in [0,T]} = B(t)$ is called a Brownian motion if:*

1. *$B_0 = 0$ almost surely (a.s.).*

2. *B has independent increments: $\{B_{t_0}, B_{t_1} - B_{t_0}, \cdots\cdots, B_{t_n} - B_{t_{n+1}}\}$ are independent random variables for $0 \leq t_0 \leq t_1 \cdots < t_n$.*

3. *B has stationary and normally distributed increments:*

$$B_t - B_s \sim \mathcal{N}(0, t - s)$$

for

$$s \leq t \in [0, T].$$

4. *B has a.s. continuous sample paths.*

Corollary 2.2.1. 1. *From 3 we can easily see that $B_1 \sim \mathcal{N}(0, 1)$.*

2. *From the second and third properties, the increments are jointly Gaussian. Adding these increments up we deduce that* $\{B_{t_1}, B_{t_2}, \cdots\cdots, B_{t_n}\}$ *is also jointly Gaussian. So, for* $s \leq t$*:*

$$\mathrm{Cov}(B_s, B_t) = \mathbb{E}[B_s B_t] = \mathbb{E}[B_s(B_t - B_s + B_s)] = \mathbb{E}[B_s^2] = s = min(t, s).$$

where we used the fact that the increments are independent.

Remark 2.2.1. *There are various ways to construct a stochastic process satisfying the above properties, however, we will only look at the construction of Brownian motion given by Donsker's theorem "See chapter 8 [Dur05]" in which Brownian motion can be constructed as the scaling limit of a random walk, or some other discrete-time stochastic processes with stationary independent increments.*

In the next section, we describe this construction in further detail.

Symmetric Random Walk

Definition 2.2.2. *To construct Brownian motion, we begin with a symmetric random walk. We repeatedly toss a fair coin (i.e., p, the probability of H (Heads) on each toss, and* $q = 1 - p$*, the probability of T (Tails) on each toss, are both equal to* $\frac{1}{2}$*).*
We represent the consecutive results of these independent coin tosses as $\omega = \omega_1\omega_2\omega_3\ldots$*. In other words,* ω *is the infinite sequence of tosses, and* ω_n *is the outcome of the n-th toss. Let*

$$X_j = \begin{cases} 1 & if\, \omega_j = H \\ -1 & if\, \omega_j = T \end{cases}$$

We define $M_0 = 0$ *and*

$$M_k = \sum_{j=1}^{k} X_j, k = 1, 2, \ldots$$

The process $M_k, k = 0, 1, 2, ..$ *is a symmetric random walk. With each toss, it either steps up one unit or down one unit, and each of the two possibilities is equally likely.*

Increments of the Symmetric Random Walk

Definition 2.2.3. *A random walk has independent increments. This means that if we choose non-negative integers* $0 = k_0 < k_1 < \cdots < k_m$*, the random variables*

$$M_{k_1} = (M_{k_1} - M_{k_0}), (M_{k_2} - M_{k_1}), \ldots, \left(M_{k_m} - M_{k_{m-1}}\right)$$

are independent. Each of these random variables,

$$M_{k_{i+1}} - M_{k_i} = \sum_{j=k_i+1}^{k_{i+1}} X_j$$

is called an increment of the random walk. It is the change in the position of the random walk between times k_i *and* k_{i+1}*. Increments over non-overlapping time intervals are independent because they depend on different independent families of random variables.*

Properties 2.2.1.

- *Each increment* $M_{k_{i+1}} - M_{k_i}$ *has mean 0 (Since* $\mathbb{E}[X_j] = 0$*).*
- *Each increment* $M_{k_{i+1}} - M_{k_i}$ *has variance* $k_{i+1} - k_i$*:*

$$\operatorname{Var}\left(M_{k_{i+1}} - M_{k_1}\right) = \sum_{j=k_i+1}^{\kappa_{i+1}} \operatorname{Var}\left(X_j\right) = \sum_{j=k_i+1}^{N+1} 1 = k_{i+1} - k_i.$$

  where we used the fact that, $\operatorname{Var}\left(X_j\right) = \mathbb{E}[X_j^2] = 1$.

- *Martingale Property for the Symmetric Random Walk with respect to its own natural filtration: To see that the symmetric random walk is a martingale, we choose non-negative integers* $k < \ell$ *and compute:*

$$\begin{aligned}
\mathbb{E}\left[M_\ell \mid \mathcal{F}_k\right] &= \mathbb{E}\left[(M_\ell - M_k) + M_k \mid \mathcal{F}_k\right] \\
&= \mathbb{E}\left[M_\ell - M_k \mid \mathcal{F}_k\right] + \mathbb{E}\left[M_k \mid \mathcal{F}_k\right] \\
(M_k \text{ is } \mathcal{F}_k \text{ measurable}) &= \mathbb{E}\left[M_\ell - M_k \mid \mathcal{F}_k\right] + M_k \\
(\mathit{Independence}) &= \mathbb{E}\left[M_\ell - M_k\right] + M_k = M_k.
\end{aligned}$$

Definition 2.2.4. *Given a discrete stochastic process* $(X_t)_{t=\{0,\ldots,n\}}$*, the quadratic variation of* X_t *is defined by:*

$$\sum_{i=0}^{n}(X_{i+1} - X_i)^2$$

Now let us calculate it for the Symmetric Random Walk:

Example 2.2.1. *The quadratic variation up to time* k *of the Symmetric Random Walk is defined as*

$$[M, M]_k = \sum_{j=1}^{k} \left(M_j - M_{j-1}\right)^2 = k.$$

Note that this is computed path-by-path i.e. for each ω*. The quadratic variation up to time* k *along a path is computed by taking all the one-step increments* $M_j - M_{j-1}$ *along that path (these are equal to* X_j*, which is either 1 or* -1*, depending on the realization), squaring these increments, and then summing them up. Since* $(M_j - M_{j-1})^2 = 1$*, regardless of whether* $M_j - M_{j-1}$ *is 1 or* -1*, the above sum in is equal to* $\sum_{j=1}^{k} 1 = k$*.*

Let us now recall Donsker´s theorem:

Theorem 2.2.1. *Let* $(X_i)_i$*,* $i \geq 0$ *be a sequence of independent and identically distributed random variables with finite zero mean and variance* $\sigma^2 > 0$*. Let* $S_n = X_1 + \cdots + X_n$ *be* *the sum of the first n-random variables X_n with $X_0 = 0$, define the rescaled symmetric random walk by*

$$B^{(n)}(t) = \begin{cases} \frac{1}{\sqrt{n}} S_{nt}, & \text{if } nt \text{ is an integer} \\ \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}, & \text{if } nt \text{ is not an integer} \end{cases}$$

then:

$$B^{(n)} \to \sigma B \text{ as } n \to \infty.$$

Where B is a standard Brownian motion. The convergence here is in law in the space of all càdlàg processes.

Donsker's invariance principle [Don51] extends this convergence to the whole function $B^{(n)} := (B^{(n)}(t))_{t\in[0,1]}$. In other words, the random function $B^{(n)}$ converges in distribution to a standard Brownian motion B "Taking values in the space of all càdlàg processes" as $n \to \infty$.

In case, nt is not an integer, we define $B^{(n)}(t)$ by linear interpolation between its values at the nearest points s and s' to the left and right of t for which ns and ns' are integers. We shall obtain a Brownian motion in the limit as $n \to \infty$

2.2.1. Filtration for Brownian Motion and the Martingale Property

Definition 2.2.5. *Let $B(t), t \geq 0$ be a Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A filtration for the Brownian motion is a filtration $(\mathcal{F}_t)_{t\geq 0}, t \geq 0$, satisfying:*

- *$(B(t))_{t\geq 0}$ is $(\mathcal{F}_t)_{t\geq 0}$-adapted, for every $t \geq 0$*
- *For $0 \leq t < u$, the increment $B(u) - B(t)$ is independent of $\mathcal{F}_t$.*

Theorem 2.2.2. *Let $(B(t))_{t\geq 0}$ be a Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration for the Brownian motion, then $B(t)$ is an:*

$$(\mathcal{F}_t)_{t\geq 0} - \textit{martingale}.$$

Proof. Following the same chain of arguments as we have done for the symmetric random walk, for $0 \leq s \leq t$, we have:

$$\begin{aligned} \mathbb{E}[B(t) \mid \mathcal{F}_s] &= \mathbb{E}[(B(t) - B(s)) + B(s) \mid \mathcal{F}_s] \\ &= \mathbb{E}[B(t) - B(s) \mid \mathcal{F}_s] + \mathbb{E}[B(s) \mid \mathcal{F}_s] \\ &= \mathbb{E}[B(t) - B(s)] + B(s) \\ &= B(s). \end{aligned}$$

□

2.2.2. Constructing First and Quadratic Variations

We try to develop some intuition through ordinary functions that are one-time continuously differentiable before we precisely define these two concepts for general stochastic processes. Firstly, let us start by formally defining the concept of an interval partition:

Definition 2.2.6.
- *A partition* Π *of* $[0,T]$ *is a family of disjoint intervals* $J = (s_j, t_j]$ *covering* $[0,T]$.

- *The mesh size of* π *is given by:* $\|\Pi\| := \sup_{J\in\Pi} |t_j - s_j|$. *A sequence of partitions* $(\Pi_n)_{n\in\mathbb{N}}$ *is called zero sequence partition if:*

1. $\|\Pi_n\| \to 0$ *as* $n \to \infty$
2. $\Pi_n \subseteq \Pi_{n+1}\ \left[J \in \Pi_n \Rightarrow J^n = \bigcup_{k=1}^n J_{n+1}^k, J_{n+1}^k \in \Pi_{n+1}\right]$
3. $\{J : J \subseteq \Pi_n$ *st:* $J \subseteq [0,t]\}$ *is finite, for all* $t \in [0,T]$.

First variation

Definition 2.2.7. *Let* $f(t)$ *be a function defined on the interval* $[0,T]$ *and let a partition* $\Pi_n = \left\{t_0^n, t_1^n, \ldots, t_{k_n}^n\right\}$ *of* $[0,T]$*, which is a set of times of the form*

$$\{0 = t_0^n < t_1^n < \cdots < t_{k_n}^n = T\}$$

The maximum step size of the partition $\|\Pi_n\| = \max_{j=0,\ldots,k_n}\left(t_{j+1}^n - t_j^n\right)$. *The first variation of* f *up to time* T *is then defined as*

$$\mathrm{FV}_T(f) = \lim_{\|\Pi_n\|\to 0} \sum_{j=0}^{k_n} \left|f\left(t_{j+1}^n\right) - f\left(t_j^n\right)\right|.$$

This limit is understood in the following sense: As we include more points t_i^n*, the number of partition points* n *increases, hence the maximum step size tends to zero.*

Remark 2.2.2. *Assume* f *is a* C^1 *function, we can use the Mean Value Theorem, which asserts that there is a point* $t_j^{*n} \in [t_j^n, t_{j+1}^n]$ *such that:*

$$\frac{f\left(t_{j+1}^n\right) - f\left(t_j^n\right)}{t_{j+1}^n - t_j^n} = f'\left(t_j^{*n}\right).$$

We can rewrite the expression of $\mathrm{FV}_T(f)$ *in the following manner:*

$$\mathrm{FV}_T(f) = \lim_{\|\Pi_n\|\to 0} \sum_{j=0}^{k_n} \left|f'\left(t_j^*\right)\left(t_{j+1}^n - t_j^n\right)\right|$$

The latter expression is exactly a Riemann sum of the function $|f'(t)|$*. Thus:*

$$\mathrm{FV}_T(f) = \int_0^T |f'(t)|dt.$$

This intuitively means that the first variation measures the amount of variability of the function f *in the interval* $[0,T]$*.*

Quadratic Variation

We follow a similar approach for the quadratic variation. We begin with a definition:

Definition 2.2.8. *Let* $f(t)$ *be a function defined on the interval* $[0,T]$*. The quadratic variation of* f *up to time* T *is*

$$[f,f](T) = \lim_{\|\Pi_n\|\to 0} \sum_{j=0}^{k_n} \left[f\left(t_{j+1}^n\right) - f\left(t_j^n\right)\right]^2,$$

where $\Pi_n = \left\{t_0^n, t_1^n, \ldots, t_{k_n}^n\right\}$ *and* $0 = t_0^n < t_1^n < \ldots < t_{k_n}^n = T$*.*

Remark 2.2.3. *We suppose that* f *is a* C^1 *function, by the Mean Value theorem and using the above definition:*

$$\sum_{j=0}^{k_n} \left[f\left(t_{j+1}^n\right) - f\left(t_j^n\right)\right]^2 = \sum_{j=0}^{k_n} \left|f'\left(t_j^{*n}\right)\right|^2 \left(t_{j+1}^n - t_j^n\right)^2 \leq \|\Pi_n\| \cdot \sum_{j=0}^{k_n} \left|f'\left(t_j^{*n}\right)\right|^2 \left(t_{j+1}^n - t_j^n\right),$$

and thus

$$\begin{aligned}
[f,f](T) &\leq \liminf_{\|\Pi_n\|\to 0} \left[\|\Pi_n\| \cdot \sum_{j=0}^{k_n} \left|f'\left(t_j^{*n}\right)\right|^2 \left(t_{j+1}^n - t_j^n\right)\right] \\
&= \liminf_{\|\Pi_n\|\to 0} \|\Pi_n\| \cdot \liminf_{\|\Pi_n\|\to 0} \sum_{j=0}^{k_n} \left|f'\left(t_j^{*n}\right)\right|^2 \left(t_{j+1}^n - t_j^n\right) \\
&= \liminf_{\|\Pi_n\|\to 0} \|\Pi_n\| \cdot \int_0^T \left|f'(t)\right|^2 dt = 0
\end{aligned}$$

In the last step, we assumed $\int_0^T \left|f'(t)\right|^2 dt$ *is finite to ensure that the whole expression tends to zero as* $\|\Pi_n\|$ *goes to zero.*

Since the paths of Brownian motion cannot be differentiated with respect to the time variable, one cannot hope to apply the Mean Value theorem as we did in this section. Consequently, a different approach is needed to define the equivalent of the above in the case of Brownian motion.

2.2.3. First Variation and Quadratic Variation

For general stochastic processes, we can now state the following generalized definition:

Definition 2.2.9. *(First Variation process) Consider a right continuous stochastic process $(X_t)_{t\in[0,T]}$, its first variation process $FV(X)_{t\in[0,T]}$ is defined by:*

$$FV_t(X) := \sup_{\Pi} \sum_{J\in\Pi} \left|\Delta_{J\cap[0,T]}X\right|, t \in [0,T]$$

with $\Delta_{(u,v]\cap[0,t]}X := X_{t\wedge u} - X_{v\wedge t}$, where the supremum is taken over Π such that point 3) in Definition 2.2.2 holds.

Remark 2.2.4. *Using the triangle inequality, $\lim\limits_{\|\Pi\|\to 0} FV_t(X)$ is well-defined.*

Lemma 2.2.1. *Let $(X_t)_{t\in[0,T]}$ be a continuous martingale of locally finite variation, i.e. $FV_t(X) < \infty, \quad \mathbb{P} - a.s., \forall t \in [0,T]$, then:*

$$\mathbb{P}\left(\{\omega \in \Omega : X_t(\omega) = X_0(\omega), \forall t \in [0,T]\}\right) = 1.$$

Proof. See lemma 2.14 [Prö21] □

Corollary 2.2.2. *(First variation of the Brownian motion) For a Brownian motion $(B_t)_{t\in[0,T]}$, we have:*

$$\mathrm{FV}_t(B) = \infty, \quad \forall t \in [0,T]$$

a.s.

Proof. Consider a Brownian motion $(B_t)_{t\in[0,T]}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we showed it is a continuous martingale. Moreover $B_t \sim \mathcal{N}(0,t)$, hence it is not a constant stochastic process. Using lemma 2.2.1 we conclude that: $\mathrm{FV}_t(B) = \infty$, $\forall t \in [0,T]$ a.s. □

Proposition 2.2.1. *(Quadratic variation for continuous martingales) Let $(X_t)_{t\in[0,T]}$ be continuous and bounded i.e. $|X_t(\omega)| \leq C$ for some $C > 0$, $t \in [0,T]$, a.e $\omega \in \Omega$ martingale. Then, there exists a continuous stochastic process $[X,X] = ([X,X]_t)_{t\in[0,T]}$ such that:*

1. *$[X,X] = \lim\limits_{n\to\infty} \sum\limits_{J\in\Pi} (\Delta_{J\cap[0,T]}X)^2$ uniformly on $[0.T]$ in L^2sense, for any zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$.*
2. *$[X,X]_0 = 0$ and $[X,X]$ is non-decreasing.*
3. *$(X^2 - [X,X]_t)_{t\in[0,T]}$ is a martingale.*

Then the stochastic process $[X,X] = ([X,X]_t)_{t\in[0,T]}$ is called the quadratic variation of X.

The above proposition is too restrictive and excludes Brownian motion, therefore, we need to generalize the quadratic variation of continuous martingales to the more general class of continuous local martingales. Before going through this generalization, let us introduce continuous local martingales.

2.2.4. Local Martingales

Definition 2.2.10. *Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration satisfying the usual conditions. An $(\mathcal{F}_t)$-adapted stochastic process $(X_t)_{t\in[0,T]}$ is called a local martingale if there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ such that $\tau_n \uparrow T, \mathbb{P}$-a.s., such that:*

$$(X_t^n)_{t\in[0,T]} := (X_{t\wedge\tau_n} - X_0)_{t\in[0,T]}$$

is a martingale for all $n \in \mathbb{N}$.
We can define local sub-martingale and local super-martingale in an analogous manner.

Remark 2.2.5. *Note that:*

$$\textit{Martingales} \subseteq \textit{Local martingales}$$

Example 2.2.2. *Let $Z \in L^1 \setminus L^2$ i.e. $\mathbb{E}[|Z|] < \infty$ and $\mathbb{E}[Z^2] = \infty$. Furthermore, let W be a random variable independent of Z such that:*

$$\mathbb{P}(W = \pm 1) = \frac{1}{2}.$$

In this example, we have two timestamps and the corresponding σ-fields are respectively:

$$\mathcal{F}_0 = \sigma(Z) \tag{2.5}$$
$$\mathcal{F}_1 = \sigma(Z, W) \tag{2.6}$$

Then, we claim that the tuple $(X_0, X_1) := (Z, Z + WZ^2)$ is a local martingale but not a martingale in the usual sense!

Claim 2.2.1. *(X_0, X_1) is not a martingale*

This is obvious, Since $\mathbb{E}[X_1|\mathcal{F}_0]$ does not exist, as WZ^2 is not integrable

Claim 2.2.2. *(X_0, X_1) is a local martingale*

Set $\tau_n := \inf\{t \in \{0,1\} : |X_t| \geq n\}, \forall n \in \mathbb{N}$, with the convention that $\inf\emptyset = \infty$. We can now write:

$$\mathbb{E}[|X_{1\wedge\tau_n}|] = \mathbb{E}[|Z|\mathbb{1}_{\{|Z|\geq n\}}] + \mathbb{E}[|Z + WZ^2|\mathbb{1}_{\{|Z|<n\}}] \leq \mathbb{E}[|Z|] + n + n^2 < \infty$$

Indeed, τ_n can only take either 0 or 1. If $\tau_n = 0$ then $|X_0| = |Z| \geq n$, otherwise, $|Z| < n$. Furthermore:

$$\begin{aligned}\mathbb{E}[X_{1\wedge\tau_n}|\mathcal{F}_0] &= \mathbb{E}[Z\mathbb{1}_{\{|Z|\geq n\}}|\mathcal{F}_0] + \mathbb{E}[\left(Z + WZ^2\right)\mathbb{1}_{\{|Z|<n\}}|\mathcal{F}_0] &(2.7)\\ &= \mathbb{E}[Z\mathbb{1}_{\{|Z|\geq n\}}|\mathcal{F}_0] + \mathbb{E}[Z\mathbb{1}_{\{|Z|<n\}}|\mathcal{F}_0] = \mathbb{E}[Z|\mathcal{F}_0] = Z = X_0 &(2.8)\end{aligned}$$

Since, Z^2, W are independent of $\mathcal{F}_0$ and $\mathbb{E}[W] = 0$. Thus, we obtain:

$$\mathbb{E}[X_{1\wedge\tau_n}|\mathcal{F}_0] = X_0$$

Now, we can generalize the proposition 2.2.1.

Proposition 2.2.2. *(Quadratic variation for local martingales, see [Prö21],Theorem 2.20) Let $(X_t)_{t\in[0,T]}$ be a continuous local martingale. Then, there exists a unique continuous stochastic process $[X,X] = ([X,X]_t)_{t\in[0,T]}$ such that:*

1. $[X,X] = \lim_{n\to\infty} \sum_{J\in\Pi} (\Delta_{J\cap[0,T]}X)^2$ *uniformly on* $[0.T]$ *in* L^2*sense, for any zero-sequence of partitions* $(\Pi_n)_{n\in\mathbb{N}}$.
2. $[X,X]_0 = 0$ *and* $[X,X]$ *is non-decreasing.*
3. $(X^2 - [X,X]_t)_{t\in[0,T]}$ *is a continuous local martingale.*

Theorem 2.2.3. *(Quadratic variation of the Brownian motion) Consider a Brownian motion $(B_t)_{t\in[0,T]}$. Then $[B,B]_T = T$ for all $T \geq 0$ a.s..*

Proof. Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0,T]$. Define the sampled quadratic variation corresponding to this partition to be:

$$Q_n = \sum_{j=0}^{n-1} \left(B\left(t_{j+1}\right) - B\left(t_j\right)\right)^2$$

This sampled quadratic variation is a random variable (i.e. it depends on the path of the Brownian motion along which it is computed). We shall show that it has an expected value of T, and its variance converges to zero. Hence, it converges to its expected value T, regardless of the path along which we are doing the computation. We have

$$\mathbb{E}\left[\left(B\left(t_{j+1}\right) - B\left(t_j\right)\right)^2\right] = \operatorname{Var}\left[B\left(t_{j+1}\right) - B\left(t_j\right)\right] = t_{j+1} - t_j$$

which implies

$$\mathbb{E}Q_n = \sum_{j=0}^{n-1} \mathbb{E}\left[\left(B\left(t_{j+1}\right) - B\left(t_j\right)\right)^2\right] = \sum_{j=0}^{n-1} \left(t_{j+1} - t_j\right) = T$$

as desired. Now we turn to proving that the variance of the sampled quadratic variation converges to 0. Indeed:

$$\begin{aligned}\operatorname{Var}\left[\left(B\left(t_{j+1}\right) - B\left(t_j\right)\right)^2\right] =& \mathbb{E}\left[\left(\left(B\left(t_{j+1}\right) - B\left(t_j\right)\right)^2 - \left(t_{j+1} - t_j\right)\right)^2\right] \\ =& \mathbb{E}\left[\left(B\left(t_{j+1}\right) - B\left(t_j\right)\right)^4\right] - 2\left(t_{j+1} - t_j\right)\mathbb{E}\left[\left(B\left(t_{j+1}\right) - B\left(t_j\right)\right)^2\right] \\ &+ \left(t_{j+1} - t_j\right)^2.\end{aligned}$$

Where we used the definition of the variance $\operatorname{Var}(X) = (X - \mathbb{E}[X])^2$ and the fact that $\mathbb{E}\left[\left(B\left(t_{j+1}\right) - B\left(t_j\right)\right)^2\right] = t_{j+1} - t_j$.. On the other hand, we have that the fourth moment of a Gaussian random variable with zero mean is equal to three times its variance squared (see A.0.1). Therefore:

$$\mathbb{E}\left[\left(B\left(t_{j+1}\right)-B\left(t_j\right)\right)^4\right]=3\left(t_{j+1}-t_j\right)^2 \tag{2.9}$$

and

$$\operatorname{Var}\left[\left(B\left(t_{j+1}\right)-B\left(t_j\right)\right)^2\right]=3\left(t_{j+1}-t_j\right)^2-2\left(t_{j+1}-t_j\right)^2+\left(t_{j+1}-t_j\right)^2=2\left(t_{j+1}-t_j\right)^2 \tag{2.10}$$

Finally, we get:

$$\begin{aligned}\operatorname{Var}(Q_\Pi) &= \sum_{j=0}^{n-1}\operatorname{Var}\left[\left(B\left(t_{j+1}\right)-B\left(t_j\right)\right)^2\right]=\sum_{j=0}^{n-1}2\left(t_{j+1}-t_j\right)^2 \\ &\leq \sum_{j=0}^{n-1}2\|\Pi\|\left(t_{j+1}-t_j\right)=2\|\Pi\|T.\end{aligned}$$

In particular, $\lim\limits_{\|\Pi\|\to 0}\operatorname{Var}\left(Q_n\right)=0$, and we conclude that $\lim\limits_{\|\Pi\|\to 0}Q_\Pi=\mathbb{E}[Q_\Pi]=T$. □

In light of the previous proposition, we can state the famous Lévy´s Characterization of Brownian Motion.

Theorem 2.2.4. *"Lévy's Characterization of Brownian Motion, [Oks13] Theorem 8.6.1" Let $X=(X_t)_{t\geq 0}$ be a continuous local martingale with $X_0=0$. Then, the following are equivalent:*

1. *X is a standard Brownian motion on the underlying filtered probability space.*
2. *X is continuous and X_t^2-t is a local martingale.*
3. *X has quadratic variation $[X,X]_t=t$*

2.3. Ito Calculus

This section brings together all of the tools that were developed throughout this chapter in order to define the Ito integral but first, let us summarize what we know so far about Brownian motion:

- $FV(B)_t=\infty$, i.e. for a partition Π_n, $\sum_{j=0}^{n-1}|B(t_{j+1})-B(t_j)|=\infty$ as $||\Pi_n||\to 0$.
- The $\lim\limits_{||\Pi_n||\to 0}\mathbb{E}[Q_n]=T$, where $Q_n=\sum_{j=0}^{n-1}|B(t_{j+1})-B(t_j)|^2$.

The fact that we do not have finite variation suggests that we cannot define $\int_0^T H(s)dB(s)$ using the standard Riemann-Stieltjes approach. However, let us assume for a moment that we could write $\int_0^T H(s)dB(s)$ as a limit (in a sense to be defined) of $\sum_{j=0}^{n-1} H(t_j^*)(B(t_{j+1}) - B(t_j))$. We can then ask which point t_j^* should one put in the H function in the previous sum. In the standard Riemann-Stieltjes integration approach, the limit is independent of the choice of points one takes since we are considering continuous functions. However, this is not true in our case and as it turns out, a different choice of points will give a different result. Let us take a closer look by considering an example illustrating the issue above.

Example 2.3.1. *Let Π_n be a partition of the interval $[0,1]$, assume we can write $\int_0^1 B(s)dB(s)$ as the limit of $\sum_{j=0}^{\infty} B\left(.\right)\left[B\left(t_{j+1}\right) - B\left(t_j\right)\right]$:*

- *Choosing the right point t_{j+1} of the interval, we have*

$$\mathbb{E}\left[\sum_{j=0}^{n-1} B\left(t_{j+1}\right)\left[B\left(t_{j+1}\right) - B\left(t_j\right)\right]\right] = \sum_{j=0}^{n-1}\left(t_{j+1} - t_j\right) = t_n - t_0 = 1$$

where we used $\mathbb{E}[B(s)B(t)] = min(t,s)$.

- *Choosing the left point t_j of the interval, we have*

$$\int_0^1 B(s)dB(s) = \lim_{||\Pi_n||\to 0} B\left(t_j\right)\left[B\left(t_{j+1}\right) - B\left(t_j\right)\right]$$

Hence

$$\mathbb{E}\left[\sum_{j=0}^{n-1} B\left(t_j\right)\left[B\left(t_{j+1}\right) - B\left(t_j\right)\right]\right] = \sum_{j=1}^{n-1} 0 = 0$$

Remark 2.3.1. *In the last line of the previous argument, we used the fact that the integrand evaluated at t_j is independent of the increment. This leads us to define an important class of integrands that basically have this independence property.*

2.3.1. Class of Integrands $\mathcal{H}$

Let B be a Brownian motion defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 2.3.1. *A process $\{X(\omega), t \geq 0, \omega \in \Omega\}$ is called* ***progressively measurable*** *with respect to $\mathcal{F}_t$, if for each $t \geq 0$:*

- *The mapping $X : [0,t] \times \Omega \to \mathbb{R}$ is measurable with respect to the σ-field $\mathcal{B}([0,t]) \times \mathcal{F}_t$.*
- *An elementary progressively measurable process is a process $H(t,\omega) = \sum_{i=1}^{k} \xi_i(t) I_{(t_i,t_{i+1}]}(\omega)$, where ξ_i is $\mathcal{F}_{t_i}$-measurable and $\{t_0 = 0 < \ldots < t_{k+1} = t\}$.*

Idea: For an arbitrary progressively measurable process H, define the integral as

$$\int_0^t H(s)dB(s) := \sum_{i=1}^{n-1} \xi_i[B(t_{i+1} \wedge t) - B(t_i \wedge t)].$$

From this, we can then proceed in the following way:

1. Approximate a general progressively measurable process with a sequence of elementary processes $(H_n)_{n\in\mathbb{N}}$.
2. Check that there exists some notion of limit "$\lim \int_0^T H_n(s)dB(s)$".
3. Make sure this limit does not depend on the choice of the approximating sequence $(H_n)_{n\in\mathbb{N}}$.

Question: How can we approximate a progressively measurable process by a sequence of elementary processes? First, define $\|H\|_2^2 := \mathbb{E}\left[\int_0^T H(s)^2 ds\right]$. Then, we can summarize our procedure in the following:

Procedure 2.3.1. *1. Let H be an arbitrary progressively measurable process. we approximate it by bounded progressively measurable processes $H_n(t,\omega) = H(t,\omega) \wedge n$. Then, as $n \to \infty$, the approximation converges in L^2-norm:*

$$\|H_n - H\|_2 \to 0$$

2. For a bounded progressively measurable process H, we construct approximation using continuous progressively measurable processes defined as $H_n(s,\omega) = \frac{1}{h}\int_{(s-h)\vee 0}^{s} H(s,\omega)ds$, where $h = \frac{1}{n}$. The resulting H_n is continuous and progressively measurable as H is bounded. Applying the Dominated Convergence Theorem yields $\|H_n - H\|_2 \to 0$

3. For a continuous, progressively measurable process H, we approximate it with elementary processes where $H_n(s,\omega) = H(\frac{j}{n},\omega)$ for $s \in (\frac{j}{n}, \frac{j+1}{n}]$

Properties of elementary processes: Elementary progressively measurable processes, also called step processes, obey some nice properties:

Properties 2.3.1. *Let H be a progressively measurable step process with:*

$$\mathbb{E}\left[\int_0^T H(s)^2 ds\right] < \infty.$$

Then:

1. $\mathbb{E}\left[\int_0^T H(s)dB(s)\right] = 0.$

2. $\mathbb{E}\left[\left(\int_0^T H(s)dB(s)\right)^2\right] = \mathbb{E}\left[\int_0^T H(s)^2 ds\right].$

Proof. 1. Indeed, $\mathbb{E}\left[\sum_{i=1}^k \xi_i(\omega)[B(t_{i+1}) - B(t_i)\right] = \sum_{i=1}^k \mathbb{E}\left[\xi_i(\omega)[B(t_{i+1}) - B(t_i)]\right].$
Using the partial averaging property from 2.4, and since ξ_i is $\mathcal{F}_{t_i}$-measurable, we get:

$$\sum_{i=1}^k \mathbb{E}\left[\xi_i \mathbb{E}\left[B(t_{i+1}) - B(t_i) \mid \mathcal{F}_{t_i}\right]\right] = 0$$

where we used 3 to conclude.

2. By definition:

$$\mathbb{E}\left[\left(\int_0^T H(s)dB(s)\right)^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^k \xi_i(\omega)[B(t_{i+1}) - B(t_i)]\right)^2\right]$$

Now we write the square using different indices and use the fact that we have symmetrical terms:

$$\mathbb{E}\left[\sum_{i,j=1}^k \xi_i(\omega)\xi_j(\omega)[B(t_{i+1})-B(t_i)][B(t_{j+1})-B(t_j)\right] = \begin{cases} 0, & \text{if } i < j \\ \mathbb{E}\left[\int_0^T H(s)^2 ds\right], & \text{if } i = j \end{cases}$$

- If $i = j$ we get $\mathbb{E}\left[\sum_{i=j}^k \xi_i(\omega)^2[B(t_{i+1}) - B(t_i)]^2\right] = \sum_{i=j}^k \mathbb{E}\left[\xi_i(\omega)^2[t_{i+1} - t_i]\right]$ which is clearly $\mathbb{E}[\int_0^T H(s)^2 ds]$, since H is a progressively measurable step process.

- Otherwise, we get $2\mathbb{E}\left[\sum_{i<j}^{k}\xi_i(\omega)\xi_j(\omega)[B(t_{i+1})-B(t_i)][B(t_{j+1})-B(t_j)]\right]=0$ using the partial averaging property from 2.4 with respect to $\mathcal{F}_{t_i}$ and 3 we conclude.

$\square$

Let us sum up all this in the following theorem:

Theorem 2.3.1. *(Ito integral) Suppose H is a progressively measurable process such that*

$$\mathbb{E}\left[\int_0^T H(s)^2 ds\right] < \infty.$$

and suppose H_n converges to H in L^2. Then, the Ito integral is defined as follows:

$$\int_0^T H(s)dB(s) = \lim_{n\to\infty}\int_0^T H_n(s)dB(s)$$

where the limit is thought of in L^2- sense and it does not depend on the choice of the sequence of progressively step process $(H_n)_n$. Moreover, we have:

1. $\mathbb{E}\left[\int_0^T H(s)dB(s)\right] = 0$
2. $\mathbb{E}\left[\left(\int_0^T H(s)dB(s)\right)^2\right] = \mathbb{E}\left[\int_0^T H(s)^2 ds\right]$

Example 2.3.2. *Let us compute the simplest integral we can think of $\int_0^t B(s)dB(s)$ using the aforementioned recipe:*

1. *Approximate the Brownian motion with the step processes $B_n(t) = \sum_{j=1}^{n-1} B(t_j)I_{[t_i,t_{i+1}]}(t)$, where $t_j = \frac{j}{n}$. Then:*

$$\begin{aligned}\|B_n - B\|_2^2 &= \mathbb{E}\left[\int_0^t [B_n(s)-B(s)]^2 ds\right] = \mathbb{E}\left[\sum_{j=0}^{n-1}\int_{\frac{j}{n}}^{j+\frac{1}{n}}[B(\frac{j}{n})-B(s)]^2 ds\right] \\ &= \sum_{j=0}^{n-1}\int_{\frac{j}{n}}^{j+\frac{j}{n}}\mathbb{E}[B(\frac{j}{n})-B(s)]^2 ds = \sum_{j=0}^{n-1}\frac{1}{2n^2} = \frac{1}{n} \to 0 \text{ as } n\to\infty\end{aligned}$$

Where we use $\mathbb{E}[B(\frac{j}{n})-B(s)]^2 = (s-\frac{j}{n})^2$. Hence we get that B_n indeed approximates B.

2. *Now define* $S_k = B(t_k)[B(t_{k+1}) - B(t_k)]$. *It is clear that*

$$S_k = \frac{1}{2}[B(t_{k+1})^2 - B(t_k)^2] - \frac{1}{2}[B(t_{k+1}) - B(t_k)]^2.$$

Now we treat the two quantities separately:

$$\frac{1}{2}\sum_{k=0}^{n}[B(t_{k+1})^2 - B(t_k)^2] = \frac{1}{2}[B(t)^2 - B(0^2)]$$

$$\frac{1}{2}\sum_{k=0}^{n}[B(t_{k+1}) - B(t_k)]^2 = Q_n \to \frac{1}{2}t \text{ as } n \to \infty$$

All together $\sum_{k=0}^{n} S_k \to \frac{1}{2}[B(t)^2 - t]$ *and we get that*

$$\int_0^t B(s)dB(s) = \frac{1}{2}[B(t)^2 - t].$$

Properties 2.3.2. *Let* H *be progressively measurable with respect to* $\mathcal{F}_t = \sigma(B(s) : s \leq t)$. *Let* H_1, H_2 *be progressively measurable with* $\|H_1\|_2$, $\|H_1\|_2 < \infty$ $\left(\|H\|_2^2 := \mathbb{E}\left[\int_0^T H(s)^2 ds\right]\right)$. *Then:*

$$\int_0^t [\alpha H_1(s) + \beta H_2(s)]\, dB(s) = \alpha \int_0^t H_1(s)dB(s) + \beta \int_0^t H_2(s)dB(s).$$

Now we come to a key result in Ito calculus, namely the Ito's formula or the Ito lemma:

2.3.2. Ito formula and Ito processes

Theorem 2.3.2. *Let* $f \in C^2(\mathbb{R})$, *then:*

$$f(B(t)) = f(0) + \int_0^t f'(B(s))dB(s) + \frac{1}{2}\int_0^t f''(B(s))ds \tag{2.11}$$

In what follows, we give an informal proof of 2.11.

sketch of proof 2.3.1. *For a partition* $\Pi_n = \{0 = t_1 < \cdots < t_n\}$, *it holds that*

$$f(B(t)) - f(0) = \sum_{i=1}^{n-1} f(B(t_{i+1})) - f(B(t_i))$$

Using a Taylor expansion

$$f(x_{j+1}) - f(x_j) = f'(x_j)[x_{j+1} - x_j] + \frac{1}{2}f''(x_j)[x_{j+1} - x_j]^2 + \text{Remainder term}$$

we can write

$$f(B(t)) - f(0) = \sum_{i=1}^{n-1} f'(B(t_i))[B(t_{i+1}) - B(t_i)] + \frac{1}{2}\sum_{i=1}^{n-1} f''(B(t_i))[B(t_{i+1}) - B(t_i)]^2 + \text{Remainder term}$$

1. *By 3, we expect* $[B(t_{i+1}) - B(t_i)]^2 \approx [t_{i+1} - t_i]$. *Thus, in the limit the second term converges to* $\frac{1}{2}\int_0^t f''(B(s))ds$.

2. *In the limit, the first term converges to* $\int_0^t f'(B(s))dB(s)$.

Remark 2.3.2. *1. The function f can have a t dependency as well i.e. $f := f(t,x)$. Then, 2.11 has the following form:*

$$f(t,B(t)) = f(0,B(0)) + \int_0^t f_s(s,B(s))ds + \int_0^t f_x(s,B(s))dB(s) + \frac{1}{2}\int_0^t f_{xx}(s,B(s))ds. \tag{2.12}$$

We can define Ito's formula for processes that are more general than the above.

Definition 2.3.2. *We call an Ito process any process X that can be written as*

$$X(t) = X(0) + \int_0^t g(s,X(s))dB(s) + \int_0^t f(s,X(s))ds$$

where, $f(s,X(s)), g(s,X(s))$ are stochastic processes, adapted to the filtration $\mathcal{F}_t$ associated to Brownian motion $B(t)$ such that they satisfy certain conditions guaranteeing the existence of the integrals defining X. Equivalently, X can be written in differential form,

$$dX(t) = g(t,X(t))dB(t) + f(t,X(t))dt \tag{2.13}$$

Similarly, an n-dimensional Ito process is a vector-valued process X_t with representation where δ,θ are vector-valued and $B(t) = \big(B^1(t),\dots,B^n(t)\big)$ is an n-dimensional Brownian motion.

Lemma 2.3.1. *If $X(t)$ is an Ito process with representation* 2.13 *(either one-dimensional or multi-dimensional), its* ***quadratic variation*** *is defined to be the process:*

$$[X,X]_t := \int_0^t |g(s)|^2 ds$$

Lemma 2.3.2. *If $X_1(t), X_2(t)$ are two Ito processes relative to the same n-dimensional Brownian motion $B(t)$, with their corresponding representations(in differential form)*

$$dX_i(t) = f_i(t,X(t))dt + \sum_{j=1}^n g_{ij}(t,X(t))dB^j(t)$$

Then the ***covariation*** *of $X_1(t)$ and $X_2(t)$ is defined by*

$$d[X_i,X_j]_t := \sum_{l=1}^n g_{il}(t,X(t))f_{jl}(t,X(t))dt$$

Let us go back to the one-dimensional situation in 2.13, the Ito formula has the following form for $t \geq 0$:

$$h(t, X(t)) = h(0, X(0)) + \int_0^t h_s(s, X(s))ds + \int_0^t h_x(s, X(s))dX(s) + \frac{1}{2}\int_0^t h_{xx}(s, X(s))[dX, dX]_s$$

Plugging in the shorthand form, we get:

$$h(t, X(t)) = h(0, X(0)) + \int_0^t h_s(s, X(s))ds + \int_0^t h_x(s, X(s))f(s)dB(s) + \int_0^t h_x(s, X(s))f(s)ds + \frac{1}{2}\int_0^t h_{xx}(s, X(s))g(s)^2 ds$$

where we have used the fact that all the cross variation terms get canceled out and only the covariation term $g^2(s, X(s))dB(s).dB(s)$, which is equal to $g^2(s, X(s))ds$, survives. This is due to the identity $d[B(s), B(s)] = [dB(s), dB(s)] = dB(s).dB(s) = ds$.

2.4. Stochastic Differential Equations

With Ito's integral and the equivalent of the chain rule for stochastic processes i.e. Ito's lemma in hand, we now turn our attention to stochastic differential equations (SDEs) driven by Brownian motion. We start with a treatment of SDEs defined using Ito's integral and then finish by looking at a class of SDEs, that resemble more ordinary differential equations (ODEs), driven by another type of stochastic integrals, namely Stratonovich integrals.

Motivation 2.4.1. *Consider the ODE:*

$$\begin{cases} dX(t) = b(X(t), t)dt \\ X(0) = X_0 \end{cases}$$

A stochastic differential equation (SDE) can be thought of as an ODE with trajectories that are perturbed by randomness that can be attributed to measurement error or fluctuations resulting from the surrounding environment. This can be modeled by adding a perturbation term, modeling the strength of perturbation, $\sigma(X(t), t)$, times a Brownian motion i.e., random noise. We can also arrive at these dynamics using a limit of a discrete approximation. To this end, let us consider the following discrete approximation:

$$X(t + \Delta t) - X(t) = b(X(t), t)\Delta t + \sigma(X(t), t)[B(t + \Delta t) - B(t)].$$

Letting $\Delta t \to 0$, we get, under suitable conditions on the coefficients, that the discrete process converges in an appropriate sense to the solution of the following equation:

$$\begin{cases} dX(t) = b(X(t), t)dt + \sigma(X(t), t)dB(t) \\ X(0) = X_0 \end{cases} \tag{2.14}$$

where, again, the differential formulation is nothing but the integral one.

Definition 2.4.1. *Assume b and σ are functions from $\mathbb{R} \times [0, \infty] \to \mathbb{R}$ that are Borel measurable. A process $X = \{X(t), t \in [0, T]\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a* ***Strong solution*** *to 2.14, if:*

1. $X(0) = X_0$.
2. *X is $\mathcal{F}_t$-adapted random process, where $\mathcal{F}_t$ is the σ-field generated by Brownian motion.*
3. $\int_0^T |b(X(t), t)| dt < \infty$ *a.s ;* $\int_0^T \sigma^2(X(t), t) dt < \infty$ *a.s.*
4. $X(t) = X_0 + \int_0^t b(X(s), s) ds + \int_0^t \sigma(X(s), s) dB(s)$, *for all* $t \geq 0$ *a.s.*

Definition 2.4.2. *SDE 2.14 has a unique strong solution if, for any strong solutions X, Y we have:*

$$\mathbb{P}\{\omega : X(\omega, t) = Y(\omega, t), \forall t \in [0, T]\} = 1.$$

Theorem 2.4.1. *Let $f : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ be a twice differentiable function. Assume $X(t)$ satisfies 2.14, then we have:*

$$df(X(t), t) = \left[f_1'(X(t), t) b(X(t), t) + \frac{1}{2} f_{22}''(X(t), t) + f_2'(X(t), t) \right] dt + f_1'(X(t), t) \sigma(X(t), t) dB(t) \tag{2.15}$$

This is also called Ito formula for the process $Y(t) := f(X(t), t)$, this formula turns out to be very useful when specifying $X(t)$.

Proof. (Sketch) Recall Ito's formula:

$$f(B(t)) = f'(B(t)) dB(t) + \frac{1}{2} f''(B(t)) dt$$

and its derivation, especially the relation $dt = dB_t \cdot dB_t$. Assume $X(t)$ satisfies 2.14, then we can in a similar way derive the following expression:

$$df(X(t), t) = f_1'(X(t), t) dX(t) + f_2'(X(t), t) dt + \frac{1}{2} f_{22}''(X(t), t) dX(t).dX(t). \tag{2.16}$$

Plugging in the expression for $dX(t)$ and using the cross-variation rules we get:

$$df(X(t), t) = f_1'(X(t), t)[b(X(t), t) dt + \sigma(X(t), t) dB(t)] + f_2'(X(t), t) dt + \frac{1}{2} f_{22}''(X(t), t) dt \tag{2.17}$$

$$= \left[f_1'(X(t), t) b(X(t), t) + \frac{1}{2} f_{22}''(X(t), t) + f_2'(X(t), t) \right] dt + f_1'(X(t), t) \sigma(X(t), t) dB(t) \tag{2.18}$$

This is also called Ito formula for the process $Y(t) := f(X(t), t)$, this formula turns out to be very useful when specifying $X(t)$. □

Let us discuss some examples of SDEs:

Example 2.4.1. *Consider*

$$\begin{cases} dX(t) = -\lambda X(t)dt + dB(t) \\ X(0) = X_0 \end{cases}$$

We can try to guess a function f and apply Ito's formula to it to check for a solution. Define $f(X(t),t) = e^{\lambda t}X(t)$, so:

$$\begin{cases} f_1'(x,t) = e^{\lambda t} \\ f_2'(x,t) = \lambda e^{\lambda t}x \\ f_{11}''(x,t) = 0 \end{cases}$$

Then by Ito formula:

$$f(X(t),t) - f(X(0),0) = e^{\lambda t}X(t) - X(0) = \int_0^t [e^{\lambda s}(-\lambda X(s)) + \lambda e^{\lambda s}x]ds + e^{\lambda s}dB(s)$$

Hence:

$$X(t) = X(0)e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)}dB(s)$$

*This process is called the **Ornstein-Uhlenbeck Process**.*

Example 2.4.2. *Consider*

$$\begin{cases} dX(t) = rX(t)dt + \sigma X(t)dB(t) \\ X(0) = X_0 > 0, \end{cases}$$

Where r, $\sigma > 0$ are constants. We try to find a function f and apply the Ito formula to it. Define $f(X(t)) = \log X(t)$, so using Ito formula as above we get

$$d\log X = \frac{dX(t)}{X(t)} - \frac{1}{X(t)^2}(\sigma X(t))^2 = rdt + \sigma dB(t) - \frac{\sigma^2}{2}dt$$

Therefore, the solution is of the form:

$$X(t) = X(0)e^{\sigma B(t)+(r-\frac{\sigma^2}{2})t}$$

*This process is called the **Geometric Brownian motion**.*

Let us now state the most important result in this section without proof.

Theorem 2.4.2. *Suppose that $b : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ and $A : \mathbb{R}^n \times [0,T] \to \mathbb{M}^{m\times n}$ are Borel-measurable and satisfy the following conditions:*

1. *They are uniformly Lipschitz:*

$$|b(x,t) - b(\hat{x},t)| \le L|x - \hat{x}|$$
$$|A(x,t) - A(\hat{x} - t)| \le L|x - \hat{x}|$$

for all $0 \le t \le T, x \in \mathbb{R}^n$, for some constant L.

2. *Their growth is linear:*

$$|b(x,t)| \leq L(1+|x|)$$
$$|A(x,t)| \leq L(1+|x|)$$

for all $0 \leq t \leq T, x \in \mathbb{R}^n$, *for some (other) constant* L.

3. *Assume* $X(0) = X_0$ *be a random variable non-dependent of the* σ*-field generated by Brownian motion, and such that:*

$$E\left(|X_0|^2\right) < \infty$$

Then there exists a unique strong solution X *to 2.14*

Note that these integrals are all in the Ito sense and a problem with such integrals is related to them not following the ordinary rules of calculus e.g., the extra term appearing when applying Ito's lemma. In what follows we will try to remedy the situation by introducing another stochastic integral that behaves in a manner that is closer to what we expect from the classical integrals. This is the notion of Stratonovich integral. We also define SDE's driven by Stratonovich integrals and investigate their relationship with their counterparts driven by Ito integrals.

2.4.1. Stratonovich integral/ODE

Definition 2.4.3. *Let* B_t *be a Brownian motion, let* $\Pi_M = \{t_0, t_1, \ldots, t_M\}$ *be a partition of* $[0,T]$. *Define,*

$$\xi_t = \sum_{i=1}^{M} \xi_{t_i} I_{(t_i, t_{i+1}]}(t)$$

a simple process, adapted to the filtration generated by Brownian motion. Moreover, the ξ_{t_i} *are* $\mathcal{F}_{t_i}$*-measurable. The Stratonovich integral of* ξ_t *is defined as follows:*

$$\int_0^T \xi_t \circ dB_t = \sum_{i=0}^{M} \frac{\xi_{t_i} + \xi_{t_{i+1}}}{2} \Delta B_i$$

where $\Delta B_i = B_{t_{i+1}} - B_{t_i}$. *The Stratonovich integral* $\int_0^T X_t \circ dB_t$ *of an adapted process* X_t *is defined as the limit of the sequence* $\int_0^T \xi_t^{(n)} \circ dB_t$ *where* $\xi_t^{(n)}$ *is a sequence of simple, progressively-measurable processes converging to* X_t *in* L^2.

Fact 2.4.1. *Ito and Stratonovich integrals differ only by the quadratic variation term.*

We provide an example to show the above fact:

Example 2.4.3. *Let $t_k = t_k(n) = k2^{-n}$ and $M = \lfloor 2^n T \rfloor$. So by the above definition:*

$$\begin{aligned}\int_0^T B_t \circ \mathrm{d}B_t &= \lim_{n\to\infty} \sum_{k=1}^{M} \frac{B_{t_k} + B_{t_{k-1}}}{2} \Delta B_k \\ &= \frac{1}{2} \lim_{n\to\infty} \sum_{k=1}^{M} \left(B_{t_k}^2 - B_{t_{k-1}}^2 \right) \\ &= \frac{1}{2} B_T^2.\end{aligned}$$

This should be compared to

$$\int_0^T B_t dB_t = \frac{1}{2} B_T^2 - T$$

in the case of Ito integral.

Remark 2.4.1. *Let $g : \mathbb{R} \to \mathbb{R}$ be a C^2 function and let X_t be an adapted process defined as $X_t = \int_0^t g(X_s) \circ dB(s)$ for all $t \in [0,1]$, Y_t is defined as an Ito integral $\int_0^t g(X_s) dB_s$. Choosing the partition $t_k(n)$ as above. The processes:*

$$\begin{aligned} X_t^{(n)} &= \sum_{k=1}^{M} \frac{g\left(X_{t_k}\right) + g\left(X_{t_{k-1}}\right)}{2} \Delta B_k \\ Y_t^{(n)} &= \sum_{k=1}^{M} g\left(X_{t_{k-1}}\right) \Delta B_k \end{aligned}$$

with $M = M(t)$ correspond to the interval partition containing t, converge respectively to X_t and Y_t as $n \to \infty$. Their difference can be written as:

$$X_t^{(n)} - Y_t^{(n)} = \frac{1}{2} \sum_{k=1}^{M} \left[g\left(X_{t_k}\right) - g\left(X_{t_{k-1}}\right) \right] \Delta B_k$$

Taylor's formula implies:

$$g\left(X_{t_k}\right) - g\left(X_{t_{k-1}}\right) = g'\left(X_{t_{k-1}}\right)\left(X_{t_k} - X_{t_{k-1}}\right) + o\left(X_{t_k} - X_{t_{k-1}}\right).$$

On the other hand,

$$\begin{aligned} X_{t_k} - X_{t_{k-1}} &= \int_{t_{k-1}}^{t_k} g\left(X_s\right) \circ \mathrm{d}B_s \\ &= g\left(X_{t_{k-1}}\right) \Delta B_k + \int_{t_{k-1}}^{t_k} \left[g\left(X_s\right) - g\left(X_{t_{k-1}}\right) \right] \circ \mathrm{d}B_s \\ &= g\left(X_{t_{k-1}}\right) \Delta B_k + o\left(\left(t_k - t_{k-1}\right)\right). \end{aligned}$$

Hence:

$$g\left(X_{t_k}\right) - g\left(X_{t_{k-1}}\right) = g'\left(X_{t_{k-1}}\right) g\left(X_{t_{k-1}}\right) \Delta B_k + o\left(\Delta B_k\right) + o\left(\left(t_k - t_{k-1}\right)\right).$$

Substituting in the expression of $X_t^{(n)} - Y_t^{(n)}$*, we obtain then*

$$X_t^{(n)} - Y_t^{(n)} = \frac{1}{2}\sum_{k=1}^{M}\left[g'\left(X_{t_{k-1}}\right) g\left(X_{t_{k-1}}\right) \Delta B_k^2 + o\left(\Delta B_k^2\right) + o\left(\left(t_k - t_{k-1}\right)\Delta B_k\right)\right].$$

Again by 3, we expect $[B(t_{i+1}) - B(t_i)]^2 \approx [t_{i+1} - t_i]$*. Thus, in the limit the difference converges to* $\frac{1}{2}\int_0^t g'(X_s)g(X_s)ds$*.*

Example 2.4.4. *Let* $B(t)$ *be an* $\mathbb{R}^n$ *Brownian motion. Let us compute:*

$$A(t) := \int_0^t e^{B(s)} \circ dB(s)$$

We have now that

$$dA(t) = e^{B(t)}dB(t) + \frac{1}{2}\left[de^{B(t)}, dB(t)\right]$$

Now we use Ito-Formula to compute $de^{B(t)}$ *and substitute. Recall that for a Brownian motion* $B(t)$ *and a smooth function* f*:*

$$df(B(t)) = \frac{1}{2}f''\left(B_t\right)dt + f'\left(B_t\right)dB_t$$

Thus, we have:

$$de^{B(t)} = \frac{1}{2}e^{B(t)}dt + e^{B(t)}dB(t)$$

Using the covariation rules established in 2.3.2 for Ito processes, and $d[B(t), B(t)] = dt$ *we conclude*

$$dA(t) = e^{B(t)}dB(t) + \frac{1}{2}e^{B(t)}dt$$

Remark 2.4.2. *Let* B_t *be an* m*-dimensional Brownian motion and suppose* $b : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$*,* $A : \mathbb{R}^n \times [0,T] \to \mathbb{M}^{n\times m}$ *where* $A = (a^{ij})_{i,j}$ *satisfy the hypotheses of the existence and uniqueness theorem 2.4.2. Then the following Ito stochastic differential equation has a solution* X *in the sense of 2.4.1*

$$\begin{cases} dX_t = b(X_t,t)dt + A(X_t,t)dB_t \\ X(0) = X_0 \end{cases}$$

or in integral form

$$X_t = \int_0^t b(X_s,s)ds + \int_0^t A(X_s,s)dB_s$$

We can also consider ***Stratonovich SDE****:*

$$\begin{cases} dX_t = b(X_t,t)dt + A(X_t,t)\circ dB_t \\ X(0) = X_0 \end{cases} \tag{2.19}$$

or in integral form

$$X_t = \int_0^t b(X_s,s)ds + \int_0^t A(X_s,s)\circ dB_s.$$

In light of the previous remark or the next lemma, we can write for $n=1$ *and* $A(X_t,t) = a(X_t,t) = aX_t$*:*

$$\int_0^t A(X_s,s)\circ dB_s - \int_0^t A(X_s,s)dB_s = \frac{1}{2}\int_0^t \frac{\partial a}{\partial x}(X_s,s)a(X_s,s)ds = \frac{1}{2}\int_0^t (aX_s)'(aX_s)ds$$

Let us now state the following crucial property of Stratonovich integrals:

Lemma 2.4.1. *(Stratonovich chain rule) Assume 2.19 holds and let* $g:\mathbb{R}^n\times[0,T]\to\mathbb{R}$ *be smooth such that* $Y_t = g(X(t),t)$*. Then:*

$$\begin{aligned} dY_t = \frac{\partial g}{\partial t}dt + \sum_{i=1}^n \frac{\partial g}{\partial x_i}\circ dX^i &= \frac{\partial g}{\partial t}dt + \sum_{i=1}^n \frac{\partial g}{\partial x_i}\circ [b^i(X,t)dt + \sum_{k=1}^n a^{ik}(X,t)\circ dB^k] \\ &= [\frac{\partial g}{\partial t} + \sum_{i=1}^n \frac{\partial g}{\partial x_i}b^i(X,t)]dt + \sum_{i=1}^n\sum_{k=1}^n \frac{\partial g}{\partial x_i}a^{ik}(X,t)\circ dB^k \end{aligned}$$

It turns out that, in the Stratonovich context we have the usual ordinary chain rule, in contrast to the Ito formula 2.12 where we get an additional term f_{xx}*.*

The previous lemma allows us to state an important fact, helping us to link the two formulations of SDEs i.e., the one with Ito integrals with the one using Stratonovich integrals:

Proposition 2.4.1. *Let* B_t *be an* m*-dimensional Brownian motion and suppose* $b:\mathbb{R}^n\times[0,T]\to\mathbb{R}^n$, $A:\mathbb{R}^n\times[0,T]\to\mathbb{M}^{n\times m}$ *where* $A=(a^{ij})_{i,j}$ *satisfy the hypotheses of the existence and uniqueness theorem 2.4.2. Then* $X(\cdot)$ *solves the Ito stochastic differential equation*

$$\begin{cases} dX_t = b(X_t,t)dt + A(X_t,t)dB_t \\ X(0) = X_0 \end{cases}$$

if and only if $X_t(\cdot)$ *solves the Stratonovich stochastic differential equation*

$$\begin{cases} dX_t = \left[b(X_t,t) - \frac{1}{2}c(X_t,t)\right]dt + A(X_t,t)\circ dB_t \\ X(0) = X_0 \end{cases}$$

for

$$c^i(x,t) = \sum_{k=1}^m\sum_{j=1}^n \frac{\partial a^{ik}}{\partial x_j}(x,t)a^{jk}(x,t) \quad (1\le i\le n).$$

Example 2.4.5. *For $m = n = 1$, X_t solves*

$$dX_t = b(X_t)dt + g(X_t)dB_t$$

iff it solves

$$dX_t = \left[b(X_t) - \frac{1}{2}g'(X_t)g(X_t)\right] dt + g(X_t) \circ dB_t$$

Example 2.4.6. *Let us consider 2.4.2 again, it has the following Stratonovich formulation:*

$$\begin{cases} dX(t) = rX(t)dt + \sigma X(t) \circ dB(t) \\ X(0) = X_0 > 0, \end{cases}$$

To solve this SDE in the Stratonovich sense we can use the conversion rule to write it in the Ito sense and use the Ito formula to solve it. By the conversion rule the above equation equivalent to

$$dX(t) = [rX(t) + \frac{1}{2}\sigma^2 X(t)]dt + \sigma X(t)dB(t)$$

This SDE is in the Ito sense and has as a solution:

$$X(t) = X_0 e^{rt+\sigma B(t)}$$

We see this just by adding the term $\frac{1}{2}(\sigma X(t))'(\sigma X(t)) = \frac{1}{2}\sigma^2 X(t)$.

Remark 2.4.3. *Note that, $t \mapsto \int_0^t \sigma_s \circ dB(s)$ is usually not a martingale.*

2.5. SDE's driven by Semi-Martingales in $\mathbb{R}^n$

As was stated in previous chapters, stochastic integration, and in turn SDEs, can be defined not only for Brownian motion but for a general class of so-called good integrators. This class of stochastic processes is called semi-martingales and this section aims at giving a brief overview of how integrals and SDEs involving semi-martingales are defined. The study of semi-martingales is a very technical and deep topic and the following aims only at gathering the most important results, without proofs, which can be found in any standard reference on the topic such as: [RY99], [HWY92]. We begin our section by stating the definition of semi-martingales on $\mathbb{R}^n$. In the next definition, we introduce the notion of a semi-martingale. Semi-martingales are good "integrators" i.e. for this class of integrators we are able to define a notion of integral against them.

Definition 2.5.1. *An adapted process $X = (X_t)_{t\geq 0}$ is called a continuous semi-martingale if it can be written in the form:*

$$X = X_0 + M + A$$

where M is a continuous local martingale and A is a continuous, adapted, with finite variation stochastic process, where $M_0 = X_0$, $A_0 = 0$. Furthermore, if X is an $\mathbb{R}^n$-valued *stochastic process, i.e.* $X_t = (X_t^1, \ldots, X_t^n)$*, we say it is a semi-martingale if each component is a semi-martingale.*

Remark 2.5.1. *(Doob-Meyer) For* $X = (X_t)_{t\geq 0}$ *a semi.martingale, the decomposition into a local martingale* M *and continuous, adapted, with finite variation stochastic process* A*, where* $M_0 = X_0$*,* $A_0 = 0$ *is unique.*

Example 2.5.1. *If* X *is an Ito process satisfying the SDE* $dX_t = \alpha_t dB_t + a_t dt$ *then in this case we have:*

$$M_t = X_0 + \int_0^t \alpha_s dB_s \text{ and } A_t = \int_0^t a_s ds$$

1. $M_t(\omega) = X_0(\omega) + \int_0^t \alpha_s(\omega) dB_s(\omega)$ *is a local martingale. Indeed, by construction of the Ito integral we have that:* $\int_a^b \alpha_s^2(\omega) ds < \infty$*. A localizing sequence of stopping times* τ_n *can be defined as follows:*

$$\tau_n = \inf \left\{ s : \int_a^s \alpha_s^2(\omega) ds > n \right\} \wedge n$$

   with the property that: $\tau_n \uparrow t$*.*

2. $A_t(\omega) = \int_0^t a_s(\omega) ds$ *has a finite variation since:*

$$\sum_{k=1}^n |A_{t_k}(\omega) - A_{t_{k-1}}(\omega)| \leq \sum_{k=1}^n \int_{t_{k-1}}^{t_k} |a_s(\omega)| ds = \int_0^t |a_s(\omega)| ds < \infty$$

   for all partitions $\{t_0 = 0 < t_1 \cdots \leq t_n = t\}$*.*

Proposition 2.5.1. *If* Z_t *is an* $\mathbb{R}^l$ *semi-martingale and* $V : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^l$ *is a smooth vector field, then:*

$$X_t = X_0 + \int_0^t V(X_s) dZ_s \text{ and } Y_t = Y_0 + \int_0^t V(Y_s) \circ dZ_s$$

are well defined and X_t *and* Y_t *are semi-martingales.*

In the following theorem, we state the Ito formula for general semi-martingales in $\mathbb{R}^n$:

Theorem 2.5.1. *Let* X_t *be an* $\mathbb{R}^n$*-valued semi-martingale. Then for* $f \in C_c^\infty(\mathbb{R}^n)$ *we have:*

$$f(X_t) - f(X_0) = \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(X_s) dX_s^i + \frac{1}{2} \sum_{i=1}^n \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) d\left[X^i, X^j\right]_s \tag{2.20}$$

Definition 2.5.2. *Let X_t and Z_t be $\mathbb{R}^d$-valued semi-martingales and let $V : \mathbb{R}^d \to \mathbb{R}^d$ be smooth. Then the following defines the Stratonovich integral:*

$$\int_0^t V(X_s) \circ dZ_s := \int_0^t V(X_s)\, dZ_s + \frac{1}{2} \sum_{i,j=1}^{n} \int_0^t \frac{\partial V_i}{\partial x_j}(X_s)\, d\left[Z^i, Z^j\right]_s .$$

Definition 2.5.3. *Let $V_1 \ldots V_l : \mathbb{R}^n \to \mathbb{R}^n$ be smooth vector fields and Z_t an $\mathbb{R}^l$ semi-martingale. Let X_t be the solution of the Stratonovich SDE*

$$dX_t = V_\alpha(X_t) \circ dZ_t^\alpha$$

Then Ito formula states that for each $f \in C^\infty(\mathbb{R}^n)$,

$$df(X_t) = V_\alpha f(X_s) \circ dZ_t^\alpha,$$

where $V_\alpha f(x)$ is the directional derivative of f in the direction V_α at the point x, i.e. $V_\alpha f(.) = \nabla f(.) \nabla V_\alpha f(.)$. Equivalently:

$$df(X_t) = V_\alpha f(X_t) dZ_t^\alpha + \frac{1}{2} \nabla_{V_\beta} V_\alpha f(X_s) d[Z^\alpha, Z^\beta]_t \tag{2.21}$$

2.6. Second order differential operators associated to SDEs

We conclude this chapter by presenting a strong relationship between SDEs and second-order differential operators. This will be useful when constructing Brownian motion on manifolds in the next chapter. We start this section with a motivating discussion related to the flows of ODEs. We know that for a given vector field V we can construct a **flow** of V. Let us formulate this in the following definition:

Definition 2.6.1. *Given a vector field V on a manifold M, consider the smooth curve $t \mapsto x(t)$ in M defined as:*

$$\forall x_0 \in M : \quad \begin{aligned} \dot{x}(t) &= V(x(t)) \\ x(0) &= x_0 \end{aligned}$$

Then the corresponding flow curve to V passing x_0 (Initial point), $(t \mapsto \Phi_t(x) := x(t))$ is given by:

$$\forall f \in C_c^\infty(M), \quad \begin{aligned} \frac{d}{dt} f(\Phi_t) &= V f(\Phi_t) \\ f(\Phi_0) &= f. \end{aligned}$$

In simple words, $\Phi_t(x)$ is the flow curve corresponding to V, if for every $f \in C_c^\infty(M)$ and every $x \in M$:

$$f(\Phi_t(x)) - f(\Phi_0(x)) - \int_0^t V f(\Phi_s(x))\, ds = 0$$

A natural question would then be to ask if we have something similar but for second-order partial differential operators on a manifold M. We first define the differential operators we have in mind:

Note 2.6.1. *L be a second order partial differential operator on M, for instance $L = V_0 + \sum_{i=1}^{r} V_i^2$, where $V_0, V_1, \ldots, V_r$ are vector fields and V_i^2 is thought of as, for $f \in C_c^\infty(M)$,*

$$V_i^2(f) = V_i(V_i(f)).$$

Example 2.6.1. *$M = \mathbb{R}^n$, the Laplace operator Δ is a second-order partial differential operator. Indeed, for a function $f \in C_c^\infty(M)$:*

$$\Delta f = \sum_{i=1}^{r} \partial_i^2 f = \sum_{i=1}^{r} \partial_i(\partial_i f)$$

A notion of a flow corresponding to L exists, however, the flow lines are no longer deterministic, i.e. they are stochastic processes. The following states what we mean more explicitly:

Definition 2.6.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, an adapted continuous M-valued process $(X_t(x))_{t\geq 0} = (X_t(x,\omega))_{t\geq 0}$ is called flow process to L or L-diffusion with a starting point $X_0(x) = x$ if for all $f \in C_c^\infty(M)$, the process:*

$$N_t^f(x) := f\left(X_t(x)\right) - f(x) - \int_0^t (Lf)\left(X_s(x)\right) ds,$$

for all $t \geq 0$, is a martingale. This means that, for $s \leq t$, we have:

$$\mathbb{E}\left[N_t^f(x) - N_s^f(x) | \mathcal{F}_s\right] = 0.$$

By the martingale property $\mathbb{E}\left[N_t^f(x) - N_0^f(x)\right] = 0$ and since, $N_0^f(x) = 0$, we have that:

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

Now we can define:

$$P_t f(x) = \mathbb{E}\left[f\left(X_t(x)\right)\right].$$

We observe:

$$\mathbb{E}\left[f\left(X_t(x)\right)\right] = f(x) + \int_0^t \mathbb{E}\left[(Lf)\left(X_s(x)\right)\right] ds.$$

Thus,

$$\frac{d}{dt} P_t f(x) = \mathbb{E}\left[(Lf)\left(X_t(x)\right)\right] = P_t(Lf)(x).$$

Note that we can "evaluate" at $t = 0$ to get:

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

Thus we see that we can recover the second-order partial differential operator L from its stochastic flow.

Example 2.6.2. *Let $M = \mathbb{R}^n, B = (B_t)$ an n-dimensional Brownian motion on $\mathbb{R}^n$ and $L = \Delta$, where Δ is the Laplacian on $\mathbb{R}^n$ i.e $\Delta = \sum_{i=1}^n \partial_i^2$. By Ito's formula, for $f \in C_c^\infty(M)$, we have that:*

$$\begin{aligned} d(f \circ B_t) &= \sum_{i=1}^n \partial_i f(B_t) dB_t^i + \frac{1}{2} \sum_{i,j=1}^n \partial_i \partial_j f(B_t) d[B^i, B^j]_t \\ &= \left\langle (\nabla f)(B_t), dB_t \right\rangle + \frac{1}{2}(\Delta f)(B_t) dt \end{aligned}$$

Thus:

$$f(B_t) - f(B_0) - \frac{1}{2}\int_0^t (\Delta f)(B_s) ds$$

is a martingale, since "$df(B_t) - \frac{1}{2}(\Delta f)(B_t)dt = d(martingale)$". This means by the above definition that $X_t(x) := x + \sqrt{2}B_t$ is an L-diffusion to $L = \Delta$.

Remark 2.6.1. *The previous example shows that Brownian motion can be seen as a stochastic flow to the Laplacian operator on $\mathbb{R}^n$. This can be extended to a general manifold M and hence we construct a Brownian motion on a manifold. This will be the goal of the next chapter.*

3. Stochastic Differential Geometry

In this chapter, we start with an introduction to frame bundles, horizontal lifts, anti-developments, and their use in the context of stochastic differential equations on manifolds to present the Eells-Elworthy-Malliavin construction of Brownian motion on manifolds. Moreover, we present different descriptions of Brownian motion on the sphere and compare them with the Eells-Elworthy approach as well as give various characterizations of Brownian motion on manifolds. Finally, we use the concepts and tools we defined to look at an essential class of semi-martingales, namely "Γ-martingales on manifolds". We will see that Brownian motion on a Riemannian manifold is an example of a manifold-valued martingale. We finish by examining this concept and provide an example of a Γ martingale which is also a martingale in the ordinary sense. In contrast to the previous chapters, we give proofs for many of the results stated in this chapter but we defer to [Hsu02] and [HT94] for the proofs of a handful of results when these are too involved.

3.1. Frame Bundles, Horizontal lifts and Development of a curve

Definition 3.1.1. ***(Frame bundles)*** *Let (M, g) be a Riemannian manifold. A frame at $x \in M$ is a vector space isomorphism $u : \mathbb{R}^d \to T_xM$. Let $\{e_1, \ldots, e_d\}$ be the coordinate unit vectors of $\mathbb{R}^d$. Then the family of tangent vectors $(ue_1, \ldots, ue_d)$ (or equivalently, a frame) form the basis for the tangent space T_xM. We use $\mathcal{F}(M)_x$ to denote the space of all frames at x. The general linear group $\mathbb{GL}(d, \mathbb{R})$ acts on $\mathcal{F}(M)_x$ by $u \to uf$, where uf denotes the composition*

$$\mathbb{R}^d \xrightarrow{f} \mathbb{R}^d \xrightarrow{u} T_xM \tag{3.1}$$

The set

$$\mathcal{F}(M) = \bigcup_{x \in M} \mathcal{F}(M)_x$$

can be turned into a differentiable manifold of dimension $d + d^2$, it can be even made into a fiber bundle over M with canonical projection $\pi : \mathcal{F}(M) \to M$. Notice that,

- *The canonical projection π is a smooth map and induces a linear mapping i.e the tangent map:*

$$T_{(x,u)}\pi : T_{(x,u)}\mathcal{F}(M) \to T_xM \tag{3.2}$$

The group $\mathbb{GL}(d, \mathbb{R})$ acts on $\mathcal{F}(M)$ freely, transitively on each fiber $\mathcal{F}(M)_x$ in the sense that:

1. Transitive means:

$$\forall u, v \in \mathcal{F}(M)_x, \exists f \in \mathbb{GL}(d, \mathbb{R}) : v = uf$$

2. Free means:

$$uf = u \iff f = id$$

Each fiber $\mathcal{F}(M)_x$ is diffeomorphic to $\mathbb{GL}(d, \mathbb{R})$.

- *$\mathcal{F}(M)$ can be viewed as a principal bundle over M. See [Hus94], or [Lee00].*

Therefore:

$$\mathbb{GL}(d, \mathbb{R}) \simeq \mathcal{F}(M)_x$$

The tangent bundle in our case is simply the associated bundle (see [Sch17]), because of the natural map given by:

$$TM := \mathcal{F}(M) \times_{\mathbb{GL}(d,\mathbb{R})} \mathbb{R}^d \to \mathbb{R}^d \tag{3.3}$$

$$(u, e) \mapsto ue. \tag{3.4}$$

3.1.1. Vertical and Horizontal subspaces of $T\mathcal{F}(M)$

In the following, we assume that M is equipped with a connection ∇.

Definition 3.1.2.
- *A tangent vector $X \in T_{(x,u)}\mathcal{F}(M)$ is called **vertical** if it is tangent to the fiber $\mathcal{F}(M)_{\pi u}$, i.e $X \in kerT\pi$*
- *The space of vertical vectors at u is denoted by $V_u\mathcal{F}(M)$*

Remark 3.1.1. *The space of vertical vectors at u is a subspace of $T_u\mathcal{F}(M)$ of dimension d^2. The canonical projection $\pi : \mathcal{F}(M) \to M$ induces the tangent map $T\pi$ and we have:*

$$kerT\pi = \left\{v \in T_{(x,u)}\mathcal{F}(M) : T_{(x,u)}\pi(v) = 0\right\} = V_{(x,u)}\mathcal{F}(M) \subseteq T_{(x,u)}\mathcal{F}(M) \tag{3.5}$$

Remark 3.1.2. *The vertical subspace represents the part of the tangent vectors that "point" along the fibers of the bundle, i.e., along the directions corresponding to the different frames associated with each point on the manifold.*

Definition 3.1.3. *A curve $(u_t)_t$ in $\mathcal{F}(M)$ is said to be over a curve x_t on M, if $\pi \circ u_t = x_t$ for all t. u_t is then said to be lifting of the curve x_t to the principle bundle $\mathcal{F}(M)$.*

Definition 3.1.4. *A tangent vector $X \in T_{(x,u)}\mathcal{F}(M)$ is called **horizontal** if it is the tangent vector of a horizontal curve starting at (x, u).*
We define below the space of horizontal vectors at (x, u) denoted by $H_{(x,u)}\mathcal{F}(M)$. By the first isomorphism theorem (see A.0.1), we get:

$$H_{(x,u)}\mathcal{F}(M) := T_{(x,u)}\mathcal{F}(M)/kerT_{(x,u)}\pi \simeq T_xM \tag{3.6}$$

It is a subspace of dimension d, and we have the decomposition:

$$T_{(x,u)}\mathcal{F}(M) = V_{(x,u)}\mathcal{F}(M) \oplus H_{(x,u)}\mathcal{F}(M) \tag{3.7}$$

By 3.6 the projection map $\pi : \mathcal{F}(M) \to M$ induces an isomorphism:

$$\pi_* : H_{(x,u)}\mathcal{F}(M) \to T_{(x,\pi u)}M$$

Thanks to this decomposition of $T\mathcal{F}(M)$ for any $X \in T_xM$ and a frame u at x there is a unique horizontal lift of X to u, such that $\pi_* X^* = X$. Thus, if X is a vector field on M, then X^* is a horizontal vector field on $\mathcal{F}(M)$.

Remark 3.1.3.

1. *We saw that, the tangent frame bundle can be decomposed into a horizontal and vertical subspace. The horizontal subspace corresponds to the tangent vectors in the tangent bundle of the base manifold. The vertical subspace, on the other hand, consists of the tangent vectors along the fibers of the bundle.*

2. *The above says that $V\mathcal{F}(M)$ consists of derivatives of paths v_t satisfying $\pi \circ v_t = x_t$ for all t, meaning that the curve takes values in a single fiber. The curve u_t is called* ***horizontal*** *if for each $e \in \mathbb{R}^d$ the vector field $u_t e$ is parallel along πu_t, as pictorially shown in 3.2*

3. *Paths in $H\mathcal{F}(M)$ are the ones along which each basis vector $u_i(t)$ is parallel-transported along the corresponding path in M, namely πu_t pictorially shown in 3.2*

Remark 3.1.4.

- *The above definitions show that a connection ∇ on M gives rise to a choice of a complement $H_{(x,u)}\mathcal{F}(M)$ of the vertical subspace $V_{(x,u)}\mathcal{F}(M)$ in $T_{(x,u)}\mathcal{F}(M)$ at each point (x,u) "See 3.2". The subspaces $H_{(x,u)}M$ satisfy the equivariance property: For $e \in \mathbb{R}^d$ and $f \in \mathbb{GL}(d,\mathbb{R})$. Then*

$$f_* H_e(u) = H_{fe}(fu), \quad \forall u \in \mathcal{F}(M),$$

  where $f_* : T_u\mathcal{F}(M) \to T_{uf}\mathcal{F}(M)$ is the action of f on the tangent bundle $T\mathcal{F}(M)$ induced by the canonical action $f : \mathcal{F}(M) \to \mathcal{F}(M)$ defined earlier.

- *The converse is also true, namely a smooth assignment $u \to H_u\mathcal{F}(M)$ of a d-dimensional subspace of $T_u\mathcal{F}(M)$ complementing $V_u\mathcal{F}(M)$ at each point $u \in \mathcal{F}(M)$ and satisfying the above equivariance property, corresponds to a connection ∇ such that $H_u\mathcal{F}(M)$ is its horizontal vector space.*

- *Given a curve $\{x_t\}$ and a frame u_0 at x_0, there is a unique horizontal curve $\{u_t\}$ such that $\pi u_t = x_t$. It is called the horizontal lift of x_t from u_0 "See 3.2". Moreover, the linear map*

$$\tau_{t_0 t_1} = u_{t_1} u_{t_0}^{-1} : T_{x_{t_0}}M \to T_{x_{t_1}}M \tag{3.8}$$

  is independent of the choice of the initial frame u_0 and is called the parallel translation (or parallel transport) along $\{x_t\}$ from x_{t_0} to x_{t_1}

- *For each $e \in \mathbb{R}^d$, the vector field H_e on $\mathcal{F}(M)$ defined at $u \in \mathcal{F}(M)$ by the relation*

$$H_e(u) = (ue)^* = \text{the horizontal lift of } ue \in T_{\pi u}M \text{ to } u \tag{3.9}$$

is a horizontal field on $\mathcal{F}(M)$. Let $e_1, \dots, e_d$ be the coordinate unit vectors of $\mathbb{R}^d$. Then $H_{e_i}(u) = H_i(u)$, $i = 1, \dots, d$, are the fundamental horizontal fields of $\mathcal{F}(M)$; they span $H_u\mathcal{F}(M)$ at each $u \in \mathcal{F}(M)$.

Figure 3.1 gives a summary of how the manifold M, frame bundle $\mathcal{F}(M)$, the horizontal subspace $H\mathcal{F}(M)$ and the vertical subspace $V\mathcal{F}(M)$ interact with each other.

$$\begin{tikzcd}
T\mathcal{F}(M) \arrow[r, "H_e(u)+v \mapsto v"'] \arrow[d, "H_e(u)+v \mapsto H_e(u)"] & V\mathcal{F}(M) \arrow[d] & \\
H\mathcal{F}(M) \arrow[r] \arrow[dr, "\pi_*"] & \mathcal{F}(M) \arrow[r, "\pi"'] & M \\
& TM &
\end{tikzcd}$$

Figure 3.1.: Relationship between the Manifold M, horizontal subspace $H\mathcal{F}(M)$ and the vertical subspace $V\mathcal{F}(M)$

The action of $\mathbb{GL}(d, \mathbb{R})$ on $\mathcal{F}(M)$ preserves the fundamental horizontal fields in the sense described in the following proposition:

Remark 3.1.5. *A local chart $x = \{x^i\}$ on a neighborhood $U \subseteq M$ induces a local chart on $\tilde{U} = \pi^{-1}(U)$ in $\mathcal{F}(M)$ as follows: Let $X_i = \dfrac{\partial}{\partial x^i}, 1 \leq i \leq d$, be the moving frame defined by the local chart $\{x^i\}$. For a frame $u \in \tilde{U}$ we have $ue_i = e_i^j X_j$ for some matrix $e = \left(e_i^j\right) \in \mathbb{GL}(d, \mathbb{R})$. Then $(x, e) = \left(x^i, e_j^i\right) \in \mathbb{R}^{d+d^2}$ is a local chart for $\widetilde{U}$. Let $z = (x, e)$ then the vertical vector fields have the following expression locally:*

$$V_i(z) = e_i^k \frac{\partial}{\partial e_j^k}$$

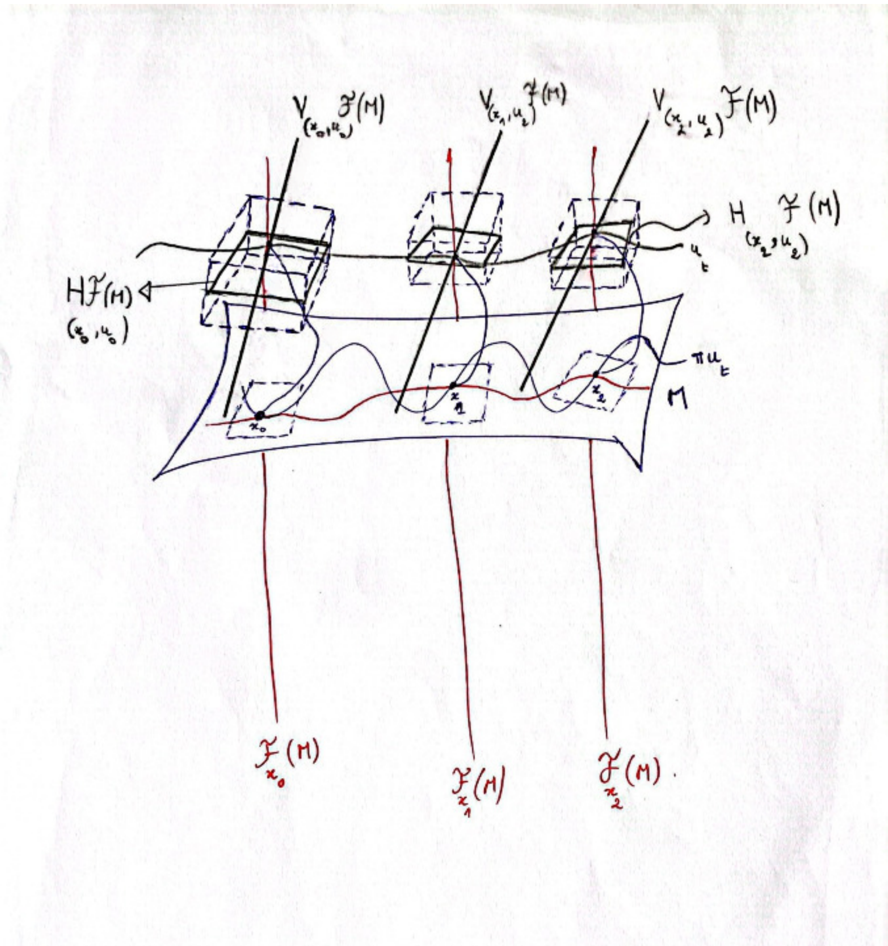

Figure 3.2.: Frame bundle and Horizontal lift of a curve on a manifold

Note that, the vector fields $\left\{ \frac{\partial}{\partial x^i}, \frac{\partial}{\partial e_j^k}, 1 \leq i, j \leq d \right\}$ *span the tangent space* $T_u\ \mathcal{F}(M)$ *for every* $u \in \tilde{U}$.

The local expression of the fundamental horizontal vector fields H_i is more relevant than the one for vertical vector fields, hence we provide this in the following proposition:

Proposition 3.1.1. *[Hsu02]Proposition 2.1.3. In terms of the local chart on* $\mathcal{F}(M)$ *described above, at* $u = (x, e) = \left(x^k, e_j^k\right) \in \mathcal{F}(M)$ *we have*

$$H_i(u) = e_i^j X_j - e_i^j e_m^l \Gamma_{jl}^k(x) X_{km},$$

where

$$X_i = \frac{\partial}{\partial x^i}, \quad X_{km} = \frac{\partial}{\partial e_m^k}.$$

Proof. $H_i(u)$ is the horizontal lift of ue_i with e_i the i-th coordinate unit vector of $\mathbb{R}^n$. Let $(x_t, e(t))$ be a horizontal curve starting at an initial frame u_0 corresponding to the tangent vector ue_i. In other words, if we project down $\dot{u_0}$ using π_*, we should recover ue_i i.e. $\pi_*\dot{u}_0 = ue_i$. Using the decomposition of $T\mathcal{F}(M)$ we established earlier, one can write an expression for the horizontal lift of ue_i:

$$H_i(u) = \dot{u_0} = \dot{x}_0^j X_j + \dot{e}_m^k(0)X_{km}$$

Since we have a horizontal curve $u_t = (x_t, e(t))$ then for each e_m the vector field $V_t = u_t e_m$ should be parallel along $\pi u_t = x_t$. By 1.4 the equation for parallel vector fields, we obtain

$$\dot{e}_m^k(t) + \Gamma_{jk}^k(x_t)\dot{x}_t^j e_m^l(t) = 0$$

On the other hand,

$$\begin{aligned}
\pi_*\dot{u}_0 &= \pi_*(\dot{x}_0^j X_j + \dot{e}_m^k(0)X_{km}) && (3.10)\\
&= \dot{x}_0^j X_j|_{\pi u} && (3.11)\\
&= u_0 e_i && (3.12)\\
&= e_i^j(0)X_j && (3.13)
\end{aligned}$$

So $\dot{x}_0^j = e_i^j(0)$. Moreover, plugging in for $t = 0$ in the equation for parallel transport we get

$$\dot{e}_m^k(0) = -\Gamma_{jl}^k(x)e_i^j e_m^l$$

In summary, we get:

$$H_i(u) = e_i^j X_j - \Gamma_{jl}^k(x)e_i^j e_m^l X_{km}.$$

□

It turns out that we will want to restrict our attention to a smaller class of frames, namely the orthonormal frames which are defined as follows:

Definition 3.1.5. *Let $\mathcal{O}_x(M)$ be the set of all orthonormal frames of the tangent space T_xM. An element u of $\mathcal{O}_x(M)$ is a Euclidean isometry from $\mathbb{R}^d$ into T_xM. If we have a manifold with a connection that is compatible with the Riemannian metric 2, then orthonormality is defined with respect to the Riemannian metric g such that:*

$$g(ue_i, ue_j) = \delta_{ij},$$

when $u \in \mathcal{O}_x(M)$

Note 3.1.1. *Everything we developed for the frame bundle $\mathcal{F}(M)$ carries over to the orthonormal frame bundle $\mathcal{O}(M)$.*

Definition 3.1.6. *Let u_t be the horizontal lift of a curve x_t on M. We know that $\dot{x}_t \in T_{x_t}M$. Thus, $u_t^{-1}\dot{x}_t \in \mathbb{R}^d$. The anti-development of the curve x_t (or the horizontal curve u_t) is a curve w_t in $\mathbb{R}^d$ defined as $\displaystyle\int_0^t u_s^{-1}\dot{x}_s ds$.*

Remark 3.1.6. *Note that, w_t is constructed from x_t and the connection endowed on M. We used the connection to construct a parallel transported frame on the new curve w_t, this suggests that the curve w_t depends on the initial frame u_0 at x_0.*

Remark 3.1.7. *A relation between the horizontal and the anti-development can be derived as follows:*
Since $u_t\dot{w}_t = \dot{x}_t$ and $w_t \in \mathbb{R}^d$ we can write:

$$\dot{w}_t = \sum_{i=1}^{d} \dot{w_t}^i e_i$$

Then:

$$H_{\dot{w}_t}(u_t) = (u_t\dot{w}_t)^* = (\dot{x}_t)^* = \dot{u}_t$$

On the other hand,

$$H_{\dot{w}_t}(u_t) = \sum_{i=1}^{d} \dot{w_t}^i H_{e_i}(u_t)$$

Therefore, the horizontal lift and the anti-development are connected by the ODE:

$$\dot{u}_t = \dot{w_t}^i H_{e_i}(u_t) \tag{3.14}$$

Now let us give some intuition of the above using our previous analogy: Let us give a more concrete example:

Example 3.1.1. *Let $M = \mathbb{R}$ with a general connection given by $\nabla_e e = \Gamma(t)e$, where e is the usual unit vector field in $\mathbb{R}$ and $\Gamma(t)$ is a given function that defines our connection. In other words, for a differentiable function f on $\mathbb{R}$, $e(f) = f'$ i.e., the derivative of f in the direction e.*
Consider the line $x_t = t$, locally the frame u_t can be expressed by $u_t^i\partial_e$. Let $u_0\partial_e$ be the initial frame, where $u_0 = (0,1)$. What is the horizontal lift of x_t from u_0? Well, let us formulate the condition that the vector field $u_t^i\partial_e$ is parallel along x_t:

$$\nabla_{\partial_e}(u_t^i\partial_e) = 0 \iff \dot{u_t}^i = -\Gamma(t)\dot{u}_t^i$$

We introduce:

$$K(x) = \int_0^x \Gamma(y)dy,$$

and

$$\psi(x) = \int_0^x e^{K(y)}dy$$

With the canonical embedding $\mathbb{GL}(\mathbb{R}) = \mathbb{R} \times \mathbb{R}$, the horizontal lift of x_t is therefore described by $\left(x_t, e^{-K(x_t)}\right)$ and the anti-development w_t is defined as $\displaystyle\int_0^x e^{K(y)}dy$.

Remark 3.1.8. *In the above example, parallel translation means a change in the length of the unique unit tangent vector. We wrote down the parallel transport of the initial vector (the unit vector) along the line $x_t = t$ under the connection defined by the function Γ and this translates to a contracting or a stretching effect.*

Next, we will extend the concepts introduced in the current section to semi-martingales on $\mathbb{R}^n$ which in turn will lead us to the notion of stochastic development on $\mathcal{F}(M)$. We will see that the projection of the stochastic development on $\mathcal{F}(M)$ will yield the corresponding notion of semi-martingales on manifolds. Before that, let us introduce the notions of SDEs and semi-martingales on manifolds.

3.2. SDE's, Ito formula and Semi-martingales on smooth manifolds

Definition 3.2.1. *Let M be a manifold and $V_1 \dots V_l : \mathbb{R}^n \to TM$ be smooth vector fields and Z_t a $\mathbb{R}^l$-semi-martingale. An SDE on a manifold M is defined by the vector fields $V_1 \dots V_l$, the $\mathbb{R}^l$-valued driving semi-martingale Z and an M-valued random variable $X_0 = x_0$ serving as the initial value of the solution. We write the SDE in Stratonovich form as:*

$$dX_t = V_\alpha(X_t) \circ dZ_t^\alpha \tag{3.15}$$

almost surely for all $t \geq 0$

Definition 3.2.2. *X_t is called a solution to 3.15 with initial condition x_0, if for all $f \in C^\infty(M)$:*

$$f(X_t) - f(X_0) = \int_0^t V_\alpha f(X_s) \circ dZ_s^\alpha.$$

almost surely for all $t \geq 0$

Remark 3.2.1. *The solution exists and is unique, as we are considering smooth vector fields*

Example 3.2.1. *Let $Z' = (Z^1, \dots, Z^r)$ be an $\mathbb{R}^r$ Brownian motion and consider $Z = (t, Z')$. Denote the standard basis of $\mathbb{R}^{r+1}$ by $(e_0, \dots, e_r)$ and the vector fields $V_i = V(.)e_i$, then the associated SDE is:*

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

That reads:

$$dX = V_0(X)dt + \sum_{i=1}^{r} V_i(X) \circ dZ^i$$

For $f \in C^\infty(\mathbb{R}^r)$*:*

$$\begin{aligned} d(f \circ X) = \sum_{i=0}^{r} (V_i f)(X) \circ dZ^i &= (V_0 f)(X)dt + \sum_{i=1}^{r} (V_i f)(X) \circ dZ'^i \\ &= (V_0 f)(X)dt + \sum_{i=1}^{r} (V_i f)(X) dZ'^i + \sum_{i=1}^{r} (V_i^2 f)(X)dt \\ &= (Lf)(X)dt + \sum_{i=1}^{r} (V_i f)(X) dZ'^i \end{aligned}$$

Thus:

$$f(X_t) - f(X_0) - \int_0^t (Lf)(X_s)ds$$

is a local martingale, since "$df(X_t) - (Lf)(X_t)dt = d(\text{local martingale})$*", where* $L = V_0 + \sum_{i=1}^{r} V_i^2$ *is a second order partial differential operator. This can be upgraded to a martingale if we consider* $f \in C_c^\infty(\mathbb{R}^r)$

The main takeaway from the previous example is that solutions to SDEs are stochastic flows of such sum of squares operator. We know that Brownian motion is an L-diffusion 2.6.2, in other words, a stochastic flow of some second order partial differential operator L i.e., Brownian motions are stochastic flows to the Laplace-Beltrami operator Δ^M. In contrast, one cannot write Δ^M as a sum of squares as we did before, this is an involved issue, thus,we refer to "[Pau98]". However, there is a lifting of Δ^M to $\mathcal{O}(M)$ which can be written as the sum of squares of $dim(M)$-vector fields on $\mathcal{O}(M)$.

Definition 3.2.3. *The operator:*

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

is called Bochner's Horizontal Laplacian on $\mathcal{O}(M)$

Theorem 3.2.1. *[Hsu02]* ***Page****75 Bochner's Horizontal Laplacian is the lift of the Laplace-Beltrami operator* Δ^M *to the orthonormal frame bundle* $\mathcal{O}(M)$*, their relation is best described as follows: Let* $f \in C^\infty(M)$*, and* $f \circ \pi$ *its lift to* $\mathcal{O}(M)$*. Then, for any* $u \in \mathcal{O}(M)$*, such that* $\pi u = x$ *we have the following identity:*

$$\Delta^M f(x) = \Delta^{hor}(f \circ \pi)(u).$$

The Eells-Elworthy-Malliavin description of Brownian motion on manifolds builds on Δ^{hor} in order to define Brownian motion over a manifold. We will provide a thorough illustration of the Eells-Elworthy-Malliavin description of Brownian motion on manifolds once we define the necessary tools. Now let us define a semi-martingale on a manifold M:

Definition 3.2.4. *We say that X_t is an M-valued semi-martingale if for every $C^\infty(M)$ coordinate function f, $f(X_t)$ is an $\mathbb{R}$-valued semi-martingale "See remark 2.5.3 in [Hsu02]", equivalently to say: If it is a solution of 3.15 in the sense presented in 3.2.2 for all $f \in C^\infty(M)$.*

In the following, we assume that M is equipped with a connection, and $\{H_{e_i}\}$ are the horizontal vector fields on $\mathcal{O}(M)$. All processes are defined on a fixed, filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and are $\mathcal{F}$-adapted. Let us begin with a definition:

Definition 3.2.5.
1. *U_t is said to be an $\mathcal{O}(M)$-valued horizontal semi-martingale if there exists an $\mathbb{R}^d$-valued semi-martingale W such the following SDE holds on $\mathcal{O}(M)$:*

$$dU_t = \sum_{i=1}^{d} H_{e_i}(U_t) \circ dW_t^i$$

   or in a more compact form

$$dU_t = H(U_t) \circ dW_t. \tag{3.16}$$

   W_t is called the anti-development of U

2. *Let W be an $\mathbb{R}^d$-valued semi-martingale. U_0 an $\mathcal{F}_0$-measurable random variable, and an initial orthonormal frame. The solution to U to 3.16 is called a "**Stochastic development of** W" in $\mathcal{O}(M)$, on the other hand, its projection to the manifold M, $\pi U = X$ is called a "**Stochastic development of** W" in M.*

3. *Let X be an M-valued semi-martingale. An $\mathcal{O}(M)$-valued horizontal semi-martingale such that its projection $\pi U = X$ is called a "**Stochastic horizontal lift of** X" in $\mathcal{O}(M)$.*

With these tools, we can naturally extend the concept of local martingales on Euclidean spaces to that on manifolds equipped with a connection.

Definition 3.2.6. *[Hsu02],* ***definition****2.5.1.*
Let M be a manifold equipped with a connection ∇. An M-valued semi-martingale X is called a Γ-martingale if its anti-development W with respect to ∇ is an $\mathbb{R}^d$-valued local martingale.

We want to give an alternative definition of Γ-martingales on M without appealing to anti-development. Before doing so, we need to define the quadratic variation of X on M. In the case where $M = \mathbb{R}^d$ equipped with the canonical metric, we have the usual concept of quadratic variation of a semi-martingale X, whereas, on a manifold, we don't have a notion of ordinary multiplication, so the idea is to replace it with an arbitrary bi-linear form i.e. $(0, 2)$-covariant tensor h. We give two definitions, one through horizontal lift and anti-development and another in a more general setting.

Definition 3.2.7. *Let ω be a 1-form on M, x_t a smooth curve on M. Fixing a frame u_0 at x_0 we know that x_t has a corresponding horizontal lift u_t from u_0. The anti-development of x_t is an n-dimensional valued curve and it is defined as follows:*

$$dx_t = u_t \circ dw_t^\alpha.$$

In this case, we define the stochastic line integral of ω along x_t as:

$$\int_{x[0,t]} \omega = \int_0^t \omega(u_s e_i) \circ dw_s^\alpha.$$

Now we derive a definition of the line integral along X in terms of its horizontal lift and anti-development.

Definition 3.2.8. *Let ω be a 1-form on M and X an M-valued semi-martingale. Let U be the horizontal lift of X and W be its anti-development (with respect to some connection). Then, the line stochastic integral of ω along X is given by:*

$$\int_{X[0,t]} \omega = \int_0^t \omega(U_s e_i) \circ dW_s^i.$$

Remark 3.2.2.

- *The analog of the above definition would be simply the definition of a line integral of a scalar field f along a curve*
- *For a thorough construction of stochastic line integral we refer to section 2.4 from [Hsu02]*

Now we define the quadratic variation of X with respect to a $(0,2)$-tensor.

Definition 3.2.9. *[Hsu02]**Definition 2.4.6**. Let h be a $(0,2)$-tensor on M and X be an M-valued semi-martingale. Let U be a horizontal lift of X and W be its anti-development. Then the h-quadratic variation of X is defined as:*

$$\int_{X[0,t]} h(dX_s, dX_s) = \int_0^t h(U_s e_i, U_s e_j) d[W^i, W^j]_s$$

Example 3.2.2. *Let $x = \{x^i\}$ be a global coordinate system on M. A $(0,2)$-tensor can be written as*

$$h|_x = h_{ij}|_x dx^i \otimes dx^j, \quad h_{ij}|_x = h|_x \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right)$$

If $X = \{X^i\}$ is a semi-martingale on M, then

$$\int_0^t h\left(dX_s, dX_s\right) = \int_0^t h_{ij}\left(X_s\right) d[X^i, X^j]_s$$

Example 3.2.3. *Let $\{u_t\}$ be a horizontal curve such that $u_0 = u$ and $H_i(u) = i u_0$ and two vector fields X and Y such that $X = u_t e_i$ and $Y = u_t e_j$. Furthermore, let $f : M \to \mathbb{R}$ be the coordinate function. Then, the Hessian $\nabla^2 f$ is a $(0,2)$-tensor, and is defined "See 1.3.5" as follows:*

$$\nabla^2 f\,(u_t e_i, u_t e_j) := u_t e_i(u_t e_j f) - (\nabla_{u_t e_i} u_t e_j) f$$

Let us relate the Hessian of f with $H_i H_j \tilde{f}$ where $\tilde{f} : \mathcal{F}(M) \to \mathbb{R}$ defined by

$$\tilde{f}(u) = f(\pi u) = \pi u = x \in M$$

is the lift of f to $\mathcal{F}(M)$. The vector field $Y(x_t) = u_t e_j$ is parallel along $x_t = \pi u_t$. In other words, by 1.3.4 $\nabla_{u_t e_i} Y = 0$; hence by definition $\nabla^2 f\,(u_t e_i, u_t e_j) = u_t e_i(u_t e_j f)$.

$$H_j \tilde{f}(u) = Y f(x_t) = (u_t e_j) f$$

Now let us look at:

$$H_i H_j \tilde{f}(u) = \frac{d\,\{Y f(x_t)\}}{dt} = \nabla_{u e_i} Y f = \nabla^2 f\,(u_t e_i, u_t e_j)$$

where f is the coordinate function on M. Thus, we have by this equality and 3.2.9

$$\int_0^t \nabla^2 f\,(dX_s, dX_s) = \int_0^t H_i H_j \tilde{f}\,(U_s)\, d[W^i, W^j]_s \tag{3.17}$$

In the following, we give a second definition of the concept of the quadratic variation of X with respect to a $(0,2)$-tensor field h on M. Note that, $\mathcal{A}$ is the set of continuous, adapted, real-valued with finite variation processes.

Proposition 3.2.1. *[HT94]**Satz 7.57**. Let X be a semi-martingale on M. Then, there exist a unique linear map $\mathcal{T}^{0,2}(T^*M, T^*M) \to \mathcal{A}$, $h \mapsto \displaystyle\int h(dX, dX)$ such that for all $f, g \in C^\infty(M)$ we have:*

$$df \otimes dg \mapsto [f(X), g(X)] \tag{3.18}$$

$$f.h \mapsto \int f(X) h(dX, dX) \tag{3.19}$$

Definition 3.2.10. *[HT94]**Definition 7.58**. The process $\displaystyle\int h(dX, dX)$ is said to be the integral of h along X or h-quadratic variation of X. The value of the h-quadratic variation of X at time t is denoted as $\displaystyle\int_0^t h(dX_s, dX_s)$*

Now we are ready to give an alternative definition of a Γ-martingale on M.

Proposition 3.2.2. *[Hsu02]* ***Proposition 2.5.2.*** *X is an M-valued semi-martingale X, it is a Γ-martingale if and only if:*

$$N^f(X)_t = f(X_t) - f(X_0) - \frac{1}{2}\int_0^t \nabla^2 f\,(dX_s, dX_s)$$

is an $\mathbb{R}$-valued local martingale for every $f \in C^\infty(M)$. A more elegant way of writing the above equation:

$$df(X_t) \stackrel{m}{=} \frac{1}{2}\nabla^2 f\,(dX_t, dX_t)$$

where m stands for modulo a continuous local martingale, i.e. $df(X_t) - \frac{1}{2}\nabla^2 f\,(dX_t, dX_t)$ is a continuous local martingale.

Proof. Let us give a sketch of the proof:
($\Longrightarrow$) To show that if X_t is a Γ-martingale, then $N^f(X)_t$ is an $\mathbb{R}$-valued local martingale for every $f \in C^\infty(M)$.

1. We lift X_t to U_t and lift $f : M \to \mathbb{R}$ to $\tilde{f} : \mathcal{O}(M) \to \mathbb{R}$
2. Apply Ito´s formula to $\tilde{f}(U_t)$ and the identity from the previous example 3.17, to get
$$f(X_t) - f(X_0) = \int_0^t H_i\tilde{f}(U_s)dW_s^i + \frac{1}{2}\int_0^t \nabla^2 f\,(dX_s, dX_s)$$
3. Therefore:
$$N^f(X)_t = \int_0^t H_i\tilde{f}(U_s)dW_s^i$$
is an $\mathbb{R}$-valued local martingale, for every $f \in C^\infty(M)$:

($\Longleftarrow$) To show that, if $N^f(X)_t$ is an $\mathbb{R}$-valued local martingale, for every $f \in C^\infty(M)$. Then, X_t is a Γ-martingale.

1. Assume M is embedded in an Euclidean space.
2. Consider f to be the coordinate function $f(x) = x$
3. We claim that, the anti-development W_t can be explicitly given by
$$W_t = \int_0^t V_s dN^f(X)_s$$
for some process V_s. If $N^f(X)_t$ is a local martingale, then so is W_t

□

Remark 3.2.3. *Note that in case M is embedded in $\mathbb{R}^d$ (M is a submanifold of $\mathbb{R}^d$) these test functions f are just the coordinate functions, and to show that X_t is a semi-martingale on M it is enough to show that each coordinate process is a $\mathbb{R}$-valued semi-martingale. "See remark 2.5.3 in [Hsu02]"*

The local version of this proposition:

Proposition 3.2.3. *[Hsu02]**Proposition 2.5.4**. Suppose that $x = \{x^i\}$ is a local chart on M and $X = \{X^i\}$ a semi-martingale on M. Then X is a Γ-martingale if and only if*

$$X_t^i = X_0^i + \text{ local martingale } - \frac{1}{2} \int_0^t \Gamma_{jk}^i (X_s)\, d[X^j, X^k]_s$$

Proof. The proof relies on relating the anti-development W of X and N, where N is the $\mathbb{R}^d$-valued semi-martingale defined by

$$N_t^i = X_t^i - X_0^i + \frac{1}{2} \int_0^t \Gamma_{jk}^i (X_s)\, d[X^j, X^k]_s.$$

using the local formulas for the horizontal vector fields. To start with, we apply the equation $dU_t = H_i(U_t) \circ dW_t^i$ to local coordinate functions $U_t = (X_t^i, e_j^i(t))$ and use the local formulas for the horizontal vector fields in 3.1.1. This gives the equations

$$\begin{cases} dX_t^i = e_j^i(t) \circ dW_t^j, \\ de_j^i(t) = -\Gamma_{kl}^i (X_t)\, e_j^l(t) \circ dX_t^k. \end{cases}$$

Let $\left\{f_j^i\right\}$ be the matrix inverse of $\left\{e_j^i\right\}$, then we have from the first equation

$$dW_t^j = f_k^j(t) \circ dX_t^k.$$

Now we have

$$dX_t^i = e_j^i(t) dW_t^i + \frac{1}{2}[de_j^i(t), dW_t^j]$$

Using the above equations for $de_j^i(t), dW_t^j$ respectively, we get:

$$dX_t^i = e_j^i(t) dW_t^i - \frac{1}{2}\Gamma_{kl}^i (X_t)\, d[X^k, X^l]_s$$

But, we defined N_t^i to be:

$$dN_t^i = dX_t^i + \frac{1}{2}\Gamma_{kl}^i (X_t)\, d[X^k, X^l]_t$$

Matching the two equations, we conclude that:

$$dN_t^i = e_j^i(t) dW_t^i \text{ and } dW_t^i = f_j^i(t) dN_t^i.$$

Thus, N is a local martingale if and only if W is a local martingale. □

Example 3.2.4. *$M = \mathbb{R}^n$ with the canonical connection ∇ (Directional derivative). Then, the Γ-martingales are nothing but all continuous local martingales. Indeed in this particular* *setting, we use Ito´s formula in 2.5.1 to argue that a continuous $\mathbb{R}^n$-semi-martingale is a local martingale if and only if, for every $f \in C^\infty(\mathbb{R}^n)$ it holds*

$$df(X_t) - \frac{1}{2}\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t)\, d\left[X^i, X^j\right]_t$$

is a local martingale. This is equivalent to:

$$df(X_t) \overset{m}{=} \frac{1}{2}\nabla^2 f\,(dX_t, dX_t)$$

In other words,

$$N^f(X)_t := \int_0^t \langle \nabla f(X_s), dX_s \rangle$$

Example 3.2.5. *Let X_t be an M-valued semi-martingale, recall our one-dimensional example 3.1.1, the horizontal lift of X_t is given by $U_t = \left(X_t, e^{-K(X_t)}\right)$. At this stage, we can formulate what it means for X_t to be a Γ-martingale (using 3.2.3 and 3.2.6). X_t is a Γ-martingale if and only if:*

1. *There is a local martingale M_t such that: $X_t = X_0 + M_t - \frac{1}{2}\int_0^t \Gamma(X_s) d[M, M]_s$.*
2. *It's anti-development $W_t = \psi(X_t)$ is an $\mathbb{R}$-valued local martingale.*

Both definitions are satisfied.

3.3. Stochastic anti-development and horizontal lift: Important theorems

Following up on the definitions we gave in 3.2.5, let us state some important theorems about stochastic anti-development and horizontal lift.

Theorem 3.3.1. *A horizontal semi-martingale U_t on the frame bundle $\mathcal{O}(M)$ has a unique anti-development W_t*

Proof. See [Hsu02] page 47 □

Theorem 3.3.2. *Suppose that X_t is a semi-martingale on M and $U_0 \in \mathcal{O}(M)$ with $\pi U_0 = X_0$. Then there is a unique horizontal lift U_t of X_t starting from U_0*

Proof. See [Hsu02] page 48 □

Proposition 3.3.1. *Let X_t be an M-semi-martingale with initial condition X_0 satisfying*

$$dX_t = V_\alpha(X_t) \circ dZ_t^\alpha.$$

Let V_α^* be the horizontal lift of V_α to $\mathcal{O}(M)$. Then:

1. *The horizontal lift U_t of X_t with initial frame U_0 is a solution to:*

$$dU_t = V^*_\alpha(U_t) \circ dZ^\alpha_t.$$

2. *The anti-development of X_t is given by:*

$$W_t = \int_0^t U_s^{-1} V_\alpha(X_t) \circ dZ^\alpha_t.$$

3.4. Eells-Elworthy Malliavin construction of Brownian motion

Theorem 3.4.1. *If W_t is a Brownian motion on $\mathbb{R}^n$, then the solution of*

$$dU_t = H_\alpha(U_t) \circ dW^\alpha_t$$

*on the orthonormal frame bundle $\mathcal{O}(M)$, where H_α are the fundamental horizontal vector field of $\mathcal{O}(M)$, is called the **horizontal Brownian motion** and its projection $X_t = \pi U_t$ is the **Brownian motion on** M*

Brownian motion on M can also be characterized using L-diffusion operators and this is summarized in the following theorem:

Theorem 3.4.2. *(Equivalent characterizations of Brownian motion on Manifolds) Suppose, X_t is a semi-martingale on M, where U_t its horizontal lift and W_t its anti-development. The following assertions are equivalent:*

1. *W_t is a Brownian motion on $\mathbb{R}^n$ i.e. W_t is a diffusion generated by the classical Laplace operator Δ in $\mathbb{R}^n$.*

2. *U_t is an L-diffusion on $\mathcal{O}(M)$ for $L = \Delta^{Hor} := \sum_{i=1}^n H_i^2$.*

3. *X_t is a Brownian motion on M i.e. X_t is a diffusion generated by the Laplace-Beltrami operator Δ^M. In other words, X is a solution to the martingale problem for $\frac{1}{2} \cdot \Delta^M$ i.e.,*

$$M^f(X)_t := f(X_t) - f(X_0) - \frac{1}{2}\int_0^t \Delta^M f(X_s) ds$$

is a local martingale for all $f \in C^\infty(M)$. This equation is equivalent to:

$$df(X_t) \overset{m}{=} \frac{1}{2}\Delta^M f(X_t) dt$$

Remark 3.4.1. *In case M is a closed submanifold of $\mathbb{R}^n$. A point $x \in M$ can be described by n-coordinates as a point in the ambient space $\mathbb{R}^n$. Then, as we pointed out in 3.2.3 the functions $f^i(x) = x^i$ are a suitable set of test functions for Ito´s formula on M. We will make use of this observation to check the third bullet of 3.4.2*

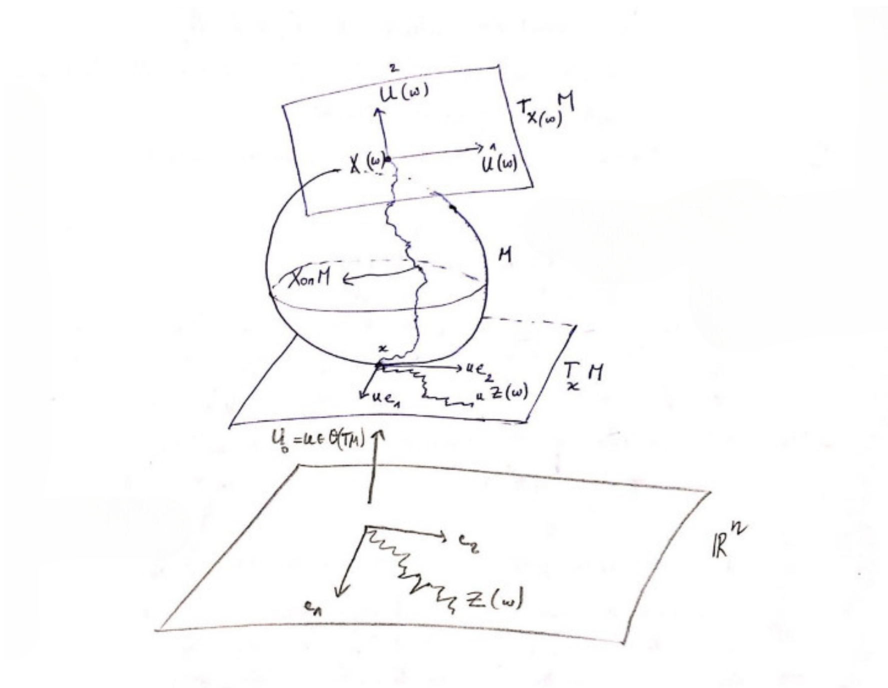

Figure 3.3.: Brownian motion construction on manifolds

Let us explain pictorially and in words the Eells-Elworthy-Malliavin construction using the following remark and 3.3:

Remark 3.4.2. *Figure 3.3 shows a visualization of the Eells-Elworthy-Malliavin construction. We start out with an ordinary Brownian motion process $Z(\omega)$ in $\mathbb{R}^n$ and an ordinary frame (e_1, e_2), we identify $\mathbb{R}^n$ with T_xM for some point $x \in M$ and get a corresponding lifted frame (ue_1, ue_2) at x. Now, we "roll" the manifold without slipping along this path to get a path traced on the manifold M. This results in a parallel translation of our initial frame (ue_1, ue_2) to (U_1, U_2) at $X(\omega)$. In the end, we get a process $X(\omega)$ on M called the development of $Z(\omega)$, i.e., a Brownian motion on Manifold M and a process $U(\omega)$ which is horizontal to the Brownian motion defined at the beginning.*

Now that we constructed the Brownian motion on M using Eells-Elworthy-Malliavin construction and gave it a characterization being the solution to a martingale problem on a Riemannian manifold "Third assertion in 3.4.2", let us summarize the general Lévy-characterization of Brownian motion on M in the following theorem:

Theorem 3.4.3. *[HT94]**Satz 7.116**. Let X be a semi-martingale on M. Then, the following are equivalent:*

1. *X is a Brownian motion on M.*
2. *X is a Γ-martingale with $[f(X), f(X)] = \int \| \operatorname{grad} f(X)\|^2 dt$ for every $f \in C^\infty(M)$*

Using 3.2.2, this is equivalent to have:

$$[N^f(X), N^f(X)]_t := \int \|\operatorname{grad} f(X)\|^2 dt$$

In this case, we have for all $(0,2)$-tensor fields h on M:

$$\int h(dX, dX) = \int \operatorname{tr} h(X) dt$$

Proof. Assume X with $[f(X), f(X)] = \int \|\operatorname{grad} f(X)\|^2 dt$ for every $f \in C^\infty(M)$. By polarization, we also have for every $f, g \in C^\infty(M)$:

$$[f(X), g(X)] = \int \langle \operatorname{grad} f(X), \operatorname{grad} g(X)\rangle dt$$

Now, take an orthonormal basis $\{F_i\}_{i=1}^n$ of $T_X M$, then:

$$\operatorname{tr} df \otimes dg = \sum_{i=1}^n (df \otimes dg)(F_i, F_i) \tag{3.20}$$

$$= \sum_{i=1}^n df(F_i) dg(F_i) \tag{3.21}$$

Using the definition we gave for the gradient in **??**, we rewrite the above expression to get:

$$\operatorname{tr} df \otimes dg = \sum_{i=1}^n \langle \operatorname{grad} f, F_i\rangle \langle \operatorname{grad} g, F_i\rangle = \langle \operatorname{grad} f, \operatorname{grad} g\rangle$$

But, by definition in 3.2.1 we have $[f(X), g(X)] = \int (df \otimes dg)(dX, dX)$. Using the above equation, we have on the other hand:

$$[f(X), g(X)] = \int \langle \operatorname{grad} f(X), \operatorname{grad} g(X)\rangle dt = \int \operatorname{tr} df \otimes dg dt$$

By uniqueness of 3.2.1 this translates to

$$\int (df \otimes dg)(dX, dX) = \int \operatorname{tr} df \otimes dg dt$$

And for an arbitrary $h \in \mathcal{T}^{0,2}(T^*M, T^*M)$:

$$\int h(dX, dX) = \int \operatorname{tr} h(X) dt \tag{3.22}$$

Let us prove ($2 \implies 1$), it suffices to set $h = \nabla df = \nabla^2 f$ in 3.22 and use the definition of a Γ-martingale in 3.2.2. It follows that:

$$df(X) \overset{m}{=} \frac{1}{2}\nabla^2 f\,(dX, dX) = \frac{1}{2}\operatorname{tr}\nabla^2 f(X)dt \tag{3.23}$$
$$= \frac{1}{2}\Delta^M f(X)dt \tag{3.24}$$

Where we used the definition of Laplace-Beltrami operator provided here 1.3.6.
i.e. X is a Brownian motion by third characterisation in 3.4.2
($1 \implies 2$), let X is a Brownian motion on M and $f \in C^\infty(M)$. Then,

$$\nabla df^2 = \nabla(2f df) = 2\nabla f df + 2f\nabla df \tag{3.25}$$
$$= 2df \otimes df + 2f\nabla df \tag{3.26}$$

Taking the trace on both sides to get:

$$\Delta^M f^2 = 2\|\operatorname{grad} f\|^2 + 2f\Delta^M f$$

Notice that:

$$df^2(X) \overset{m}{=} \frac{1}{2}\Delta^M f^2(X)dt = \|\operatorname{grad} f(X)\|^2 dt + f(X)\Delta^M f(X)dt$$

Whereas, applying Ito´s formula to f^2 gives:

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

Now by the uniqueness of the decomposition of a semi-martingale into a local martingale and a process of finite variation, we get:

$$[f(X), f(X)] = \int \|\operatorname{grad} f(X)\|^2 dt$$

Now making use of 3.22 to see:

$$\nabla^2 f(dX, dX) = \nabla df(dX, dX) \overset{3.22}{=} \operatorname{tr}\nabla df(X)dt \overset{1.3.6}{=} \Delta^M f(X)dt$$

Thus, X is a Γ-martingale. $\square$

The following proposition gives a useful description of the SDE appearing in the Eells-Elworthy-Malliavin construction on M in local coordinates:

Proposition 3.4.1. *[Hsu02]Example 3.3.5 . In the setting of 3.4.1, the equations for the Brownian motion X_t on M in local coordinates are given by*

$$dX_t^i = \sigma_j^i(X_t)dW_t^j - \frac{1}{2}g^{kl}(X_t)\Gamma_{kl}^i(X_t)dt \tag{3.27}$$

where W_t is a d-dimensional Euclidean Brownian motion.

Proof. The equation for the horizontal Brownian motion is $dU_t = H_\alpha(U_t) \circ dW_t^\alpha$ where W_t is an d-dimensional Euclidean Brownian motion. By 3.1.1 the horizontal vector fields H_i are locally given by:

$$H_i(u) = e_i^j X_j - e_i^j e_m^l \Gamma_{jl}^k(x) X_{km}$$

where $X_i = \dfrac{\partial}{\partial x^i}$. The ordinary differential equation relating the anti-development of the process X_t and the horizontal process U_t established in 3.14, as well as the local expression of H_i allow us to obtain the following two equations for $U_t = (X_t^i, e_j^i(t))$:

$$\begin{cases} dX_t^i = e_j^i(t) \circ dW_t^j \\ de_j^i(t) = -\Gamma_{kl}^i\left(X_t\right) e_j^l(t) e_m^k(t) \circ dW_t^m \end{cases} \tag{3.28}$$

From the first equation, we can find an equation for our Brownian motion X_t given by

$$dX_t^i = e_j^i(t) dW_t^j + \frac{1}{2}[de_j^i, dW^j]_t. \tag{3.29}$$

First, let $dM_t^i := e_j^i(t) dW_t^j$ be the martingale part, then

$$d[M^i, M^j]_t = \sum_{k=1}^d e_k^i(t) e_k^j(t) dt.$$

Let u be a frame, by definition $ue_l = e_l^i X_i$. Hence

$$\delta_{lm} = \langle ue_l, ue_m \rangle = e_l^i g_{ij} e_m^j,$$

or in other words,

$$\sum_{k=1}^d e_k^i e_k^j = g^{ij}.$$

This means that

$$d[M^i, M^j]_t = g^{ij}\left(X_t\right) dt.$$

If we let σ be the positive definite matrix square root of g^{-1}, then we have

$$W_t = \int_0^t \sigma\left(X_s\right)^{-1} dM_s.$$

and

$$dM_t = \sigma\left(X_s\right) dW_s.$$

On the other hand, from the second equation of 3.28, the last term in 3.29 becomes

$$d[e_j^i, dW^j]_t = -\Gamma_{kl}^i(X_t) e_j^l(t) e_m^k(t) = -g^{kl}(X_t) \Gamma_{kl}^i(X_t).$$

Putting everything together, the equation for X_t in local coordinates is:

$$dX_t^i = \sigma_j^i(X_t) dW_t^j - \frac{1}{2} g^{kl}(X_t) \Gamma_{kl}^i(X_t) dt.$$

□

Remark 3.4.3. *A different approach to the one presented in this section can be realized by working with the manifold using an embedding in* $\mathbb{R}^l$*, for this purpose, we assume that* M *is a closed submanifold of* $\mathbb{R}^l$ *and regard* $X = \{X^\alpha\}$ *as an* $\mathbb{R}^l$*-valued semi-martingale. For each* $x \in M$*, let* $P(x) : \mathbb{R}^l \to T_xM$ *be the orthogonal projection from* $\mathbb{R}^l$ *onto the subspace* $T_xM \subseteq \mathbb{R}^l$*. Then, theorem 3.1.4 in [Hsu02] which says that for a given submanifold* M *of* $\mathbb{R}^l$*, constructing Brownian motion on* M *amounts to solving the following SDE driven by an* l *dimensional Euclidean standard Brownian motion* B_t*:*

$$dX_t = P_\alpha(X_t) \circ dB_t^\alpha \tag{3.30}$$

In other words, the solution of the above SDE is a diffusion generated by

$$\frac{1}{2}\sum_{\alpha=1}^{l} P_\alpha^2 = \frac{1}{2}\Delta^M$$

Furthermore, Using Ito's formula 2.21 we have:

$$f(X_t) - f(X_0) = M^f(X)_t + \frac{1}{2}\int_0^t \Delta^M f(X_s)ds$$

Where,

$$M^f(X)_t = \int_0^t \langle P_\alpha f(X_s), dB_s^\alpha \rangle$$

Example 3.4.1. *To illustrate this approach, we give a short description of how this can be done in the case of the sphere. We start first by embedding the sphere* $\mathbb{S}^n$ *in* $\mathbb{R}^{n+1}$*. For an* $x \in \mathbb{S}^n$*, the projection of the unit vector* $\nu \in \mathbb{R}^{n+1}$ *to* $T_x\mathbb{S}^n$ *is given by:*

$$P_x(\nu) = \nu - \langle \nu, x\rangle x.$$

Hence the matrix $P = \{P_1, \ldots, P_n\}$ *has entries* P_{ij} *that can be expressed as*

$$P_{ij}(x) = \delta_{ij} - x_i x_j.$$

Now 3.30 takes the shape:

$$dX_t^i = (\delta_{ij} - X_t^i X_t^j) \circ dB_t^j. \tag{3.31}$$

This is the so-called Stroock representation of spherical Brownian motion. Two remarks are worth mentioning given the above:

Remark 3.4.4.
- *We can view* X_t *as the result of an extrinsic approach resulting from viewing it as an embedded process since it depends on the embedding of* M *into some Euclidean space* $\mathbb{R}^l$*. It has the drawback that the equation 3.30 is driven by a Brownian motion* B *whose dimension* l *is larger than the dimension* n *of the manifold* M*, whereas one would expect that Brownian motion on* M *should also be an object of dimension* n*. This should be contrasted with the intrinsic approach of Eells-Elworthy-Malliavin introduced earlier which reveals the full strength of Brownian motion on* M *by writing it as an* n*-dimensional object, i.e., as the solution of a stochastic differential equation driven by an* n*-dimensional Euclidean Brownian motion 3.27*

- *From now on and unless explicitly stated otherwise, we will follow the convention of denoting the driving Euclidean Brownian motion by W_t in the case of the intrinsic Brownian motion and by B_t in the case of the extrinsic Brownian motion.*

3.5. Brownian motion on Torus and Sphere, equivalent constructions and characterizations

3.5.1. Brownian motion on circle

Before treating the main manifolds, let us give a quick example using the simplest compact manifold i.e., the circle $\mathbb{S}^1 = \{e^{i\alpha}, \alpha \in [0, 2\pi]\} \subseteq \mathbb{R}^2$.

Example 3.5.1. *Firstly, we give an extrinsic description of Brownian motion on $\mathbb{S}^1$. It turns out that an application of Ito formula to the function:*

$$f(B_t) = (\cos B_t, \sin B_t)$$

where $X_t := (X_t^1, X_t^2) = (\cos B_t, \sin B_t) = f(B_t)$ and B_t is a one-dimensional Brownian motion, give us indirectly a 3.30 type system of SDEs. This goes as follows:

$$\begin{aligned} dX_t &= \frac{\partial f}{\partial t}|_{(B_t)} dt + \frac{\partial f}{\partial x}|_{(B_t)} dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}|_{(B_t)} (dB_t)^2 \\ &= 0dt + (-\sin B_t, \cos B_t) dB_t + \frac{1}{2}(-\cos B_t, -\sin B_t) dt \end{aligned}$$

Since $X_t^1 = \cos B_t$, $X_t^2 = \sin B_t$:

$$dX_t = (-X_t^2, X_t^1) dB_t + \frac{1}{2}(-X_t^1, -X_t^2) dt$$

Equivalently,

$$\begin{cases} dX_t^1 = -X_t^2 dB_t - \frac{1}{2} X_t^1 dt \\ dX_t^2 = X_t^1 dB_t - \frac{1}{2} X_t^2 dt \end{cases}$$

or

$$dX_t = \frac{-1}{2} X_t.dt + K.X_t.dB_t \tag{3.32}$$

with

$$K = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

In Stratonovich form

$$dX_t = K.X_t \circ dB_t$$

As a side note, let X be a solution to 3.32 and set $\sigma(x_1, x_2) = K.(x_1, x_2) = (-x_2, x_1)$, then write, for a C^2-function f:

$$Af(x_1, x_2) = \frac{1}{2}\left[x_2^2 \frac{\partial^2 f}{\partial x_1^2} + x_1^2 \frac{\partial^2 f}{\partial x_2^2} - 2x_1 x_2 \frac{\partial^2 f}{\partial x_1 x_2} - x_1 \frac{\partial f}{\partial x_1} - x_2 \frac{\partial f}{\partial x_2} \right]$$

the infinitesimal generator of X, where $\frac{1}{2}\sigma\sigma^T = \frac{1}{2}\begin{pmatrix} x_2^2 & -x_1x_2 \\ -x_1x_2 & x_1^2 \end{pmatrix}$

Remark 3.5.1.
- *In contrast, when following an intrinsic approach, Brownian motion on $\mathbb{S}^1$ is given by $X_t = e^{iW_t} = (\cos W_t, \sin W_t)$, where W_t is a 1-dimensional Brownian motion, furthermore, the anti-development of X_t is just the 1-dimensional Brownian motion W_t.*
- *Note that, for the circle, both the intrinsic and extrinsic approaches are identical, as for both constructions, we only needed a one-dimensional Brownian motion, either B_t or W_t.*

3.5.2. Brownian motion on Torus $\mathbb{T}$

Construction of Brownian motion on $\mathbb{T}$

We provide an intrinsic construction of Brownian motion on $\mathbb{T}$. Before that, we need to give a parametrization, provide the Riemannian metric, and finally compute the Levi-Civita connection. Let r and a be positive numbers such that $r > a$ and define a circle by $\{(x, 0, z) : (x-r)^2 + z^2 = a^2\}$. We rotate this circle around the xy-axis to form the $\mathbb{T}$. We parameterize the $\mathbb{T}$ by angles (θ, ϕ) which have the following (x, y, z)-coordinates:

- $x = (r - a\cos\phi)\cos\theta$
- $y = (r - a\cos\phi)\sin\theta$
- $z = a\sin\phi$

We compute the Riemannian metric as we did for other manifolds in the first chapter to get the matrix:

$$g = (g_{ij}) = \begin{pmatrix} (r - a\cos\phi)^2 & 0 \\ 0 & a^2 \end{pmatrix}$$

and its inverse:

$$g^{-1} = \left(g^{ij}\right) = \begin{pmatrix} \frac{1}{(r-a\cos\phi)^2} & 0 \\ 0 & \frac{1}{a^2} \end{pmatrix}.$$

The Levi-Civita connection on the $\mathbb{T}$ is given by:

$$\Gamma^{\theta}_{\theta\phi} = \Gamma^{\theta}_{\phi\theta} = \frac{\sin\phi}{\frac{r}{a} - \cos\phi} \tag{3.33}$$

$$\Gamma^{\phi}_{\theta\theta} = -\sin\phi\left(\frac{r}{a} - \cos\phi\right) \tag{3.34}$$

or using $\beta := \frac{r}{a} > 1$:

$$\Gamma^{\theta}_{\theta\phi} = \Gamma^{\theta}_{\phi\theta} = \frac{\sin\phi}{\beta - \cos\phi} \tag{3.35}$$

$$\Gamma^{\phi}_{\theta\theta} = -\sin\phi\,(\beta - \cos\phi) \tag{3.36}$$

Using 3.27, we can find an expression for $d\theta_t$ and $d\phi_t$. Indeed, we have:

$$dX_t^1 = d\theta_t = \sigma_j^1(X_t)dW_t^j - \frac{1}{2}g^{kl}(X_t)\Gamma_{kl}^1(X_t)dt$$

and

$$dX_t^2 = d\phi_t = \sigma_j^2(X_t)dW_t^j - \frac{1}{2}g^{kl}(X_t)\Gamma_{kl}^2(X_t)dt.$$

g^{kl} vanish for k,l that are different i.e.

$$g^{12}(X_t)\Gamma_{12}^1(X_t) = 0, \qquad g^{11}(X_t)\Gamma_{11}^2(X_t) = -\frac{\sin\phi_t\,(\beta - \cos\phi_t)}{a^2(\beta - \cos\phi_t)^2}$$

Moreover, Since σ is the square root of g^{-1}:

$$\sigma_1^1(X_t) = \sqrt{\frac{1}{(r - a\cos\phi_t)^2}}, \qquad \sigma_2^2(X_t) = \sqrt{\frac{1}{a^2}}$$

To sum up:

$$d\theta_t = \frac{1}{r - a\cos\phi_t}dW_t^\theta, \qquad d\phi_t = \frac{1}{a}dW_t^\phi + \frac{1}{2}\frac{\sin\phi_t}{a^2(\beta - \cos\phi_t)}dt.$$

Remark 3.5.2. *The Brownian motion on $\mathbb{T}$ can also be obtained by running two independent Brownians on $\mathbb{S}^1$ each and taking their product.*

Characterization of Brownian motion on $\mathbb{T}$ Now we check that the Brownian motion we arrived at using 3.27 is indeed a Brownian motion on $\mathbb{T}$ using the characterization given by the third assertion in 3.4.2. As discussed in 1.3.6 we have the following local expression of the Laplace-Beltrami operator:

$$\Delta^M f = g^{ij}\frac{\partial}{\partial x^i}\frac{\partial f}{\partial x^j} + b^i\frac{\partial f}{\partial x^i}.$$

Let the process (θ_t, ϕ_t) be the solution of the above SDEs, and let us verify whether it defines a Brownian motion on the $\mathbb{T}$. We verify the third bullet in 3.4.2 to do so. Thus, we compute b^1 and b^2 to get:

$$b^1 = \frac{1}{\sqrt{G}}\frac{\partial(\sqrt{G}g^{1j})}{\partial x^j} = \frac{1}{a(r - a\cos\phi)}\frac{\partial(a(r - a\cos\phi)\cdot\frac{1}{(r - a\cos\phi)^2})}{\partial\theta} = 0 \tag{3.37}$$

$$b^2 = \frac{1}{\sqrt{G}}\frac{\partial(\sqrt{G}g^{2j})}{\partial x^j} = \frac{1}{a(r - a\cos\phi)}\frac{\partial(\frac{r - a\cos\phi}{a})}{\partial\phi} = \frac{\sin\phi}{a(r - a\cos\phi)} \tag{3.38}$$

So:

$$\Delta f = \frac{1}{(r - a\cos\phi)^2}\frac{\partial^2 f}{\partial\theta^2} + \frac{1}{a^2}\frac{\partial^2 f}{\partial\phi^2} + \frac{\sin\phi}{a(r - a\cos\phi)}\frac{\partial f}{\partial\phi}$$

For $f_1(\phi) = a\sin\phi$, we compute $\Delta^{\mathbb{T}} f_1$. In summary, we obtain:

$$\Delta^{\mathbb{T}} f_1(\phi_t) = \frac{-\sin\phi_t}{a} + \frac{\sin\phi_t\cos\phi_t}{r - a\cos\phi_t}$$

Now we apply Ito formula to f_1 to get:

$$df_1(\phi_t) = a\cos\phi_t d\phi_t + \frac{1}{2}(-a\sin\phi_t)d[\phi,\phi]_t$$

Now we plug in the expression $d\phi_t$ to get that the quadratic variation term is $\frac{1}{a^2}dt$. In summary, we get:

$$df_1(\phi_t) = a\cos\phi_t \frac{1}{a} dW_t^{\phi} + \frac{a\cos\phi_t\sin\phi_t}{2a^2(\beta - \cos\phi_t)}dt - \frac{1}{2}(a\sin\phi_t)\frac{1}{a^2}dt.$$

Now we look at:

$$M^{f_1}(\phi_t) := df_1(\phi_t) - \frac{1}{2}\Delta^{\mathbb{T}} f_1(\phi_t)dt = \cos\phi_t dW_t^{\phi}$$

is a local martingale.

- For $f_2(\theta,\phi) = (r - a\cos\phi)\sin\theta$, we apply Ito formula to get:

$$\begin{aligned} df_2(\theta,\phi) = (r - a\cos\phi_t)\cos\theta_t d\theta_t + \sin\theta_t(a\sin\phi_t)d\phi_t - \frac{1}{2}(r - a\cos\phi_t)\sin\theta_t d[\theta,\theta]_t \\ + \cos\theta_t(a\sin\phi_t)d[\theta,\phi]_t + \frac{1}{2}\sin\theta_t(a\cos\phi_t)d[\phi,\phi]_t \end{aligned}$$

Let us compute each term and plug in the SDEs $d\theta_t, d\phi_t$ to get:

- $(r - a\cos\phi_t)\cos\theta_t d\theta_t = \cos\theta_t dW_t^{\theta}$
- $\sin\theta_t(a\sin\phi_t)d\phi_t = \sin\theta_t\sin\phi_t dW^{\phi} + \dfrac{\sin{\phi_t}^2\sin\theta_t}{2(r - a\cos\phi_t)}dt$
- $\frac{1}{2}(r - a\cos\phi_t)\sin\theta_t d[\theta,\theta]_t = \dfrac{\sin\theta_t}{2(r - a\cos\phi_t)}dt$
- $\cos\theta_t(a\sin\phi_t)d[\theta,\phi]_t = 0$
- $\frac{1}{2}\sin\theta_t(a\cos\phi_t)d[\phi,\phi]_t = \dfrac{\sin\theta_t\cos\phi_t}{2a}dt$

On the other hand:

$$\Delta^{\mathbb{T}} f_2(\theta,\phi) = \frac{-\sin\theta}{(r - a\cos\phi)} + \frac{\sin\theta\cos\phi}{a} + \frac{\sin\phi^2\sin\theta}{2(r - a\cos\phi)}.$$

In summary:

$$M^{f_2}(\theta_t,\phi_t) := df_2(\theta_t,\phi_t) - \frac{1}{2}\Delta^{\mathbb{T}} f_2(\theta_t,\phi_t)dt = \cos\theta_t dW_t^\theta + \sin\theta_t \sin\phi_t dW^\phi$$

is a local martingale, hence $\dfrac{\Delta^{\mathbb{T}}}{2}$ is a generator of our process θ_t, ϕ_t, which is a solution of:

$$d\theta_t = \frac{1}{r - a\cos\phi_t} dW_t^\theta, \qquad d\phi_t = \frac{1}{a} dW_t^\phi + \frac{1}{2}\frac{\sin\phi_t}{a^2(\beta - \cos\phi_t)}dt$$

Since $f_2(\frac{\pi}{2} - \theta_t, \phi_t) = (r - a\cos\phi_t)\sin(\frac{\pi}{2} - \theta_t) = (r - a\cos\phi_t)\cos\theta_t := f_3(\theta_t,\phi_t)$, we verified the martingale property for all test functions $f_i(x) = x^i$, where $i = 1, 2, 3$

Corollary 3.5.1. *According to 3.4.3 the previously constructed Brownian motion is a Γ-martingale on $\mathbb{T}$.*

3.5.3. Constructions of Brownian motion on Sphere

First construction of Brownian motion on Sphere

The following construction relies mainly on a study of the process $u(B) := \frac{B}{\|B\|}$ for an n-dimensional Brownian $B = (B^1, \dots, B^n)$ and a function $u : \mathbb{R}^n \setminus \{0\} \to \mathbb{S}^n$. Firstly, let us apply Ito's formula 3.40 to this process. To do so, we need to compute the first, and second derivatives of u and the term that appears beside the quadratic variation:

- $\partial_k u_i(B_t) = \dfrac{\delta_{ik}}{\|B_t\|} - \dfrac{B_t^i B_t^k}{\|B_t\|^3}$, where δ is the Kronecker delta function.
- $\partial_k^2 u_i(B_t) = -2\delta_{ik}\dfrac{B_t^k}{\|B_t\|^3} - \dfrac{B_t^i}{\|B_t\|^3} + 3\dfrac{B_t^i (B_t^k)^2}{\|B_t\|^5}$.
- $\displaystyle\sum_{k=1}^n \partial_k^2 u_i(B_t) = -2\frac{B_t^i}{\|B_t\|^3} - n\frac{B_t^i}{\|B_t\|^3} + 3\frac{B_t^i}{\|B_t\|^5}\|B_t\|^2 = -(n-1)\frac{B_t^i}{\|B_t\|^3}$.

Putting everything together, Ito's formula for $Y_t := u_i(B_t)$ gives:

$$dY_t = \left[\frac{1}{\|B_t\|} - \frac{(B_t^i)^2}{\|B_t\|^3}\right] dB_t^i - \sum_{j\neq i}\frac{B_t^j B_t^i}{\|B_t\|^3} dB_t^j - \frac{(n-1)}{2}\frac{B_t^i}{\|B_t\|^3}dt. \tag{3.39}$$

Or in a more compact way,

$$dY_t = \frac{1}{\|B_t\|}\sigma(Y_t)dB_t + \frac{1}{\|B_t\|^2}b(Y_t)dt. \tag{3.40}$$

where $\sigma = [\sigma_{ij}] \in \mathbb{R}^{n\times n}$, with $\sigma_{ij}(Y_t) = \delta_{ij} - Y_t^i Y_t^j$. And $b(Y_t) = \dfrac{(n-1)}{2}[Y_t^1, \dots, Y_t^n]$ Using the notion of "time change" to find the diffusion process, We define our process Y_t as in Example 8.5.8 ([Oks13]) where a random time change ([Oks13] 8.5 Random Time

Change) argument is used to drive the SDE satisfied by a Brownian motion on the unit sphere in $\mathbb{R}^n$. The process Y_t satisfies an Ito type SDE "up to a time change" and hence it is again an Ito process. One can then read the generating operator for the SDE, for a smooth f, as given by:

$$\Delta f(y) := \frac{1}{2}\Delta f(y) - \frac{1}{2}\sum_{i,j} y^i y^j \frac{\partial^2 f}{\partial y^i y^j} - \frac{(n-1)}{2}\sum_i y_i \frac{\partial f}{\partial y^i}.$$

Thus, Y_t is a diffusion process that lives on $\mathbb{S}^n$.

Note 3.5.1. *We can avoid making use of the notion of a random time change by guessing a diffusion and showing that it indeed lives on the sphere. We illustrate this in what follows:*
Let's start from the n-dimensional SDE:

$$dX_t^i = -\frac{n-1}{2}X_t^i dt + \sum_{j=1}^n \sigma_{ij}(X_t)dB_t^j, \quad i = 1, 2, ..., n, \tag{*}$$

inspired by equation 3.40 for the process X_t. Our claim is, the process X_t lies in $\mathbb{S}^n$ and to show this we rewrite the above equation in the Stratonovich sense to get:

$$dX_t = \sum_{j=1}^n V_j(X_t) \circ dB_t^j, \quad j = 1, \dots, n$$

where V_j's are vector fields. If we manage to show that these vector fields are tangent to $\mathbb{S}^n$ (See 1.1.2), then the claim follows. Firstly, we know that:

$$\sigma_{ij}(X_t) = \delta_{ij} - X_t^i X_t^j \tag{3.41}$$

$$\partial_k \sigma_{ij}(X_t) = -(\delta_{ik}X_t^j + \delta_{jk}X_t^i) \tag{3.42}$$

Thanks to what we have established about the Stratonovich formulation in 2.4.1 we can rewrite our SDE in the following form:

$$\begin{aligned} dX_t^i &= -\frac{n-1}{2}X_t^i dt + \sum_{j=1}^n \sigma_{ij}(X_t) \circ dB_t^j - \frac{1}{2}\sum_{j,k=1}^n \sigma_{kj}(X_t)\partial_k\sigma_{ij}(X_t)dt \\ &= \left(1 - \sum_{j=1}^n (X_t^j)^2\right) X_t^i dt + \sum_{j=1}^d (\delta_{ij} - X_t^i X_t^j) \circ dB_t^j. \end{aligned} \tag{3.43}$$

Define the following vector fields:

$$V_0 = \left(1 - \sum_{j=1}^n (X_t^j)^2\right) X_t^i \frac{\partial}{\partial X^i} \tag{3.44}$$

$$V_j = (\delta_{ij} - X_t^i X_t^j)\frac{\partial}{\partial X^i} \tag{3.45}$$

and let $Z_t := (t, B_t)$*. We can then write:*

$$dX_t = \sum_{j=0}^{n} V_j(X_t) \circ dZ_t^j$$

$$V_0 X = \sum_{i=1}^{n} V_0^i X^i = \sum_{i=1}^{n} \left(1 - \sum_{j=1}^{n} (X^j)^2\right) X^i X^i \tag{3.46}$$

$$= \left(1 - \sum_{j=1}^{n} (X^j)^2\right) \|X\|^2, \tag{3.47}$$

$$V_j X = \sum_{i=1}^{n} V_j^i X^i \tag{3.48}$$

$$= \sum_{i=1}^{n} (\delta_{ij} - X^i X^j) X^i \tag{3.49}$$

$$= X^j \left(1 - \|X\|^2\right) \tag{3.50}$$

When restricted to $\mathbb{S}^n$ *these vector fields vanish (1.1.2) and hence are tangent to the sphere. Proposition 1.2.8 in [Hsu02] gives then that*

$$X_t \in \mathbb{S}^n$$

for all times t up to explosion time.

Second construction of Brownian motion using extrinsic, intrinsic descriptions

Construction 3.5.1. *Similar to the torus case, we use 3.27 and provide an intrinsic construction using frames to construct Brownian motion on the unit sphere. To this aim, we need to collect information about the* g^{ij} *and* Γ *terms. For this, we recall the calculations that were performed in 1.3.4. Performing the exact same calculations as we did there for the torus case, it turns out that the Brownian motion on the two dimensional sphere obeys the SDE:*

$$d\theta_t = dW_t^\theta + \frac{1}{2}\cot(\theta_t)dt, \qquad d\varphi_t = \frac{1}{\sin(\theta_t)} dW_t^\varphi$$

As a second approach, we recall the extrinsic approach discussed in remark 3.4.2 where we embed the unit sphere in $\mathbb{R}^3$ *and construct Brownian motion as a solution to 3.31. In the following, we will see that simple computations using Ito's lemma enable us to transform the SDE in cartesian coordinates (i.e., an extrinsic approach using Stroock's representation) to an SDE in spherical coordinates* (θ_t, φ_t) *(i.e., an intrinsic approach using the Eells-Elworthy-Malliavin construction). Concretely, all we need to do is to write the process in terms of* θ *and* φ *and apply Ito's formula. Then, we can see that the two*

constructions are equivalent.
To this end, first define

$$Z_t = (\sin\theta_t \cos\varphi_t, \sin\theta_t \sin\varphi_t, \cos\theta_t), \theta \in [0, \pi]$$

and assume the coordinate process (θ_t, φ_t) *satisfies an SDE in the form of 3.27*

$$\begin{cases} d\varphi_t = Cdt + DdW_t^{\varphi} \\ d\theta_t = Adt + \mathbf{B}dW_t^{\theta} \end{cases}$$

where W_t^{θ}, W_t^{φ} *are independent Brownian motions.*
The SDE for Z_t *using Stroocks representation of spherical Brownian motion in Ito formulation 3.31 reads*

$$dZ_i = d\tilde{B}_i - \sum_{j=1}^{3} Z_i Z_j d\tilde{B}_j - Z_i dt.$$

The plan is to apply Ito's formula to each component of Z_t*, plug in the expression of* $d\theta_t, d\varphi_t$*, and equate with the above SDE. Let us do this for* $Z_3 = \cos\theta_t$*:*

$$d\cos\theta_t = -\sin\theta_t d\theta_t - \frac{1}{2}\cos\theta_t d[\theta_t, \theta_t]_t \tag{3.51}$$

$$= -\sin\theta_t d\theta_t - \frac{1}{2}\mathbf{B}^2 \cos\theta_t dt \tag{3.52}$$

$$= d\tilde{B}_3 - \sin\theta_t \cos\varphi_t \cos\theta_t d\tilde{B}_1 - \sin\theta_t \sin\varphi_t \cos\theta_t d\tilde{B}_2 - (\cos\theta_t)^2 d\tilde{B}_3 - \cos\theta_t dt \tag{3.53}$$

$$\Rightarrow \begin{cases} -A\sin\theta_t - \frac{1}{2}\cos\theta_t \cdot \mathbf{B}^2 = -\cos\theta_t \\ -\mathbf{B}dW_t^{\theta} = \sin\theta_t d\tilde{B}_3 - \cos\varphi_t \cos\theta_t d\tilde{B}_1 - \sin\varphi_t \cos\theta_t d\tilde{B}_2 \end{cases}$$

We omit the tedious computations and only provide the end equations:

$$\begin{cases} -A\sin\theta_t - \frac{1}{2}\cos\theta_t \mathbf{B}^2 = -\cos\theta_t \\ -\mathbf{B}dW_t^{\theta} = -\cos\theta_t \cos\varphi_t d\tilde{B}_1 - \cos\theta_t \sin\varphi_t d\tilde{B}_2 + \sin\theta_t d\tilde{B}_3 \\ -DdW_t^{\varphi} = \dfrac{\sin\varphi_t}{\sin\theta_t} d\tilde{B}_1 - \dfrac{\cos\varphi_t}{\sin\theta_t} d\tilde{B}_2 \\ -1 = -C\tan\varphi_t - \frac{1}{2}D^2 + A\cot\theta_t - \frac{1}{2}\mathbf{B}^2 \\ C\cot\varphi_t - \frac{1}{2}D^2 + A\cot\theta_t - \frac{1}{2}\mathbf{B}^2 = -1 \end{cases}$$

We combine the fourth and last equations to get $C = 0$ *and the first with the fourth to get:*

$$\frac{1}{2}D^2 = A\frac{1}{\cos\theta_t}\frac{1}{\sin\theta_t}$$

If we look at the second and third equations we get expressions for dW^{θ}, dW^{φ}*:*

$$dW_t^{\theta}.dW_t^{\varphi} = \frac{1}{\mathbf{B}} \begin{bmatrix} \cos\varphi_t \cos\theta_t \\ \sin\varphi_t \cos\theta_t \\ \sin\theta_t \end{bmatrix} . \frac{1}{D} \begin{bmatrix} -\dfrac{\sin\varphi_t}{\sin\theta_t} \\ \dfrac{\cos\varphi_t}{\sin\theta_t} \\ 0 \end{bmatrix} = 0$$

Hence, the covariation vanishes and dW^θ and dW^φ are independent. To determine D, we require $\mathbb{E}[(dW_t^\varphi)^2] = dt$. A quick calculation yields:

$$D = \frac{1}{\sin\theta_t}, \qquad \mathbf{B} = 1, \qquad A = \frac{1}{2}\cot\theta_t$$

Hence,

$$\begin{cases} d\varphi_t = \dfrac{1}{\sin\theta_t} dW_t^\varphi \\ d\theta_t = \dfrac{1}{2}\cot\theta_t dt + dW_t^\theta \end{cases}$$

Construction 3.5.2. *This construction is inspired by the paper of [PW]. Let B_t be a Brownian motion on $\mathbb{R}^3$, we will give an extrinsic type description, in the spirit of (3.31), of Brownian motion on the unit sphere embedded in $\mathbb{R}^3$. To this end, consider the following SDE:*

$$dX_t = X_t \otimes dB_t$$

where $\otimes$ denotes a Stratonovich cross product and B_t is a 3-dimensional Brownian motion. In other words:

$$\begin{aligned} dX_t &= Y_t \circ dB_t^{(3)} - Z_t \circ dB_t^{(2)} \\ dY_t &= Z_t \circ dB_t^{(1)} - X_t \circ dB_t^{(3)} \\ dZ_t &= X_t \circ dB_t^{(2)} - Y_t \circ dB_t^{(1)} \end{aligned}$$

First of all, we can check that by the Stratonovich chain rule

$$d\left(X_t^2 + Y_t^2 + Z_t^2\right) = 2\left(X_t \circ dX_t + Y_t \circ dY_t + Z_t \circ dZ_t\right) = \ldots = 0$$

Hence (X_t, Y_t, Z_t) is on $\mathbb{S}^2$ for all $t \geq 0$ iff (X_0, Y_0, Z_0) is on $\mathbb{S}^2$. Then by the Stratonovich chain rule, we obtain:

$$dX = \frac{\partial X}{\partial\theta} \circ d\theta_t + \frac{\partial X}{\partial\varphi} \circ d\varphi_t$$

where $X = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$. This corresponds to the following set of equations:

$$\begin{aligned} dX_t &= (\cos\theta_t\cos\varphi_t) \circ d\theta_t - (\sin\theta_t\sin\varphi_t) \circ d\varphi_t \\ dY_t &= (\cos\theta_t\sin\varphi_t) \circ d\theta_t + (\sin\theta_t\cos\varphi_t) \circ d\varphi_t \\ dZ_t &= (-\sin\theta_t) \circ d\theta_t - 0 \circ d\varphi_t \end{aligned}$$

Using algebra, expanding everything, and matching with the expressions above, we can solve for $d\theta_t$ in terms of the Brownian motions:

$$\begin{aligned} d\theta_t &= \sin\varphi_t \circ dB_t^{(1)} - \cos\varphi_t \circ dB_t^{(2)} \\ d\varphi_t &= \cot\theta_t\left(\cos\varphi_t \circ dB_t^{(1)} + \sin\varphi_t \circ dB_t^{(2)}\right) - dB_t^{(3)} \end{aligned}$$

We get this back into Itô form, which leads to a drift term in θ_t.

$$
\begin{aligned}
d\theta_t &= \frac{1}{2}\cot\theta_t dt + \sin\varphi_t dB_t^{(1)} - \cos\varphi_t dB_t^{(2)} \\
d\varphi_t &= \cot\theta_t \left(\cos\varphi_t dB_t^{(1)} + \sin\varphi_t dB_t^{(2)}\right) - dB_t^{(3)}
\end{aligned}
$$

This follows for $d\theta_t$, and similarly for $d\varphi_t$, from:

$$d\theta_t = \sin\varphi_t dB_t^{(1)} - \cos\varphi_t dB_t^{(2)} + \frac{1}{2}[d\sin\varphi_t, dB_t^{(1)}] - \frac{1}{2}[d\cos\varphi_t, dB_t^{(2)}]$$

Now we apply Ito's formula to each $\sin\varphi_t$, $\cos\varphi_t$, then plug it in the above equation to get:

$$d\theta_t = \sin\varphi_t dB_t^{(1)} - \cos\varphi_t dB_t^{(2)} + \frac{1}{2}(\cot\theta_t \cos\varphi_t^2)dt + \frac{1}{2}(\cot\theta_t \sin\varphi_t^2)dt.$$

Putting:

$$
\begin{aligned}
dB_t^{\theta} &= \sin\varphi_t dB_t^{(1)} - \cos\varphi_t dB_t^{(2)}, \\
dB_t^{\phi} &= \cos\theta_t \left(\cos\varphi_t dB_t^{(1)} + \sin\varphi_t dB_t^{(2)}\right) - \sin\theta_t dB_t^{(3)}
\end{aligned}
$$

and checking that these are two independent Brownian motions, one can rewrite the process as

$$
\begin{aligned}
d\theta_t &= \frac{1}{2}\cot\theta_t dt + dB_t^{\theta} \\
d\varphi_t &= \frac{1}{\sin\theta_t} dB_t^{\phi}.
\end{aligned}
$$

A third Brownian motion $dB_t^N = \mathbf{X}_t \cdot d\mathbf{B}_t$ is normal to the sphere and is thus canceled out.

Characterization of Brownian motion on the unit sphere $\mathbb{S}^2$ Again as a sanity check, we define what a Brownian motion is in terms of the Laplace-Beltrami 1.14 generator (This is our starting point). Then we check that the SDEs we obtained earlier do have the expected generator by using Ito's lemma on a specific test function f.
In other words, let our process (θ_t, φ_t) obey the above SDEs, we verify the third bullet in 3.4.2 i.e. we take a C^2 function $f(\theta_t, \varphi_t)$ and use Itô's lemma to check that

$$df(\theta_t, \varphi_t) - \frac{\Delta^{\mathbb{S}^2} f(\theta_t, \varphi_t)}{2} dt$$

is a local martingale. Doing so we show that the generator of the process is $\frac{\Delta^{\mathbb{S}^2}}{2}$. In the case of $\mathbb{S}^2$, using 1.14 with $b^1 = \cot\theta$, $b^2 = 0$ the Laplace-Beltrami is expressed as follows:

$$\Delta^{\mathbb{S}^2} f = \frac{\partial^2 f}{\partial\theta^2} + \frac{1}{\sin\theta^2}\frac{\partial^2 f}{\partial\varphi^2} + \cot\theta\frac{\partial f}{\partial\theta}$$

For the function $f_1(\theta) = \cos\theta$, we have $\Delta^{\mathbb{S}^2} f_1(\theta) = -2\cos\theta$. Now we apply Ito's formula to f_1 to get:

$$df_1(\theta_t) = -\sin\theta_t d\theta_t - \frac{1}{2}(\cos\theta_t)[d\theta, d\theta]_t$$

Now we plug in the expression $d\theta_t$ to get that the quadratic variation term is simply dt. In summary, we get:

$$df_1(\theta_t) = -\sin\theta_t dW_t^\theta - \cos\theta_t dt$$

Now we look at:

$$M^{f_1}(\theta_t) := df_1(\theta_t) - \frac{\Delta^{\mathbb{S}^2} f_1(\theta_t)}{2} dt = \cos\theta_t dW_t^\theta$$

This is a local martingale since it is free of dt, hence $\dfrac{\Delta^{\mathbb{S}^2}}{2}$ is a generator of our process θ_t, which is a solution of $d\theta_t = \frac{1}{2}\cot\theta_t dt + dW_t^\theta$

- For $f_2(\theta,\varphi) = \sin\theta\sin\varphi$, we apply Ito formula to get:

$$\begin{aligned} df_2(\theta,\varphi) = \cos\theta_t\sin\varphi_t d\theta_t + \sin\theta_t\cos\varphi_t d\varphi_t - \frac{1}{2}\sin\theta_t\sin\varphi_t d[\theta,\theta]_t \\ + \cos\theta_t\cos\varphi_t d[\theta,\varphi]_t - \frac{1}{2}\sin\theta_t\sin\varphi_t d[\varphi,\varphi]_t \end{aligned}$$

Let us compute each term and plug in the SDEs $d\theta_t, d\phi_t$ to get:

- $\cos\theta_t\sin\varphi_t d\theta_t = \dfrac{\cos{\theta_t}^2}{2\sin\theta_t}\sin\varphi_t dt + \cos\theta_t\sin\varphi_t dW_t^\theta$
- $\sin\theta_t\cos\varphi_t d\varphi_t = \cos\varphi_t dW^\varphi$
- $\frac{1}{2}\sin\theta_t\sin\varphi_t d[\theta,\theta]_t = \frac{1}{2}\sin\theta_t\sin\varphi_t dt$
- $\cos\theta_t\cos\varphi_t d[\theta,\varphi]_t = 0$
- $\frac{1}{2}\sin\theta_t\sin\varphi_t d[\varphi,\varphi]_t = \dfrac{\sin\varphi_t}{2\sin\theta_t} dt$

On the other hand:

$$\Delta f_2(\theta,\varphi) = -\sin\theta\sin\varphi - \frac{\sin\varphi}{\sin\theta} + \frac{\cos\theta^2\sin\varphi}{\sin\theta}$$

In summary:

$$M^{f_2}(\theta_t,\phi_t) := df_2(\theta_t,\varphi_t) - \frac{\Delta^{\mathbb{S}^2} f_2(\theta_t,\varphi_t)}{2} dt = \cos\theta_t\sin\varphi_t dW_t^\theta + \cos\varphi_t dW^\varphi$$

is a local martingale, hence $\frac{\Delta^{\mathbb{S}^2}}{2}$ is a generator of our process (θ_t,φ_t), which is a solution of:

$$d\theta_t = dW_t^\theta + \frac{1}{2}\cot(\theta_t)dt, \qquad d\varphi_t = \frac{1}{\sin(\theta_t)} dW_t^\varphi$$

Since $f_2(\theta_t, \frac{\pi}{2} - \varphi_t) = \sin\theta_t\sin(\frac{\pi}{2} - \varphi_t) = \sin\theta_t\cos\varphi_t := f_3(\theta_t,\varphi_t)$, we verified the martingale property for all test functions $f_i(x) = x^i$, where $i = 1,2,3$

Corollary 3.5.2. *We deduce once again by 3.4.3 that the constructed Brownian motion is a Γ-martingale on $\mathbb{S}^2$.*

3.6. Discussion on Γ-martingale on $\mathbb{T}$

As we pointed out in 3.5.1 the Brownian motion on $\mathbb{T}$ is a Γ-martingale. However, its coordinate representation in $\mathbb{R}^3$ cannot be a martingale because each coordinate process is bounded. Indeed, if the coordinate processes were martingales, they would have to converge "Doob's martingale convergence theorem", whereas it is known that the Brownian motion on a compact Riemannian manifold is recurrent simply by compactness of M or "See [MBL14] p99", which means that Brownian motion on a $\mathbb{T}$ is a recurrent process that does not converge. In the remainder of this chapter, we construct a Γ-martingale $\Psi = (\gamma, \phi)$, for which its coordinate processes are martingales in the ordinary sense i.e. a martingale in $\mathbb{R}^3$ based on the exposition in [Dar83]. Note that (γ, ϕ) is a point on $\mathbb{T}$ so the following computation is valid only locally.

Example 3.6.1. *(A Γ-martingale on $\mathbb{T}$) Let (γ, ϕ) be a point on $\mathbb{T}$. The Levi-Civita connection on $\mathbb{T}$ has the following form:*

$$\Gamma^{\gamma}_{\gamma\phi} = \Gamma^{\gamma}_{\phi\gamma} = \frac{\sin\phi}{\beta - \cos\phi} \tag{3.54}$$

$$\Gamma^{\phi}_{\gamma\gamma} = -\sin\phi(\beta - \cos\phi) \tag{3.55}$$

Assume $\Psi = (\gamma, \phi)$ is a semi-martingale on $\mathbb{T}$, then using the definition of Γ-martingale 3.2.3, locally the expressions are given by:

$$d\gamma_t + \Gamma^1_{12} d[\gamma, \phi]_t, \qquad d\phi_t + \frac{1}{2}\Gamma^2_{11} d[\gamma, \gamma]_t$$

which are nothing but the differentials of $\mathbb{R}$-valued local martingales, say M_t and N_t. Now we decompose the coordinate processes γ and ϕ using Doobs decomposition:

$$d\gamma_t = dM_t + dV^1_t \tag{3.56}$$

$$d\phi_t = dN_t + dV^2_t \tag{3.57}$$

Now the bounded variation process by assumption has the following representation:

$$dV^1_t = -\Gamma^1_{12} d[M, N]_t \tag{3.58}$$

$$dV^2_t = -\frac{1}{2}\Gamma^2_{11} d[M, M]_t. \tag{3.59}$$

Thus,

$$d\gamma_t = dM_t - \frac{\sin\phi}{\beta - \cos\phi} d[M, N]_t \tag{3.60}$$

$$d\phi_t = dN_t + \frac{1}{2}\sin\phi(\beta - \cos\phi) d[M, M]_t \tag{3.61}$$

Using the parametrization of the torus, our process Ψ has (x,y,z)-coordinates given as

$$((r - a\cos\phi)\cos\gamma, (r - a\cos\phi)\sin\gamma, a\sin\phi)$$

and so Ψ is naturally a process in the ambient space. Therefore, using Ito's formula we verify that each component is a martingale in the usual sense. The lemma in [Dar83] relates the quadratic variations of the martingale parts (M_t, N_t) in the following manner:

$$d[N,N]_t = \cos\phi_t(\beta - \cos\phi_t)d[M,M]_t$$

This means that the Euclidean coordinate process $((r - a\cos\phi)\cos\gamma, (r - a\cos\phi)\sin\gamma, a\sin\phi)$ is a martingale, in the sense that each component is a martingale in the ordinary sense. Let us check this fact using Ito's formula applied to each component:

1. *Apply Ito formula to $z_t = a\sin\phi_t$:*

$$\begin{aligned} dz_t &= a\cos\phi_t d\phi_t + \frac{1}{2}(-a\sin\phi_t)d[\phi,\phi]_t & (3.62)\\ &= a\cos\phi_t[dN_t + \frac{1}{2}(\frac{\sin\phi}{\beta - \cos\phi})d[M,M]_t] + \frac{1}{2}(-a\sin\phi_t)d[\phi,\phi]_t & (3.63)\\ &= a\cos\phi_t dN_t + \frac{a}{2}\frac{\cos\phi_t\sin\phi_t}{\beta - \cos\phi_t}d[M,M]_t - \frac{1}{2}a\sin\phi_t d[N,N]_t & (3.64)\\ &= a\cos\phi_t dN_t & (3.65)\end{aligned}$$

   Since N_t is the $\mathbb{R}$-valued martingale part, z_t is a martingale.

2. *Apply Ito formula to $x_t = a(\beta - \cos\phi_t)\cos\gamma_t$:*

$$\begin{aligned} dx_t = -a(\beta - \cos\phi_t)\sin\gamma_t d\gamma_t + a(\sin\phi_t)\cos\gamma_t d\phi_t - \frac{a}{2}(\beta - \cos\phi_t)\cos\gamma_t d[\gamma,\gamma]_t \\ -\frac{a}{2}\sin\phi_t\sin\gamma_t d[\gamma,\phi]_t + \frac{a}{2}\cos\phi_t\cos\gamma_t d[\phi,\phi]_t\end{aligned}$$

   Proceeding in the same manner as in the previous bullet point, we get:

$$dx_t = -a(\beta - \cos\phi_t)\sin\gamma_t dM_t + a\sin\phi_t\cos\gamma_t dN_t$$

   Since M_t, N_t are $\mathbb{R}$-valued martingale parts, therefore x_t is a martingale.

3. *For y_t, observe:*

$$x_t(\gamma_t - \frac{\pi}{2}, \phi_t) = a(\beta - \cos\phi_t)\cos(\gamma_t - \frac{\pi}{2}) = y_t$$

   Thus, y_t is a martingale too.

4. Conclusions

This thesis has delved deep into the interplay between different disciplines in mathematics, namely, stochastic calculus and geometry with a special emphasis on the equivalent constructions of Brownian motions on non-Euclidean geometries. Our journey began by establishing the necessary foundation in Riemannian geometry along with essential concepts from stochastic analysis in Euclidean space. The main third chapter of this work navigated through the concept of frame bundles, horizontal lifts, anti-developments, and their use in the realm of SDE´s on manifolds. The Eells-Elworthy construction emerged as a milestone and a fundamental tool for giving a mathematically satisfying construction of Brownian motion on Riemannian manifolds, and indeed, its efficacy was demonstrated through a comparison with simpler constructions, particularly in the case of the sphere. At a later stage, we provided an equivalent characterization of Brownian motion on a manifold using the Laplace-Beltrami operation. We showed that Brownian motion on the sphere and torus fulfills this criterion. This paper ended with an examination of the concept of Γ-martingales on a manifold, offering an illustrative example of a special Γ-martingales on the torus. In essence, this thesis provides valuable insights into the different construction as well as characterizations of Brownian motion of compact Riemannian manifolds and hence, it lays a foundation for future research paths at the intersection of probability theory, Riemannian geometry, finance, and mathematical physics. A few possible research topics are:

- Martingale Optimal Transport problems: Investigating the geometric interpretation of some martingale optimal transport problems that arise from finance. One can ask the question, how does the curvature of the underlying space affect the existence, uniqueness, and properties of optimal martingale transport maps?
- Brownian motion and Heat Semigroup: Heat kernels and heat semigroups are objects of intensive research in geometry. As the minimal heat kernel $p_M(t, x, y)$ is the transition density function of Brownian motion, a link to stochastic analysis can be established. When a geometric problem lends itself to a probabilistic interpretation, it can be investigated by stochastic analysis techniques.
- Numerical methods for Stochastic Analysis on Manifolds: Develop numerical algorithms of certain stochastic processes taking values in non-flat geometry.
- Application to Finance: Investigate whether the developed framework can contribute to modeling more complex financial markets with non-Euclidean geometries, as a result leading to more realistic and accurate models for asset price dynamics.

Bibliography

[And] Ben Andrews. *Lectures on Differential Geometry.*

[Dar83] R. W. R. Darling. A Martingale on the Imbedded Torus. *Bulletin of the London Mathematical Society*, 15(3):221–225, 05 1983.

[Don51] M.D. Donsker. *An Invariance Principle for Certain Probability Limit Theorems.* American Mathematical Society. Memoirs. 1951.

[Dur05] R. Durrett. *Probability: Theory and Examples.* Duxbury advanced series. Thomson Brooks/Cole, 2005.

[Eva13] L.C. Evans. *An Introduction to Stochastic Differential Equations.* Miscellaneous Books. American Mathematical Society, 2013.

[GN12] Leonor Godinho and José Natário. An introduction to riemannian geometry. *With Applications*, 2012.

[Hsu02] E.P. Hsu. *Stochastic Analysis on Manifolds.* Graduate studies in mathematics. American Mathematical Society, 2002.

[HT94] W. Hackenbroch and A. Thalmaier. *Stochastische Analysis: Eine Einführung in die Theorie der stetigen Semimartingale.* Mathematische Leitfäden. Vieweg+Teubner Verlag, 1994.

[Hus94] D. Husemöller. *Fibre Bundles.* Springer, second edition, 1994. Fibre Bundles (Third ed.). New York: Springer.

[HWY92] S. He, J. Wang, and J.A. Yan. *Semimartingale Theory and Stochastic Calculus.* Science Press, 1992.

[KS17] Michael Kunzinger and Roland Steinbauer. Riemannian geometry. *Lecture Notes, https://www. mat. univie. ac. at/ stein/teaching/SoSem16/rg-2017-03-06. pdf*, 2017.

[Lee00] John M. Lee. Introduction to smooth manifolds. 2000.

[MBL14] H.P. McKean, Z.W. Birnbaum, and E. Lukacs. *Stochastic Integrals.* Probability and mathematical statistics. Elsevier Science, 2014.

[Mil63] J. Milnor. *Morse theory.* Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.

[Oks13] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

[Pau98] Wilfried H. Paus. Sum of squares manifolds: The expressibility of the Laplace-Beltrami operator on pseudo-Riemannian manifolds as a sum of squares of vector fields. *Transactions of the American Mathematical Society*, 350(10):3943–3966, 1998.

[Prö21] David Prömel. *Stochastic Calculus*. HWS. Universität Mannheim, 2021.

[PW] G. C. PRICE and D. WILLIAMS. Rolling with slipping, seminaire de probabilités (paris, 1983).

[RY99] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*. Number 293 in Grundlehren der mathematischen Wissenschaften. Springer, Berlin [u.a.], 3. ed edition, 1999.

[Sch17] Frederic Schuller. *Notes on the Geometric Anatomy of Theoretical Physics*. Springer, first edition, 2017. Notes on the Geometric Anatomy of Theoretical Physics.

[Shr04] Steven Shreve. *Stochastic calculus for finance II: Continuous-time models*, volume 11. Springer, 2004.

A. Appendix

Theorem A.0.1. *Let G_1, G_2 be two groups, let $\phi : G_1 \mapsto G_2$ be a group homomorphism and let $\pi : G_1 \mapsto {}^{G_1}/_{\ker\phi}$ be the canonical (surjective) homomorphism $g \mapsto g\ker\phi$. Then, there exist a unique isomorphism:*

$$\psi : {}^{G_1}/_{\ker\phi} \to \phi(G_1),$$

so that $\psi \circ \pi = \phi$

Example A.0.1. *Let $X \sim \mathcal{N}(\mu, \sigma^2)$, so $X - \mu$ has mean zero.*
***The moment generating function** of $X - \mu$ is say $\psi(u)$:*

$$\psi(u) = \mathbb{E}[e^{u(X-\mu)}] \tag{A.1}$$

$$= e^{u^2\sigma^2\frac{1}{2}} \tag{A.2}$$

Where u is a real variable. For proof of the above equality see "3.2.13 in [Shr04]".
Now, taking derivatives with respect to u and using the product rule for differentiation, we get the following:

1. $\psi'(u) = \mathbb{E}[(X-\mu)e^{u(X-\mu)}] = \sigma^2 u e^{u^2\sigma^2\frac{1}{2}}$
2. $\psi''(u) = \mathbb{E}[(X-\mu)^2 e^{u(X-\mu)}] = [\sigma^2 + u\sigma^4]e^{u^2\sigma^2\frac{1}{2}}$
3. $\psi'''(u) = \mathbb{E}[(X-\mu)^3 e^{u(X-\mu)}] = [3u\sigma^4 + u^3\sigma^6]e^{u^2\sigma^2\frac{1}{2}}$
4. $\psi''''(u) = \mathbb{E}[(X-\mu)^4 e^{u(X-\mu)}] = [3\sigma^4 + 6u^2\sigma^6 + u^4\sigma^8]e^{u^2\sigma^2\frac{1}{2}}$

$\implies \psi''''(0) = 3\sigma^4$
For our proof, we set $X := B(t_{j+1}) - B(t_j)$ so:

$$B(t_{j+1}) - B(t_j) \sim \mathcal{N}(0, t_{j+1} - t_j)$$

And we get our desired result:

$$\psi''''(0) = \mathbb{E}\left[\left(B\left(t_{j+1}\right) - B\left(t_j\right)\right)^4\right] = 3\left(t_{j+1} - t_j\right)^2$$