ENS Paris-Saclay

M1 Jacques Hadamard end of year Research project

Some ground work for the analysis of Lie-Trotter type splittings for Langevin diffusions on Lie Groups.

Amar AHMANE

supervised by
Sonja Cox
Erwin Luesink

July 2025

Abstract

Due to applications in several fields, such as physics, chemistry, and machine learning, the study of stochastic differential equations (SDEs) on Riemannian manifolds has gained significant interest. More generally, SDEs on smooth manifolds is an active area of reasearch of itself (see [4], [8] and [9]).

SDEs on Riemannian manifolds arise naturally in many problems of physics and chemistry where dynamics is constrained to curved configuration spaces. For instance, the diffusion of lipids or proteins on biological membranes is modeled by Brownian motion on curved surfaces, governed by the Laplace–Beltrami operator. In general relativity, relativistic Brownian motion is formulated as an SDE on a pseudo-Riemannian manifold, capturing thermal fluctuations in curved spacetime. In molecular dynamics, holonomic constraints (e.g. fixed bond lengths) restrict atomic motion to a submanifold of the Euclidean space, and Langevin dynamics on this manifold describes thermal fluctuations of molecules. Further examples include rotational diffusion of colloids, modeled by SDEs on the sphere S^2 or the rotation group SO(3), stochastic thermodynamics on curved state spaces, and spin dynamics in condensed matter physics, where the magnetization follows a stochastic Landau–Lifshitz–Gilbert equation on S^2. These applications highlight the crucial role of geometry in shaping stochastic behavior in complex physical and chemical systems.

While working with such equations, some challenges appear, such as the need for Riemannian Brownian motion, which is not easy to simulate in general, but also the fact that in order to numerically integrate such systems, one needs to work with local charts, which can become a problem regarding the choice of step-size and transition between charts. This is why we may require an additional algebraic structure to the Riemannain manifold in order to have access to additional powerful tools that will allow us to work with the geometry of the manifold in a more global way. The best candidates are Lie groups, which include many well known matrix groups such as $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{SO}(n)$, $\mathrm{O}(n)$. Lie groups do not have a natural Riemannain geometry, hence the need for a notion of a Riemannian metric compatible with the group structure, which is called a bi-invariant metric. Compact Lie groups in particular always admit such a metric. Sections 2.1 and 2.2 provide a discussion of the necessary preliminariy notions from differential geometry and Lie groups leading the result of the existence of bi-invariant metrics on compact Lie groups. Section 2.3 discusses stochastic differential equations on manifolds while mentionning examples on Riemannain manifolds and Lie groups.

As in [11], we will consider Langevin diffusion on compact Lie groups as special cases of stochastic Hamiltonian dynamics. Considering a compact Lie group G and $M = T^*M$ its cotangent bundle, which is a symplectic manifold, consider a Hamiltonian $H : G \times \mathfrak{g}^* \simeq M \to \mathbb{R}$, then hamiltonain vector fields X_H satisfy $i_{X_H}\omega = dH$, where ω is the symplectic form on M. Then deterministic Hamiltonian mechanics are expressed as $\frac{\mathrm{d}S}{\mathrm{d}t} = X_H(S)$, where S describes the state of a system. By the

means of Malliavin's transfer priple, we lift deterministic Hamiltonian dynamics to stochastic Hamiltonian dynamics, which are expressed as $\mathrm{d}S = X_{H_\alpha}(S) \circ Z_t^\alpha$, where the Z^α is a family of driving semi-martingales and H_α Hamiltonians. Deterministic Hamilotinian dynamics have the Gibbs measure as their invariant measure, this does not transfer directly to the stochastic case, and we need to add a double bracket dissipation term (see [2] and [3]) in order to recover the Gibbs measure an an invariant measure. More precisely, if $Z_t^0 = t$ and the rest of the semi-martingales are Brownian motions ($Z^\alpha = W^\alpha$ $\alpha \geq 1$), then, applying an Itô correction gives the following Itô interpretation of the stochastic Hamiltonian dynamics

$$\mathrm{d}S_t = \left(X_{H_0}(S_t) + \frac{1}{2} X_{H_\alpha} X_{H_\alpha}(S_t) \right) \mathrm{d}t + X_{H_\alpha}(S_t)\, \mathrm{d}W_t^\alpha$$

where we have taken the Einstein summation convention on α. Correcting the previous equation by adding a double bracket dissipation term as in [11] yields, in the Itô form,

$$\mathrm{d}S_t = \left(X_{H_0}(S_t) + \frac{\beta}{2} \{H_0, H_\alpha\} X_{H_\alpha}(S_t) + \frac{1}{2} X_{H_\alpha} X_{H_\alpha}(S_t) \right) \mathrm{d}t + X_{H_\alpha}(S_t)\, \mathrm{d}W_t^\alpha$$

and this equation has the Gibbs measure as an invariant measure. This is discussed in section 3.2.

The idea of the research project was then to build on the work done by Luesink and Street in [11] and try to provide an analysis of splitting methods for the special case of Langevin diffusions on compact Lie groups. In the Euclidian case, work has already been carried out in [1] where the authors provide interesting sufficient conditions to achieve an arbitrary order r in the approximation of the invariant measure using Lie-Trotter splitting methods. The hope was that in the case of compact Lie groups, the ideas of Abdulle and Vilmart would carry over to the setting studied in our project but with a much easier analysis due to the compactness assumption. Before doing so, some gorund work needed to be clarrified, and an important property that we needed to be true was the ergodicity of the corrected equation above. This is the main result of this paper which is proved in section 3.3.

Résumé

En raison des diverses applications qu'elles présentent en physique, en chimie et en machine learning„ l'étude des équations différentielles stochastiques (EDS) sur les variétés riemanniennes gagnent en intérêt. Les EDS sur les variétés riemanniennes apparaissent naturellement dans de nombreux problèmes de physique et de chimie où la dynamique est contrainte à des espaces de configurations courbes. Plus généralement, les EDS sur les variétés lisses constituent en elles-mêmes un domaine de recherche actif (voir [4], [8] et [9]).

En travaillant sur de tels problèmes, plusieurs difficultés apparaissent, telles que le besoin de simuler le mouvement brownien riemannien, ce qui n'est pas aisé en général, mais aussi le fait que, pour intégrer numériquement de tels systèmes, il faut travailler avec des cartes locales, ce qui complique les choix de pas dans les méthodes numériques, les changements de cartes étant aussi des opérations délicates. C'est pourquoi une structure algébrique supplémentaire peut être requise sur la variété riemannienne afin de disposer d'outils plus puissants permettant de travailler avec la géométrie de la variété de manière plus globale. Les meilleurs candidats sont les groupes de Lie, qui incluent de nombreux groupes matriciels connus tels que $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{SO}(n)$, $\mathrm{O}(n)$. Les groupes de Lie ne possèdent pas naturellement de géométrie riemannienne, d'où la nécessité d'une notion de métrique riemannienne compatible avec la structure de groupe, appelée métrique bi-invariante. En particulier, les groupes de Lie compacts admettent toujours une telle métrique. Les sections 2.1 et 2.2 présentent une discussion des notions préliminaires nécessaires de géométrie différentielle et de groupes de Lie, conduisant au résultat d'existence de métriques bi-invariantes sur les groupes de Lie compacts. La section 2.3 discute des équations différentielles stochastiques sur les variétés tout en mentionnant des exemples sur des variétés riemanniennes et des groupes de Lie.

Comme dans [11], nous considérerons les diffusions de Langevin sur les groupes de Lie compacts comme des cas particuliers de problèmes hamiltoniens stochastiques. Considérons un groupe de Lie compact G et $M = T^*M$ son fibré cotangent, qui est une variété symplectique, ainsi qu'un Hamiltonien $H : G \times \mathfrak{g}^* \simeq M \to \mathbb{R}$. Les champs hamiltoniens X_H satisfont alors $i_{X_H}\omega = dH$, où ω est la forme symplectique sur M. La mécanique hamiltonienne déterministe s'écrit alors $\frac{\mathrm{d}S}{\mathrm{d}t} = X_H(S)$, où S décrit l'état d'un système. Au moyen du principe de transfert de Malliavin, nous élevons la dynamique hamiltonienne déterministe en dynamique hamiltonienne stochastique, qui s'exprime par $\mathrm{d}S = X_{H_\alpha}(S) \circ Z_t^\alpha$, où les Z^α forment une famille de semi-martingales directrices et les H_α des hamiltoniens. La dynamique hamiltonienne déterministe admet la mesure de Gibbs comme mesure invariante, ce qui ne se transpose pas directement au cas stochastique, et l'on doit ajouter un terme de dissipation à double crochet (voir [2] et [3]) afin de retrouver la mesure de Gibbs comme mesure invariante. Plus précisément, si $Z_t^0 = t$ et que les autres semi-martingales sont des mouvements browniens ($Z^\alpha = W^\alpha$, $\alpha \geq 1$), alors, en appliquant une correction d'Itô, on obtient l'interprétation d'Itô suivante du problème

hamiltonien stochastique :

$$\mathrm{d}S_t = \left(X_{H_0}(S_t) + \tfrac{1}{2} X_{H_\alpha} X_{H_\alpha}(S_t)\right) \mathrm{d}t + X_{H_\alpha}(S_t)\, \mathrm{d}W_t^\alpha$$

où nous avons utilisé la convention de sommation d'Einstein sur l'indice α. En corrigeant l'équation précédente par l'ajout d'un terme de dissipation à double crochet comme dans [11], on obtient, sous la forme d'Itô :

$$\mathrm{d}S_t = \left(X_{H_0}(S_t) + \tfrac{\beta}{2}\{H_0, H_\alpha\} X_{H_\alpha}(S_t) + \tfrac{1}{2} X_{H_\alpha} X_{H_\alpha}(S_t)\right) \mathrm{d}t + X_{H_\alpha}(S_t)\, \mathrm{d}W_t^\alpha$$

et cette équation admet la mesure de Gibbs comme mesure invariante. Ceci est discuté en section 3.2.

L'idée du projet de recherche était donc de partir du travail effectué par Luesink et Street dans [11] et de tenter de fournir une analyse des méthodes de séparation pour le cas particulier des diffusions de Langevin sur les groupes de Lie compacts. Dans le cas euclidien, un travail a déjà été réalisé dans [1], où les auteurs fournissent des conditions suffisantes pour atteindre un ordre r arbitraire dans l'approximation de la mesure invariante en utilisant des méthodes de séparation de Lie–Trotter. L'espoir était que, dans le cas des groupes de Lie compacts, les idées d'Abdulle et Vilmart se transposent au cadre étudié dans notre projet, mais avec une analyse beaucoup plus simple grâce à l'hypothèse de compacité. Avant cela, certains préliminaires devaient être clarifiés, et une propriété importante dont nous avions besoin était l'ergodicité de l'équation corrigée ci-dessus. Ceci constitue le résultat principal de ce travail, qui est démontré en section 3.3.

Contents

1 Introduction

Stochastic differential equations on manifolds have been a topic of growing interest in the past years, particularly SDEs on Riemannian manifolds. It is interesting in the purposes of developping numerical integrators to work on manifolds that have an additional algebraic structure, such as Lie groups. At the intersection of Riemannian geometry and Lie groups, we find Lie groups equipped with a bi-invariant metric, this is the case for compact Lie groups in particular. In this setting, we can view Langevin diffusions as a particular case of stochastic Hamiltonian mechanics on the cotangent bundle of a Lie group. We prove, under the assumption of Hörmander's condition, that the Gibbs measure is the unique invariant measure of such systems and that it is ergodic.

Mathematics Subject Classification : 60H10, 65C30, 37A99, 37M25

2 Prelimnaries

2.1 Some vocabulary from differential geometry

In this document, we study special families of stochastic differential equations on manifolds. Recall that a topological manifold M of dimension n is a Hausdorff σ-compact topological space such that every point $x \in M$ admits an open neighbourhood V and a homeomorphism $\varphi : V \to U \subseteq \mathbb{R}^n$. In more simple words, these are spaces that locally resemble $\mathbb{R}^n$, the Hausdorff and σ-compact assumptions are here to garantee the existence of what we call partitions of unity, essential for example to prove Whitney's embedding theorem, mentionned later on.

Still, no differential calculus is possible at this stage, so we need to require more things for the local homeomorphisms. These local homeomorphisms are called charts and a family of charts $\mathcal{A} : (\varphi_i : V_i \to \mathbb{R}^n)$ such that (V_i) is an open cover of M is called an atlas. An atlas $\mathcal{A}$ is called smooth if the change of charts φ_{ij} : $\varphi_j \circ \varphi_i^{-1} : \varphi_i(V_i \cap V_j) \to \varphi_j(V_i \cap V_j)$ are $\mathcal{C}^\infty$ diffeomorphisms. Topological manifolds equipped with a smooth atlas are called smooth manifolds. Hence smooth functions $f : M \to N$ between two manifolds are maps such that for each point $x \in X$, the map between euclidian spaces $\psi \circ f \circ \varphi^{-1}$ is $\mathcal{C}^\infty$ at point $\varphi(x)$ where φ and ψ are suitable charts around x and $f(x)$ respectively. Considering smooth paths $\gamma : I \to M$, we are able to give a notion of the differential of a smooth function f between manifolds by considering the map that sends a smooth path γ to the smooth path $f \circ \gamma$. In fact, in local charts, the derivative at 0 of smooth paths does not depend on the charts used, hence we can identify smooth paths passing through $x \in M$ at 0 that have the same derivative at 0. The quotient obtained is then the tangent space at x, denoted T_xM. The induced map $\gamma \to f \circ \gamma$ is the differential of f at x. The tangent bundle is the set $\cup(\{x\} \times T_xM)$, denoted TM, which is also a smooth manifold, and vector fields over M are smooth sections of the canonical projection $\pi : TM \to M$. The set of vector fields is denoted $\Gamma(TM)$. Vector fields act on smooth functions $f : M \to \mathbb{R}$ in the following way : if X is a vector field and f is such a function, then Xf is a new smooth function $M \to \mathbb{R}$ defined by $Xf(x) = \mathrm{d}f_x(X(x))$. The set $\mathcal{C}^\infty(M)$ of smooth real-valued functions over M is a ring when equipped with usual pointwise addition and multiplication, and $\Gamma(TM)$ with the external composition law $(f, X) \mapsto (x \mapsto f(x) \cdot X(x))$ and pointwise addition is a $\mathcal{C}^\infty(M)$-module.

For any open set $U \subset M$, an n-tuple of vector fields (X_i) is called a frame over U iff $(X_i(x))$ spans T_xM for every $x \in M$. If $U = M$, the X_i are global sections and $(X_1, \ldots, X_n)$ is called simply a frame. For example, if $M = \mathbb{R}^n$ and $(e_1, \ldots, e_n)$ is a basis of M, then the maps $X_i : x \in M \mapsto (x, e_i) \in TM$ are vector fields, usually denoted $\partial/\partial x_i$, and the n-tuple $(\partial/\partial x_1, \ldots, \partial/\partial x_n)$ is a frame, called the global coordinate frame.

Now let M and N be two smooth manifolds, $f : M \to N$ a local diffeomorphism and $Y \in \Gamma(TN)$. We define a vector field over M, which we denote f^*Y and call the pullback of Y by f, by the following

$$f^*Y : x \mapsto (\mathrm{d}f_x)^{-1}(Y(f(x)))$$

We also have the notion of a pushforward, in the case where f is a global diffeomorphism. If $X \in \Gamma(TM)$, then $f_*X \in \Gamma(TN)$ is defined by

$$f_*X : y \mapsto \mathrm{d}f_{f^{-1}(y)}f(X(f^{-1}(y)))$$

If (φ, U) is a chart of M, $(e_1, \ldots, e_n)$ is a basis of $\mathbb{R}^n$ ($n = \dim M$) and $X \in \Gamma(TM)$, then $\varphi_* X_{|U}$ is a vector field over $\varphi(U)$, which is an open subset of $\mathbb{R}^n$, hence there exists smooth real-valued functions $g_1, \ldots, g_n$ such that

$$\varphi^* X = \sum_{i=1}^{n} g_i \frac{\partial}{\partial x_i}$$

Putting $f_i = g_i \circ \varphi$, we get the following local expression of X:

$$X_{|U} = \sum_{i=1}^{n} f_i \varphi^* \frac{\partial}{\partial x_i}$$

And the n-tuple $(\varphi^* \frac{\partial}{\partial x_1}, \ldots, \varphi^* \frac{\partial}{\partial x_n})$ is a local frame over U.

2.1.1 Connections

Let M be a smooth manifold and $\Gamma(TM)$ be the $\mathcal{C}^\infty(M)$-module of smooth vector fields over M. Recall that $\Gamma(TM)$ acts on $\mathcal{C}^\infty(M)$ in the following sense: if f is a smooth real-valued function over M, Xf is the smooth real-valued function over M defined by $Xf(x) = \mathrm{d}f_x(X(x))$. It is also denoted $\mathcal{L}_X f$. $\mathcal{L}_X$ is a derivation of $\mathcal{C}^\infty(M)$ in the sense that $f \mapsto \mathcal{L}_X f$ is linear and that it verifies a Leibniz property

$$\mathcal{L}_X(fg) = \mathcal{L}_X fg + f\mathcal{L}_X g$$

for all smooth real-valued functions f, g over M.

For example, if M is an open set of $\mathbb{R}^n$ and $(e_1, \ldots, e_n)$ is the canonical basis of $(\mathbb{R}^n)$, then, since $\partial f / \partial x_i(x) = \mathrm{d}f_x(e_i)$, we get that $\mathcal{L}_{X_i}$ is the partial derivation with respect to x_i. If U is an open set of M, is straightforward that $(\mathcal{L}_X)_U = \mathcal{L}_{X_{|U}}$.

If M is a smooth manifold and X and Y vector fields over M, we denote by $[X, Y]$ the smooth vector field over M that verifies

$$\mathcal{L}_{[X,Y]} = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X$$

We call this new vector field the Lie bracket of the vector fields X and Y. This will be useful later to state Hörmander's condition.

Let's define the notion of connection now and give some of its properties.

Definition 2.1. *Let M be a smooth manifold and ξ a vector bundle with projection $p : E \to M$. A connection on ξ is an $\mathbb{R}$-bilinear map $\nabla : \Gamma(TM) \times \Gamma(\xi) \to \Gamma(\xi)$, denoted $(X, \sigma) \mapsto \nabla_X \sigma$, such that for every $X\Gamma(TM)$ and $\sigma \in \Gamma(\xi)$ and every smooth real valued function f over M, we have*

1. $\nabla_{fX}\sigma = f\nabla_X\sigma$;
2. $\nabla_X(f\sigma) = X(f)\sigma + f\nabla_X\sigma$.

When $\xi = TM$, then ∇ is called connection over M.

Example 2.1. If ξ is the trivial vector bundle over M (denote F its fiber, a vector space), then there exists a unique connection ∇^0 over ξ, called the trivial connection, and it is such that for each smooth vector field X over M and each smooth section $\sigma : M \to M \times F$ of constant second componant, we have $\nabla^0_X\sigma = 0$. In fact, we can identify $\Gamma(\xi)$ with $\mathcal{C}^\infty(M, F)$, and since $T_yF = F$ for each $y \in F$, we can put $\nabla^0_X\sigma = X(\sigma)$. The uniqueness is a direct consequence of the second property in the above definition and from the fact that the family of sections $(x \mapsto (x, e_i))$ is a basis of the $\mathcal{C}^\infty(M)$-module $\Gamma(\xi)$ whenever $(e_1, \dots, e_n)$ is a basis of F.

Let U be an open subset of a smooth n-dim manifold M and ∇ a connection over M. Recall that the n-tuple $(\partial/\partial x_1, \dots, \partial/\partial x_n)$ is a local frame over U. This means that for indexes i, j, the vector fields $\nabla_{\partial/\partial x_i}\partial/\partial x_j$ are expressed as linear combinations of elements of $(\partial/\partial x_1, \dots, \partial/\partial x_n)$. This yields the following definition

Definition 2.2. *Let (U, φ) be a chart of an n-dim manifold M and ∇ a connection over M. We have*

$$\nabla_{\varphi^* \frac{\partial}{\partial x_i}} \varphi^* \frac{\partial}{\partial x_j} = \sum_{k=1}^{n} \Gamma^k_{ij} \varphi^* \frac{\partial}{\partial x_k}$$

where the Γ^k_{ij} are smooth real valued functions over U. We call these functions the Christoffel symbols.

Let (U, φ) be a chart, then for any vector fields

$$X = \sum_{i=1}^{n} f_i^X \varphi^* \frac{\partial}{\partial x_i}, \quad Y = \sum_{i=1}^{n} f_i^X \varphi^* \frac{\partial}{\partial x_i}$$

We have

$$\nabla_X Y = \sum_{i=1}^{n} f_i^{\nabla_X Y} \varphi^* \frac{\partial}{\partial x_i}, \quad f_i^{\nabla_X Y} = \sum_{j=1}^{n} f_j^X \varphi^* \frac{\partial}{\partial x_i} f_i^Y + \sum_{1 \le j,k \le n} \Gamma^i_{jk} f_j^X f_k^Y$$

In fact, bilinearity of the connection gives

$$\nabla_X Y = \sum_{1 \le j,k \le n} X_j \nabla_{\varphi^* \frac{\partial}{\partial x_j}} \left(Y_k \varphi^* \frac{\partial}{\partial x_k} \right)$$

By applying the second axiom in the definition of a connection over a manifold, we have, for every $j, k = 1, \dots, n$

$$\nabla_{\varphi^* \frac{\partial}{\partial x_j}} \left(Y_k \varphi^* \frac{\partial}{\partial x_k} \right) = \varphi^* \frac{\partial}{\partial x_j} Y_k \varphi^* \frac{\partial}{\partial x_k} + Y_k \nabla_{\varphi^* \frac{\partial}{\partial x_j}} \left(\varphi^* \frac{\partial}{\partial x_k} \right)$$
$$= \left[\varphi^* \frac{\partial}{\partial x_j} Y_k \right] \varphi^* \frac{\partial}{\partial x_k} + Y_k \sum_{i=1}^{n} \Gamma^i_{jk} \varphi^* \frac{\partial}{\partial x_i}$$

Arranging the sums gives the desired result.

We will now see an important example of connections on Riemannian manifolds.

Expressing the equations of geometric mechanics require us to endow the manifolds on which they are written with metrics, so that we are able to have a notion of gradient, Hessian and Laplacian of a function. We do so by adding a family of inner products (b_p) to the manifold that is studied M, one for each tangent space T_pM, in such a way that b_p "varies smoothly" with p. Such a family is called a Riemannian metric on M, and when equipped with it, M is a Riemannian manifold.

More precisely,

Definition 2.3. *Let M be a smooth n-dimensional manifold. A family (b_p) of inner products on each tangent space T_pM is called a Riemannian metric if for every chart (φ, U), for every frame $(X_1, \dots, X_n)$ on U, the maps $p \in U \mapsto b_p(X_i(p), X_j(p))$ are smooth for every $i, j = 1, \dots, n$.*

A smooth manifold with a Riemannian metric is a Riemannnian manifold.

Let M be a smooth manifold of dimension n with a metric $(\langle \cdot, \cdot \rangle_x)$. Given two vector fields $X, Y \in \Gamma(TM)$, we can define a smooth real-valued function over M by putting $\langle X, Y \rangle(x) = \langle X(x), Y(x) \rangle_x$ for all $x \in M$.

We want to consider connections on M that are compatible with the metric in the following sense:

Definition 2.4. *A connection ∇ on M is compatible with the metric iff for all vector fields X, Y, Z on M, the following holds*

$$\mathcal{L}_X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla X Z \rangle$$

Proposition 2.1. *Riemannian manifolds admit compatible connections.*

To obtain unicity of such a connection, we need to require some kind of symmetry properties. To do so, we will define several error terms. First, we have the curvature term, for any two vector fields X, Y,

$$R(X, Y) = \nabla_{[X,Y]} + \nabla_Y \nabla_X - \nabla_X \nabla_Y$$

Then, we have the torsion term, given by, for any two vector fields X, Y,

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$$

We say that the connection ∇ is torsion free if the torsion term is 0 everywhere. We show that there exist a canonical connection on Riemannian manifolds.

Proposition 2.2. *Let $(M, \langle \cdot, \cdot \rangle_*)$ a Riemannian manifold. There exists a unique metric compatible torsion free connection ∇ on M. This connection is called the Levi-Civita connection, or canonical connection on M.*

Proof. Let us begin by proving unicity. Let ∇ be a metric compatible torsion free connection. Then for every $X, Y, Z \in \Gamma(TM)$, we have

$$\begin{aligned}
\mathcal{L}_X\langle Y, Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla X Z\rangle, \\
\mathcal{L}_Y\langle X, Z\rangle &= \langle \nabla_Y X, Z\rangle + \langle X, \nabla Y Z\rangle, \\
\mathcal{L}_Z\langle X, Y\rangle &= \langle \nabla_Z X, Y\rangle + \langle X, \nabla Z Y\rangle
\end{aligned}$$

Adding the first two equations and subtracting the third we get

$$\begin{aligned}
2\langle \nabla_X Y, Z\rangle = {} & \mathcal{L}_X\langle Y, Z\rangle \mathcal{L}_Y\langle X, Z\rangle - \mathcal{L}_Z\langle X, Y\rangle \\
& - \langle [X, Z], Y\rangle - \langle [Y, Z], X\rangle + \langle [X, Y], Z\rangle
\end{aligned}$$

Hence, we get unicity. For all vector fields X, Y, Z over M, we denote by $\ell_{X,Y}(Z)$ the right-hand side of the equality above. When X and Y are fixed, we can see that the map $\ell_{X,Y}$ is $\mathcal{C}^\infty$-linear. Thanks to the non-degenerecy of the scalar products $\langle \cdot, \cdot \rangle_*$, the map $\Phi : X \mapsto (Y \mapsto \langle X, Y\rangle)$ is an isomorphism of $\mathcal{C}^\infty$-modules. Then the connection defined by $\nabla_X Y = \Phi^{-1}(\ell_{X,Y}/2)$ works. $\square$

2.2 Lie Groups

2.2.1 General definitions

Generally, a (real) Lie group is a group endowed with the structure of a smooth manifold such that the multiplication and inverse of an element define smooth maps. Relevant examples of Lie Groups include well known matrix groups such as $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{SO}(n)$, $\mathrm{O}(n)$... These groups can be seen as smooth submanifolds of $\mathcal{M}_n(\mathbb{R})$, and the multiplication and inversion maps are smooth thanks to a restriction argument since multiplication and inversion define smooth maps from and into open sets of $\mathcal{M}_n(\mathbb{R})$.

As we'll see, one important object to consider, given a Lie group G, is the tangent space to G at the identity element, we will denote $\mathfrak{g} = T_1G$. equipped with a non-associative multiplication operation that we call the Lie bracket, the vector space $\mathfrak{g}$ has the structure of a Lie algebra.

Definition 2.5. *A (real) Lie algebra A is a real vector space together with a bilinear map $[\cdot, \cdot] : A \times A \to A$ called the Lie bracket on A such that for all elements a, b, c of A, we have the following:*

1. $[a, b] = -[b, a]$;
2. $[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0$ *(Jacobi identity).*

Remark. It is possible that, instead of (1), we ask that for every vector a, we have $[a, a] = 0$, and then the skew-symetricity follows from that and (2)

Note that we are given smooth functions, thanks to the smoothness of the multiplication, which are the left and right translations $L_g : h \in G \mapsto gh$ and $R_g : h \in G \mapsto hg$, where g is an element of the group. These maps are actually diffeomorphisms, and the study of their derivatives plays an important role. We consider the map $\psi : G \to \mathrm{Aut}(G)$ defined by $\psi_g(h) = ghg^{-1}$. We will denote Ad_g the derivative at 1 of the map ψ_g, which is an isomorphism of Lie algebras. We can prove that the induced map $\mathrm{Ad} : g \in G \to \mathrm{GL}(\mathfrak{g})$ is smooth. Thus, we are able to get the derivative of this map at the identity, which is a map of Lie algebras, and which we denote $\mathrm{ad} : \mathfrak{g} \to T_1 \mathrm{GL}(\mathfrak{g})$, this is called the adjoint representation of the Lie group G. Hence, we are able to give a definition of the Lie bracket for a general abstract Lie group.

Definition 2.6. *Given a Lie group G, the tangent space at the identity $\mathfrak{g} = T_1G$ with the Lie bracket defined by, for all elements g, h of the group,*

$$[g, h] = \mathrm{ad}(g)(h)$$

is the Lie algebra of the Lie group G.

2.2.2 The exponential map

The exponential map for matrices can be defined with a series as shown below

$$\exp(M) = \sum_{n\geq 0} \frac{M^n}{n!}$$

We want to construct an exponential map for general abstract Lie groups in a way that is consistent with this definition for linear Lie groups. To do so, we will need to go through a canonical example of what we call left and right-invariant vector fields on a given Lie group.

Let G be a Lie group. A vector field X over G is said to be left-invariant if for every elements $g, h \in G$, we have

$$(\mathrm{d}L_h)_g(X(g)) = X(hg)$$

Whenever $v \in \mathfrak{g}$, setting $X_v(g) = (\mathrm{d}L_g)_1(v)$ gives a left-invariant vector field. Actually, the condition to be left-invariant shows, if we put $g = 1$ above, that X is determined by its value in 1. There is correspondance between left-invariant vector fields and the Lie algebra $\mathfrak{g}$, and we can construct a trivialisation of the bundle TG (show that it is isomorphic to $G \times \mathfrak{g}$).

Another important fact to notice is that left translations send integral curves for X to integral curves for X whenever X is a left-invariant vector field. In fact, let φ

be an integral curve for a left-invariant vector field X. Put $\psi(t) = L_g(\varphi(t))$ where g is a fixed element of the group. Applying the chain rule while differentiating ψ, we get

$$\psi'(t) = (\mathrm{d}L_g)_{\varphi(t)}(\varphi'(t)) = (\mathrm{d}L_g)_{\varphi(t)}(X(\varphi(t))) = X(L_g\varphi(t)) = X(\psi(t))$$

Furthermore, if φ is taken to be the maximal integral curve, then by a uniqueness argument, we get, denoting by Ψ the flow of X, $g\varphi(t) = g\Psi(1,t) = \Psi(g,t)$, because ψ is an integral curve with initial point g.

Of course, we have an analogous notion for right-invariant vector fields with analogous properties.

Now, the following result will enable us to define the exponential in the general case.

Proposition 2.3. *Let G be a Lie group. Let $v \in \mathfrak{g}$ and X_v be the left-invariant vector field defined in the discussion above. Then, the following holds*

1. *X_v is the infinitesimal generator for a smooth $\mathbb{R}$-action $\alpha : \mathbb{R} \times G \to G$;*

2. *There exists a smooth group homomorphism $\varphi_v : \mathbb{R} \to G$ such that $\alpha(t,g) = g\varphi_v(t)$.*

Proof.

1. Denote by Ψ the flow of X and φ the maximal integral curve for X s.t $\varphi(0) = 1$. The idea is to show that φ is defined for all times. Let s and t be times such that φ is defined for s, t and $s+t$. We write $\Psi(s+t,1) = \Psi(s,\Psi(t,1))$. By proposition 2.1, we get $\Psi(s+t,1) = \Psi(s,1)\Psi(t,1)$. Hence, by a left translation, we are able to extend φ further and further, which means it is defined for all times. The action $\alpha : (t,g) \mapsto \Psi(g,t) = g\varphi(t)$ is the desired one. It is smooth thanks to the smoothness of X_v.

2. Putting $\varphi_v(t) = \alpha(t,1)$ works.

□

The homomorphism φ_v is referred to as a 1-parameter subgroup.

Definition 2.7. *Given a Lie group G, the exponential map* $\exp : \mathfrak{g} \to G$ *is given by*

$$\exp(v) = \varphi_v(1)$$

By noticing that $\mapsto (\alpha(t,g),v)$ is the flow of the vector field $(g,v) \mapsto (X_v(g),0)$, which smooth, we show that exp is smooth. Furthermore, differentiating exp at $0 \in \mathfrak{g}$ gives $\mathrm{d}\exp_0 = \mathrm{id}_{\mathfrak{g}}$. Hence, by the inverse function theorem, exp is a local diffeomorphism around 0.

2.2.3 Metrics compatible with the structure of a Lie group

Recall the definition of a metric over a smooth manifold.

Definition 2.8. *Let M be a smooth n-dimensional manifold. A family (b_p) of inner products on each tangent space T_pM is called a Riemannian metric if for every chart (φ, U), for every frame $(X_1, \dots, X_n)$ on U, the maps $p \in U \mapsto b_p(X_i(p), X_j(p))$ are smooth for every $i, j = 1, \dots, n$.*

A smooth manifold with a Riemannian metric is a Riemannnian manifold.

Since we are working in the setting of Lie groups, we want to add a metric to Lie groups but in such a way that it is compatible with the Lie group structure.

The important notions are those of left, right and bi-invariant metrics on Lie groups. In a Lie group G, a metric (b_p) is called left-invariant (resp. right-invariant) if for every $g \in G$, the operation $\mathrm{d}L_g$ is an isometry (resp. $\mathrm{d}R_g$ is an isometry). More precisely, for the left-invariant case, we want

$$b_h(X, Y) = b_{gh}((\mathrm{d}L_g)_h X, (\mathrm{d}L_g)_h Y)$$

A metric on a Lie group is called bi-invariant if it is left and right-invariant at the same time. The existence of left and right-invariant metrics is straightforward since they can be constructed when given an inner product $\langle \cdot, \cdot \rangle$ on the Lie algebra $\mathfrak{g}$. In fact, it suffices to put, for every element g of the group,

$$b_g(X, Y) = \langle (\mathrm{d}L_{g^{-1}})_g X, (\mathrm{d}L_{g^{-1}})_g Y \rangle X, Y \in T_h G$$

the idea being to go back to the Lie algebra using $(\mathrm{d}L_{g^{-1}})_g : T_g G \to T_{L_{g^{-1}}(g)} G = \mathfrak{g}$. This verifies the definition, it is a Riemannian metric thanks to the smoothness of the left translation operation, and it is left-invariant since for every $X, Y \in T_h G$

$$b_{gh}((\mathrm{d}L_g)_h X, (\mathrm{d}L_g)_h Y) = \langle (\mathrm{d}L_{(gh)^{-1}})_{gh}((\mathrm{d}L_g)_h X), (\mathrm{d}L_{(gh)^{-1}})_{gh}((\mathrm{d}L_g)_h Y) \rangle$$

But, since left translations are group isomorphisms, we get $(\mathrm{d}L_{(gh)^{-1}})_{gh}((\mathrm{d}L_g)_h X) = (\mathrm{d}(L_{(gh)^{-1}} \circ L_g))_h = (\mathrm{d}L_{h^{-1}})_h$, which concludes. Constructing a right invariant metric is analogous. But remains the question of the existence of bi-invariant metrics. This is not an easy matter; the following theorem gives a simple condition that is actually equivalent to the existence of such metrics for connected Lie groups:

Theorem 2.4 ([5], Theorem 21.9)**.** *A connected Lie group G admits a bi-invariant metric iff it is isomorphic to the product of a compact Lie group and a vector space $\mathbb{R}^N$ for some $N \geq 0$.*

In particular, compact Lie groups admit bi-invariant metrics.

2.3 Stochastic analysis on manifolds

2.3.1 Stochastic differential equations on manifolds

In this section we discuss stochastic differential equations on manifolds. Solutions of stochastic differential equations will live in the space of manifold-valued semimartingales. Recall that a real semimartingale is a real stochastic process that can be decomposed as the sum of a local martingale and a continuous adapted finite-variation process.

Definition 2.9 ([8], definition 1.2.1)**.** *Let M be a smooth manifold and $(\Omega, \mathcal{F}, P)$ a filtered probability space. Let τ be an $\mathcal{F}$-stopping time. A continuous, M-valued process X defined on $[0, \tau)$ is called an M-valued semimartingale if $f(X)$ is a real-valued semimartingale for all smooth functions $f : M \to \mathbb{R}$.*

A stochastic differential equation on a manifold M is defined by l vector fields $V_1, \ldots, V_l$ on M, an $\mathbb{R}^l$ valued driving semimartingale Z, and an M valued random variable $X_0 \in \mathcal{F}_0$, serving as the initial value of the solution.

It is written as

$$\mathrm{d}X_t = V_\alpha(X_t) \circ \mathrm{d}Z_t^\alpha \tag{1}$$

Now, we need to give an interpretation to the solution(s) of equation (1).

Definition 2.10 ([8], definition 1.2.3)**.** *An M-valued semimartingale X defined up to a stopping time τ is a solution of equation (1) up to τ if for all smooth function $f : M \to \mathbb{R}$, the following holds*

$$f(X_t) = f(X_0) + \int_0^t V_\alpha f(X_s) \circ \mathrm{d}Z_s^\alpha, \quad 0 \leq t < \tau$$

Notice that we made the choice of choosing Stratonovich integration over Itô's for the good reason that, under the Stratonovich formulation, apllying diffeomorphisms to solutions of an SDE still gives solutions to an SDE. Let $\varphi : M \to N$ be a diffeomorphism between two smooth manifolds. When we denote by $\Gamma(TM)$ the space of smooth vector fields on M, the same for $\Gamma(TN)$, we notice that φ induices a map $\varphi_* : \Gamma(TM) \to \Gamma(TN)$ through the following

$$(\varphi_* V) f(y) = V(f \circ \varphi)(x)$$

where $V \in \Gamma(TM)$, f is a smooth function $f : N \to \mathbb{R}$ and x is such that $y = \varphi(x)$.

Proposition 2.5 ([8], propostion 1.2.4)**.** *If X is a solution of the SDE (1) on M, then $\varphi(X)$ is a solution to*

$$dY_t = (\varphi_* V_\alpha)(Y_t) \circ Z_t^\alpha$$

the initial condition being $\varphi(X_0)$.

Proof. Put $Y = \varphi(X)$. Let $f : N \to \mathbb{R}$ be a smooth function. The map $g = f \circ \varphi : M \to \mathbb{R}$ is also smooth, we then apply the definition 1.5 to g and X to obtain

$$g(X_t) = g(X_0) + \int_0^t V_\alpha g(X_s) \circ \mathrm{d}Z_s^\alpha$$

But since $(\varphi_* V)f(Y_s) = V(f \circ \varphi)(X_s)$, we see that Y is indeed a solution of the equation formulated in the above proposition. $\square$

At this point, one might ask the very relevant question of the existence and uniqueness of solutions to such equations. In the current setting, we are able to prove that equation (1) has a unique solution up to its explosion time.

The proof will require embedding a manifold into a euclidian space.

Theorem 2.6 (Whitney's embedding theorem)**.** *Suppose that M is a smooth manifold. Then there exists an embedding $i : M \to \mathbb{R}^N$ for some N such that the image $i(M)$ is a closed subset of $\mathbb{R}^N$.*

Proof. We will prove the theorem for compact smooth manifolds. Let M be a compact smooth manifold. Let $(\varphi_i : U_i \to \mathbb{R}^n)$ be an atlas of M. Since M is compact, we can extract a finite covering family from (U_i), which covers M by definition of an atlas, and hence we can suppose withtout loss of generality that the atlas considered is finite ($i \leq p$). Let (ρ_i) be a partition of unity subordinate to the open cover (U_i). Since $\sum \phi_i = 1$, there exists some index $i(x)$ such that $\rho_{i(x)}(x) \geq 1/(2p)$. Consider a smooth function $h : \mathbb{R} \to [0, 1]$ that vanishes for all $t \leq 1/(8p)$ and is equal to 1 for all $t \geq 1/(4p)$. Let, for every index i, $\lambda_i = h \circ \rho_i$. Notice that the family $(V_i = \rho_i^{-1}((1/(3p), 1]))$ is an open cover of M, and furthermore that $V_i \subseteq U_i$. Each λ_i is of compact support and is constant equal to 1 on V_i. Consider the maps $f_i : M \to \mathbb{R}^n$ defined by $f_i(x) = \mathbf{1}_{U_i}(x)\lambda_i(x)\varphi_i(x)$. These are smooth maps, and local diffeomorphisms on the corresponding open sets V_i. Finally, the map $g = (g_1, \ldots, g_p) : M \to \mathbb{R}^{p(n+1)}$ where $g_i(x) = (f_i(x), \lambda_i(x))$. We can easily verify that the map g is an immersion, and that it is injective. But since M is compact, g is an embedding. $\square$

Let M be a closed submanifold of $\mathbb{R}^N$ without boundary. Every point x of M has N coordinates (x^i) in $\mathbb{R}^N$. The functions $f^i(x) = x^i$ are called coordinate functions. An important result allowing us to prove existence and uniqueness of solutions is the following:

Proposition 2.7 ([8], proposition 1.2.7)**.** *Let M be a closed submanifold of $\mathbb{R}^N$ without boundary. Let (f^i) be the coordinate functions. Let X be an M-valued continuous process.*

1. *X is a semimartingale iff it as an $\mathbb{R}^N$-valued semimartingale;*

2. *X is a semimartingale iff $f^i(X)$ is a real-valued semimartingale for all i;*

3. X is a solution to the equation (1) up to stopping time τ iff for each i,

$$f^i(X_t) = f^i(X_0) + \int_0^t V_\alpha f^i(X_s) \circ \mathrm{d}Z_s^\alpha, \quad 0 \leq t < \tau$$

Let's get back to the case of a general smooth manifold M. According to theorem 1.3, we are able to embed M into a euclidian space $\mathbb{R}^N$. We can now regard M as a closed submanifold of $\mathbb{R}^N$. Identifying the tangent spaces with $\mathbb{R}^n \times \{0\} \subseteq \mathbb{R}^N$, each vector field V_α can be regarded as a smooth $\mathbb{R}^N$ valued vector field and can be extended to a vector field $\tilde{V_\alpha}$ on $\mathbb{R}^N$. According to results from classical stochastic analysis (see Appendix), the extended equation

$$X_t = X_0 + \int_0^t \tilde{V}_\alpha(X_s) \circ \mathrm{d}Z_s^\alpha \tag{2}$$

has a unique solution X up to its explosion time $e(X)$. Before that time, the solution does not leave the manifold M.

Proposition 2.8 ([8], propostion 1.2.8)**.** *Let X be the solution of the extended equation (2) up to its explosion time $e(X)$ and $X_0 \in M$. Then $X_t \in M$ for all $0 \leq t < e(X)$.*

A direct consequence of this result is the following theorem.

Theorem 2.9 ([8], theorem 1.2.9)**.** *There is a unique solution of equation (1) up to its explosion time.*

Proof. We apply proposition 1.6 and get a solution to (2) that stays in M up to its explosion time, but (2) is a rewriting of (1), we conclude with proposition 1.5. $\square$

2.3.2 Diffusion processes

Let M be a smooth n-dimensional manifold. A mapping $\mathcal{L} : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ is an elliptic differential operator of second order if it can be expressed in local coordinates (x_i) as

$$\mathcal{L} = \frac{1}{2} \sum_{i,j} a^{i,j} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i v^i \frac{\partial}{\partial x_i}$$

such that the coefficient functions $a^{i,j}$ and b^i are smooth, and such that at any point the matrix a is symetric postive definite.

Let M be a finite dimensional smooth manifold. Recall that we denote by $W(M)$ the path space (continuous paths on the Alexandrov compactification of M with explosion time). On this space, we take a filtration $(\mathcal{B}_t)$ generated by coordinate functions up to t. The filtered measurable space $(W(M), \mathcal{B}_t)$ is the standard filtered path space of M. We take $\mathcal{L}$ to be a smooth second order elliptic as defined above. The goal is to give a notion of stochastic motion goverened by $\mathcal{L}$.

Definition 2.11 ([8], definition 1.3.1). *A diffusion process generated by $\mathcal{L}$, or $\mathcal{L}$-diffusion, given a filtered probability space $(\Omega, (\mathcal{F}_t), P)$, is an M-valued $(\mathcal{F}_t)$-semimartingale up to $e(X)$ such that the process*

$$M^f(X)_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,\mathrm{d}s, \quad 0 \le t < e(X)$$

is a local $(\mathcal{F}_t)$-martingale for all smooth functions $f : M \to \mathbb{R}$.

A probability measure μ on the standard filtered path space of M is called a diffusion measure generated by $\mathcal{L}$, or $\mathcal{L}$-diffusion measure, if the process

$$M^f(x)_t = f(x_t) - f(x_0) - \int_0^t \mathcal{L}f(x_s)\,\mathrm{d}s, \quad 0 \le t < e(x)$$

is a local $(\mathcal{B}_t)$-martingale for all paths with explosion time x and all smooth functions $f : M \to \mathbb{R}$.

It is immediate to see the correspondance between $\mathcal{L}$-diffusions and $\mathcal{L}$-diffusion measures. In fact, if X is an $\mathcal{L}$-diffusion then the law of X is an $\mathcal{L}$-diffusion measure, and if μ is an $\mathcal{L}$-diffusion measure then $X_t(\omega) = \omega_t$ is an $\mathcal{L}$-diffusion on $(W(M), (\mathcal{B}_t), \mu)$.

Important propoerties of such measures and processes include uniqueness give an initial distribution and strong Markov properties. These are the results of the following theorems.

Theorem 2.10 ([8], theorems 1.3.4 and 1.3.6). *An $\mathcal{L}$-distribution measure with a given initial distribution exists and is unique.*

Theorem 2.11 ([8], theorem 1.3.7). *Suppose that X is an $\mathcal{L}$ diffusion process on a filtered probability space $(\Omega, (\mathcal{F}_t), P)$ and let τ be an $(\mathcal{F}_t)$-stopping time. Then for every $C \in \mathcal{B}_\infty$, the following holds P-a.s. on $(\tau < e(X))$:*

$$P(X_{\tau+\cdot}|\mathcal{F}_\tau) = P_{X_\tau}(C)$$

In particular, the process X shifted by τ is still an $\mathcal{L}$-diffusion with respect to the same filtration shifted by τ.

Remark. Carrying the proofs of the above theorems (see [8]), we encover the very interesting fact that the solution of a stochastic differental equation of the form

$$\mathrm{d}X_t = V_\alpha(X_t) \circ \mathrm{d}W_t^\alpha + V_0(X_t)\,\mathrm{d}t$$

where W is a euclidian Brownian motion, is an L-diffusion generated by a Hörmander type second order elliptic

$$L = \frac{1}{2}\sum_{i=1}^{n} V_i^2 + V_0$$

2.3.3 Brownian motion on Riemannian manifolds

Let $(M, \langle\cdot,\cdot\rangle_*)$ be a Riemannian manifold with its canonical Levi Civita connection ∇.

Whenever $f : M \to \mathbb{R}$ is a smooth function, then we can define the gradient of f, denoted $\operatorname{grad} f$, as the algebraic dual of the differential, in other terms $\operatorname{grad} f$ satisfies

$$\langle \operatorname{grad} f, X\rangle = \mathcal{L}_X f$$

for every vector field X over M.

Whenever X is a vector field, using the connection ∇, we define the divergence of X as the contraction of the covariant derivative of X.

Definition 2.12. *Let $(M, \langle\cdot,\cdot\rangle_*)$ be a Riemannian manifold with its canonical Levi Civita connection ∇. The Laplace Baltrami operator is the map $\Delta_M : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ defined as $\Delta_M f = \operatorname{div}\operatorname{grad} f$.*

We can show that $\frac{1}{2}\Delta_M$ is a second order elliptic. In fact, for every smooth function $f : M \to \mathbb{R}$, we have

$$\Delta_M f = \sum_i \nabla^2 f(X_i, X_i)$$

whenver (X_i) is a local orthonormal frame of M ([8], proposition 3.1.1). We then are able to give a definition of the Brownian motion on M.

Definition 2.13. *Let $(M, \langle\cdot,\cdot\rangle_*)$ be a Riemannian manifold with its canonical Levi Civita connection ∇. A Brownian motion on M with respect to a filtration $(\mathcal{F}_t)$ is an M-stochastic process $X : \Omega \to W(M)$ that is strong Markov with respect to $(\mathcal{F}_t)$ and it is a $\frac{1}{2}\Delta_M$ diffusion process.*

3 Formulation of Langevin dynamics on Lie groups and sampling from the Gibbs measure

3.1 Markov processes and invariant measures

Results of this section are due to works of Martin Hairer on the ergodic theory of stochastic PDEs (see [6]).

3.1.1 Markov processes

Markov processes are stochastic processes such that the future and the past are independant given the present. More precisely, we have

Definition 3.1 ([6], definition 2.1)**.** *A stochastic process (X_t) taking values in a state space E is called a Markov process if for any times $t_{-N} < \cdots < t_0 < \cdots < t_N$ and any two functions $f, g : E^N \to \mathbb{R}$, the following holds almost surely*

$$\begin{aligned}&E[f(X_{t_1}, \dots, X_{t_N})g(X_{-t_1}, \dots, X_{-t_N})|X_0]\\=&E[f(X_{t_1}, \dots, X_{t_N})|X_0]E[g(X_{-t_1}, \dots, X_{-t_N})|X_0]\end{aligned}$$

In general, Markov processes are understood from their transition probabilities, the following definition gives some insight into this.

Definition 3.2 ([6], definition 2.2)**.** *A Markov Kernel over a Polish space (completely metrizable separable topological space) E is a map $\mathcal{P} : E \times \mathcal{B}(E) \to \mathbb{R}_+$ such that*

1. *For every set $A \in \mathcal{B}(E)$, the map $x \mapsto \mathcal{P}(x, A)$ is measurable;*
2. *For every $x \in E$, the map $A \mapsto \mathcal{P}(x, A)$ is a probability measure.*

Definition 3.3 ([6], definition 2.3)**.** *A Markov operator over a Polish space E is a bounded linear operator $\mathcal{P} : \mathcal{B}_b(X) \to \mathcal{B}_b(X)$ such that:*

1. $\mathcal{P}1 = 1$*;*
2. *$\mathcal{P}\varphi$ is positive whenever φ is;*
3. *Let (φ_n) be a sequence of elements of $\mathcal{B}_b(X)$ converging pointwise to φ, then $\mathcal{P}\varphi_n$ converges pointwise to $\mathcal{P}\varphi$.*

There is a one-to-one correspondance between Markov transition kernels over E and Markov operators over E, which is given by $\mathcal{P}(x, A) = (\mathcal{P}\mathbf{1}_A)(x)$. By the linearity of Markov operators, it suffices to define a Markov operator given a Markov transition kernel on elementary functions, and extend this definition to all $\mathcal{B}_b(E)$ by Lebesgue's dominated convergence theorem. Conversly, given a Markov operator, to verify that the $(x, A) \mapsto (\mathcal{P}\mathbf{1}_A)(x)$ is in fact a Markov transition kernel, we only need to verify that $A \mapsto (\mathcal{P}\mathbf{1}_A)(x)$ is a probability measure for every $x \in E$. The only tricky axiom to verify is countable additivity, which follows from the third axiom in the definition of a Markov operator and the fact that $\sum_{i \leq n} \mathbf{1}_{A_i}$ converges pointwise to $\mathbf{1}_{\cup A_i}$ whenever (A_i) is a countable family of measurable sets. Hence, we will use the same symbol $\mathcal{P}$ to talk about a Markov transition kernel and its corresponding Markov operator. We will also use the same symbol to talk about the operator acting on signed measures by

$$\mathcal{P}\mu(A) = \int_E \mathcal{P}(x, A)\mu(\mathrm{d}x)$$

When (X_t) is a Markov process with continuous time, we consider the familiy $(\mathcal{P}_t)$ of corresponding Markov operators $(\mathcal{P}_t)$. If $\mathcal{P}_{t+s} = \mathcal{P}_t \circ \mathcal{P}_s$, (X_t) is called a Markov semi-group.

3.1.2 Invariant measures

The important notion of this section is that of invariant measures.

Definition 3.4 ([6]). *A probability measure on E is invariant for the markov operator $\mathcal{P}$ if the equality*

$$\int_E (\mathcal{P}\varphi)(x)\mu(\mathrm{d}x) = \int_E \varphi(x)\mu(\mathrm{d}x)$$

holds for every bounded measurable function φ. In other words, $\mathcal{P}_t\mu = \mu$ for all times t.

Let us now suppose that a given Markov semigroup $(\mathcal{P}_t)$ has the Feller property, meaning that it maps continuous bounded functions into continuous bounded functions.

In this setting, we have

Theorem 3.1 (Krylov-Bogolioubov, [6], theorem 6.1). *Let $(\mathcal{P}_t)_{t\geq 0}$ be a Feller Markov semigroup over a Polish space E. Suppose that there exists a bounded signed measure μ_0 such that $\mu_0(X) = 1$ and such that the sequence $(\mathcal{P}_t\mu_0)$ is tight in the sence that for each $\varepsilon > 0$, the exists a compact subset K_ε of X such that for all times t we have*

$$\mathcal{P}_t\mu_0(K_\varepsilon) > 1 - \varepsilon$$

Under these conditions, there exists an invariant measure for $(\mathcal{P}_t)$.

Proof. Put, for all times t

$$\mu_t(A) = \frac{1}{t}\int_0^t (\mathcal{P}_s\mu_0)(A)\,\mathrm{d}s$$

We immediatly get that the family (μ_t) is tight by tightness of the family $(\mathcal{P}_t\mu_0)$. The family (μ_t) viewed as a set is then weak*-relativelty compact in the sapce of probability measures (with its weak* topology), which means that there exists an accumulation point $\mu*$ of the sequence of measures defined above and a sequence $t_n \to \infty$ such that $\mu_{t_n} \to \mu_*$ in the weak sense. For every bounded continuous function, we have

$$\begin{aligned}|(\mathcal{P}_t\mu_*)(\varphi) - \mu_*(\varphi)| &= \lim_n |(\mu_{t_n}(\mathcal{P}_t\varphi) - \mu_{t_n}(\varphi))| \\ &\leq \frac{1}{t_n}\lim_n \left| \int_{t_n}^{t_n+t} \mu_0(\mathcal{P}_s\mu_0)\,\mathrm{d}s - \int_0^t \mu_0(\mathcal{P}_s\mu_0)\,\mathrm{d}s \right|\end{aligned}$$

The function φ being bounded, the integral terms are bounded and hence $|(\mathcal{P}_t\mu_*)(\varphi) - \mu_*(\varphi)| = 0$. This being true for arbitrary test functions φ and times t, we have shown that μ_* is an invariant measure. □

3.1.3 Ergodic invariant measures and a uniqueness criterion for invariant measures

Since we are proving an ergodicity result later on, we need to define this notion for invariant measures of Markov semigroups. One way to define this property is through the analogous notion of ergodic invariant measures of dynamical systems. Let E be a polish space. A dynamical system on E is a collection $(\theta_t)_{\mathbb{R}_+}$ of $E \to E$ maps such that $\theta_{s+t} = \theta_t \circ \theta_s$ for all times s, t and such that $(t, x) \mapsto \theta_t(x)$ is jointly measurable. We say that the dynamical system is continuous if the maps θ_t are continuous for all times t. Denote by $\theta_t^*\mu$ the pushforward of the measure μ under θ_t, or image measure, and define the set of invariant measures of the dynamical system (θ_t), denoted $\mathcal{J}(\theta)$, as the set of bounded normalized measures μ such that $\theta_t^*\mu = \mu$. The invariant subsets of E, which are all the subsets A s.t. $\theta_t^{-1}(A) = A$ for all times t, forms a σ-algebra.

Definition 3.5 ([6], definition 3.4)**.** *An invariant measure for a dynamical system (θ_t) is ergodic if $\mu(A) = 0, 1$ for all invariant subsets A of E.*

Given a Markov semigroup $\mathcal{P}_t$ over a polish space E and an invariant measure μ of $\mathcal{P}_t$, a construction found in [6] section 4 concludes that there exists a time-homogenous Markov process (X_t) on $E^{\mathbb{R}}$ such that the law of X_0 is μ and such that the transition operator of (X_t) is $\mathcal{P}_t$. We can then define the law of the process (X_t) as a probability measure P_μ. Define θ as the dynamical system of shift maps $(\theta_t x)(s) = x(t+s)$. We are able to, using the invariance of μ, prove that P_μ is an invariant measure for the dynamical system θ. In fact, it is straightforward to check that the measure P_μ is stationary, that is $\theta_t^* P_\mu = P_\mu$ for all times t, thanks to the invariance of μ.

Definition 3.6 ([6], definition 4.1)**.** *Let $\mathcal{P}_t$ be a Markov semigroup on a Polish space and μ an invariant measure of $\mathcal{P}_t$. Then we say that μ is ergodic if P_μ is ergodic for the dynamical system of shift maps (θ_t).*

Let $\mathcal{P}_t$ be a Markov semigroup on a Polish space. We denote by $\mathcal{J}(\mathcal{P})$ the set of all invariant measures for the Markov semigroup $\mathcal{P}_t$. We have the following characterisation of the set of invariant measures:

Theorem 3.2 ([6], theorem 5.1)**.** *Let $\mathcal{P}_t$ be a Markov semigroup on a Polish space. An invariant measure of $\mathcal{P}_t$ is ergodic iff it cannot be decomposed as a non trivial convex linear combination of invariant measures. Any two ergodic invariant measures are either identical or mutually singular.*

Corollary 3.3 ([6], corollary 5.6)**.** *If a Markov process with transition operator $\mathcal{P}$ has a unique invariant measure, then it is ergodic.*

We are now able to state a uniqueness criterion for the invariant measure of a Markov process.

Definition 3.7 ([6], definition 7.6)**.** *A Markov operator $\mathcal{P}$ over a polish space E has the strong Feller property if, for every function bounded measurable function φ, $\mathcal{P}\varphi$ is bounded continuous.*

Proposition 3.4 ([6], proposition 7.7)**.** *If a Markov operator over a Polish space E has the strong Feller property, then the topological supports of any two mutually singular invariant measures are disjoint.*

3.2 Stochastic hamiltonian dynamics and Langevin diffusions on Lie groups

We follow the discussions from [11] to formulate stochastic hamiltonian mechanics. In the following section, $M = T^*G \simeq G \times \mathfrak{g}^*$ will denote the cotangent bundle of a compact Lie group G. The manifold M is a symplectic manifold with symplectic form $\mathrm{d}g \wedge \mathrm{d}m$. On symplectic manifolds M with symplect form ω, a non-degenerate 2-form, we can define hamiltonian vector fields as the vector field X_H satisfying $\omega(X_H, \cdot) = \mathrm{d}H$. Deterministic hamiltonian mechanics are then formulated as

$$\frac{\mathrm{d}S}{\mathrm{d}t} = X_H(S)$$

where $S_t \in M$ describes the state of the system.

In [11], the authors lift the previously described deterministic mechanics to stochastic geometric mechanics by the means of Malliavin's transfer principle. Given driving real-valued semimartingales $(Z^\alpha)_{\alpha=0,\dots,n}$, with the convetion that $Z_t^0 = t$, and hamiltonians $(H_\alpha)_{\alpha=0,\dots,n}$ (real-valued smooth functions), the stochastic version of Hamiltion's equations is given by

$$\mathrm{d}S = X_{H_\alpha} \circ \mathrm{d}Z_t^\alpha \tag{SHM}$$

Our hope at this stage it's that the invariant measure for such dynamics is the Gibbs measure, since deterministic Hamiltonian systems on symplectic manifolds have the Gibbs measure as their invariant measure. Unfortunately, that is not the case and we need to account for correction terms in order to obtain equations with the desired invariant measure.

Let $H : G\times\mathfrak{g}^* \to \mathbb{R}$ a smooth function, the Gibbs measure P_∞ is the probability measure defined as

$$P_\infty = Z^{-1}e^{-\beta H}\mu(\mathrm{d}g)\lambda(\mathrm{d}m), Z = \int_{G\times\mathfrak{g}^*} e^{-\beta H}\mu(\mathrm{d}g)\lambda(\mathrm{d}m)$$

where the measure μ is the Haar measure (unique left-translation-invariant measure, with finite value on compact sets) on G and λ is the Lebesgue measure on the dual space $\mathfrak{g}^*$.

Whenever F and G are two hamiltonians, define the poisson bracket of F and G as $\{F, G\} = \omega(X_F, X_G)$. We introduce a dissipation term in the equation (SHM) and consider the following stochastic differential equation on M

$$\mathrm{d}S_t = \left(X_{H_0}(S_t) + \frac{\beta}{2}\{H_0, H_i\}X_{H_i}(S_t)\right)\mathrm{d}t + X_{H_i}(S_t) \circ W_i^t \tag{MSHM}$$

where the functions H_i are hamiltonian vector fields and the W_i are linear brownian motions.

Theorem 3.5 ([11], proposition 5.2)**.** *Equation (MSHM) has the Gibbs measure as an invariant measure.*

3.3 An ergodicity result

In this section, we prove that the invariant measure of equation (MSHM) is ergodic under assumptions on the hamiltonian vector fields. We will mainly use the result of the previous discussion about invariant measures of Markov semi-groups. In fact, in section 3.1 we stated that if there is only one invariant measure to a Markov semi-group, then the invariant measure is ergodic, under strong Feller assumptions. Hence, to carry out our proof, we need to make sure that, given a solution to the studied equation, it satisfies the strong Markov property and that its corresponding Markov semi-group is strong Feller. To do so, we will assume the following regarding the hamiltonian vector fields in the equation (MSHM)

Assumption 1. *Consider the hamiltonians $H_1, \dots, H_n$ involved in the equation mentionned above. Recall that the bracket of two vector fields X and Y is the vector field $[X, Y]$ such that $\mathcal{L}[X, Y] = [\mathcal{L}_X, \mathcal{L}_Y]$. We assume that the hamiltonian vector fields $X_i = X_{H_i}$ for $i = 1, \dots, n$ and X_0 the vector field against $\mathrm{d}t$ are such that the successive Lie brackets $[X_{i_1}, X_{i_2}], \cdots, [X_{j_1}, [X_{j_2}, X_{j_3}]], \cdots$ span the tangent spaces at each point.*

This is called Hörmander's condition.

Assumption 2. *Assume that the hamiltonain vector fields verify a growth condition such that solutions to the equation (MSHM) are well-defined for all times.*

The second assumption is, for the examples that we generally consider, not difficult to obtain. In fact, in some examples that we will expose later on, we will be able to write them in local coordinates as $T^*G \simeq G \times \mathfrak{g}^*$, where G is a compact Lie group and $\mathfrak{g}$ its Lie algebra. We will see that some terms only depend on elemnts of the group, which is compact so satisfies the growth condition, and for the terms depending on the dual space, they will be linear, hence also satisfying the growth condition.

We are then able to prove the following ergodicity result

Proposition 3.6. *Under the previous assumptions, equation (MSHM) has a unique ergodic invariant measure.*

Proof. Applying Hörmander's theorem (see [7]), we have several deductions. First, the generator of equation (MSHM) is an elliptic second order operator, the unique solution (S_t) to this equation up to explosion time $e(S)$ is a diffusion with respect to this generator, it has in particular the strong Markov property. Under the assumption 2, we have that $e(S) = \infty$, and denote by $(\mathcal{P}_t)$ the Markov semi-group associated with (S_t). Then, thanks to Hörmander's theorem, the Markov semi-group $(\mathcal{P}_t)$ maps bounded measurable functions to smooth bounded functions, meaning that it is strong Feller. We know that the Gibbs measure is an invariant measure to $\mathcal{P}$. We show that the Gibbs measure is the only invariant measure and apply Corrollary 3.3 to conclude that it is ergodic. In fact, if there exists another invariant measure μ, thanks to the Feller property of the studied Markov semi-group, then the topological supports of the Gibbs measure and μ are disjoint, since they are mutually singular because of Theorem 3.2. Hence, because the topological support of the Gibbs measure is the entire space, μ cannot exist and the Gibbs measure is the only invariant measure, which concludes the proof. □

Let's apply this result to examples from [11] :

Example 3.1. 1. **Momentum Langevin on compact Lie groups**. Let G be a compact semisimple Lie group with bi-invariant metric Q. Choose the semimartingale terms as $(Z^\alpha)_{\alpha=0}^n = (t, W_t^1, \dots, W_t^n)$. Let the drift Hamiltonian be $\mathcal{H}_0(g,m) = \frac{1}{2}Q(m,m) + V(g)$ where $V : G \to \mathbb{R}$ is the potential. Let $\{X_i\}_{i=1}^n$ be a basis for $\mathfrak{g}$ and take the diffusion Hamiltonians to be $\mathcal{H}_i(g,m) = -\sqrt{2\gamma}\,\mathrm{Tr}(gX_i)$ with $\gamma > 0$. Substituting this into (MSHM) yields, in Darboux coordinates,

$$
\begin{aligned}
\boldsymbol{d}m &= \big(-\beta\gamma m - (dL_{g^{-1}})_g \nabla V\big)\boldsymbol{d}t + \sqrt{2\gamma}\, X_i\, \boldsymbol{d}W_t^i \\
\boldsymbol{d}g &= (dL_g)_e\, m\, \boldsymbol{d}t,
\end{aligned}
$$

we notice that terms depending on m are linear in m, such that the growth condition is satisfied and with that assumption 2. Assumption 1 is also satisfied as the succesive Lie brackets of the vector fields include the elements of the basis of $\mathfrak{g}$, and this is sufficient to span the tangent space at each point. Hence, our result applies and this equation has the Gibbs measure as a unique ergodic invariant measure.

2. **Position Langevin on compact Lie groups**. Let G be a compact Lie group with bi-invariant metric Q. Choose the semimartingale terms as $(Z^\alpha)_{\alpha=0}^n = (t, W_t^1, W_t^2, \ldots, W_t^n)$. Let the drift Hamiltonian be $\mathcal{H}_0(g,m) = \frac{1}{2}Q(m,m) + V(g)$ where $V : G \to \mathbb{R}$ is a potential. Let $\{X_i\}_{i=1}^n$ be a basis for $\mathfrak{g}$ and take the diffusion Hamiltonians to be $\mathcal{H}_i(g,m) = \sqrt{2\gamma}\, Q(m, X_i)$. Substituting these expressions into (MSHM) leads to, in Darboux coordinates,

$$\begin{aligned}\boldsymbol{d}m &= \big(-(dL_{g^{-1}})_g \nabla V + \beta\gamma Q(X_i, (dL_{g^{-1}})_g \nabla V)[X_i, m]\big)\, \boldsymbol{d}t + \sqrt{2\gamma}[X_i, m] \circ \boldsymbol{d}W_t^i, \\ \boldsymbol{d}g &= (dL_g)_e \Big(\big(m - \beta\gamma\, \nabla V\big)\, \boldsymbol{d}t + \sqrt{2\gamma}\, X_i\, \boldsymbol{d}W_t^i \Big),\end{aligned}$$

Here too, the terms depending on m are the $[X_i, m]$, which are linear in m, hence the growth condition is satisfied. Assumption 1 is also satisfied, because the elements of the basis of $\mathfrak{g}$ are also included in the successive Lie brackets of the vector fields. Hence, our result applies and this equation has the Gibbs measure as a unique ergodic invariant measure.

4 Conclusions

Scientific conclusion

Building on the work of Luesink and Street in [11], we have shown that under the assumption of Hörmander's condition, the Gibbs measure is the unique invariant measure of the Langevin diffusion on compact Lie groups and that it is ergodic. We are then confident that we should have all the to tools to carry out the analysis of Lie-Trotter type splittings for Langevin diffusions as done by Abdulle and Vilmart in [1] in the Euclidean case but in the setting of compact Lie groups. Although some complications may arise due to the curvature of the manifolds we are working with. In particular, work in [1] concerns energy preserving integrators for the deterministic part of the splitting, which is something hard to achieve on manifolds, all the simulations carried out seem to either guaratnee energy conservation but without the solution exiting the manifold, or do not guarantee energy conversation but providing a solution that stays on the manifold.

Personal conclusion

This project was a great opportunity to work on a topic that is at the intersection of several fields of mathematics that I am passionate about. It was an ambitious project because it required to catch up with a lot of litterature in different fields, but it was rewarding at the end and it got me to discuss interesting ideas with people working in different fields. My supervisors were of great help throughout this internship, Erwin Luesink never hesitated to answer my questions or have long discussions about one of my problems and it helped me a lot to gain of insight on the subject, especially for the geometry part of the project. The same goes for Sonja Cox, who was very helpful in our meetings and who provided me with precious advice. Both also provided with a lot of freedom, I was allowed to go in any direction I wanted building on the results of Luesink and Street.

Ultimately, this project got me to experience the life of a reasercher inside an important institute of mathematics and go through the process of research in mathematics, from litterature review to ideas and finally some results and outlooks.

5 Appendix

5.1 Basic elements of stochastic analysis and stochastic differential equations

5.1.1 Brownian Motion

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Recall that a real random variable X is said to be a standard Gaussian (or normal) variable if its law has density

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

with respect to the Lebesgue measure on $\mathbb{R}$.

The complex Laplace transform of X is then given by, for $z \in \mathbb{C}$

$$E[e^{zX}] = e^{\frac{z^2}{2}}$$

The calculation is done by first verifying the result for all $z \in \mathbb{R}$ and then extending it to the complex plane by analytic continuation, since $z \mapsto E[e^{zX}]$ is well defined for all $z \in \mathbb{C}$ and holomorphic.

We also get the characteristic function of X by taking $z = i\theta$ for $\theta \in \mathbb{R}$ in the above equality:

$$E[e^{i\theta X}] = e^{-\frac{\theta^2}{2}}$$

Definition 5.1. *We say that a random variable Y is a Gaussian with $\mathcal{N}(\mu, \sigma^2)$ distribution if it satisfies $Y = \sigma X + \mu$ where X is a standard Gaussian variable.*

Remark. The above definition is equivalent to any of the following properties:

1. The law of Y has density:

$$p_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$$

2. The characteristic function of Y is given by:

$$E[e^{i\theta Y}] = e^{i\theta\mu - \frac{\sigma^2\theta^2}{2}}$$

Proposition 5.1 ([10], propostion 1.1)**.** *Let (X_n) be a sequence of real random variables such that for all $n \in \mathbb{N}$, X_n follows the $\mathcal{N}(m_n, \sigma_n^2)$ distribution. Suppose that X_n converges in L^2 to X. Then:*

1. *The random variable X follows the $\mathcal{N}(m, \sigma^2)$ distribution where $m = \lim_{n\to\infty} m_n$ and $\sigma = \lim_{n\to\infty} \sigma_n$.*

2. *The convergence also holds in all L^p spaces, $1 \leq p < \infty$.*

Proof. 1. First, we notice that

$$|E[X_n] - E[X]| \leq \|X_n - X\|_{L^2} \to 0$$

This means that the sequence (m_n) admits a limit that we choose to call m. We can also write

$$\sigma_n^2 = E[|X_n - m_n|^2] = \|X_n - m_n\|_{L^2}^2 \leq (\|X_n - X\|_{L^2} + \|X - m\|_{L^2} + \|m - m_n\|_{L^2})^2$$

which shows that (σ_n^2) admits a limit $\sigma^2 = \|X - m\|_{L^2}^2$. We now show that X follows the $\mathcal{N}(m, \sigma^2)$ distribution. Given our previous remark, it sufficies to show that X has the right characteristic function. For every $\theta \in \mathbb{R}$, we have

$$E[e^{i\theta X_n}] = e^{i\theta m_n - \frac{\sigma_n^2\theta^2}{2}} \to e^{i\theta m - \frac{\sigma^2\theta^2}{2}}$$

but also

$$E[|e^{i\theta X_n} - e^{i\theta X}|] \leq E[|X_n - X|] + \mathcal{O}(E[|X_n - X|^2]) \to 0$$

2. For every n, X_n has same distribution as $m_n N + \sigma_n$. The sequences (m_n) and (σ_n) being bounded, it follows that the sequence $(|X_n - X|)$ is bounded in L^p for every $p \geq 1$. But, since X_n converges to X in L^2, the sequence $(|X_n - X|)$ converges to 0 in probability, which means that it converges in L^p for every $p \geq 1$.

□

Now we generalize this notion to random vectors. If E is a a d-dimentional euclidien space, we say that a random vector X is a Gaussian vector if for every $x \in E$, the random variable $\langle x, X\rangle$ is a Gaussian variable. The map $x \mapsto E[\langle x, X\rangle]$ is an element of E^*, we denote by m_X the unique vector in E such that $E[\langle x, X\rangle] = \langle x, m_X\rangle$ for every $x \in E$. The map $x \mapsto \mathrm{Var}(\langle x, X\rangle)$ is a quadratic form on E, we denote it by Q_X. Notice that $\langle x, X\rangle$ follows the $\mathcal{N}(\langle x, m_X\rangle, Q_X(x))$ distribution.

Furthermore, there exists a unique endomorphism A_X of E such that for every $x \in E$, $\mathbb{Q}_X(X) = \langle Ax, x\rangle$. We can show that if Q_X is definite, then the distribution of X is absolutely continuous with respect to the Lebesgue measure on E, and the density is given by:

$$P_X(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det A_X}} \exp\left(-\frac{1}{2}\langle A_X^{-1}x, x\rangle\right)$$

Before we could construct the Brownian motion, let's first introduce Gaussian spaces and Gaussian processes.

We will consider only centered Gaussian variables.

Definition 5.2 ([10], definition 1.4)**.** *A centered Gaussian space is a closed linear subspace of L^2 which contains only centered Gaussian variables.*

Example 5.1. If $(X_1, \ldots, X_n)$ is a centered Gaussian vector in $\mathbb{R}^n$, then the space generated by $(X_1, \ldots, X_n)$ is a centered Gaussian space.

Definition 5.3 ([10], definition 1.5)**.** *If $(E, \mathcal{E})$ is a measurable space and T an arbitrary index set, a random process indexed by T with values in E is a collection of random variables $(X_t)_{t \in T}$ of random variables with values in E.*

Most of the time, T will be $\mathbb{R}_+$.

Definition 5.4 ([10], definition 1.6)**.** *A real-valued process $(X_t)_{t \in T}$ is said to be a centered Gaussian process if any finite linear combination of the variables X_t is centered Gaussian.*

Remark. It is clear that the linear subspace generated by $(X_t)_{t \in T}$ is a centered Gaussian space.

Definition 5.5 ([10], definition 1.12)**.** *Let $(E, \mathcal{E})$ be a measurable space, and let μ be a σ-finite measure on $(E, \mathcal{E})$. A Gaussian white noise with intensity μ is an isometry G from $L^2(E, \mathcal{E}, \mu)$ to a centered Gaussian space.*

Remark. For $f \in L^2(E, \mathcal{E}, \mu)$, $G(f)$ is centered Gaussian with variance

$$E[G(f)^2] = \|G(f)\|^2_{L^2(\Omega, \mathcal{F}, P)} = \|f\|^2_{L^2(E, \mathcal{E}, \mu)} = \int_E f^2 \, \mathrm{d}\mu$$

For $f, g \in L^2(E, \mathcal{E}, \mu)$, the covariance of $G(f)$ and $G(g)$ is given by

$$E[G(f)G(g)] = \langle G(f), G(g) \rangle_{L^2(\Omega, \mathcal{F}, P)} = \langle f, g \rangle_{L^2(E, \mathcal{E}, \mu)} = \int_E fg \, \mathrm{d}\mu$$

We now construct the Brownian motion and inverstigate some of its properties. We start by introducing the pre-brownian motion.

Definition 5.6 ([10], definition 2.1)**.** *Let G be a Gaussian white noise on $\mathbb{R}_+$ with intensity the Lebesgue measure. The random process $(B_t)_{t \in \mathbb{R}_+}$ defined by*

$$B_t = G(\mathbf{1}_{[0,t]})$$

is called pre-Brownian motion.

Proposition 5.2 ([10], propostion 2.2)**.** *Pre-Brownian motion is a centered Gaussian process with covariance*

$$K(s, t) = s \wedge t$$

Proof. G being a white noise, the variables B_t all belong to some Gaussian space, which means that the process $(B_t)_{t \in \mathbb{R}_+}$ is a centered Gaussian process. Also, for $s, t \in \mathbb{R}_+$, we have

$$E[B_s B_t] = E[G(\mathbf{1}_{[0,s]})G(\mathbf{1}_{[0,t]})] = \langle G(\mathbf{1}_{[0,s]}), G(\mathbf{1}_{[0,t]}) \rangle_{L^2(\Omega, \mathcal{F}, P)} = \int_0^{s \wedge t} \mathrm{d}\mu = s \wedge t$$

□

The above properties also characterize pre-Brownian motions. We can also give different ways of doing so:

Proposition 5.3 ([10], proposition 2.3). *Let $(X_t)_{t\in\mathbb{R}_+}$ be a real-valued random process. The following properties are equivalent:*

1. *(X_t) is pre-Brownian motion;*
2. *(X_t) is a centered Gaussian process with covariance $K(s,t) = s \wedge t$;*
3. *$X_0 = 0$ a.s., and, for every $s < t$, the random variable $X_t - X_s$ is independant of $\sigma(X_r, r \leq s)$ and distributed according to $\mathcal{N}(0, t-s)$*

This gives us a practical way of verifying that a process is a pre-Brownian motion.

Given a process $(X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, P)$ with values in a metric space E equipped with its Borel σ-field, an important notion is that of the sample paths of said process. The sample paths of X are simply the mappings $t \mapsto X_t(\omega)$ for $\omega \in \Omega$.

In general, we cannont state anything about these mappings, not even that they are measurable, even for a pre-brownian motion. This leads us to the definition of a linear Brownian motion started from 0.

Definition 5.7 ([10], definition 2.12). *A process $(B_t)_{t\in\mathbb{R}_+}$ is said to be a Brownian motion (BM) if it is a pre-Brownian motion with continuous sample paths.*

The existence of such a process is not trivial but is guaranteed by Kolmogorov's lemma, which provides a way to construct a process with continuous paths starting from a given pre-Brownian motion such that these two processes have a.s the same sample paths.

5.1.2 Itô integration and SDEs

In the following, we will consider a probability space $(\Omega, \mathcal{A}, P)$ on which we define a real Brownian motion $(B_t)_{t\in\mathbb{R}_+}$. We consider the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ generated by the Brownian motion, that is, $\mathcal{F}_t = \sigma(B_s, s \leq t)$.

Our goal is to define a stochastic integral with respect to the BM for functions $f : \Omega \times [0,T] \to \mathbb{R}$ which are measurable with respect to $\mathcal{F} \otimes \mathcal{B}([0,T])$. To do so, we first describle a special class of such functions:

Definition 5.8. *Let H^2 be the class of such functions f that verify the following:*

1. *For every $t \in [0,T]$, $f(\cdot,t)$ is $\mathcal{F}_t$-measurable;*
2. *$E[\int_0^T f(\cdot,t)^2 \,\mathrm{d}t] < \infty$.*

The strategy to define the stochastic integral is to find a subclass of H^2 simple enough to give a first construction and "big enough" to extend to all of H^2.

Consider then the following class of elementary functions of H^2:

$$H_0^2 = \{f \in H_0^2 | f(\omega, t) = \sum_{i=1}^{n} \mathbf{1}_{[t_i, t_{i+1})}(t) a_i(\omega), 0 = t_0 < t_1 < \cdots < t_n = T, a_i \in L^2(\Omega, \mathcal{F}_{t_i}, P)\}$$

We want to have $\int_s^t 1 \, \mathrm{d}B_t = B_s - B_t$ and, naturally, the stochastic integral to be a linear operator, we are then left with the following choice:

$$I(f)(\omega) = \sum_{i=1}^{n} a_i(\omega)(B_{t_{i+1}} - B_{t_i})(\omega)$$

I maps then an element of H_0^2 to a random variable.

The final and most difficult step is now to extend this definition to all of H^2. We will do so by a density argument.

We will indeed prove that H_0^2 is dense in H^2 for the norm $\|f\| = \sqrt{E[\int_0^T f(\cdot, t)^2 \, \mathrm{d}t]}$. A first key result is Itô's isometry for elementary functions:

Proposition 5.4. *Let $f \in H_0^2$, then:*

$$E[I(f)^2] = E\left[\int_0^T f(\cdot, t)^2 \, \mathrm{d}t\right]$$

Proof. Indeed, we have

$$E[I(f)^2] = E\left[\left(\sum_{i=1}^{n} a_i(B_{t_{i+1}} - B_{t_i})\right)^2\right] = \sum_{1 \le i, j \le n} E[a_i a_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})]$$

By noticing that $a_i a_j (B_{t_{i+1}} - B_{t_i})$ and $B_{t_{j+1}} - B_{t_j}$ are independant random variables for $i < j$, and since $E[B_{t_{i+1}} - B_{t_i}] = 0$, we get

$$E[I(f)^2] = \sum_{i=1}^{n} E[a_i^2] E[(B_{t_{i+1}} - B_{t_i})^2] = \sum_{i=1}^{n} E[a_i^2](t_{i+1} - t_i) = E\left[\int_0^T f(\cdot, t)^2 \, \mathrm{d}t\right]$$

□

Proposition 5.5. *The class H_0^2 is dense in H^2 for the norm precedently mentioned.*

Proof. The idea is to approximate a function $f \in H^2$ step by step. It is easy to see that f can be approximated by a sequence of bounded functions of H^2, this is done by truncating f for values exceeding n in absolute value for every n and applying Lebesgue's dominated convergence theorem.

Then, any bounded function h of H^2 can be approximated by a sequence of bounded t-continuous functions of H^2 in the following way: take a mollifier $(\varphi_n)_{n\in\mathbb{N}}$ such that φ_n is equal to 0 outside of $(-1/n, 0)$ and set for all ω, $g_n(\omega, t) = \int_0^t \varphi_n(s-t)h(\omega, s)\,\mathrm{d}s$. If follows that $\int_0^T (g_n(\omega, t) - h(\omega, t))^2\,\mathrm{d}t \to 0$ for each ω. Since we also have, $|g_n(\omega, t)| \leq M$, we conclude by the dominated convergence theorem.

Finally, we can approach any bounded t-continuous function g of H^2 by a sequence of elementary functions in the following way: set

$$h_n(\omega, t) = \sum_{i=1}^{n} \mathbf{1}_{[t_i, t_{i+1})}(t) g(\omega, t_i)$$

On $[0, T]$, $g(\omega, \cdot)$ is uniformly continuous since $[0, T]$ is compact, which gives the uniforme convergence $h_n(\omega, \cdot) \to g(\omega, \cdot)$ for every $\omega \in \Omega$, hence

$$\int_0^T (g - h_n)^2\,\mathrm{d}t \to 0$$

and we conclude by bounded convergence. □

We are now able to define the following

Definition 5.9. *Let $f \in H^2$. Then the Itô integral of f (from 0 to T) is defined by*

$$\int_0^T f(t, \omega)\,\mathrm{d}B_t(\omega) = \lim_{n\to\infty} \int_0^T h_n(t, \omega)\,\mathrm{d}B_t(\omega)$$

where (h_n) is a sequence of elementary such that

$$E\left[\int_0^T |f - h_n|^2\,\mathrm{d}t\right] \to 0$$

The limit defining the integral exists thanks to the previous proposition and Itô's isometry, and it does not depend on the sequence (h_n). In fact, since $h_n \to f$ for the norm $\|\cdot\|$, this sequence is a Cauchy sequence, which in turns gives us that $(I(h_n))$ is a Cauchy sequence in $L^2(\Omega, \mathcal{F}, P)$ thanks to Itô's isometry, hence it admits a limit since the considered space is complete. The fact that it does not depend on the approximating sequence is proven by putting $f_{2n} = h_n$ and $f_{2n+1} = g_n$ where (h_n) and (g_n) are two approximating sequences.

Note that the Itô isometry extends to H^2 as well.

We directly deduce the following properties of the Itô integral:

Proposition 5.6. *Let $f, g \in H^2$ and $0 \leq u \leq T$, then:*

1. $\int_0^T f\,\mathrm{d}B_t = \int_0^u f\,\mathrm{d}B_t + \int_u^T f\,\mathrm{d}B_t$ *for a.e. ω;*
2. $\int_0^T (\lambda f + g)\,\mathrm{d}B_t = \lambda \int_0^T f\,\mathrm{d}B_t + \int_0^T g\,\mathrm{d}B_t$ *for a.e. ω;*

3. $E[\int_0^T f\,\mathrm{d}B_t = 0]$*;*

4. $\int_0^T f\,\mathrm{d}B_t$ *is* $\mathcal{F}_T$ *measurable.*

Remark. 1. We can also define an integral from S to T rather than from 0 to T and the above proofs do not change;

2. It is possible to extend the Itô integral to higher finite dimensions. First, it is possible to allow hte integrand to depend on more than just the filtration $\mathcal{F}_t$ as long as the brownian motion remains a martingale with respect to the history of f. This allows us for example to define the Itô integral for integrands adapted to the filtration generated by a n-dimensional Brownian motion with respect to one of the components of the BM. Hence, if f is a $m\times n$ matrix such that for each (i,j), $f_{i,j}$ is a process meeting the requirements to be an element of H^2, except for the adaptability condition which we replace by the weaker condition stated above. Then, using matrix notation, we define

$$\int_0^T f\,\mathrm{d}B = \int_0^T \begin{pmatrix} f_{1,1} & \dots & f_{1,n} \\ \vdots & \ddots & \vdots \\ f_{m,1} & \dots & f_{m,n} \end{pmatrix} \begin{pmatrix} \mathrm{d}B_1 \\ \vdots \\ \mathrm{d}B_n \end{pmatrix}$$

to be the m sized vector where the i'th component is the 1-dim Itô integral $\sum_j \int_0^T f_{i,j}\,\mathrm{d}B_j$. There are other extensions of the Itô integral which mainly imply weeking some of the conditions on the integrand.

3. Of course, the Itô integral is not the only way to define a stochastic integral. We have shown that if $f \in H^2$ is t-continuous, then

$$\int_0^T f\,\mathrm{d}B = \lim_{\Delta t_j} \sum_j f(t_j,\omega)(B_{t_{j+1}} - B_{t_j})$$

It is possible to define the Stratonovich integral of f with respect to B, which we denote $\int_0^T f \circ \mathrm{d}B_t$, by taking the same limit as above in L^2 but with replacing $f(t_j,\omega)$ by $f(\frac{1}{2}(t_j + t_{j+1}),\omega)$, whenever it exists. Differrent stochastic integrals don't have any reason to produce the same results. For example, one would want or expect the integral $\int_0^T B\,\mathrm{d}B$ to produce $B_T^2/2$ if $B_0 = 0$, but when we apply Itô's formula with $f(t,x) = x^2/2$, we get $\int_0^T B\,\mathrm{d}B = \frac{1}{2}B_T^2 - \frac{1}{2}t$, but in the Stratonovich sense, the chain rules gives the desired result $\int_0^T B \circ \mathrm{d}B = \frac{1}{2}B_T^2$.

In the case of deterministic integration, we usually resorst to a chain rule in order to compute an integral. We are able to write a chain rule for the Itô integral as well.

Definition 5.10. *An Itô process is a stochastic process X such that there exists $\mathcal{F}_t$ adapted processes u and v such that the following holds and makes sense:*

$$X_t = X_0 + \int_0^t u_s \,\mathrm{d}s + \int_0^t v_s \,\mathrm{d}B_s$$

We usually write $\mathrm{d}X_t = u\,\mathrm{d}t + v\,\mathrm{d}B_t$.

We then have the Itô formula as follows:

Theorem 5.7 (Itô's formula, [10] section 5.2)**.** *Let X be an Itô process and let $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^2$. Then, for every $t \in \mathbb{R}_+$, we have a.s.*

$$f(t, X_t) = f(0, X_0) + \int_0^t \left(\partial_s f(s, X_s) + \frac{1}{2} \partial_{x^2} f(s, X_s) v_s^2 \right) \mathrm{d}s + \int_0^t \partial_x f(s, X_s) v_s \,\mathrm{d}B_s$$

We formally write this as:

$$\boxed{\mathrm{d}f(t, X_t) = \partial_t f \,\mathrm{d}t + \partial_x f(s, X_s) v_s \,\mathrm{d}X_t + \frac{1}{2} \partial_{x^2} f (\mathrm{d}X_t)^2}$$

5.1.3 Stochastic differential equations

In the Itô interpretation, we call stochastic differential equation (SDE) an equation of the form

$$X_t = Z + \int_0^t b(s, X_s) \,\mathrm{d}s + \int_0^t \sigma(s, X_s) \,\mathrm{d}B_s$$

which is also written as

$$\mathrm{d}X_t = b(t, X_t) \,\mathrm{d}t + \sigma(t, X_t) \,\mathrm{d}B_t$$

Existence and uniqueness of solutions to such equations obviously depend on the properties of the functions b and σ. We can formulate the following theorem stating sufficient conditions for the existence and uniqueness of "strong" solutions to such equations:

Theorem 5.8 ([10], section 8.2)**.** *Let $T > 0$, $b : [0, T] \times \mathbb{R} \to \mathbb{R}$ and $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ be measurable functions such that, on $[0, T] \times \mathbb{R}$, we have*

$$|b(t, x)| + |\sigma(t, x)| \leq C(1 + |x|),$$
$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq C|x - y|$$

for some strictly positive constant C, and Z an L^2 random variable independant from the σ-algebra generated by (B_s). Then, the SDE

$$\mathrm{d}X_t = b(t, X_t) \,\mathrm{d}t + \sigma(t, X_t) \,\mathrm{d}B_t, \quad X_0 = Z$$

admits a unique t-continuous solution X such that X_t is adapted to the filtration $\mathcal{F}_t^Z$ generated by Z and B_s $s \leq t$ and

$$E\left[\int_0^T |X-t|^2 \,\mathrm{d}t\right] < \infty$$

Remark. The previous theorem gives a unique solution that satisfies the differential equation for all times. As we can see in the proof, this is due to global Lipschitz and growth assumptions on the coefficients of the SDE. When the global Lipschitz condition is not met, we may have the possibility of explosion. Simple examples can be constructed if we take the diffusion coefficient to be 0, for example the deterministic equation $\mathrm{d}X_t = X_t^2 \,\mathrm{d}t$, the solution of which obviously blows up at 1. Other examples include the SDE $dX_t = X_t^3 \,\mathrm{d}t - X_t^2 \,\mathrm{d}B$, the solution of which is $X_t = 1/(1+B_t)$.

We will now allow solutions to explode. Recall that every locally compact metric space M admits an Alexandrov (one-point) compactification $\hat{M} = M \cup \{\infty_M\}$.

If M is a closed submanifold of some euclidian space, we have the follwing

Proposition 5.9. *Let M be a non-compact closed submanifold of $\mathbb{R}^N$ and $\hat{M} = M \cup \{\infty_M\}$ be the Alexandrov compactification of M. A sequence of points (x_n) in M converges to ∞_M in $\hat{M}$ iff it converges to infinity in norm in $\mathbb{R}^N$.*

Proof. If the sequence (x_n) converges to infinity in norm in $\mathbb{R}^N$, this means that it leaves every compact set eventually, but by definition of the topology of the Alexandrov compactification, $\{K^c \cup \{\infty_M\}, \quad K \text{ compact set}\}$ is a neighbourhood basis of ∞_M, hence the sequence (x_n) enters every neighbourhood of ∞_M eventually. Conversly, if the sequence (x_n) converges to ∞_M in $\hat{M}$, enters every neighbourhood of ∞_M eventually, which means that for every compact set K, it enters $K^c \cup \{\infty_M\}$ eventually, but since every x_n is an element of M and hence is not equal to ∞_M, the sequence leaves every compact set eventually, meaning that it converges to infinity in norm in $\mathbb{R}^N$. □

This means that if we take $x : [0,a) \to M$ to be a continuous path in M where $a > 0$ is a real number such that $|x(t)| \to \infty$ as $t \to \infty$, then the extended map $x : \mathbb{R}_+ \to \hat{M}$ that agrees with x on $[0,a)$ and verifies $x(t) = \infty_M$ for $t > a$ is a continuous map. This motivates the following definition:

Definition 5.11 ([8], definition 1.1.4)**.** *Let M a locally compact metric space. An M valued path x with explosion time $e(x) > 0$ is a continuous map $x : \mathbb{R}_+ \to M$ such that $x(t) \in M$ for all $t \in [0, e(x))$ and $x(t) = \infty_M$ for $t > e(x)$. The space of M valued paths with explosion time is called the path space of M and is denoted by $W(M)$.*

Recall that if $(\Omega, (\mathcal{F}_t))$ is a filtered measurable space, a stopping time with respect to the filtration $(\mathcal{F}_t)$ is a random variable $\tau : \Omega \to [0,\infty]$ such that $(\tau \leq t) \in \mathcal{F}_t$ for every $t \geq 0$.

Now, M is a locally compact metric space. Over $W(M)$, we consider the filtration (B_t) where B_t is generated by the coordinate maps up to t for every $t \geq 0$. Then $(W(M), (B_t))$ is a filtered measurable space and the map

$$e : x \in W(M) \mapsto e(x)$$

is a stopping time with respect to the filtration (B_t) by construction.

Solutions to our SDEs will now take the form of semimartingales defined up to some stopping time.

Definition 5.12 ([8], definition 1.1.6)**.** *Let $(\Omega, (\mathcal{F}_t), \mathcal{P})$ be a filtered probability space and τ a stopping time with respect to $(\mathcal{F}_t)$. Let X be a continuous process defined on the stochastic interval $[0, \tau)$. The process X is said to be an $(\mathcal{F}_t)$-semimartingale up to τ if there exists of non-decreasing $(\mathcal{F}_t)$-stopping times (τ_n) such that $\tau_n \to \tau$ and such that the process defined by $X^n_t = X_{t \wedge \tau_n}$ is a semimartingale in the usual sense, for every n.*

Example 5.2. If (B_t) is an $(\mathcal{F}_t)$ brownian motion, then it is an $(\mathcal{F}_t)$-semimartingale up to ∞. Taking $\tau_n = n$, the stopped processes (B^n_t) are semimartingales as continuous local martingales, because (B_t) is a martingale with continuous sample paths.

We will now consider the following stochastic differential equation:

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t \tag{E}$$

Definition 5.13 ([8], definition 1.1.7)**.** *A semimartingale X up to a stopping time τ is a solution of equation (E) if there is a sequence of stopping times $\tau_n \uparrow \tau$ such that for each integer n, the process $X^n_t = X_{t \wedge \tau_n}$ is a semimartingale and*

$$X^n_t = X_0 + \int_0^{t \wedge \tau_n} \sigma(X_s)\,\mathrm{d}B_s$$

In this setting, we are able to weaken the conditions on the coefficient functions in order to have existence and uniqueness of solutions, but up to some explosion time.

Theorem 5.10 ([8], theorem 1.1.8)**.** *Let $(\Omega, (\mathcal{F}_t), \mathcal{P})$ be a filtered probability space. Suppose that the coefficient function σ in (E) is locally Lipschitz and let X_0 be an $\mathcal{F}_0$-measurable random variable. Then there exists a unique $W(\mathbb{R})$-valued random variable solution to equation (E) up to its explosion time $\mathrm{e}(X)$ with initial condition X_0.*

References

[1] Assyr Abdulle, Gilles Vilmart, and Konstantinos C. Zygalakis. “Long Time Accuracy of Lie–Trotter Splitting Methods for Langevin Dynamics”. In: *SIAM Journal on Numerical Analysis* 53.1 (2015). DOI: 10.1137/140962644. eprint: https://doi.org/10.1137/140962644. URL: https://doi.org/10.1137/140962644.

[2] Anthony M. Bloch et al. “The Euler-Poincaré equations and double bracket dissipation”. In: *Communications in Mathematical Physics* 175.1 (1996).

[3] Goffredo Chirco and Isha Kotecha. “Generalized Gibbs Ensembles in Discrete Quantum Gravity”. In: *Geometric Science of Information: 4th International Conference, GSI 2019, Toulouse, France, August 27–29, 2019, Proceedings*. Toulouse, France: Springer-Verlag, 2019, pp. 638–646. ISBN: 978-3-030-26979-1. DOI: 10.1007/978-3-030-26980-7_66. URL: https://doi.org/10.1007/978-3-030-26980-7_66.

[4] Kenneth David Elworthy. “Stochastic differential equations on manifolds”. In: *Probability Towards 2000*. Ed. by Luigi Accardi and Christopher C. Heyde. Vol. 128. Lecture Notes in Statistics. New York, NY: Springer, 1998.

[5] Jean Gallier and Jocelyn Quaintance. *Differential Geometry and Lie Groups: A Computational Perspective*. Vol. 12. Geometry and Computing. Cham: Springer, 2020. ISBN: 978-3-030-46039-6.

[6] Martin Hairer. *Ergodic theory for Stochastic PDEs*. LMS-EPSRC Short Course on Stochastic Partial Differential Equations, Imperial College London. July 2008.

[7] Martin Hairer. “On Malliavin’s proof of Hörmander’s theorem”. In: *Bulletin des sciences mathématiques* 135.6-7 (2011).

[8] Elton P. Hsu. *Stochastic Analysis on Manifolds*. Vol. 38. Graduate Studies in Mathematics. Providence, RI: American Mathematical Soc., 2002. ISBN: 978-0821808023.

[9] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*. 2nd. Vol. 24. North-Holland Mathematical Library. Amsterdam: Elsevier, 2014.

[10] Jean-François Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Cham: Springer, 2016. ISBN: 3319310887.

[11] Erwin Luesink and Oliver D. Street. *Symplectic techniques for stochastic differential equations on reductive Lie groups with applications to Langevin diffusions*. 2025. arXiv: 2504.02707 [math.PR]. URL: https://arxiv.org/abs/2504.02707.