Strong Interpolation Between Brownian Motion and the Geodesic Flow

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Mathematics

By

Heng Guo

EVANSTON, ILLINOIS

June 2019
Abstract

Strong Interpolation Between Brownian Motion and the Geodesic Flow

Heng Guo

This dissertation concerns the probabilistic aspects of diffusion processes generated by a family of differential operators, which is similar to the family of hypoelliptic Laplacian operators, acting on the tangent bundle of a compact Riemannian manifold. By lifting the processes to the product of the frame bundle and the euclidean space, we show that this family of stochastic processes interpolates between Riemannian Brownian motion and the geodesic flow.
Acknowledgements

There are many people to whom I owe my gratitude.

First of all, my family. I want to thank my parents and my wife for their constant support of my everyday life. They have always been standing by me through my ups and downs. Without their sacrifice and tolerance, I wouldn’t be able to pursue my career in mathematics or anything in general.

Secondly, my advisor. I owe my greatest gratitude to my advisor, Elton Hsu. Elton is a wonderful mentor in mathematics. Before I had met Elton, my understanding of probability theory was superficial. It was Elton who shared his insightful thoughts and ideas in his fields of interest and taught me how to think as a probabilist and eventually become a probabilist. As an extremely patient person, Elton guided me through every step of my academic life. He was never reluctant in offering advice for improvement as well as explaining technical details to me. On top of that, he has shown me how to produce a general result merely from heuristic intuition and sheer interest in understanding a phenomenon. Aside from his mathematical teachings, Elton is a person with life wisdom. He is more like a friend to me who set an example and taught me how to achieve a balance between life and mathematics.

Thirdly, Yimin Xiao, Antonio “Tuca” Auffinger and Steve Zelditch. Yimin is an expert on Gaussian fields. With his help on properties of Gaussian processes, the result
of this dissertation improved greatly. Tuca has seemingly infinite knowledge of probability theory. His teaching is masterful and I have benefited greatly from his classes. He also provided great explanations about core concepts in my research project, which serve as guidance as I tackle the problems in my thesis. Steve really serves as a role model for me as a mathematician. His relentless desire for knowledge is contagious. He has been constantly sharing his knowledge, suggesting interesting problems to not only me, but everyone in this department. He kept offering different points of view on probability theory that I find very inspiring.

Finally, I want to thank all of the students and faculty members in my department. They’ve been extremely friendly and helpful to me. It was a great privilege for me to be able to discuss math or anything in general with them over the past six years. Special thanks to Xavier Garcia, with whom I began my journey into the probability world. Zhenan Wang, a former student of Elton’s, who offered me infinite help at the beginning of my Ph.D. study. Robert Chang and Guchuan Li, two close friends of mine, who provided help beyond research.
# Table of Contents

Abstract 3

Acknowledgements 4

Table of Contents 6

Chapter 1. Introduction 8
  1.1. History and motivation 10
  1.2. Statement of the main result 11

Chapter 2. Background Materials 14
  2.1. Geometric preliminaries 15
  2.2. Geodesic flow 23
  2.3. Riemannian Brownian motion 25
  2.4. Wong-Zakai’s theorem 35

Chapter 3. The Euclidean Case 39
  3.1. Bismut’s hypoelliptic diffusions 39
  3.2. Convergence in euclidean case 42
  3.3. Convergence of the attendant Gaussian motion 48
  3.4. From uniform motion to Brownian motion 53
  3.5. Ito’s formula revisited 56
CHAPTER 1

Introduction

The connection of geodesics with Brownian motion arises in many pure and applied sciences. Brownian motion, first observed by Robert Brown in 1827 and later made mathematically rigorous by Albert Einstein and Marian Smoluchowvski in 1905, is a stochastic process modeling the random motion of a particle in a fluid subjected to random collisions with fluid molecules. The geodesic flow is a fundamental subject in Hamiltonian mechanics that describes the classical trajectory of a free particle in phase space. These two mathematical objects can be put under an overarching framework and studied simultaneously by careful analysis of a family of stochastic differential operators that is similar to the family of hypoelliptic Laplacian operators.

The approximation of Brownian motion is not a new subject. The classical result of Donsker [3] states that if we connect a symmetric simple random walk by line segments and rescaled time and space properly, the resulting path converges to Brownian motion in the weak sense. On a compact Riemannian manifold, Pinsky [8] showed that if we walk repeatedly along the geodesic in a uniformly chosen random direction for an exponentially distributed duration of time, then the resulting path-wise geodesic path converges weakly to a Brownian motion on the manifold. More recently, Gordina [4] used a similar scheme to construct a class of hypoelliptic diffusion processes on a smooth manifold. These constructions of Brownian motion have two aspects that could be improved. The first aspect is that the approximating processes are not smooth
enough. It would be nice if the approximating processes could be at least continuously differentiable so that they can be treated by the classical calculus. More importantly, the convergence is in the weak sense. It would be desirable if the convergence is in some stronger sense (for example, see the proof of classical Ito’s formula (3.5) from strong $C^1$ approximation of Brownian motion or at least in probability) so that stochastic calculus based on Brownian motion could be regarded as a limiting form of classical calculus. Based on these concerns, we are led to the following question: Is it possible to find a natural path that is smooth enough and could approximate Brownian motion on a compact Riemannian manifold? One of the purposes of this dissertation is to answer this question and do so in an elegant way.

The organization of this dissertation is as follows. In chapter two, we will review some background materials to the extent that is necessary for the rest of this dissertation. We will talk about the geometric settings that are essential to define semimartingale, which is an invariant class of stochastic processes under smooth operations, on any Riemannian manifold. We shall also review necessary probability theory and stochastic analysis, especially with connection to Riemannian geometry. In particular, we discuss Riemannian Brownian motion and its construction on a Riemannian manifold. Eventually, we will mention without proving the Wong-Zakai theorem in its simplest form, as it provides key insight into my research. In chapter three, we will mostly focus on the euclidean case. We discuss Bismut’s family of diffusion operators, namely hypoelliptic Laplacian, and discuss the interpolation property of this family of stochastic diffusion processes. We will also use this simple case of one-dimensional Brownian
motion to explain the required modification of this family of differential operators introduced in [2] to make it more natural in stochastic analysis. In the last chapter, we define our family of diffusion processes as the solutions of geometrically invariant stochastic differential equations and show that the family of diffusion processes indeed connects the geodesic flow to Riemannian Brownian motion in the strong sense. Connection with Ito’s formula will also be discussed as one application of the result.

1.1. History and motivation

Before we dive into any details, I would like to introduce some of the history of the problem I was working on, as well as explain why the main result should be expected. My research is inspired by the recent work [2] by Jean-Michel Bismut. In [2], Bismut provided a detailed explanation on the connection between hypoelliptic Laplacian operators and the standard elliptic Laplacian. Based on that, he proved the projections of the stochastic diffusion processes generated by the family of hypoelliptic Laplacian operators onto the manifold converge to a Riemannian Brownian motion in the weak sense. In addition, the associated driven Gaussian processes converge to independent Gaussian random variables. His method of study was the heat semigroup method, which relies on the characteristic function, or Fourier transform, of a given process at some fixed time \( t \). One theorem in [2] states for some fixed times \( t_i \)'s, the distributions of the stochastic diffusion processes generated by the hypoelliptic Laplacian family converge to the law of a Riemannian Brownian motion restricted at times \( t_i \)'s. In his proof, the approximating processes are smooth and this family of processes interpolates between a standard Brownian motion and a scaled geodesic flow. For
characteristic functions could only serve the weak convergence purpose, the mode of convergence is in the weak sense. Nevertheless, his work provides sufficient insight about the problem and we are convinced that, with some reasonable modifications, it would be possible to construct a family of stochastic diffusion processes that interpolates between Brownian motion and a genuine geodesic flow. Moreover, we would be able to derive a strong approximation of Riemannian Brownian motion on a proper geometric setting.

Our work starts from the basic euclidean case, where we wish to approximate the standard euclidean Brownian motion using a smooth Gaussian process, i.e., a Gaussian process with continuously differentiable sample paths. This step could serve as a fundamental building block of our research because, once we accomplish this, we wouldn’t be too far from the general result thanks to the theory developed by Eugene Wong and Moshe Zakai in [9]. According to the Wong-Zakai philosophy, any strong approximation of Brownian motion by continuously differentiable processes will eventually lead towards the Stratonovich stochastic integral, which is exactly how Riemannian Brownian motion is defined on a Riemannian manifold.

1.2. Statement of the main result

Let $M$ be a compact Riemannian manifold of dimension $n$, $\mathcal{O}(M)$ be the orthonormal frame bundle over $M$, $\mathbb{R}^n$ be the trivial bundle over $M$ and $\mathcal{O}(M) \times \mathbb{R}^n$ be the product bundle. Consider the operator $L_\epsilon$ on the product bundle defined as

$$L_\epsilon = \frac{\Delta V}{2\epsilon^2} - \frac{Y}{\epsilon} + U,$$  

(1.2.1)
where $\Delta^Y$ is the Laplacian operator, $Y$ is the radial vector field, both defined on the trivial bundle $\mathbb{R}^n$ over $\mathcal{O}(M) \times \mathbb{R}^n$, and $U$ is the lifting of the vector field on $TM$ that generates the geodesic flow to the product bundle.

Pick local coordinate system $(x, e, y)$ on the product bundle, $L_\epsilon$ could be expressed as

$$L_\epsilon = \frac{1}{2\epsilon^2} \sum_{i=1}^{n} \frac{\partial}{\partial y^i} \frac{\partial}{\partial y^i} - \frac{1}{\epsilon} \sum_{i=1}^{n} y^i \frac{\partial}{\partial y^i} + \sum_{i=1}^{n} y^i H_i,$$

where $H_i$ is the lifting of the fundamental horizontal vector field over $\mathcal{O}(M)$ to the product bundle. This family of operators is the modified version of Bismut’s family of hypoelliptic Laplacian operators. Consider the family of stochastic diffusion processes $Z_t^\epsilon$ generated by $L_\epsilon$, we have

**Theorem 1.2.1 (main result).** Let $X_t^\epsilon$ be the projections of the processes $Z_t^\epsilon$ onto the manifold $M$. As $\epsilon \to 0$, this family of processes $X_t^\epsilon$ converges to a Riemannian Brownian Motion almost surely. As $\epsilon \to \infty$, the family $X_t^\epsilon$ converges to a geodesic on $M$.

What the theorem says is that the projections of the processes $Z_t^\epsilon$ onto the manifold $M$ interpolate between the geodesic and Riemannian Brownian motion. The proof of this theorem has two key ingredients. Firstly, we follow Hsu [6] and lift the processes $Z_t^\epsilon$ generated by $L_\epsilon$ properly to the product bundle $\mathcal{O}(M) \times \mathbb{R}^n$, where $\mathcal{O}(M)$ is the orthonormal frame bundle over $M$. This is primarily because we need a space large enough that both Brownian motion and geodesic flow could live there naturally. In addition, from Bismut’s work, the euclidean component is also needed to accommodate the Gaussian field, a byproduct of the construction. The advantage of doing this is that the restriction of the lifted process to $\mathbb{R}^n$, as a subspace of $\mathcal{O}(M) \times \mathbb{R}^n$, can be
solved independently. This allows us to get an explicit expression for the process $Z_t^\varepsilon$. After that, we follow the method used by Wong and Zakai in [9] and prove the result by showing that $\pi Z_t^\varepsilon$ satisfies the same stochastic differential equation as a Riemannian Brownian Motion. Although Brownian Motion itself has significant meanings in probability theory and has been studied extensively, the proof of the theorem provides methods that could potentially lead to deeper understandings of Riemannian Brownian motion. More importantly, it connects ordinary integrations to stochastic integrations in an intuitive way.
CHAPTER 2

Background Materials

In this chapter, we review some aspects of probability theory and Riemannian geometry that are related to my research. Riemannian geometry is the study of spaces, called manifolds, that are endowed with a Riemannian metric. Using this metric, one can make measurements about distances between points, lengths of curves, and areas/volumes of domains. In particular, the metric gives rise to the notion of a geodesic, namely the shortest path between two points. For example, geodesics on euclidean spaces are straight lines, while those on a round sphere are great circles. Under the framework of Hamiltonian mechanics, a free particle on a manifold travels along geodesics. Its equation of motion is a second-order ordinary differential equation, called the geodesic flow. Meanwhile, the concept of (local) martingale on a euclidean space can be extended to a differential manifold equipped with a connection. This relies on the concept that is called anti-development. For a smooth manifold equipped with a connection, stochastic anti-development gives a one-to-one correspondence between semimartingales on the manifold and those on a euclidean space. The procedure of passing from euclidean space to manifold is referred to as “rolling without slipping”. In the end, We state the Wong-Zakai theorem in its simplest form and explain its significance in our research.
2.1. Geometric preliminaries

In this section, we introduce what is necessary to talk about stochastic processes on any manifold. We introduce horizontal lift and stochastic development of a manifold-valued semimartingale, two concepts central to the construction of Brownian motion on a Riemannian manifold. In differential geometry, it is possible to lift a smooth curve on \( M \) to a horizontal curve on the frame bundle \( \mathcal{F}(M) \) by solving an ordinary differential equation, and this horizontal curve corresponds uniquely to a smooth curve in the euclidean space of the same dimension. Up to an action by the general linear group, there is a one-to-one correspondence between the set of smooth curves on the manifold and their anti-development in the euclidean space. Unsurprisingly, we show that an analogous construction can be carried over for semimartingales on a manifold equipped with a connection. This construction is realized by solving stochastic differential equations on manifolds.

2.1.1. Frame bundle and connection

Let \( M \) be a differentiable manifold of dimension \( n \). The tangent space at \( x \) is denoted by \( T_x M \) and the tangent bundle is denoted by \( TM \). The space \( \Gamma(TM) \) of smooth sections of the tangent bundle is just the set of vector fields on \( M \). A connection on \( M \) is a convention of differentiating a vector field along another vector field. \( \nabla_X Y \) is called covariant differentiation of \( Y \) along \( X \). In local coordinations, a connection is expressed in terms of its Christoffel symbols. Let \( x = \{x^i\} \) be a local chart on an open subset \( O \) of \( M \). Then the vector fields \( X_i = \frac{\partial}{\partial x^i} \) span the tangent space \( T_x M \) at each point \( x \in O \),
and the Christoffel symbols $\Gamma^k_{ij}$ are functions on $O$ defined uniquely by the relation

$$\nabla_{X_i}X_j = \Gamma^k_{ij}X_k.$$ 

Suppose $M$ is a manifold equipped with a connection. A vector field $V$ along a curve $\{x_t\}$ on $M$ is said to be parallel along the curve if $\nabla_{\dot{x}}V = 0$ at every point of the curve. In local coordinates, if $x_t = \{x^i_t\}$ and $V_{x_t} = v^i(t)X_i$, then $V$ is parallel along the curve $\{x_t\}$ if and only if its components $v^i(t)$ satisfy the system of first order differential equations

$$
(2.1.1) \quad \dot{v}^k(t) + \Gamma^k_{jl}\dot{x}^j_t v^l(t) = 0.
$$

Hence locally a parallel vector field $V$ along a curve $\{x_t\}$ is uniquely determined by its initial value $V_{x_0}$.

Let us now see how the connection $\nabla$ manifests itself on the frame bundle $\mathcal{F}(M)$ of $M$. A frame at $x$ is an $\mathbb{R}$-linear isomorphism $u : \mathbb{R}^d \to T_x M$. Let $e_i$ be the $i$-th coordinate unit vectors of $\mathbb{R}^d$. Then the tangent vectors $ue_i$ make up a basis for the tangent space $T_x M$. We use $\mathcal{F}(M)_x$ to denote the space of all frames at $x$. The general linear group $GL(d, \mathbb{R})$ acts on $\mathcal{F}(M)_x$ by $u \mapsto ug$, where $ug$ denotes the composition

$$
\mathbb{R}^d \overset{g}{\to} \mathbb{R}^d \overset{u}{\to} T_x M.
$$

The frame bundle

$$
\mathcal{F}(M) = \bigcup_{x \in M} \mathcal{F}(M)_x
$$
can be made into a differentiable manifold of dimension \( d + d^2 \), and the canonical projection \( \pi: \mathcal{F}(M) \to M \) is a smooth map. The group \( GL(d, \mathbb{R}) \) acts on \( \mathcal{F}(M) \) fibre-wise; each fibre \( \mathcal{F}(M)_x \) is diffeomorphic to \( GL(d, \mathbb{R}) \), and \( M = \mathcal{F}(M) / GL(d, \mathbb{R}) \). In differential geometry terminology, these facts make \( (\mathcal{F}(M), M, GL(d, \mathbb{R})) \) into a principal bundle with the structure group \( GL(d, \mathbb{R}) \). With the standard action of \( GL(d, \mathbb{R}) \) on \( \mathbb{R}^d \), the tangent bundle is simply the associated bundle

\[
\mathcal{F}(M) \times_{GL(d, \mathbb{R})} \mathbb{R}^d = TM,
\]

given by

\[
(u, e) \mapsto ue.
\]

The tangent space \( T_u\mathcal{F}(M) \) of the frame bundle is a vector space of dimension \( d + d^2 \). A tangent vector \( X \in T_u\mathcal{F}(M) \) is called vertical if it is tangent to the fibre \( \mathcal{F}(M)_{\pi u} \).

Now assume that \( M \) is equipped with a connection \( \nabla \). A curve \( \{u_t\} \) in \( \mathcal{F}(M) \) is just a smooth choice of frames at each point of the curve \( \pi u_t \) on \( M \). The curve \( \{u_t\} \) is called horizontal if, for each \( e \in \mathbb{R}^d \) the vector field \( u_t e \) is parallel along \( \pi u_t \). A tangent vector \( X \in T_u\mathcal{F}(M) \) is called horizontal if it is the tangent vector of a horizontal curve from \( u \). The space of horizontal vectors at \( u \) is denoted by \( H_u\mathcal{F}(M) \).

It follows that the canonical projection \( \pi: \mathcal{F}(M) \to M \) induces an isomorphism \( \pi_*: H_u\mathcal{F}(M) \to T_{\pi u}M \), and for each \( X \in T_xM \) and a frame \( u \) at \( x \), there is a unique horizontal vector \( X^* \), the horizontal lift of \( X \) to \( u \), such that \( \pi_* X^* = X \).
For each $e \in \mathbb{R}^d$, the vector field $H_e$ on $F(M)$ defined at $u \in F(M)$ by the relation

$$H_e(u) = (ue)^*$$

is a horizontal vector field on $F(M)$. Let $\{e_i\}$ be the coordinate unit vectors of $\mathbb{R}^d$. Then $H_i = H_{e_i}$ are the fundamental horizontal vector fields of $F(M)$.

A local chart $x = \{x^i\}$ on a neighborhood $O \in M$ induces a local chart on $\tilde{O} = \pi^{-1}(O)$ in $F(M)$ as follows. Let $X_i = \partial / \partial x_i$ be the moving frame defined by the local chart. For a frame $u \in \tilde{O}$ we have $ue_i = e_j^i X_j$ for some matrix $e = (e_j^i) \in GL(d, \mathbb{R})$. Then $(x, e) = (x^i, e^i_j) \in \mathbb{R}^{d+2}$ is a local chart in $\tilde{O}$. We will need the local expression for the fundamental horizontal vector field $H_i$.

In terms of the local chart on $F(M)$ described above, at $u = (x, e) \in F(M)$ we have

$$H_i(u) = e_j^i X_j - e_j^l e_l^m \Gamma_{jl}^k(x) X_{km},$$

(2.1.2)

where

$$X_i = \frac{\partial}{\partial x^i}, \text{ and } X_{km} = \frac{\partial}{\partial e^k_m}.$$  

If $M$ is a Riemannian manifold and $g$ its Riemannian metric, then we can restrict ourselves to a smaller set of frames, namely the orthonormal frames. Let $O(M)$ be the orthonormal frame bundle. By definition, an element in $O(M)$ is a euclidean isometry $u: \mathbb{R}^d \to T_xM$. The action group is correspondingly reduced from $GL(d, \mathbb{R})$ to the orthogonal group $O(d)$, and $O(M)$ is a principal fibre bundle with the structure group $O(d)$. 
Given a connection $\nabla: \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$, the associated parallel transport need not preserve the orthogonality of $u \in \mathcal{O}(M)$. If it does, the connection is said to be compatible with the Riemannian metric. This happens if and only if for every triple of vector fields $X, Y, Z$ on $M$

$$\nabla_X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle X, \nabla_Y Z \rangle.$$ 

If a Riemannian manifold is equipped with such a connection, everything we have said so far in this section about the general linear frame bundle carries over to the orthonormal frame bundle. In particular, the formula for the fundamental horizontal vector fields in (2.1.2) is still valid with the caveat that now $\{x^i, e^j\}$ is a set of local coordinates for $\mathcal{O}(M)$.

2.1.2. Horizontal lift and stochastic development

Stochastic differential equations on a manifold are a convenient and useful tool for generating semimartingales on a manifold $M$ from ones on $\mathbb{R}^N$. If a manifold $M$ is equipped with a connection, then there are invariantly defined fundamental horizontal vector fields $H_i$ on the frame bundle $\mathcal{F}(M)$, and many things we have said about smooth curves on $M$ can be generalized to semimartingales on $M$. In particular, we can lift a semimartingale $X$ on $M$ to a horizontal semimartingale $U$ on $\mathcal{F}(M)$ once an initial frame $U_0$ at $X_0$ is fixed, and then to a semimartingale $W$ on $\mathbb{R}^d$. Once a horizontal lift $U_0$ of the initial value $X_0$ is fixed (i.e., $\pi U_0 = X_0$), the correspondence $W \leftrightarrow X$ is one-to-one. Because euclidean semimartingales are easier to handle than manifold-valued semimartingales, we can use this geometrically defined correspondence to our
advantage. Later we will see that a connection also gives rise to the notion of manifold-valued martingales. For semimartingales, stochastic development and horizontal lift are obtained by solving stochastic differential equations driven by either $\mathbb{R}^d$-valued or $M$-valued semimartingales as expected.

Consider the following SDE on the frame bundle $\mathcal{F}(M)$:

\begin{equation}
    dU_t = H(U_t) \circ dW_t,
\end{equation}

where $W$ is an $\mathbb{R}^d$-valued semimartingale and the integral is in the Stratonovich sense. Whenever necessary, we will use the more precise notation

\[ dU_t = H_i(U_t) \circ dW^i_t. \]

In writing the above equation, we have implicitly assumed that $M$ has been equipped with a connection, and $H_i$ are the corresponding fundamental horizontal vector fields on $\mathcal{F}(M)$. We now give a few definitions. Keep in mind that all processes are defined on a fixed filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and are $\mathcal{F}_*$-adapted.

**Definition 2.1.1.** (i) An $\mathcal{F}(M)$-valued semimartingale $U$ is said to be horizontal if there exists an $\mathbb{R}^d$-valued semimartingale $W$ such that (2.1.3) holds. The unique $W$ is called the anti-development of $U$.

(ii) Let $W$ be an $\mathbb{R}^d$-valued semimartingale and $U_0$ an $\mathcal{F}(M)$-valued, $\mathcal{F}_0$-measurable random variable. The solution $U$ of the SDE in (2.1.3) is called a (stochastic) development $W$ in $\mathcal{F}(M)$. Its projection $X = \pi U$ is called a (stochastic) development of $W$ in $M$. 
(iii) Let $X$ be an $M$-valued semimartingale. An $\mathcal{F}(M)$-valued horizontal semimartingale $U$ such that its projection $\pi U = X$ is called a (stochastic) horizontal lift of $X$.

In the correspondence $W \leftrightarrow U \leftrightarrow X$, the only transitions which are non-trivial are $X \mapsto U$ and $U \mapsto W$. We will prove the existence of a horizontal lift by deriving stochastic differential equation for it on the frame bundle $\mathcal{F}(M)$ driven by $X$. For this purpose, we assume $M$ is a closed submanifold of $\mathbb{R}^N$ and regard $X = \{X^a\}$ as an $\mathbb{R}^N$-valued semimartingale. For each $x \in M$, let $P(x): \mathbb{R}^N \mapsto T_xM$ be the orthogonal projection from $\mathbb{R}^N$ onto the subspace $T_xM \subset \mathbb{R}^N$. Then intuitively we have, on $\mathbb{R}^N$,

$$X_t = X_0 + \int_0^t P(X_s) \circ dX_s.$$

Rewriting this more explicitly, we have

$$dX_t = \sum_{a=1}^N P_a(X_s) \circ dX_t^a.$$

Once we are convinced that this identity holds, the obvious candidate for the horizontal lift $U$ of $X$ is the solution of the following equation on $\mathcal{F}(M)$:

$$dU_t = \sum_{a=1}^N P^*_a(U_t) \circ dX_t^a,$$

where $P^*_a(u)$ is the horizontal lift of $P_a(\pi u)$. For this purpose, we need the following lemma.

**Lemma 2.1.2.** Suppose that $M$ is a closed submanifold of $\mathbb{R}^N$. For each $x \in M$, let $P(x)$ be the orthogonal projection from $\mathbb{R}^N$ to the tangent space $T_xM$ as described above. If $X$ is an
If $X_t$ is an $M$-valued semimartingale, then

\begin{equation}
X_t = X_0 + \int_0^t P(X_s) \circ dX_s.
\end{equation}

**Proof.** Let $\xi_\alpha$ be the canonical basis for $\mathbb{R}^N$. Define

\[ P_\alpha(x) = P(x)\xi_\alpha, \quad Q_\alpha(x) = \xi_\alpha - P_\alpha(x). \]

Then $P_\alpha$ is tangent to $M$ and $Q_\alpha$ is normal to $M$ with the property $P_\alpha + Q_\alpha = \xi_\alpha$. Let

\[ Y_t = X_0 + \int_0^t P_\alpha(X_s) \circ dX_s^\alpha. \]

We need to verify two things. First, we need $Y_t$ to live on $M$. Let $f$ be a smooth nonnegative function on $\mathbb{R}^N$ which vanishes only on $M$. By Ito’s formula,

\[ f(Y_t) = f(X_0) + \int_0^t P_\alpha f(X_t) \circ dX_t^\alpha. \]

But if $x \in M$, then $P_\alpha \in T_xM$ and $P_\alpha f(x) = 0$. Hence $P_\alpha f(X_t) = 0$ and $f(Y_t) = 0$, which proves $Y_t \in M$.

For each $x \in \mathbb{R}^N$, let $h(x)$ be the point on $M$ that is closest to $x$. Since $M$ is a closed submanifold, $h$ is well-defined and is constant on each line segment perpendicular to $M$. This means that $Q_\alpha h(x) = 0$ for $x \in M$. As a consequence, we have

\[ P_\alpha h(x) = P_\alpha h(x) + Q_\alpha h(x) = \xi_\alpha h(x). \]
Now we have
\[
Y_t = h(Y_t) = X_0 + \int_0^t P_{h} h(X_s) \circ dX^s_t = X_0 + \int_0^t \xi_{h} h(X_s) \circ dX^s_t = h(X_t) = X_t,
\]
which completes the proof. 

\[\square\]

2.2. Geodesic flow

In this section, we review some basic facts about the geodesic flow. To be more specific, we talk about the vector field on the tangent bundle $TM$ over a Riemannian manifold $M$ that generates the geodesic flow. As will be seen, regarded as a differential operator, this vector field is one of the key components of the differential operator $L_\epsilon$ defined in (4.3.1) and it shouldn’t be a surprise since our family of diffusion processes converges to the geodesic flow as $\epsilon$ goes to infinity.

We start with geodesics on a Riemannian manifold. A curve $\{x_t\}$ in $M$ is called a geodesic if

\[(2.2.1) \quad \nabla_{\dot{x}_{t}} \dot{x}_{t} = 0\]

along $x_t$, i.e., if the tangent vector field itself is parallel along the curve. Let $x = \{x^i\}$ be a local chart on an open subset $O \subset M$, $X_i = \partial / \partial x^i$ be the basis of the tangent space $T_x M$ at $x$, and write the curve $\{x_t\}$ in local coordinates as $\dot{x}_t = \dot{x}_t^i X_i$. Using the geodesic equation (2.2.1), we obtain the ordinary differential equation for geodesics, which is

\[(2.2.2) \quad \ddot{x}_t^k + \Gamma^k_{ij}(x_t) \dot{x}_t^i \dot{x}_t^j = 0.\]
One example of geodesics could be, consider a sphere centered at the origin in $\mathbb{R}^3$, the
great circles on the sphere.

Now let’s imagine we are standing on the north pole of this sphere, and we want to
travel to the south pole along a great circle. We immediately see a problem: there are
infinitely many great circles passing through both north and south poles. The obvious
solution to this problem is to specify a direction. Thus a geodesic is uniquely deter-
mined by its initial position $x_0$ as well as an initial direction $\dot{x}_0$. This gives rise to the
geodesic flow, as it is defined on the tangent bundle $TM$ over a manifold.

Let $TM$ be the tangent bundle over $M$ of dimension $d$ and treat it as a differentiable
manifold of dimension $2d$. Assume the vector field on the tangent bundle $TM$ that
generates the geodesic flow is $U$, it is often useful to get an expression for $U$ in terms
of the local coordinate system $\{x^i, y^i\}$ on an open subset of $TM$.

**Proposition 2.2.1.** Let $\{x, y\} = \{x^i, y^i\}$ be a local chart on an open subset of $TM$ given
by the relation

$$
\mathbb{R}^{2d} \to TM, \quad (x^i, y^i) \mapsto (x^i, y^i \frac{\partial}{\partial x^i}),
$$

then the vector fields $\partial/\partial x^i$ and $\partial/\partial y^i$, $i = 1, 2, ..., d$, span the tangent space $T_{(x,y)}(TM)$ at
each point $(x, y) \in TM$. The generator of the geodesic flow can be expressed as

$$
U(x^i, y^i) = y^i \frac{\partial}{\partial x^i} - \Gamma^i_{jk}(x)y^j y^k \frac{\partial}{\partial y^i}
$$

**PROOF.** Since $\partial/\partial x^i$ and $\partial/\partial y^i$, $i = 1, 2, ..., d$ span the tangent space $T(TM)$, we
assume $U(x^i, y^i) = \alpha^i \frac{\partial}{\partial x^i} + \beta^i \frac{\partial}{\partial y^i}$ for some $\alpha^i$’s and $\beta^i$’s. Our goal is to prove $\alpha^i = y^i$ and
$\beta^i = -\Gamma^i_{jk} y^j y^k$. 

Let $x_t$ be the geodesic that starts at point $(x, y)$ with $\dot{x}^i(0) = y^i$ such that it is the projection of the geodesic flow generated by $U$ onto the manifold $M$, by definition this means $\alpha^i = y^i$. $\beta^i$ by definition is $\dot{\beta}^i = \dot{x}^i(0) = \dot{y}^i(0)$. Using equation (2.2.2) we have,

$$\ddot{x}^i(0) + \Gamma^i_{jk} \dot{x}^j(0) \dot{x}^k(0) = 0,$$

or

$$\beta^i + \Gamma^i_{jk} y^j y^k = 0$$

which finishes that proof. □

Later on, we shall see how this operator fits in $L_\epsilon$ and plays a role in the interpolation of the corresponding diffusion processes.

2.3. Riemannian Brownian motion

In this section, we discuss Brownian motion and how to carry this stochastic diffusion process over to a Riemannian manifold using the geometric setting we have mentioned. As Brownian motion is the core subject in stochastic analysis and probability theory, we would like to cover as much properties as possible and achieve this goal gradually.

2.3.1. Martingale theory

Stochastic analysis is the study of stochastic, or random, processes using probabilistic techniques. The most important stochastic process is Brownian motion because of its nice properties and the connections it has with other stochastic processes. Brownian motion has many equivalent characterizations, such as the continuous limit of simple
random walks, or a Gaussian process with certain covariance properties, or a continuous martingale which is Markovian and Gaussian. In fact, a large class of random processes can be treated as scaled version of Brownian Motion. In this dissertation, we are mostly concerned about the martingale property of Brownian motion, as it is the foundation of stochastic calculus.

To start with, we review basic facts about martingale theory. We start with discrete-time parameter martingales and extend the results from discrete-time to continuous time.

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space and \(\mathcal{G}\) a \(\sigma\)-algebra of measurable events contained in \(\mathcal{F}\). Suppose that \(X \in L^1(\Omega, \mathcal{F}, \mathbb{P})\) is an integrable random variable. There exists a unique random variable \(Y\) which have the follow two properties:

1. \(Y \in L^1(\Omega, \mathcal{G}, \mathbb{P})\) and is integrable.
2. For any \(C \in \mathcal{G}\), we have

\[ \mathbb{E}\{X; C\} = \mathbb{E}\{Y; C\}. \]

This random variable \(Y\) is called the conditional expectation of \(X\) with respect to \(\mathcal{G}\) and is denoted by \(\mathbb{E}\{X|\mathcal{G}\}\).

A sequence \(\mathcal{F}_* = \{\mathcal{F}_n, n \in \mathbb{Z}_+\}\) of increasing \(\sigma\)-algebra is called a filtration. A sequence of random variables \(X = \{X_n\}\) is said to be adapted to the filtration \(\mathcal{F}_*\) if \(X_n\) is measurable with respect to \(\mathcal{F}_n\) for all \(n\). A sequence of integrable random variables

\[ X = \{X_n, n \in \mathbb{Z}_+\} \]
on a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P})\) is called a martingale with respect to \(\mathcal{F}_n\) if \(X\) is adapted to \(\mathcal{F}_n\) and
\[
\mathbb{E}\{X_n | \mathcal{F}_{n-1}\} = X_n
\]
for all \(n\).

The definition of martingale extends in an obvious way to the case of continuous-time parameters. For example, a real-valued stochastic process \(X = \{X_t\}\) is a martingale with respect to a filtration \(\mathcal{F}_t = \{\mathcal{F}_t\}\) if \(X_t \in L^1(\Omega, \mathcal{G}, \mathbb{P})\) and
\[
\mathbb{E}\{X_t | \mathcal{F}_s\} = X_s
\]
for \(s \leq t\).

The quadratic variation is an important characterization of a continuous martingale. In fact, Levy’s criterion shows that Brownian motion \(W\) is a martingale completely characterized by its quadratic variation.

Let \(M = \{M_t\}\) be a continuous local martingale. Then \(M_t^2\) is a continuous local submartingale. By the Doob-Meyer decomposition theorem, there is a continuous increasing process, which we will denote by \(\langle M, M \rangle = \{\langle M, M \rangle_t\}\) or simply \(\langle M \rangle = \{\langle M \rangle_t\}\), with \(\langle M, M \rangle_0 = 0\), such that \(M_t^2 - \langle M, M \rangle\) is a continuous local martingale. This increasing process \(\langle M, M \rangle\) is uniquely determined by \(M\) and is called the quadratic variation process of the local martingale \(M\).
2.3.2. Brownian motion and Ito’s formula

In stochastic analysis, we deal with two important classes of stochastic processes: Markov processes and martingales. Brownian motion is the most important example for both classes, and is also the most thoroughly studied stochastic process.

**Definition 2.3.1.** A stochastic process \( W = \{W_t, t \in \mathbb{R}_+ \} \) defined on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) is called a Brownian motion if it has the following two properties:

(1) \( W \) has independent increments;

(2) for any \( t > s \geq 0 \), the distribution of the increments \( W_t - W_s \) is a Gaussian distribution with mean zero and variance \( t - s \).

Let \( X \) and \( Y \) be two stochastic processes on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\). We say that \( X \) is a version of \( Y \) if

\[
\mathbb{P}\{X_t = Y_t\} = 1
\]

for all \( t \geq 0 \). The following theorem is always helpful since we are usually interested in continuous stochastic processes.

**Theorem 2.3.2** (Wiener’s theorem). Every Brownian motion has a continuous version with continuous sample paths.

Sometimes continuity of sample path is a part of the definition of a Brownian motion to force the Brownian paths to be continuous as well.

Recall that \( W \) is a Brownian motion then its quadratic variation process \( \langle W, W \rangle_t = t \). The following theorem shows that this property characterizes Brownian motion completely.
Theorem 2.3.3 (Levy’s characterization of Brownian motion). Suppose that $W$ is a continuous local martingale with respect to a filtration $\mathcal{F}_t$ whose quadratic variation process is $\langle W, W \rangle_t$, then it is a Brownian motion with respect to $\mathcal{F}_t$.

The proof heavily relies on the most important formula in stochastic analysis, namely Itô’s formula. Itô’s formula is the fundamental theorem for stochastic calculus.

Theorem 2.3.4 (Itô’s formula). Suppose that $F \in C^2(\mathbb{R})$ and $N$ is a semimartingale. Then

\begin{equation}
F(N_t) - F(N_0) = \int_0^t F'(N_s) \, dN_s + \frac{1}{2} \int_0^t F''(N_s) \, d\langle N, N \rangle_s.
\end{equation}

In general, this formula gives us the Doob-Meyer decomposition of a submartingale, which roughly says that a submartingale $S$ can be uniquely written as the sum of a martingale $M$ and an increasing predictable process $P$. For a semimartingale, there is a unique decomposition as well. To be more precise,

Proposition 2.3.5. A continuous semimartingale $N$ can be uniquely decomposed into the sum $N = M + A$ of a continuous local martingale $M$ and a continuous process of bounded variation $A$ with $A_0 = 0$.

Itô’s formula plays an essential role in stochastic calculus and we use this formula extensively as well in this research.

2.3.3. Martingale on manifolds and Riemannian Brownian motion

The concept of (local) martingale on a euclidean space can be extended to a differentiable manifold equipped with a connection.
Definition 2.3.6. Suppose that $M$ is a differentiable manifold equipped with a connection $\nabla$. An $M$-valued semimartingale $X$ is called a $\nabla$-martingale if its anti-development $W$ with respect to the connection $\nabla$ is an $\mathbb{R}^d$-valued local martingale.

Next, we will focus on Brownian motion on a Riemannian manifold. It is defined as a diffusion process generated by the Laplace-Beltrami operator $\Delta_M/2$. Let $L$ be a smooth second order elliptic differential operator on a differentiable manifold. An $\mathcal{F}_s$-adapted stochastic process $X$ defined on a filtered probability space $(\Omega, \mathcal{F}_s, \mathbb{P})$ is called a diffusion process generated by $L$ if

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

is a local martingale for all function $f \in C^\infty(M)$.

Let $\mathcal{O}(M)$ be the orthonormal frame bundle of $M$ and $\pi: \mathcal{O}(M) \mapsto M$ the canonical projection. Recall that the fundamental horizontal vector fields $H_i$ are the unique horizontal vector fields on $\mathcal{O}(M)$ such that $\pi_* H_i(u) = e_i$, where $e_i$ is the canonical basis for $\mathbb{R}^d$. Bochner’s horizontal Laplacian is the second order elliptic operator on $\mathcal{O}(M)$ defined by

$$\Delta_{\mathcal{O}(M)} = \sum_{i=1}^d H_i^2.$$ 

Its relation to the Laplace-Beltrami operator is by the following theorem.

Theorem 2.3.7. Bochner’s horizontal Laplacian $\Delta_{\mathcal{O}(M)}$ is the lift of the Laplace-Beltrami operator $\Delta_M$ to the orthonormal frame bundle $\mathcal{O}(M)$. More precisely, let $f \in C^\infty(M)$, and
\( \tilde{f} = f \circ \pi \) its lift to \( \mathcal{O}(M) \). Then for any \( u \in \mathcal{O}(M) \),

\[
\Delta_M f(x) = \Delta_{\mathcal{O}(M)} \tilde{f}(u),
\]

where \( x = \pi u \).

**Proof.** We need to find the corresponding operations for grad and div in \( \mathcal{O}(M) \).

Recall that the scalarization of a 1-form \( \theta \) is defined as \( \tilde{\theta}(u) = \{ \theta(ue_i) \} \). Thus the scalarization of \( df \) is given by

\[
\tilde{d}f_i = df(ue_i) = (ue_i)f = H_i \tilde{f}(u),
\]

that is,

\[
\tilde{d}f = \tilde{\nabla} f.
\]

The scalarization \( \tilde{\nabla} f \) is given by the same vector:

\[
\tilde{\nabla} f = H_i \tilde{f}(u).
\]

On the another hand, if \( u \in \mathcal{O}(M) \) and \( \pi u = x \), then \( ue_i \) is an orthonormal basis for \( T_{\pi u} M \). Recall divergence of a vector field is

\[
\text{div} X = \sum_{i=1}^{d} \langle \nabla_{X_i} X, X_i \rangle
\]

and the scalarization of the covariant derivative \( \nabla_X \theta \) is given by

\[
\tilde{\nabla}_X \theta = X^* \theta,
\]
where $X^*$ is the horizontal lift of $X$. We have
\[
\text{div} X = \langle \nabla_{ue_i} X, ue_i \rangle = \langle u^{-1} \nabla_{ue_i} X, e_i \rangle = \langle H_i \tilde{X}(u), e_i \rangle,
\]
where $\tilde{X}$ is the scalarization of $X$. The above relation can be written equivalently as
\[
\text{div} X(x) = \sum_{i=1}^{d} (H_i \tilde{X})^i(u).
\]
It follows that
\[
\Delta_M f(\pi u) = \sum_{i=1}^{d} (H_i \nabla f)^i(u) = \sum_{i=1}^{d} H_i H_i \tilde{f}(u) = \Delta_{\mathcal{E}(M)} \tilde{f}(u).
\]
\[\square\]

From now on, we assume that $M$ is a Riemannian manifold equipped with the Levi-Civita connection $\nabla$, and $\Delta_M$ the Laplace-Beltrami operator on $M$. Let $W(M)$ be the path space over $M$, we can define Brownian motion on a given manifold $M$ as following.

**Definition 2.3.8.** Let $X: \Omega \to W(M)$ be a measurable map defined on a probability space $(\Omega, \mathcal{F}, P)$. Let $\mu = P \circ X_0^{-1}$ be its initial distribution. Then the following three definitions are equivalent.

(1) $X$ is a $\Delta_M/2$-diffusion process, i.e.
\[
f(X_t) - f(X_0) - \frac{1}{2} \int_0^t \Delta_M f(X_s) \, ds,
\]
is an $\mathcal{F}$-local martingale for all $f \in C^\infty(M)$. 

(2) The law $\mathbb{P}^X = \mathbb{P} \circ X^{-1}$ is a Wiener measure on the path space $W(M)$.

(3) $X$ is a $\mathbb{F}^X_*$-semimartingale on $M$ whose anti-development is a standard euclidean Brownian motion.

An $M$-valued process $X$ satisfying any of the above conditions is called a Riemannian Brownian motion on $M$.

It is often useful to have a stochastic differential equation of Riemannian Brownian motion in local coordinates. Recall the equation for a horizontal Brownian motion on $\mathcal{O}(M)$ is
\[ dU_t = H_i(U_t) \circ dW^i_t, \]
where $W$ is a $d$-dimensional euclidean Brownian motion. Pick local coordinate system $\{x^i, e^i_j\}$ on $\mathcal{O}(M)$ as described before, we have shown that the horizontal vector fields are given by

\[ H_i(u) = e^i_j X_j - e^i_k e^k_m \Gamma^j_{ij}(x) X_m, \quad (2.3.2) \]

where
\[ X_i = \frac{\partial}{\partial x^i} \quad \text{and} \quad X_{km} = \frac{\partial}{\partial e^k_m}. \]

Hence the equation for $U_t = \{X^i_t, e^i_j(t)\}$ is

\[ dX^i_t = e^i_j(t) \circ dW^j_t \quad \text{and} \quad de^i_j(t) = -\Gamma^i_{kj}(X_t)e^j_k(t) e^k_m(t) \circ dW^m_t, \quad (2.3.3) \]

Using (2.3.3), we can find an equation for the Riemannian Brownian motion $X$ itself. Consider the first equation in (2.3.3). Rewrite this Stratonovich differential equation as
an Ito type differential equation, we have

\[(2.3.4) \quad dX^i_t = e^i_j(t) \, dW^j_t + \frac{1}{2} d\langle e^i_j, dW^j \rangle_t.\]

If we let \(dM^i_t = e^i_j(t) \, dW^j_t\) be its martingale part, then

\[(2.3.5) \quad d\langle M^i, M^j \rangle_t = \sum_{k=1}^d e^i_k(t) e^j_k(t) \, dt.\]

At a frame \(u\), by definition \(ue^i_l = e^i_j(X^j_t)\). Using the definition of orthonormal frame bundle, we have

\[\delta^m_l = \langle ue^i_l, ue^i_m \rangle = e^i_l g^i_j e^j_m,\]

or \(e g e^T = I\) in matrix notation. This shows that

\[\sum_{k=1}^d e^i_k e^j_k = g^{ij},\]

where \(g^{ij}\) is the inverse matrix of \(g_{ij}\). Using this identity, we can rewrite (2.3.5) as the following form:

\[d\langle M^i, M^j \rangle_t = g^{ij}(X_t) \, dt.\]

If we let \(\sigma\) be the positive definite matrix square root of \(g^{-1}\), then by Levy’s characterization of Brownian motion,

\[W_t = \int_0^t \sigma(X_s)^{-1} \, dM_s\]

is a euclidean Brownian motion, and we have

\[dM_t = \sigma(X_s) \, dW_s.\]
From the second equation if (2.3.3), the last term in (2.3.4) becomes
\[ d\langle e^j_i, dW^l_i \rangle_t = -\Gamma^i_{kl}(X_t)e^l_m(t)e^k_m(t) = -g^{lk}(X_t)\Gamma^i_{kl}(X_t). \]

Therefore the equation for the Brownian motion \( X \) in local coordinates is
\[ (2.3.6) \quad dX^i_t = \sigma^j_i(X_t) dW^j_t - \frac{1}{2} g^{lk}(X_t)\Gamma^i_{kl}(X_t) \, dt, \]
where \( W \) is a \( d \)-dimensional euclidean Brownian motion.

We can also obtain the equation for \( X \) directly from its generator in local coordinates
\[ \Delta_M = g^{ij} \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} + b^i \frac{\partial}{\partial x^i}, \]
where
\[ b^i = -g^{ik}\Gamma^i_{jk}. \]
Using this, we can verify that the generator of the solution of (2.3.6) is indeed \( \Delta_M/2 \).

### 2.4. Wong-Zakai’s theorem

In this section, we shall discuss the Wong-Zakai theorem in [9], we will focus on one specific type of approximation of Brownian motion. We will have a family of smooth processes denoted by \( X_n(t) \) that converges to a standard Brownian motion \( W_t \), and we will see how integrations against \( X_n(t) \) could converge to a stochastic integral against \( W_t \).

Let \( W_t \) be the real Brownian motion and \( X_n(t) \) be a sequence of approximations to \( W_t \) with the following properties. For each \( n \), \( X_n(t) \) is of bounded variation, continuous
and converges almost surely to \( W_t \) as \( n \to \infty \). On one hand, we have

\[
\lim_{n \to \infty} \int_0^t X_n(t) \, dX_n(t) = \lim_{n \to \infty} X_n(t)^2 / 2 = W_t^2 / 2;
\]
on the other hand, for the stochastic integral we have

\[
\int_0^t W_s \, dW_s = (W_t^2 - t) / 2.
\]
This direct calculation shows that there is a difference between ordinary integral and stochastic integral due to the nature of stochastic process.

Let \( Z_t \) be the solution to the stochastic differential equation

\[
(2.4.1) \quad dZ_t = \sigma(Z_t) \, dW_t + b(Z_t) \, dt,
\]
where \( Z_0 \) is a random variable independent of \( W_t - W_0 \). It is well known that

**Theorem 2.4.1.** Assume that \( \sigma(x) \) and \( b(x) \) satisfy the global Lipschitz condition: there exists a constant \( K \) such that

\[
|\sigma(x) - \sigma(y)| + |b(x) - b(y)| \leq K|x - y|
\]
for all \( x \) and \( y \). Then the stochastic differential equation

\[
dZ_t = \sigma(Z_t) \, dW_t + b(Z_t) \, dt
\]
has a unique solution.
Assume that $X_n(t)$ has continuous derivative almost surely and $Z_n(t)$ be the solutions of integrals of the ordinary differential equations

$$dZ_n(t) = \sigma(Z_n(t)) \, dX_n(t) + b(Z_n(t)) \, dt,$$

where $Z_n(0) = Z_0$. The purpose of the Wong-Zakai’s theorem is to derive relations between $Z_t$ and the limit of $Z_n(t)$.

To be more precise, we require $X_n(t)$ to have the following properties

1. For all $t \geq 0$, $X_n(t) \to W_t$ almost surely and $X_n(t)$ are continuous and of bounded variation.
2. For almost all $\omega \in \Omega$ there exist $N(\omega)$ and $k(\omega)$ both finite such that for all $n \geq N(\omega)$, $X_n(t, \omega) \leq k(\omega)$.
3. $X_n(t)$ has a piecewise continuous derivative almost surely.

The statement of the theorem is as follows

**Theorem 2.4.2 (Wong-Zakai theorem).** Assume that $Z_t$ and $Z_n(t)$ satisfy the following equations

$$dZ_t = \sigma(Z_t) \circ dW_t + b(Z_t) \, dt$$

and

$$dZ_n(t) = \sigma(Z_n(t)) \, dX_n(t) + b(Z_n(t)) \, dt.$$

If

1. $\sigma(x)$, $b(x)$, $\partial \sigma(x)/\partial x$ are continuous for all $x$ and $t \geq 0$ and satisfy the Lipschitz condition.
2. $\sigma(x)$, $b(x)$ and $\partial \sigma^2(x)/\partial x$ satisfy the global Lipschitz condition, and $\inf_x \sigma(x) > 0$. 


(3) \( \inf_x \sigma(x) > 0. \)

(4) \( Z_0 = Z_n(0) \) is a random variable independent of the increment \( W_t - W_s \) for all \( 0 \leq s \leq t. \)

Then \( Z_n(t) \to Z_t \) almost surely as \( n \to \infty. \)

In addition if \( X_n(t) \to W_t \) uniformly then \( Z_n(t) \to Z_t \) uniformly as well.

At this moment, if one further simplify the equation by setting \( b(x) = 0, \) the equations become

\[
 dZ_t = \sigma(Z_t) \circ dW_t, \quad dZ_n(t) = \sigma(Z_n(t)) \, dX_t(t).
\]

And \( Z_n(t) \to Z_t \) if \( X_t \to W_t \) and the above conditions are also satisfied.

This theorem is important as it states the relation between the process \( Z_t \) and \( Z_n(t). \)

It basically says the ordinary integrals converge to a Stratonovich integral. For the detailed proof of this theorem, the reader is referred to [9].
CHAPTER 3

The Euclidean Case

In this chapter, we discuss the approximation for the simplest case. We assume the underlying manifold to be $\mathbb{R}^n$, and we show convergence of the diffusion processes to a standard Brownian Motion. The purpose of this chapter is to partially illustrate what we should expect for the general case of Riemannian Brownian Motion on a Riemannian manifold.

Throughout this chapter, our manifold will always be the simplest one-dimensional manifold $M = \mathbb{R}$. In general, everything can be carried over higher dimensional euclidean space $\mathbb{R}^n$ without too much extra work. We will first explain the construction of Bismut’s hypoelliptic Laplacian operator in the euclidean case, then prove the convergence in this case. Properties of the processes along the tangent fibre will also be covered. Eventually, we can discuss what modifications we can make to improve our result.

3.1. Bismut’s hypoelliptic diffusions

Consider the real line $M = \mathbb{R}$ as a Riemannian manifold. We have the unit vector field $e = \partial/\partial x$. A general tangent vector at $x \in \mathbb{R}$ can be written as $y \cdot \partial/\partial x$, which can be identified with the point $(x, y) \in \mathbb{R}^2$. Hence the tangent bundle $TM$ is canonically identified with $\mathbb{R}^2$. 

Let $M$ be the euclidean space $\mathbb{R}$ and $T\mathbb{R} = \mathbb{R}^2$ be the tangent bundle over $M$. Consider the Ornstein-Uhlenbeck operator on the tangent fibre

$$A = \frac{1}{2}(-\Delta^V + 2R^V)$$

where $\Delta^V$ is the Laplacian operator along the fibres of the tangent bundle $TM$ with respect to the Riemannian metric $g^{TM}$, and $R^V$ is the radial vector field along the fibres of $TM$. Here we use the letter $V$ to differentiate these vertical operators from the standard Laplacian and radial vector field. Using the standard coordinates system, we have

$$A = \frac{1}{2}(-\frac{\partial^2}{\partial y^2} + 2y \frac{\partial}{\partial y}).$$

Let $U$ be the vector field on $TM$ that generates the geodesic flow. In euclidean case, we have

$$U = y \frac{\partial}{\partial x}.$$

For $\epsilon > 0$, we are interested in the second order differential operator

(3.1.1) $$M_\epsilon = \epsilon^{-2}A - \epsilon^{-1}U$$

on $TM$. In 1967, Hormander proved in [5] that the operator $L$ defined as

$$L = \frac{1}{2} \sum_{i=1}^{n} X_i^2 + X_0,$$

where $X_0, X_1, \ldots, X_n$ denote a collection of smooth vector fields, is hypoelliptic if the Lie algebra generated by $X_0, X_1, \ldots, X_n$ has dimension $d$ throughout the domain of $L$. This hypothesis is known as Hormander’s condition. By this condition, we claim $M_\epsilon$
is a hypoelliptic operator acting on the tangent bundle $TM$. Indeed, using the above coordinates system, we can express $M_\epsilon$ as

$$ M_\epsilon = -\frac{1}{2\epsilon^2} \frac{\partial^2}{\partial y^2} + \frac{y}{\epsilon^2} \frac{\partial}{\partial y} - \frac{y}{\epsilon} \frac{\partial}{\partial x}. $$

Let

$$ X_1 = \frac{1}{\epsilon} \frac{\partial}{\partial y} $$

and

$$ X_0 = \frac{y}{\epsilon^2} \frac{\partial}{\partial y} - \frac{y}{\epsilon} \frac{\partial}{\partial x}, $$

we have

$$ M_\epsilon = \frac{1}{2} X_1^2 + X_0. $$

Note the Lie bracket

$$ [X_1, X_0] = X_1 X_0 - X_0 X_1 = \frac{1}{\epsilon^3} \frac{\partial}{\partial y} - \frac{1}{\epsilon^2} \frac{\partial}{\partial x} $$

together with $X_0$ does span $\mathbb{R}^2$, therefore $M_\epsilon$ is a hypoelliptic operator. As a deformation of the standard elliptic Laplacian operator, this family of operators is Bismut’s hypoelliptic Laplacian in its simplest form.

Fix any $\epsilon > 0$, a diffusion process $Z_\epsilon^i = (X_1^\epsilon_i, Y_1^\epsilon)$ generated by $-M_\epsilon$ is a solution to the stochastic differential equation

$$ dX_1^\epsilon = e^{-1} Y_1^\epsilon \, dt, \quad dY_1^\epsilon = e^{-1} dW_t - e^{-2} Y_1^\epsilon \, dt, $$

(3.1.2)
where $W$ is a standard Brownian motion. For the initial condition, we take $X_0^\varepsilon = 0$ and $Y_0^\varepsilon$ to be a square integrable random variable that is independent of $W$.

### 3.2. Convergence in euclidean case

In this section, we show properties of the stochastic process $Z^\varepsilon = (X^\varepsilon, Y^\varepsilon)$ in the tangent bundle. In particular, we show $X_t^\varepsilon$, the projection of $Z^\varepsilon$ onto $M$, converges to a standard Brownian motion almost surely. We will give two different approaches based on different properties of Brownian motion and we will discuss the advantages and disadvantages of each approach. We begin with the weakest type of convergence.

**Theorem 3.2.1.** For fixed $t > 0$, the process $X_t^\varepsilon$ converges in distribution to $W_t$ as $\varepsilon$ goes to zero, where $W$ is a standard Brownian motion.

The proof is based on the following equivalent characterization of standard Brownian motion.

**Proposition 3.2.2.** A stochastic process is a Brownian motion if

1. it is a Gaussian process
2. its covariance matrix $\mathbb{E}X_sX_t = s \wedge t$ for $0 \leq s \leq t$.

From the first equation in (3.1.2), it’s obvious that the process $X_t$ is a Gaussian process since it is a linear combination of Ornstein-Uhlenbeck processes. Therefore we only need to prove the second condition, namely $\mathbb{E}X_sX_t = s \wedge t$ where $X_t$ has same distribution as $\lim_{\varepsilon \to 0} X_t^\varepsilon$. 
PROOF. Assume $X_0 = 0$ and $0 \leq s \leq t$, using the first equation in (3.1.2) we have
\[
X^\epsilon_t = \epsilon^{-1} \int_0^t Y^\epsilon_s \, ds.
\]

First, we can compute $Y^\epsilon_t$ using the second equation in (3.1.2). It is just a scaled Ornstein-Uhlenbeck process.

(3.2.1)  
\[
Y^\epsilon_t = \exp(-t/\epsilon^2) Y_0 + \epsilon^{-1} \int_0^t \exp(-s/\epsilon^2) \exp(s/\epsilon^2) \, dW_s.
\]

Having this better expression for $Y^\epsilon_t$, we can compute the mean of the process $X^\epsilon_t$.

Using Fubini’s theorem, we have
\[
\mathbb{E}X^\epsilon_t = \epsilon^{-1} \int_0^t \mathbb{E}Y^\epsilon_s \, ds
\]
\[
= \epsilon^{-1} \int_0^t \exp(-s/\epsilon^2) Y_0 \, ds
\]
\[
= -\epsilon Y_0 \left( \exp(-t/\epsilon^2) - 1 \right),
\]
which goes to zero as $\epsilon$ goes to zero.

Using a similar approach, we can as well get the covariance matrix of the process $X^\epsilon_t$. For $0 \leq s \leq t$,
\[
\mathbb{E}X^\epsilon_s X^\epsilon_t = \mathbb{E} \left( \int_0^s Y^\epsilon_u / \epsilon \, du \right) \left( \int_0^t Y^\epsilon_v / \epsilon \, dv \right)
\]
\[
= \epsilon^{-2} \int_0^s \int_0^t \mathbb{E} Y^\epsilon_u Y^\epsilon_v \, dv \, du.
\]
Using the assumption that $0 \leq s \leq t$, we can write the last integral as a sum of two integrals

$$E X^\epsilon_s X^\epsilon_t = e^{-2} \int_0^s \int_0^u E Y^\epsilon_u Y^\epsilon_v \, dv \, du + e^{-2} \int_0^s \int_u^t E Y^\epsilon_u Y^\epsilon_v \, dv \, du.$$ 

These two integral can be solved using Fubini’s theorem and expression for $Y^\epsilon_t$ in (3.2.1). Eventually, we have

$$\lim_{\epsilon \to 0} E X^\epsilon_s X^\epsilon_t = s,$$

for $0 \leq s \leq t$.

In general, this implies $\lim_{\epsilon \to 0} E X^\epsilon_s X^\epsilon_t = s \land t$, which finishes the proof. \qed

This approach is very natural as it is purely based on the definition of a Brownian motion. It is considered straightforward in the sense that everything can be solved explicitly thanks to (3.2.1). However, this approach relies heavily on the covariance matrix, which can be easily computed in the euclidean space but extremely hard to carry out on a Riemannian manifold. Even worse than that, the convergence is in the weak sense since we are essentially using characteristic functions behind the scene. As a result, we cannot adopt this approach in the Riemannian manifold case. With that being said, this approach gives us some guidance of the next approach to this problem.

The second approach is based on stochastic calculus. In this approach, we will start by showing an $L^2$ convergence result. From there, we will be able to get the almost surely convergence as well.

**Theorem 3.2.3** ($L^2$ convergence in euclidean space). For fixed $t > 0$, the process $X^\epsilon_t$ converges in $L^2$ to $W_t$ as $\epsilon$ goes to zero, where $W$ is a standard Brownian motion.
PROOF. Again, we use the fact that $Y_t^ε$ can be solved independently and start from there. Recall that

$$Y_t^ε = \exp(-t/ε^2)Y_0 + ε^{-1} \exp(-t/ε^2) \int_0^t \exp(s/ε^2) dW_s.$$  

From this point on, we assume $Y_0 = 0$ since it has no effect on our result. Using the above expression, we can get the expression for the process $X_t^ε$. In the integral form, we have

$$(3.2.2) \quad X_t^ε = ε^{-2} \int_0^t \exp(-s/ε^2) \left( \int_0^s \exp(r/ε^2) dW_r \right) ds.$$  

In order to simplify the expression for $X_t^ε$, we apply integration by parts for stochastic integral. Ito’s formula on the product of two semimartingales tells us for any semimartingales $M$ and $N$,

$$dM_t N_t = M_t dN_t + N_t dM_t + d\langle M, N \rangle_t.$$  

Rearrange the terms, we can get the integration by parts formula for stochastic integrations, namely

$$M_t dN_t = dM_t N_t - N_t dM_t - d\langle M, N \rangle_t.$$  

In order to take advantage of the integration by parts formula, we let

$$M_t = \int_0^t \exp(s/ε^2) dW_s$$

and

$$dN_t = ε^{-2} \exp(-t/ε^2) dt.$$
The process \( N_t \) is smooth hence the quadratic co-variation term \( \langle M, N \rangle_t = 0 \). It follows that

\[
M_t \, dN_t = dM_t N_t - N_t \, dM_t,
\]

or, in the integration form,

\[
X^\epsilon_t = \int_0^t M_s \, dN_s = M_t N_t - \int_0^t N_s \, dM_s.
\]

Using the expression for \( M_t \) and \( dN_t \), we can easily get \( dM_t = \exp(t/\epsilon^2) \, dW_t \) and \( N_t = -\exp(-t/\epsilon^2) \). Therefore, by the integration by parts formula,

\[
(3.2.3) \quad X^\epsilon_t = W_t + \epsilon (1 - \exp(-t/\epsilon^2)) - \exp(-t/\epsilon^2) \int_0^t \exp(s/\epsilon^2) \, dW_s.
\]

As \( \epsilon \to 0 \), it is clear that the second term goes to zero. The third term has mean zero and variance

\[
\exp(-2t/\epsilon^2) \int_0^t \exp(2s/\epsilon^2) \, ds = \frac{\epsilon^2}{2} (1 - \exp(-t/\epsilon^2)),
\]

which also goes to zero. Hence \( X_t \to W_t \) (in \( L^2 \), for example) as \( \epsilon \to 0 \). \( \Box \)

Although we cannot prove \( X^\epsilon_t \to 0 \) almost surely at this point, equation (3.2.3) gives us an explicit decomposition of the stochastic process \( X^\epsilon_t \) as the sum of a Brownian motion and an error term. It is clear at this point that \( X^\epsilon_t \) converges to zero almost surely if the error terms goes to zero almost surely. Luckily, this could be achieved. In the last chapter, we will prove a much stronger result. For now, we prove the almost surely convergence first.
**Theorem 3.2.4** (almost surely convergence in euclidean space). For fixed $t > 0$, the process $X^\epsilon_t$ converges almost surely to $W_t$ as $\epsilon$ goes to zero, where $W$ is a standard Brownian motion.

**Proof.** The only thing left to prove is to show that that last term

$$
\exp(-t/\epsilon^2) \int_0^t \exp(s/\epsilon^2) dW_s
$$

converges to zero almost surely as $\epsilon \to 0$. We will prove this fact by choosing a countable subsequence of the original sequence.

Let $X_n(t) = \exp(-tn^2) \int_0^t \exp(sn^2) dW_s$ be a countable subsequence of the original sequence, it is sufficient to prove $X_n(t) \to 0$ almost surely as $n \to \infty$. Fix $\delta > 0$, from the above proof and Chebyshev’s inequality we have

$$
P(|X_n(t)| > \delta) < \frac{1}{2\delta^2 n^2} (1 - \exp(-tn^2)),
$$

which implies

$$
\sum_{n=1}^{\infty} P(|X_n(t)| > \delta) < \sum_{n=1}^{\infty} \frac{1}{2\delta^2 n^2} < \infty.
$$

Apply the Borel-Cantelli Lemma, we have

$$
P(|X_n(t)| > \delta \text{ i.o.}) = 0.
$$

This statement is equivalent to $X_n(t) \to 0$ almost surely since $\delta$ can be chosen arbitrarily.
From there, we assume $X^\epsilon_1 \not\to 0$. Then we can find a subsequence of $X^\epsilon_1$, call it $X^\epsilon_{n_m}(t)$, such that $n_m > n$ for each $m$ and $X^\epsilon_{n_m}(t) \not\to 0$. By a similar calculation, we have

$$P(\|X^\epsilon_{n_m}(t)\| > \delta \text{ i.o.}) = 0$$

for any $\delta > 0$, which is a contradiction. Consequently, $X^\epsilon_1 \to 0$ almost surely for fixed $t > 0$. \hfill \square

We can derive an even stronger result, namely path-wise convergence.

**Theorem 3.2.5** (path-wise convergence in euclidean space). The family of processes $X^\epsilon$ converges almost surely to a standard Brownian motion $W$. To be more precise,

$$\sup_t |X^\epsilon_t - W_t| \to 0$$

for almost all $\omega \in \Omega$.

This is a direct consequence of lemma 4.4.3 and we will save the proof to the last chapter.

Lastly, from (3.2.3), we see that $X^\epsilon_t \to t/\epsilon$ as $\epsilon \to \infty$, which is a scaled geodesic in $\mathbb{R}$. Before we discuss what modifications could be applied to make the processes $X^\epsilon_t$ interpolate between a standard Brownian motion and a genuine geodesic flow, we cover some of the properties of the processes $Y^\epsilon_t$.

### 3.3. Convergence of the attendant Gaussian motion

We state the result first. As $\epsilon \to 0$ and for a fixed $t \geq 0$, $Y^\epsilon_t$ converges to a Gaussian random variable with mean zero and variance $1/2$. For two different times $s \neq t$, $Y^\epsilon_t$
and \( Y_\epsilon \) eventually become independent. In order to prove this result, we need the following lemma.

**lemma 3.3.1.** Let \( X \) and \( Y \) be two random variables. Then \( X \) and \( Y \) are independent if and only if, \( \forall a \in \mathbb{R} \) and \( \forall b \in \mathbb{R} \), we have

\[
\mathbb{E} \exp(iaX + ibY) = \mathbb{E} \exp(iaX) \mathbb{E} \exp(ibY).
\]

**Proof.** Let \((\tilde{X}, \tilde{Y})\) be a two tuple such that \( \tilde{X} \) has the same distribution as \( X \), \( \tilde{Y} \) has the same distribution as \( Y \) and \( \tilde{X} \), \( \tilde{Y} \) are independent. Then

\[
\mathbb{E} \exp (i(X, Y) \cdot (a, b)) = \mathbb{E} \exp(iaX) \mathbb{E} \exp(ibY) = \mathbb{E} \exp \left( i(\tilde{X}, \tilde{Y}) \cdot (a, b) \right).
\]

The result follows from the uniqueness of characteristic functions. \( \square \)

Now, we are ready to prove the following theorem about the process \( Y_\epsilon \).

**Theorem 3.3.2.** As \( \epsilon \to 0 \), the random processes \( Y_\epsilon \) have the following two properties:

1. for fixed positive \( t > 0 \), the family of random variables \( Y_\epsilon \) converges in distribution to a centered Gaussian random variable with variance \( \frac{1}{2} \);

2. for each fixed pair \( 0 \leq s < t \), \( Y_\epsilon^s \) and \( Y_\epsilon^t \) become asymptotically independent.

3. for all \( s \neq t \), \( X_\epsilon^s \) and \( Y_\epsilon^t \) become asymptotically independent.

**Proof.** As we mentioned above, we assume \( Y_0 = 0 \). Recall that the process \( Y_\epsilon \) has the following expression

\[
Y_\epsilon^t = \epsilon^{-1} \exp(-t/\epsilon^2) \int_0^t \exp(s/\epsilon^2) \, dW_s.
\]
We will prove (1) by showing the characteristic function of $Y^\epsilon_t$ converges to the one of a Gaussian random variable with mean zero and variance $1/2$.

Let $\xi \in \mathbb{R}$ and $t > 0$, the characteristic function of $Y^\epsilon_t$ is

$$f(\xi) = \mathbb{E}\exp(i\xi Y^\epsilon_t) = \mathbb{E}\exp\left(i\xi e^{-1} \exp(-t/e^2) \int_0^t \exp(s/e^2) dW_s\right).$$

Let $M_t = \int_0^t \exp(s/e^2) dW_s$ be the martingale part and let $c$ be the constant

$$c = i\xi e^{-1} \exp(-t/e^2).$$

We can rewrite the above characteristic function $f(\xi)$ as

$$f(\xi) = \mathbb{E}\exp(cM_t).$$

Using the fact that $\exp(cM_t - c^2 \langle M \rangle_t / 2)$ is an exponential martingale, we have

$$\mathbb{E}\exp(cM_t) = \exp(c^2 \langle M \rangle_t / 2) = \exp\left(-\xi^2 (1 - \exp(-2t/e^2) / 4\right).$$

As $\epsilon \rightarrow 0$, the above expression converges to $\exp(-\xi^2 / 4)$, which is the characteristic function of a Gaussian random variable with mean zero and variance $1/2$.

To prove (2), we will use the above lemma. Let $a, b \in \mathbb{R}$, fix $t > s > 0$, we need

$$\lim_{\epsilon \rightarrow 0} \mathbb{E}\exp(iaY^\epsilon_t + ibY^\epsilon_s) = \exp(-a^2 / 4 - b^2 / 4).$$

Let $\mathcal{F}_t$ be the filtration generated by the Brownian motion $W_t$, we have

$$\mathbb{E}\exp(iaY^\epsilon_t + ibY^\epsilon_s) = \mathbb{E}(\mathbb{E}(\exp(iaY^\epsilon_t + ibY^\epsilon_s)|\mathcal{F}_s))$$
Since \( Y_t^\epsilon \in \mathcal{F}_s \), we can pull this part out of the conditional expectation and get

\[
(3.3.2) \quad \mathbb{E}\exp(iaY_t^\epsilon + ibY_s^\epsilon) = \mathbb{E}\left(\exp(ibY_s^\epsilon)\mathbb{E}(\exp(iaY_t^\epsilon) | \mathcal{F}_s)\right).
\]

In order to calculate \( \mathbb{E}(\exp(iaY_t^\epsilon) | \mathcal{F}_s) \), we use the same exponential martingale trick. Let \( M_t = \int_0^t \exp(s/e^2) dW_s \) be the martingale part and \( c = iae^{-1} \exp(-t/e^2) \), we have

\[
\mathbb{E}(\exp(iaY_t^\epsilon) | \mathcal{F}_s) = \mathbb{E}(\exp(cM_t) | \mathcal{F}_s) = \exp(cM_s + c^2 \langle M \rangle_t / 2 - c^2 \langle M \rangle_s / 2).
\]

Having all the ingredients we need, we can plug in this expression into (3.3.2) and get

\[
\mathbb{E}\exp(iaY_t^\epsilon + ibY_s^\epsilon) = \mathbb{E}\exp(ibY_s^\epsilon + cM_s + c^2 \langle M \rangle_t / 2 - c^2 \langle M \rangle_s / 2)
\]

Doing a similar calculation as (3.3.1), we have

\[
\mathbb{E}\exp(cM_s) = \exp(c^2 \langle M \rangle_s / 2) = \exp(-a^2(1 - \exp(-2s/e^2) / 4).
\]

Hence as \( \epsilon \to 0 \), we have

\[
\lim_{\epsilon \to 0} \mathbb{E}\exp(iaY_t^\epsilon + ibY_s^\epsilon) = \mathbb{E}\exp(ibY_s^\epsilon) \cdot \exp(c^2 \langle M \rangle_t / 2 - c^2 \langle M \rangle_s / 2)
\]

Apply the exponential martingale trick again to solve the expectation part, and use the fact that

\[
\langle M \rangle_t = \int_0^t \exp(2s/e^2) ds
\]
we eventually get

\[
\lim_{\epsilon \to 0} \mathbb{E} \exp \left( i a Y^\epsilon_t + i b Y^\epsilon_s \right) = \lim_{\epsilon \to 0} \exp \left( -a^2 \left( 1 - \exp \left( -2t / \epsilon^2 \right) \right) / 4 - b^2 / 4 \right) = \exp \left( -a^2 / 4 - b^2 / 4 \right),
\]

which completes the proof of (2). Statement (3) could be proved using a similar calculation. □

Unlike the process \( X^\epsilon_t \), which converges almost surely, the process \( Y^\epsilon_t \) does not converge (for example in probability) to any random variable. This fact could be proved by contradiction using a standard trick.

**Proposition 3.3.3.** The family of processes \( Y^\epsilon_t \) does not converge in probability to any random variable.

**Proof.** The proof is similar to the proof of the fact that \( S_n / \sqrt{n} \) does not converge in probability, where \( S_n \) is the sum of \( n \) i.i.d random variables with mean zero and variance one.

Again, we assume \( Y_0 = 0 \) for simplicity. For a fixed \( t > 0 \), assume in contradiction that the processes \( Y^\epsilon_t \) restricted to time \( t \) converges in probability to some random variable \( Y \) as \( \epsilon \to 0 \). Consider the random process given by a similar differential equation

\[
Z^\epsilon_t = 2e^{-1} \exp \left( -t / 4\epsilon^2 \right) \int_0^t \exp \left( s / 4\epsilon^2 \right) dW_s.
\]

As \( \epsilon \to 0 \), the random variable \( Z_t \) should converge to \( Y \) in probability as well by the assumption. As a result, we have the difference \( (Y^\epsilon_t - Z^\epsilon_t) \) converges to zero in
probability as $\epsilon \to 0$. Since $Y^\epsilon_t - Z^\epsilon_t$ is Gaussian for each fixed $\epsilon > 0$, the variance of $Y^\epsilon_t - Z^\epsilon_t$ should also converge to zero as $\epsilon \to 0$. This is not the case as the variance of the difference of the two random processes restricted at time $t$ is

$$E(Y^\epsilon_t - Z^\epsilon_t)^2 = E(Y^\epsilon_t)^2 + E(Z^\epsilon_t)^2 - 2EY^\epsilon_t Z^\epsilon_t.$$ 

Using the stochastic differential equations for $Y^\epsilon_t$ and $Z^\epsilon_t$ we can compute each part in the above expression easily and get

$$\lim_{\epsilon \to 0} E(Y^\epsilon_t - Z^\epsilon_t)^2 = 1/5,$$

which is a contradiction to the assumption that $Y^\epsilon_t - Z^\epsilon_t$ converges to zero in probability. As a consequence, $Y_t$ does not converge in probability.□

3.4. From uniform motion to Brownian motion

In this section, we discuss a new family of operators, which is a variation of Bismut’s family of hypoelliptic Laplacian operators. This family of operators shares similar properties, i.e. the projection of the diffusion processes generated by this family of operators converge to a standard Brownian motion. In this case, unlike Bismut’s stochastic diffusion processes, these processes have nicer interpolation properties.

Instead of constructing operators before deriving the corresponding diffusion processes, we adopt a different approach. We try to construct a natural family of Gaussian processes with continuously differentiable sample paths which connects the uniform motion on the real line to the one-dimensional Brownian motion. We will explain the natural role that Ornstein-Uhlenbeck process plays in this construction.
We set ourselves the task of finding a natural family of Gaussian processes $X_\epsilon^t$ with continuously differentiable paths connecting the uniform motion $u(t) = t$ and Brownian motion $W_t$. Writing the process as

\begin{equation}
X_\epsilon^t = \int_0^t K_\epsilon(s) \, ds.
\end{equation}

For an appropriate Gaussian process $K_\epsilon^t$, our task becomes, at least at heuristic level, to find $K_\epsilon^t$ such that

\[
\lim_{\epsilon \to \infty} K_\epsilon^t(t) = 1 \quad \text{and} \quad \lim_{\epsilon \to 0} K_\epsilon^t(t) = dW_t / dt,
\]

where $dW_t / dt$ could be interpreted as the while noise.

It is well known that the natural way to connect a Gaussian random variable to one is through the so-called Ornstein-Uhlenbeck process

\[
K_\epsilon(t) = \exp(-1/\epsilon) + \sqrt{1 - \exp(-2/\epsilon)} \cdot \frac{dW_t}{dt}.
\]

The next step would be to use the approximation

\[
\frac{dW_t}{dt} \sim \frac{W_t - W_{t-\epsilon}}{\epsilon},
\]

which leads to the definition

\[
K_\epsilon(t) = \exp(-1/\epsilon) + \sqrt{1 - \exp(-2/\epsilon)} \cdot \frac{W_t - W_{t-\epsilon}}{\epsilon}.
\]

This representation could lead to a plausible theory but not an elegant one, mainly because this approximation of white noise is not smooth enough.
It turns out that an approximation more carefully chosen based on the intuition we’ve got from Bismut’s approximation could lead to an elegant theory. We note that

$$\frac{W_t - W_{t-\epsilon}}{\epsilon} = \int_0^t \frac{I_{[t-\epsilon,t]}(s)}{\epsilon} dW_s.$$  

The integrant is an approximation of the delta function at $t$ and it is clear that any approximation of the delta function would serve our purpose. An equally natural but smoother choice would be $\epsilon^{-1} \exp((s - t)/\epsilon)$. This leads to the following definition of the integrand in (3.4.1):

(3.4.2) \[ K_{\epsilon}(t) = \exp(-t/\epsilon) + \frac{1}{\epsilon} \int_0^t \exp(s - t)/\epsilon dW_s, \]

which we realize is $Y^\epsilon_t$ in (3.1.2). Based on the discussion above, we consider a modifies version of Bismut’s hypoelliptic Laplacian operator $L_{\epsilon}$ given by

(3.4.3) \[ L_{\epsilon} = \frac{1}{2\epsilon^2} \frac{\partial^2}{\partial y^2} - \frac{y}{\epsilon} \frac{\partial}{\partial y} + y \frac{\partial}{\partial x}. \]

The stochastic diffusion process $Z^\epsilon_t$ generated by $L_{\epsilon}$ satisfies the following stochastic differential equations

(3.4.4) \[ dX^\epsilon_t = Y^\epsilon_t dt, \quad dY^\epsilon_t = \epsilon^{-1} dW_t - \epsilon^{-1} Y^\epsilon_t dt. \]

At first glance, this process looks very similar to (3.1.2). Without too much surprise, these two processes share same property as $\epsilon \to 0$. 
**Theorem 3.4.1.** Let $X^ε_t$ be the projections of the processes generated by the family of operators $L^ε$ onto the real line $\mathbb{R}$. Then $X^ε_t$ converge almost surely to a standard Brownian motion as $ε$ goes to zero.

**Proof.** The proof is very similar to the previous one, except in this case, the process $X^ε_t$ is expressed as

$$X^ε_t = ε(1 - \exp(-t/ε)) + \int_0^t (1 - \exp((s - t)/ε)) \, dW_s.$$  

□

Although these two families of processes share a common property, one can easily see that as $ε \to \infty$, the stochastic processes generated by $L^ε$ converge to the genuine geodesic flow, i.e. $X^ε_t \to t$ as $ε \to \infty$. As a result, we see $L^ε$ as a more natural choice for our purpose.

### 3.5. Ito’s formula revisited

In this section, we give another proof of Ito’s formula of Brownian motion. Unlike the standard proof, we use our smooth approximation of Brownian motion and standard fundamental theorem of calculus to derive Ito’s formula directly.

Let $W$ be a standard one-dimensional Brownian motion and $f \in C^2(\mathbb{R})$ a twice continuously differentiable function on $\mathbb{R}$. Ito’s formula says

$$f(W_t) = f(W_0) + \int_0^t f'(W_s) \, dW_s + \frac{1}{2} \int_0^t f''(W_s) \, ds.$$
We assume \( f(W_0) = 0 \) and, by the same stopping time argument, \( f \) together with its first three derivatives are all bounded. We approximate \( W \) by the continuously differentiable Gaussian process

\[
X_t = \frac{1}{\epsilon} \int_0^t V_s \, ds,
\]

where \( V_t = \int_0^t \exp((s-t)/\epsilon) \, dW_s \). Using this expression for \( V_t \), we can get the expression for the process \( X_t \). In the integral form, we have

\[
X_t = \epsilon^{-1} \int_0^t \exp(-s/\epsilon) \left( \int_0^s \exp(r/\epsilon) \, dW_r \right) \, ds. \tag{3.5.1}
\]

We apply integration by parts for stochastic integral to simplify this expression. Recall for semimartingales \( M \) and \( N \), we have

\[
M_t \, dN_t = dM_t N_t - N_t \, dM_t - d\langle M, N \rangle_t.
\]

Let \( M_t = \int_0^t \exp(s/\epsilon) \, dW_s \) and \( dN_t = \epsilon^{-1} \exp(-t/\epsilon) \, dt \). The process \( N_t \) is smooth hence the quadratic co-variation term \( \langle M, N \rangle_t = 0 \). It follows that

\[
M_t \, dN_t = dM_t N_t - N_t \, dM_t,
\]

or, in the integration form,

\[
X_t = \int_0^t M_s \, dN_s = M_t N_t - \int_0^t N_s \, dM_s.
\]
Using the expression for $M_t$ and $dN_t$, we can easily get $dM_t = \exp(t/\epsilon)\,dW_t$ and $N_t = -\exp(-t/\epsilon)$. Therefore, by the integration by parts formula,

\[ X_t = W_t - \exp(-t/\epsilon) \int_0^t \exp(s/\epsilon)\,dW_s = W_t - V_t. \]

Since $X_t$ is smooth, we could apply the fundamental theorem of calculus to $f(X_t)$ and get

\[ df(X_t) = f'(X_t)\,dX_t. \]

Rewrite the above differential form into integral form and assume $f(X_0) = 0$, we have

\[ f(X_t) = \int_0^t f'(X_s)\,dX_s = \int_0^t f'(X_s)\,dW_s - \int_0^t f'(X_s)\,dV_s. \]

Keep in mind that our goal is to somehow create a term that involves the second derivative of $f$. In order to achieve this, we use integration by parts on the second term and assume there exists a process $H$ such that

\[ f(X_t) = \int_0^t f'(X_s)\,dW_s - f'(X_t)V_t + \frac{1}{2} \int_0^t f''(X_s)\,dH_s. \]

This process $H$ is actually defined as

\[ dH_t = 2V_t\,dX_t \]

and can be identified by letting $f(x) = x^2$. Use this specific function $f$ we have

\[ X_t^2 = 2 \int_0^t X_s\,dW_s - 2X_tV_t + H_t. \]
Rearrange the terms, we have an explicit expression for $H_t$, which is

$$H_t = X_t^2 - 2 \int_0^t X_s \, dW_s + 2X_t V_t.$$

Eventually, we need a term $t$ in the expression of $H_t$. With that in mind, we rewrite the above equation using the relation $X_t = W_t - V_t$ and

$$W_t^2 - t = 2 \int_0^t W_s \, dW_s.$$

After some simplification, we end up with

$$H_t = t + 2 \int_0^t V_s \, dW_s - V_t^2.$$

It follows the last term of $f(X_t)$ is

$$\frac{1}{2} \int_0^t f''(X_s) \, dH_s = \frac{1}{2} \int_0^t f''(X_s) \, ds + \int_0^t f''(X_s) V_s \, dW_s - \frac{1}{2} \int_0^t f''(X_s) \, dV_t^2$$

where $-\frac{1}{2} \int_0^t f''(X_s) \, dV_t^2$ could be further simplified using $dX_t = e^{-1}V_t \, dt$ and integration by parts. To be more precise, we have

$$-\frac{1}{2} \int_0^t f''(X_s) \, dV_t^2 = \frac{1}{2e} \int_0^t f''(X_s) V_s^3 \, ds - \frac{1}{2} f''(X_t) V_t^2.$$

At this point, $f(X_t)$ can be written as

$$f(X_t) = \int_0^t f'(X_s) \, dW_s + \frac{1}{2} \int_0^t f''(X_s) \, ds + R_t,$$
where \( R_t \) is the error term and it is

\[
R_t = \int_0^t f''(X_s) V_s dW_s + \frac{1}{2\epsilon} \int_0^t f'''(X_s) V_s^3 ds - \frac{1}{2} f''(X_t) V_t^2 - f'(X_t) V_t.
\]

The three \( X_s \) in the expression for \( f(X_t) \) could all be replaced by \( W_t \) since their difference \( V_t \) is Gaussian with mean zero and variance \( \mathbb{E} V_t^2 < \epsilon \). First of all, we have

\[
f(X_t) - f(W_t) \leq \|f'\|_\infty |X_t - W_t|,
\]

which implies

\[
\mathbb{E} |f(X_t) - f(W_t)|^2 \leq \|f'\|_\infty^2 \mathbb{E} |X_t - W_t|^2 < \epsilon \|f'\|_\infty^2.
\]

Next,

\[
\mathbb{E} \left| \int_0^t (f'(X_s) - f'(W_s)) dW_s \right|^2 = \mathbb{E} \int_0^t |f'(X_s) - f'(W_s)|^2 ds \leq \frac{\epsilon t}{2} \|f''\|_\infty^2.
\]

Eventually,

\[
\mathbb{E} \left( \int_0^t \left( f''(X_s) - f''(W_s) \right) ds \right)^2 \leq \|f'''\|_\infty^2 \mathbb{E} \left( \int_0^t V_s ds \right)^2,
\]

where the last term could be solved using the following trick.

\[
\mathbb{E} \left( \int_0^t V_s ds \right)^2 = \mathbb{E} \int_0^t V_s ds \int_0^t V_r dr = \int_0^t \int_0^t \mathbb{E} V_s V_r ds dr.
\]

Using \( V_t = \int_0^t \exp((s - t)/\epsilon) dW_s \), we have

\[
\mathbb{E} V_s V_r = \exp(-(s + r)/\epsilon) \int_0^{s \wedge r} \exp(2w/\epsilon) dw < \exp((2s \wedge r - s - r)/\epsilon) = \exp(-c/\epsilon),
\]

where \( c \) is a constant.
for some positive constant number $c > 0$. Therefore,

$$\mathbb{E}(\int_0^t V_s \, ds)^2 < t^2 \exp(-c/\epsilon),$$

and

$$\mathbb{E}\left| \int_0^t (f''(X_s) - f''(W_s)) \, ds \right|^2 \leq t^2 \exp(-c/\epsilon) \|f''\|_\infty^2.$$ 

Using the moment generating function of Gaussian random variable to get the third-moment of $V_s$, we can bound the expectation of the error term $R_t$ by

$$\mathbb{E}|R_t| \leq \sqrt{\frac{\epsilon}{2}} \|f'\|_\infty + \left(\frac{\epsilon}{4} + \sqrt{\frac{\epsilon t}{2}}\right) \|f''\|_\infty + \sqrt{\epsilon t^2} \frac{\pi}{4} \|f'''\|_\infty.$$ 

To be more precise, we have

$$R_t = \int_0^t f''(X_s) V_s \, dW_s + \frac{1}{2\epsilon} \int_0^t f'''(X_s) V_s^3 \, ds - \frac{1}{2} f''(X_t) V_t^2 - f'(X_t) V_t,$$

where

$$\mathbb{E} \int_0^t f''(X_s) V_s \, dW_s \leq \|f''\|_\infty \left(\int_0^t \mathbb{E} V_s^2 \, ds\right)^{1/2} \leq \sqrt{\frac{t\epsilon}{2}} \|f''\|_\infty,$$

and

$$\mathbb{E} \frac{1}{2\epsilon} \int_0^t f'''(X_s) V_s^3 \, ds \leq t\sqrt{\epsilon/4\pi} \|f'''\|_\infty,$$

and

$$\mathbb{E} \frac{1}{2} f''(X_t) V_t^2 \leq \frac{\epsilon}{4} \|f''\|_\infty,$$

and

$$\mathbb{E} f'(X_t) V_t \leq \sqrt{\frac{\epsilon}{2}} \|f'\|_\infty,$$

all of which converge to zero as $\epsilon \to 0$ and finishes the proof of Ito's formula.
CHAPTER 4

The Compact Riemannian Manifold Case

In this chapter we define our family of diffusion processes as solution of a geometrically invariant stochastic differential equation and show that it indeed connects geodesic flow to Riemannian Brownian motion in the strong sense as the controlling parameter runs through its natural range. This chapter is organized as follows. In the first section, we discuss Bismut’s construction of hypoelliptic operators on Riemannian manifold, as well as the diffusion processes generated by this family of operators. In the second section, we explain the geometric settings that are needed. Also, we construct our family of operators $L_\epsilon$ and lift each component of it to proper geometric space. In particular, we show that this new operator $L_\epsilon$ is well-defined. after that, We derive the stochastic differential equations for the family of diffusions. In the third section, we discuss another proof of Ito’s formula in the one-dimensional case, based on the smooth approximation of Brownian motion. This will assist us in understanding the connection between ordinary integrals and stochastic integrals. The climax of this dissertation is reached in the last section of this chapter, in which we provides two proofs of the main theorem based on the Wong-Zakai theorem.

4.1. Geometric hypoelliptic Laplacian

In this section, we will discuss Bismut’s geometric hypoelliptic diffusion process in its general form.
Consider any compact Riemannian manifold $M$ of dimension $n$, and let $TM$ be the tangent bundle of $M$. Pick local coordinates $(x^i)$ on $M$, we have vector fields $\partial / \partial x^i$ as basis vector fields. A general tangent vector at $(x^i) \in M$ can be expressed as $y^i \cdot \partial / \partial x^i$, which can be identified with the point $(x^i, y^i) \in \mathbb{R}^n$. Hence the tangent bundle $TM$ is identified with $\mathbb{R}^{2n}$.

The Ornstein-Uhlenbeck operator on the tangent fibre is

$$A = \frac{1}{2} (\frac{-g^{ij}}{\partial y^i} \frac{\partial}{\partial y^j} + 2y^i \frac{\partial}{\partial y^i}).$$

where $(g^{ij})$ is the inverse matrix of the positive-definite matrix $(g_{ij})$. The vector field on $TM$ that generates the geodesic flow is

$$(4.1.1) \quad U = y^i \frac{\partial}{\partial x^i} - \Gamma^i_{jk} y^j y^k \frac{\partial}{\partial y^i}.$$ 

Consider the second order differential operator

$$M_\epsilon = e^{-2} A - e^{-1} U = -\frac{y^i}{\epsilon} \frac{\partial}{\partial x^i} - \frac{g^{ij}}{2\epsilon^2} \frac{\partial}{\partial y^i} \frac{\partial}{\partial y^j} + \left( \frac{y^i}{\epsilon^2} + \frac{1}{\epsilon} \Gamma^i_{jk} y^j y^k \right) \frac{\partial}{\partial y^i},$$

where $b$ is a positive real number.

For fixed $\epsilon > 0$, a diffusion process $Z = (X, Y)$ generated by $-M_\epsilon$ is a solution to the system of stochastic differential equations

$$(4.1.2) \quad \begin{cases} dX^i_t = \frac{\gamma^i_t}{\epsilon} \, dt, \\ dY^i_t = \frac{1}{\epsilon} \sigma^j(X_t) \, dW^j_t - \left( \frac{\gamma^i_t}{\epsilon^2} + \frac{1}{\epsilon} \Gamma^i_{jk}(X_t) Y^j_t Y^k_t \right) \, dt, \end{cases}$$
where $W$ is a standard euclidean Brownian motion. For the initial condition we take $X_0$ to be zero and $Y_0$ a square integrable random variable that is independent of $W$.

4.2. Geometric setting

From now on, we assume that $M$ is a compact Riemannian manifold. Our goal is to find a naturally defined family of diffusion processes that connects the geodesic flow to Riemannian Brownian motion. As we mentioned before, in the sense of weak convergence this goal has been accomplished in [2]. However, we aim at a stronger sense of convergence. As you will see later in this chapter, once we find the correct geometric setup, we could obtain a strong result and achieve this goal in much less space and with much less technical considerations.

Our goal requires that we find a geometric setting in which both Riemannian Brownian motion and geodesic flow are defined by intrinsically defined stochastic differential equations. Let $x = \{x^i\}$ be a local coordinate system on $M$. A tangent vector $v$ at $x$ can be written as

$$v = y^i \frac{\partial}{\partial x^i}.$$  

Then $(x^i, y^i)$ gives a local coordinate system for $TM$ in a neighborhood of $v$. The geodesic vector field lives on the tangent bundle intrinsically. Recall in terms of these coordinates, we have

$$U(x^i, y^i) = y^i \frac{\partial}{\partial x^i} - \Gamma^i_{jk}(x) y^j y^k \frac{\partial}{\partial y^i}.$$  

On the other hand, a Riemannian Brownian motion $X$ on $M$ itself in general is not the solution of any intrinsically defined stochastic differential equation on $M$. Instead, we need to lift it to a horizontal Brownian motion $U$ on the orthonormal frame bundle
\( \mathcal{O}(M) \) on \( M \) so that \( U \) is the solution of the stochastic differential equation

\[
dU_t = H_i(U_t) \circ dW^i_t.
\]

Here \( W = \{W^i\} \) is a standard \( n \)-dimensional euclidean Brownian motion and \( H_i \) are the fundamental horizontal vector fields on \( \mathcal{O}(M) \). For our purpose, it is clear that we need a space that include both \( \mathcal{O}(M) \) and \( TM \) as quotient spaces. The product bundle \( \mathcal{O}(M) \times \mathbb{R}^N \) will serve our purpose.

Let \( \mathcal{O}(M) \) be the orthonormal frame bundle of \( M \). It is a principal bundle with the structure group \( O(n) \), the group of orthogonal matrices. The tangent bundle is the associated bundle

\[
TM = \mathcal{O}(M) \times_{O(n)} \mathbb{R}^n, \quad (e, y) \to ey.
\]

where the actions of \( O(n) \) on \( \mathcal{O}(M) \) and \( \mathbb{R}^n \) are as usual.

Next, we want to construct out new family of operators \( L_{\epsilon} \) similar to (3.4.3) in the product bundle. Recall in the euclidean case, \( L_{\epsilon} \) is a differential operator acting on the tangent bundle \( TM \) with the following expression

\[
L_{\epsilon} = \frac{1}{2\epsilon^2} \frac{\partial^2}{\partial y^2} - \frac{y}{\epsilon} \frac{\partial}{\partial y} + y \frac{\partial}{\partial x}.
\]

As we have discussed, this family of operators is the scaled Bismut’s hypoelliptic Laplacian operators \( M_{\epsilon} \) in (3.1.1). To be more precise, we have

\[
M_{\epsilon} = -\frac{\Delta^V}{2\epsilon^2} + \frac{R}{\epsilon^2} - \frac{U}{\epsilon},
\]
whereas
\[ L_\epsilon = \frac{\Delta^V}{2\epsilon^2} - \frac{R}{\epsilon} + U. \]

Since \( L_\epsilon \) is defined on the tangent bundle \( TM \), our task is to lift each component of \( L_\epsilon \) to the product bundle. Let \( \mathbb{R}^n \) denote the trivial bundle of \( M \) and consider the product bundle \( \mathcal{O}(M) \times \mathbb{R}^n \). The radial vector field

\[
Y = \sum_{i=1}^{n} y^i \frac{\partial}{\partial y^i}
\]

on \( \mathbb{R}^n \) and the canonical horizontal vector field \( H_i \) on \( \mathcal{O}(M) \) can be lifted to the product bundle \( \mathcal{O}(M) \times \mathbb{R}^n \) in an obvious way. By an abuse of notation, we still use the same notations for these lifted vector fields. For the geodesic vector field, we define a vector field on the product bundle by

\[
U(e, y) = \sum_{i=1}^{n} y^i H_i.
\]

Consider the projection map \( \pi: \mathcal{O}(M) \times \mathbb{R}^n \to \mathcal{O}(M) \otimes_{\mathcal{O}(n)} \mathbb{R}^n = TM \). We claim \( U \) is the lift of the geodesic flow vector field on the tangent bundle \( TM \) to the product bundle. To be more precise, we have the following proposition.

**Proposition 4.2.1.** Let \( U \) be the vector field on the product bundle as in (4.2.2), \( V \) be the vector field on \( TM \) that generates the geodesic flow, \( f \) be a smooth map on \( TM \). Then

\[
U \tilde{f} = V f,
\]

where \( \tilde{f} = f \circ \pi \).
PROOF. We will prove this result using local coordinates. Pick local coordinates system \((x,e,y)\) on \(\mathcal{O}(M) \times \mathbb{R}^n\), local coordinates system \((x,z)\) on \(TM\). Here we use \(z\) to distinguish the local coordinates system on the tangent bundle from the coordinate system on \(\mathcal{O}(M) \times \mathbb{R}^n\). Using the local expression for geodesic vector field in (4.1.1), We have

\[
V f = z^i \frac{\partial}{\partial x^i} - \Gamma^i_{jk} z^j z^k \frac{\partial f}{\partial z^l}.
\]

On the other hand we consider \(U \tilde{f} = y^i H_i \tilde{f}\), where

\[
H_i \tilde{f} = e^i_j \frac{\partial \tilde{f}}{\partial x^j} - e^i_j e^m_l \Gamma^k_{jl} \frac{\partial \tilde{f}}{\partial e^m_k}.
\]

By the definition of the projection map \(\pi\), we have \(\pi(x,e,y) = (x,z)\), or \(z^i = e^i_j y^j\). Apply chain rule, we can deduce

\[
\frac{\partial \tilde{f}}{\partial x^j} = \frac{\partial f}{\partial x^j}, \quad \text{and} \quad \frac{\partial \tilde{f}}{\partial e^m_k} = y^k \frac{\partial f}{\partial z^k}.
\]

Substitute these expressions back into \(U \tilde{f}\), we have

\[
U \tilde{f} = y^i H_i \tilde{f} =
\]

\[
= y^i e^j_i \frac{\partial f}{\partial x^j} - y^i y^m e^j_i e^m_l \Gamma^k_{jl} \frac{\partial f}{\partial z^k}
\]

\[
= z^j \frac{\partial f}{\partial x^j} - z^j z^l \Gamma^k_{jl} \frac{\partial f}{\partial z^k}.
\]

This is the same as (4.2.3) and this proposition is proved. \(\square\)
Next, we deal with the radial vector field $R$ along the tangent fibre of $TM$ and claim $R$ is lifted to $Y$. Using the same notation as in the above proof, $R$ is

$$R = \sum_{i=1}^{n} z^i \frac{\partial}{\partial z^i}.$$ 

By an abuse of notation, let $Y$ be the lift of (4.2.1) to the product bundle. Using the same local coordinates system in the previous proof,

$$Y = \sum_{i=1}^{n} y^i \frac{\partial}{\partial y^i}.$$ 

We claim that $Y$ is the lift of $R$ to the product bundle $\mathcal{O}(M) \times \mathbb{R}^n$.

**Proposition 4.2.2.** Let $Y$ be the vector field on the product bundle $R$ be the radial vector field along the fibre of $TM$ and $f$ be a smooth map on $TM$. Then

$$Y\tilde{f} = Rf,$$

where $\tilde{f} = f \circ \pi$.

**PROOF.** Using the same local coordinate system we have,

$$Y\tilde{f} = y^i \frac{\partial \tilde{f}}{\partial y^i}.$$ 

Using the fact that $\pi(x,e,y) = (x,z)$, we have

$$(4.2.4) \quad \frac{\partial \tilde{f}}{\partial y^i} = \frac{\partial f}{\partial z^k} e^k_i \frac{\partial f}{\partial z^k} = e^k_i \frac{\partial f}{\partial z^k}.$$
Therefore
\[ Y \tilde{f} = y^i e_i^k \frac{\partial f}{\partial z^k} = z^k \frac{\partial f}{\partial z^k} = Rf, \]
which finishes the proof of this proposition. \( \square \)

What is left now is the Laplacian along the tangent fibre. Denoted by \( \Delta^{TM} \), the Laplacian operator along the tangent fibre is
\[ \Delta^{TM}(x, z) = g^{ij}(x) \frac{\partial}{\partial z^i} \frac{\partial}{\partial z^j}, \]
where \( g^{ij} \) is the inverse matrix of the Riemannian metric \( g_{ij} = \langle \frac{\partial}{\partial z^i}, \frac{\partial}{\partial z^j} \rangle \). Let \( \Delta^V \) be the laplacian along the trivial bundle \( \mathbb{R}^N \) of \( M \), then the lifting of \( \Delta^V \) to the product bundle \( O(M) \times \mathbb{R}^n \) is, with the same abuse of notation,
\[ \Delta^V(x, e, y) = \sum_{i=1}^n \frac{\partial}{\partial y^i} \frac{\partial}{\partial y^i}. \]
(4.2.5)

The last claim is that \( \Delta^V \) is the lift of \( \Delta^{TM} \) to the product bundle. More precisely,

**Proposition 4.2.3.** Let \( \Delta^V \) be the operator in (4.2.5), \( \Delta^{TM} \) be the Laplacian along the tangent fibre, \( f \) be a smooth map on \( TM \).

\[ \Delta^V \tilde{f} = \Delta^{TM} f, \]

where \( \tilde{f} = f \circ \pi \).

**Proof.** We have established an expression for \( \frac{\partial \tilde{f}}{\partial y^i} \) in (4.2.4). To be more precise, we express (4.2.4) as
\[ \frac{\partial \tilde{f}}{\partial y^i}(x, e, y) = e_i^k \frac{\partial f}{\partial z^k}(x, z), \]
where, \( z = ey \). Treat \( z \) as a function of \( y \), we could apply the chain rule and do the following computation:

\[
\Delta^V \tilde{f} = \sum_i^n \frac{\partial}{\partial y^i} \left( e^i_k \frac{\partial f}{\partial z^k} (x, z) \right) = \sum_i^n e^i_k \frac{\partial^2 f}{\partial z^k \partial z^l} \frac{\partial z^l}{\partial y^i} = \sum_i^n e^i_k e^l_i \frac{\partial^2 f}{\partial z^k \partial z^l}.
\]

Using the definition of \( \mathcal{O}(M) \), we have \( \langle ue_i, ue_j \rangle = e^k_i e^l_j g_{kl} = \langle e_i, e_j \rangle = \delta_{ij} \), where \( \delta \) is the dirac delta function. Therefore, \( \sum_i^n e^i_k e^l_i = g^{kl} \) and we can simplify our expression using this relation to get

\[
\Delta^V \tilde{f} = g^{ij} \frac{\partial^2 f}{\partial z^i \partial z^j} = \Delta^T_M f,
\]

which completes the proof. \( \square \)

Now, we have lifted each component of \( L_\epsilon \) to the product bundle. With an abuse of notation, it follows that the family of stochastic diffusion processes could be lifted to the product bundle and it is generated by the following family of operators

\[
L_\epsilon = \frac{\Delta^V}{2\epsilon^2} - \frac{Y}{\epsilon} + U.
\]

It is in the form of a sum of squared vector fields plus a vector field, hence the corresponding diffusion can be generated as the solution of an intrinsically defined system of stochastic differential equations on \( \mathcal{O}(M) \times \mathbb{R}^n \) driven by a standard \( \mathbb{R}^n \)-valued Brownian motion.

**4.3. Strong interpolation from the geodesic flow to Brownian motion**

In this section, we shall focus on the product bundle and prove the main theorem. To begin with, we recap the setup briefly. Let \( \mathcal{O}(M) \) be the orthonormal frame bundle
of $M$, let $\mathbb{R}^n$ be the trivial bundle of $M$, let $\mathcal{O}(M) \times \mathbb{R}^n$ be the product bundle. Consider the operator $L_\epsilon$ on the product bundle defined in as

\begin{equation}
L_\epsilon = \frac{\Delta^V}{2\epsilon^2} - \frac{Y}{\epsilon} + U,
\end{equation}

where $Y$ is the same as (4.2.1) and $U$ is the same as (4.2.2).

Pick local coordinate system $(x, e, y)$ on the product bundle, $L_\epsilon$ could be expressed as

\begin{equation*}
L_\epsilon = \frac{1}{2\epsilon^2} \sum_{i=1}^n \frac{\partial}{\partial y_i} \frac{\partial}{\partial y_i} - \frac{1}{\epsilon} \sum_{i=1}^n y_i \frac{\partial}{\partial y_i} + \sum_{i=1}^n y_i H_i.
\end{equation*}

First, we consider the one-dimensional case. In this case, the operator $L_\epsilon$ looks like

\begin{equation*}
L_\epsilon = \frac{1}{2\epsilon^2} \frac{\partial^2}{\partial y^2} - \frac{y}{\epsilon} \frac{\partial}{\partial y} + yH,
\end{equation*}

where $H$ is the horizontal lift of the vector field $e \cdot \partial/\partial x$. The high-dimensional case is very similar.

\textbf{Lemma 4.3.1}. A diffusion process $Z^\epsilon = (U^\epsilon, Y^\epsilon)$ generated by $L_\epsilon$ is a solution of the stochastic differential equation

\begin{equation}
dU^\epsilon_t = H(U^\epsilon_t)Y^\epsilon_t dt \quad \text{and} \quad dY^\epsilon_t = \epsilon^{-1}dW_t - \epsilon^{-1}Y^\epsilon_t dt,
\end{equation}

where $W$ is a one-dimensional euclidean Brownian motion.

\textbf{Proof}. This could be checked directly by showing that

\begin{equation*}
f(Z^\epsilon_t) - f(Z^\epsilon_0) - \int_0^t L_\epsilon f(Z^\epsilon_s) ds
\end{equation*}
is a local martingale. For example, using the fact that $dU^\epsilon_t = H(U^\epsilon_t)Y^\epsilon_t \, dt$ and Ito’s formula, we have

$$f(U^\epsilon_t) - f(U^\epsilon_0) = \int_0^t \frac{\partial f}{\partial x}(U^\epsilon_s) \, dU^\epsilon_s = \int_0^t L_\epsilon f(U^\epsilon_s) \, ds,$$

or

$$f(U^\epsilon_t) - f(U^\epsilon_0) - \int_0^t L_\epsilon f(U^\epsilon_s) \, ds = 0,$$

which is a special martingale. Using $Y^\epsilon_t = \epsilon^{-1} \, dW_t - \epsilon^{-1}Y^\epsilon_t \, dt$ and apply Ito’s formula, we have

$$f(Y^\epsilon_t) - f(Y^\epsilon_0) = M_t - \int_0^t \frac{\partial f}{\partial y}(Y^\epsilon_s) \frac{Y^\epsilon_s}{\epsilon} \, ds + \frac{1}{2\epsilon^2} \int_0^t \frac{\partial^2 f}{\partial y^2}(Y^\epsilon_s) \, ds,$$

where

$$M_t = \frac{1}{\epsilon} \int_0^t \frac{\partial f}{\partial y}(Y^\epsilon_s) \, dW_s$$

is a martingale. Therefore,

$$f(Y^\epsilon_t) - f(Y^\epsilon_0) - \int_0^t L_\epsilon f(Y^\epsilon_s) \, ds$$

is also a local martingale. □

With the stochastic differential equations for $U^\epsilon$ and $Y^\epsilon$, we can prove the following theorem.

**Theorem 4.3.2.** For fixed $t > 0$, the process $U^\epsilon_t$ converges to $B_t$ in $L^2$ as $\epsilon$ goes to zero, where $B$ is a horizontal Brownian motion.
**Proof.** Using the same notation in (3.4.4) we have \( dX_t^\epsilon = Y_t^\epsilon \, dt \). Adopting this notation, we have

\begin{equation}
(4.3.3) \quad dU_t^\epsilon = H(U_t^\epsilon) \, dX_t^\epsilon, \quad Y_t^\epsilon = e^{-1} \, dW_t - e^{-1} Y_t^\epsilon \, dt.
\end{equation}

Recall that the stochastic differential equation for horizontal Brownian motion \( B_t \) is

\[ dB_t = H(B_t) \circ dW_t, \]

where \( W \) is a standard euclidean Brownian motion. This looks very similar to the differential equation for \( U_t^\epsilon \). The major difference between these two equations is, the differential equation for \( U_t^\epsilon \) is in Ito sense, whereas the differential equation for \( B_t \) is in Stratonovich sense.

At this point, it is natural to try to prove the strong convergence using a suitable version of the Wong-Zakai theorem. However, it is difficult to find a version of the theorem to quote directly, mainly because the theorem requires the approximating process to converge to Brownian motion in a very strong sense, namely uniformly. In our case we have

\[ X_t^\epsilon = W_t - \exp(-t/\epsilon) \int_0^t \exp(s/\epsilon) \, dW_s, \]

where the error term \( V_t^\epsilon = \exp(-t/\epsilon) \int_0^t \exp(s/\epsilon) \, dW_s \). In order to quote the Wong-Zakai theorem directly, all the conditions of the theorem have to be satisfied. There are two conditions that are non-trivial in our case, both of them require \( V_t^\epsilon \) to satisfy the following condition:

\[ \max_{0 \leq t \leq T} |V_t^\epsilon| \rightarrow 0, \]
as \( \epsilon \to 0 \) for fixed \( T > 0 \). While this condition is valid, it’s a non-trivial result that needs to be justified. As a consequence, we first present a proof that relies on the fact that we have an explicit approximation of Brownian motion, namely \( X_t^\epsilon \). After that, we will prove the fact that the error term does converge to zero uniformly and improve our result.

Before we do that, we would like to remind the reader that in the multi-dimensional case, there is an extra term that involves the brackets of \( \sigma_i \)'s in the general Wong-Zakai theorem. The result is in [7]. To be more precise,

**Theorem 4.3.3** (Wong-Zakai theorem). Assume that \( B_t \) and \( U_n(t) \) satisfy the following equations

\[
\begin{align*}
    dB_t &= \sum_{i=0}^{d} \sigma_i(B_t) \circ dW^i_t + \sum_{0 \leq i \leq j \leq d} s_{ij}(t)[\sigma_i, \sigma_j](B_t) \, dt \\
    dU_n(t) &= \sigma(U_n(t)) \, dX_n(t),
\end{align*}
\]

where

\[
s_{ij}(t) = \lim_{n \to \infty} \frac{1}{2t} \mathbb{E}(\int_0^t (U_n^i(t)\tilde{U}_n^j(t) - U^i_n(t)U^j_n(t)) \, ds)
\]

and

\[
U^i_n(t) = \frac{\partial}{\partial t} U^i_n(t).
\]

Under same conditions of theorem 2.4.2, we have \( U_n(t) \to B_t \) almost surely.

In our case, this term vanishes since when \( i = j \), this term is clearly zero. Otherwise, \( U^i_n \) and \( U^j_n \) are independent hence the expectation of the product is the product of the expectations, which is also zero. By embedding the smooth manifold \( \mathcal{O}(M) \times \mathbb{R}^n \) into
a euclidean space, it is enough to deal with flat case with a general diffusion process. Nash’s embedding theorem asserts that such an embedding always exists. Let us give a precise statement of this embedding theorem.

**Theorem 4.3.4** (Nash’s embedding theorem). *Every Riemannian manifold can be isometrically embedded in some euclidean space with the standard metric.*

If we regard \( \mathcal{O}(M) \times \mathbb{R}^n \) as a submanifold of the euclidean space, we are dealing with the following problem. On one hand, we have a family of stochastic diffusion processes that satisfies the following stochastic differential equation

\[
(4.3.4) \quad dU^\epsilon_t = \sigma(U^\epsilon_t) dX^\epsilon_t.
\]

On the other hand, we have a horizontal Brownian motion given by

\[
\frac{dB_t}{\epsilon} = \sigma(B_t) \circ dW_t,
\]

where \( \sigma = H \) and \( X^\epsilon_t \) converge to \( W_t \). Our goal is to derive an explicit relationship between \( U^\epsilon_t \) and the horizontal Brownian motion \( B_t \).

Since we are working on a compact manifold and \( \sigma \) is smooth, we can assume that \( \sigma \) has bounded derivatives of all orders. We claim that

\[
\mathbb{E} |U^\epsilon_t - B_t|^2 \to 0
\]

as \( \epsilon \to 0 \).
Consider the horizontal Brownian motion $B_t$ first. We expand this Stratonovich integral to Ito type integral using the following relationship between Stratonovich integral and Ito Integral

$$M_t \circ dN_t = M_t \ dN_t + \frac{1}{2} \ d\langle M, N \rangle_t.$$ 

In Ito’s sense, horizontal Brownian motion is the solution to the following stochastic differential equation:

$$dB_t = \sigma(B_t) \ dW_t + \frac{1}{2} \sigma(B_t) \sigma'(B_t) \ dt.$$ 

Next, we expand the process $U^e$ using (4.3.4). Our goal is to show that $U^e_t$ satisfies the same stochastic differential equation as $B_t$. Using the expressions for $X^e_t$ and $V^e_t$ from last section, we have

$$dU^e_t = \sigma(U^e_t) \ dX^e_t = \sigma(U^e_t) \ dW_t - \sigma(U^e_t) \ dV^e_t.$$ 

Since $\sigma$ is smooth, we apply integration by parts for the first time and get

$$dU^e_t = \sigma(U^e_t) \ dW_t - d(\sigma(U^e_t) \ V^e_t) + V^e_t \ d\sigma(U^e_t).$$ 

We rewrite the last term in the above expression using (4.3.4) and we have the following expression for $V^e_t \ d\sigma(U^e_t)$:

$$V^e_t \ d\sigma(U^e_t) = V^e_t \sigma'(U^e_t) \ dU^e_t = \sigma'(U^e_t) \sigma(U^e_t) \ V^e_t \ dX^e_t.$$ 

Recall in the previous section, we set $H^e_t$ to be the process such that $dH^e_t = 2V^e_t \ dX^e_t$. Using the same process $H^e$ defined in the last section, the equation for $U^e_t$ could be
written as

\[ dU_t^\varepsilon = \sigma(U_t^\varepsilon) \, dW_t - d(\sigma(U_t^\varepsilon)V_t^\varepsilon) + \frac{1}{2}\sigma'(U_t^\varepsilon)\sigma(U_t^\varepsilon) \, dH_t^\varepsilon. \]

Recall that

\[ H_t^\varepsilon = t + 2 \int_0^t V_s^\varepsilon \, dW_s - (V_t^\varepsilon)^2. \]

As a result, we split the last term in (4.3.5) into three separate parts and denote the sum of all the error terms by \( R_t^\varepsilon \) to get the following expression for \( dU_t^\varepsilon \):

\[ dU_t^\varepsilon = \sigma(U_t^\varepsilon) \, dW_t + \frac{1}{2}\sigma'(U_t^\varepsilon)\sigma(U_t^\varepsilon) \, dt + dR_t^\varepsilon, \]

where

\[ dR_t^\varepsilon = -d(\sigma(U_t^\varepsilon)V_t^\varepsilon) + \sigma'(U_t^\varepsilon)\sigma(U_t^\varepsilon) V_t^\varepsilon \, dW_s - \frac{1}{2}\sigma'(U_t^\varepsilon)\sigma(U_t^\varepsilon) \, d(V_t^\varepsilon)^2. \]

Again, we use the fact that both \( \sigma \) and \( \sigma' \) are smooth and apply integration by parts for the second time to compute the last term in \( R_t^\varepsilon \). Since \( U_t^\varepsilon \) is also smooth, this is just the integration by parts for ordinary integrals and

\[ dR_t^\varepsilon = -d\{\sigma(U_t^\varepsilon)V_t^\varepsilon\} \]

\[ + \sigma'(U_t^\varepsilon)\sigma(U_t^\varepsilon) V_t^\varepsilon \, dW_s \]

\[ - d\{\frac{1}{2}(V_t^\varepsilon)^2\sigma(U_t^\varepsilon)\sigma'(U_t^\varepsilon)\} \]

\[ + \epsilon^{-1}\sigma\sigma'(\sigma\sigma')' (U_t^\varepsilon)(V_t^\varepsilon)^3 \, dt. \]
Intuitively, the error term $R^c_t$ vanishes as $\epsilon \to 0$ since the order of $V^c_t$ is $V^c_t \sim \sqrt{\epsilon}$. To make it concrete, we compute the second moment of the difference between $B_t$ and the process $U^c_t$.

First of all, the difference between $B_t$ and $U^c_t$ is

(4.3.6) \[ |B_t - U^c_t| \leq \int_0^t |\sigma(B_s) - \sigma(U^c_s)| \, dW_s + \frac{1}{2} \int_0^t |\sigma\sigma'(B_s) - \sigma\sigma'(U^c_t)| \, dt + |R^c_t|. \]

Using the fact that $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$, we have

\[
\mathbb{E}|B_t - U^c_t|^2 \leq 3\mathbb{E} \left( \int_0^t |\sigma(B_s) - \sigma(U^c_s)| \, dW_s \right)^2 + \frac{3}{2} \mathbb{E} \left( \int_0^t |\sigma\sigma'(B_s) - \sigma\sigma'(U^c_t)| \, ds \right)^2 + 3\mathbb{E}(R^c_t)^2.
\]

Using the definition of quadratic variation of the martingale $\int_0^t |\sigma(B_s) - \sigma(U^c_s)| \, dW_s$ and apply Jensen’s Inequality to the second term we get

\[
\mathbb{E}|B_t - U^c_t|^2 \leq 3 \int_0^t \mathbb{E}|\sigma(B_s) - \sigma(U^c_s)|^2 \, ds + \frac{3}{2} \int_0^t \mathbb{E}|\sigma\sigma'(B_s) - \sigma\sigma'(U^c_t)|^2 \, dt + 3\mathbb{E}(R^c_t)^2.
\]

By the assumption that both $\sigma$ and $\sigma\sigma'$ satisfy the global Lipschitz condition, there exists a constant $K$ such that

\[
\mathbb{E}|B_t - U^c_t|^2 \leq K \int_0^t \mathbb{E}|B_s - U^c_s|^2 \, ds + 3\mathbb{E}(R^c_t)^2.
\]
At this point, it is clear that we can apply the Gronwall’s lemma and bound the second moment of the difference between $B_t$ and $U^\epsilon_t$. Apply Gronwall’s lemma, we have
\[ \mathbb{E}|B_t - U^\epsilon_t|^2 \leq C\mathbb{E}(R^\epsilon_t)^2 \leq C(1 + t)\epsilon, \]
for some positive constant number $C > 0$. The proof is finished by letting $\epsilon \to 0$. □

What we have done is essentially proving the Wong-Zakai theorem for a specific case. Consequently, it is not a surprise that we can quote the Wong-Zakai theorem to prove a stronger result. In fact, the processes $U^\epsilon$ converge to a horizontal Brownian motion uniformly as $\epsilon \to 0$. In order to apply the Wong-Zakai theorem to get the uniform convergence, we need a strong version of convergence from $X_t^\epsilon$ to $W_t$. Recall
\[ X_t^\epsilon = W_t - \exp\left(-\frac{t}{\epsilon}\right) \int_0^t \exp\left(\frac{s}{\epsilon}\right) dW_s, \]
where the error term is
\[ V_t^\epsilon = \exp\left(-\frac{t}{\epsilon}\right) \int_0^t \exp\left(\frac{s}{\epsilon}\right) dW_s. \]
We have two conditions in the Wong-Zakai theorem that are not obvious at this moment. The first challenge is to find a almost surely bounded random variable $k$ independent of $t$, such that $\max_{0 \leq t \leq T} X_t^\epsilon$ is bounded by $k$ almost surely. The second challenge is to prove $X_t^\epsilon \to W_t$ uniformly, or $V_t^\epsilon \to 0$ uniformly. While these two challenges seem to be two separate problems, they actually could be solved simultaneously if the condition $V_t^\epsilon \to 0$ uniformly were true. This is because we can simply pick the random variable $k = \max_{0 \leq t \leq T} W_t + \delta$ for any small positive $\delta > 0$. Once the error term
is bounded uniformly by \(\delta\), we could have \(X^e_t\) is bounded by \(k\) for \(\epsilon\) small enough. As a result, we digress from the proof of the main theorem for a moment and prove the following lemma.

**Lemma 4.3.5.** The difference between \(X^e_t\) and the standard Brownian motion \(W_t\), namely \(V^e_t\), converges to zero uniformly as \(\epsilon\) goes to zero.

**Proof.** Recall the second moment of \(V^e_t\) is

\[
\mathbb{E}(V^e_t)^2 = \int_0^t \exp\{2(t - s/\epsilon)\} \, ds = \frac{\epsilon}{2} (1 - \exp(-2t/\epsilon)).
\]

Fix any \(T > 0\), we realize that the second moment is an increasing function of \(t\), thus if we let \(\sigma^2_T = \max_{0 \leq t \leq T} \mathbb{E}(V^e_t)^2\), we have

\[
\sigma^2_T = \frac{\epsilon}{2} (1 - \exp(-2T/\epsilon)).
\]

By an immediate consequence of the Borell-TIS inequality in [1], the supremum of a general centered Gaussian process \(V^e_t\) satisfies the following inequality. For \(u > \mathbb{E}\|V^e_t\|\), where we use \(\|V^e_t\|\) to denote \(\max_{0 \leq t \leq T} V^e_t\),

\[
\mathbb{P}(\|V^e_t\| > u) \leq \exp(-(u - \mathbb{E}\|V^e_t\|)^2/2\sigma^2_T).
\]

In order to get an estimation of \(\mathbb{E}\|V^e_t\|\) and prove the convergence result, we use Dudley inequality and Jensen’s inequality and get

\[
(\mathbb{E}\|V^e_t\|)^2 \leq C \left( \int_0^{\sqrt{\epsilon}} \sqrt{\log \sqrt{\epsilon}/\delta} \, d\delta \right)^2 \leq C' \int_0^{\sqrt{\epsilon}} \log \sqrt{\epsilon}/\delta \, d\delta = C\sqrt{\epsilon}
\]
for some positive constant number $C > 0$ and $C' > 0$. Pick $u$ in (4.3.7) to be

$$u = \mathbb{E}\|V^\epsilon_t\| + \sqrt{\epsilon \log 1/\epsilon},$$

we have

$$(4.3.8) \quad \mathbb{P}(\|V^\epsilon_t\| > \mathbb{E}\|V^\epsilon_t\| + \sqrt{\epsilon \log 1/\epsilon}) \leq \epsilon.$$

This implies that as $\epsilon \to 0$,

$$\|V^\epsilon_t\| = \max_{0 \leq t \leq T} V^\epsilon_t \to 0$$

in probability and hence $L^p$ for $\forall p > 0$ since $\mathbb{E}\|V^\epsilon_t\| + \sqrt{\epsilon \log 1/\epsilon}$ goes to zero as well.

Next, we show that $V^\epsilon_t$ converges to zero uniformly. The idea is to find a suitable $n$ large enough such that the right side of (4.3.8) becomes summable so we can apply the Borel-Cantelli lemma. For example, we can take $\epsilon_m = m^{-r}$ for some $r > 1$. Apply the Borel-Cantelli Lemma we have

$$\limsup_{m \to \infty} \frac{\|V^\epsilon_m\|}{\mathbb{E}\|V^\epsilon_m\| + \sqrt{\epsilon_m \ln 1/\epsilon_m}} \leq C,$$

for some constant $C > 0$. By an argument that is similar to the proof of theorem 3.2.4, we have

$$\limsup_{\epsilon \to 0} \frac{\|V^\epsilon_t\|}{\mathbb{E}\|V^\epsilon_t\| + \sqrt{\epsilon \ln 1/\epsilon}} \leq C.$$  

Let $\epsilon \to 0$, we see

$$\limsup_{\epsilon \to 0} \|V^\epsilon_t\| \to 0.$$  

The proof of the lemma is complete now. \qed
Having this lemma, we can apply the Wong-Zakai theorem directly to the process $U^\epsilon_t$ to get the following uniform convergence result.

**Theorem 4.3.6.** The family of processes $U^\epsilon_t$ converges to a horizontal Brownian motion uniformly (in the path space) as $\epsilon$ goes to zero.

As a consequence, the projections of $U^\epsilon_t$ converge uniformly to a Riemannian Brownian motion on compact Riemannian manifold $M$.

Lastly, When $\epsilon \to \infty$, the process $X^\epsilon_t$ becomes the uniform motion in $\mathbb{R}$ and the projections of the processes $U^\epsilon_t$ onto the manifold converges to the geodesic. This is because as $\epsilon \to \infty$,

$$dU^\epsilon_t = H_t(U^\epsilon_t) \, dt,$$

which implies $U^\epsilon_t$ is stochastic development of the uniform motion $t$ in $\mathbb{R}$. The proof of the main theorem is finally complete.
Bibliography


