The Expected Signature of a Stochastic Process

Hao Ni
Lady Margaret Hall
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity Term, 2012

This thesis is dedicated to my beloved parents.

Acknowledgements

First of all, I would like to thank my supervisor Prof. Terry Lyons, who has supported me throughout my PhD study with his patience and knowledge. He is the one who introduced Rough Paths Theory to me, and always came up with uncountable new ideas and interesting topics to me. It is great pleasure and honor for me to work with him.

In my daily work, I have been blessed with a friendly and cheerful group of Stochastic Analysis Group. In particular, I would like to express my deep appreciation to Horatio Boedihardjo for valuable discussion on the expected signature of a diffusion process and SLE. Also I have benefited from helpful suggestions on the weak convergence of expected signature of approximated Brownian motion from Weijun Xu. Moreover, thanks are also due to Zichen Zhang, Danyu Yang, Maria Tchernychova for their patiently reading my thesis and giving me useful comments and feedback on it. Furthermore I would like to extend my thanks to Oxford-Man Institute of Quantitative Finance, which has provided a nice working environment and partially funded my research.

Furthermore, I would like to thank all the friends in my life for their friendship and help. Special thanks have to go to Mitch Chen for his support and encouragement during the process of the thesis writing up.

Last but not least, I would like to express my deepest gratitude to my parents and my grandparents. I can't express how lucky and happy I feel to have such parents, who give me so much love, care and trust. They are always supporting and encouraging me to follow my heart and pursue my dream.

Abstract

The signature of the path provides a top down description of a path in terms of its effects as a control. It is a group-like element in the tensor algebra and is an essential object in rough path theory. When the path is random, the linear independence of the signatures of different paths leads one to expect, and it has been proved in simple cases, that the expected signature would capture the complete law of this random variable. It becomes of great interest to be able to compute examples of expected signatures. In this thesis, we aim to compute the expected signature of various stochastic process solved by a PDE approach. We consider the case for an Itô diffusion process up to a fixed time, and the case for the Brownian motion up to the first exit time from a domain. We manage to derive the PDE of the expected signature for both cases, and find that this PDE system could be solved recursively. Some specific examples are included herein as well, e.g. Ornstein-Uhlenbeck (OU) processes, Brownian motion and Lévy area coupled with Brownian motion.

Contents

Glossary of Basic Notations

$\mathbb{N}$: The set of all natural numbers.

$\mathbb{R}^m$: Euclidean space of dimension m, where $\mathbb{R}$ denotes the field of real numbers, and $m \in \mathbb{N}$.

$\mathbb{R}^+$: $\{x \in \mathbb{R} : x \geq 0\}$.

$\langle \cdot, \cdot \rangle$: quadratic covariation.

$\langle \cdot \rangle$: quadratic variation.

$|x| = \sqrt{\sum_{i=1}^{d} (x^{(i)})^2}$, where $x = (x^{(1)}, \ldots, x^{(d)}) \in \mathbb{R}^d$.

$|I|$: the length of a word I.

$\mathbb{D}$: the unit disk centered at the origin, i.e.

$$\mathbb{D} = \{z : |z| < 1; z \in \mathbb{R}^2\}.$$

$z + \varepsilon\mathbb{D}$: the disk of radius ε centered at z.

$\partial\Gamma$: the boundary of a domain Γ.

$\bar{\mathbb{D}}$: the closure of a domain Γ;

$dist(z, \partial\Gamma)$ is the distance from the point z to the boundary of the domain Γ, i.e.

$$dist(z, \partial\Gamma) = \inf_{y \in \partial\Gamma} \{|y - z|\}$$

where $z \in \mathbb{R}^2$ and $\Gamma \subset \mathbb{R}^2$.

$re^{i\theta} = (r\cos(\theta), r\sin(\theta)) \in \mathbb{R}^2$.

$\Delta f = \sum\limits_{i=1}^{d} \frac{\partial^2 f}{\partial x_i^2}$.

$n!! = \begin{cases} \prod_{k=0}^{\frac{n+1}{2}} (2k+1), & \text{if } n \text{ is odd;} \\ \prod_{k=1}^{\frac{n}{2}} (2k), & \text{if } n \text{ is even;} \\ 1, & \text{if } n = 0, -1. \end{cases}$

$\binom{m}{n} = \frac{m!}{n!(m-n)!}$, where $m \geq n$ and $m, n \in \mathbb{N}$.

Chapter 1

Introduction and Motivation

1.1 Introduction

The theory of rough paths can be regarded as a non-linear extension of the classical theory of controlled differential equations, which is robust enough to treat the stochastic differential equation in a deterministic way, or even controlled differential equations driven by much rougher signals than semi-martingales([14],[16]). The signature of a path is a power series whose coefficients are definite iterated integrals of the path, which is an essential object in rough path theory, and has been developed for a long time in separate fields, e.g. geometry, control theory and stochastic analysis. In the context of the controlled differential equation, the signature of a path arises naturally in Picard iteration, and it allows one to obtain a pathwise Taylor series of arbitrary order for the solution y to the vector equation

$$dy_t = f(y_t)dx_t.$$

Moreover, the importance of iterated integrals has been recognised by geometers, e.g. K.T. Chen observed the multiplicative property of the signature([2]), while the multiple Brownian iterated integral has been already studied by Wiener in the early 1930's. The signature of a path over a fixed time interval, which can be read out at the terminal time, should be a top down description of the information about the flow of events over time, recorded in its path during this time interval. As an extension of Chen's result B. Hambly and T. Lyons [6] demonstrated that it is possible to recover the whole path (up to tree-like components of the path which are not counted in its signature) by reading its signature. Z. Qian and Y. LeJan [13] proved that in dimension $d \geq 2$, almost all Brownian motion sample paths (running up to the fixed time) are determined by their Stratonovich signature over $[0, T]$.

T. Lyons realized that the key element for defining an integration theory along a continuous path is the sequence of iterated integrals, which must be specified, instead of the piecewise smoothness, and this idea was the starting point of the theory of rough paths([14]).

The expected signature of a random path plays a role similar to the moment generating function of a random variable. The expected signature of a stochastic process might capture the essential information of the measure on this stochastic process itself. For example, we can read the characteristic function of the Brownian motion and its Lévy area by reading its expected signature. The expected signature is not only a conceptually interesting object, but also has various applications in practice. For example, the computation of the expected signature of Brownian motion on $[0, 1]$ leads to "cubature on Wiener Space" methodology ([17]). As an alterative to the Monte-Carlo method, it is an effective way to evaluate certain functionals on Wiener Space and it allows us to develop high order numerical schemes valid for high dimensional SDEs and semi-elliptic PDEs. It would be of great interest to compute the expected signature of various stochastic processes. In this thesis the underlying process we are interested in is the time homogeneous Itô diffusion process. There is more than one choice to define the iterated integral of diffusion processes, e.g. the Itô integral or Stratonovich integral. But in order to define the signature of the diffusion process properly, we must specify the integral, and herein we choose the Stratonovich iterated integral. The reason for choosing the Stratonovich integral comes from the fact that the Stratonovich signature of Brownian motion is a geometric rough path([14]).

It is well known that there is a mathematical equivalence between solving parabolic partial differential equations (PDEs) and the integration of certain functionals on Wiener Space by the celebrated Feynman-Kac theorem. It is natural to ask what kind of PDE the expected signature of a diffusion process should satisfy if one is interested in computing the expected signature of the stochastic process assuming that the expected signature function is smooth enough. Moreover it would also be interesting to understand the existence and uniqueness of these particular PDEs, subject to specific boundary conditions. Furthermore we would like to study whether the solution to these PDEs, subject to suitable boundary conditions is indeed the expected signature of the process and how to solve these PDEs as well. This thesis is mainly devoted to answering these questions for diffusion processes up to a fixed time, and Brownian motion up to the first exit time from a domain. We explore the recurrence relation between higher degree terms and lower degree terms of the expected signature, and by solving the PDE with certain constraints recursively, the expected signature can be calculated, and admits a corresponding stochastic representation. Furthermore several specific examples are included, e.g. the Ornstein-Uhlenbeck (OU) process up to a fixed time, the Brownian motion up to the first exit time from a unit disk, and the Lévy area up to a fixed time or the first exit time from the unit disk.

1.2 Main contribution of the thesis

1.2.1 The expected signature of an Itô diffusion process up to a fixed time t

Let us consider a time homogeneous Itô diffusion process X taking value in $E := \mathbb{R}^d$, which could be described by the following stochastic diffusion equation (SDE) driven by a $k-$dimensional standard Brownian motion B,

$$dX_t = \mu(X_t)dt + V(X_t) \cdot dB_t \tag{1.1}$$

where $\mu : E \to E$ and $V : E \to E \otimes \mathbb{R}^k$ satisfy the globally Lipschitz condition. Itô developed a theory of integration in order to give meaning to the SDE (1.1) and obtain a solution to it given data at the starting time 0. Here we use $\cdot$ to emphasize that the integral is in the sense of the Itô stochastic integral.

We could lift the diffusion process X to a rough path, by considering the sequence of Stratonovich iterated integrals of X as its signature. Let us denote the corresponding expected signature of the process $X_{[0,t]}$ starting at x by $\Phi(t, x)$ (see Definition 4.1.3). Φ could be regarded as the function mapping from $\mathbb{R}^+ \times E$ to the tensor algebra space $T((E))$, which is defined as follows. Let $\{e_i\}_{i=1}^d$ be the canonical basis of E. Then

$$\{e_{i_1} \otimes e_{i_2} \cdots \otimes e_{i_n}\}_{i_1,\dots,i_n \in \{1,2,\dots,d\}}$$

forms the basis of $E^{\otimes n}$ and $T((E))$ is the direct sum of $E^{\otimes n}$ over $n = 0, 1, 2, \dots$.
The first main result of the thesis is summarized in the following theorem. For simplicity, here we do not specify what the regularity conditions are, which can be found in Theorem 4.7.

Theorem 1.1 *Under certain regularity conditions on μ, V and Φ, Φ should satisfy the following PDE:*

$$\left(-\frac{\partial}{\partial t}+A\right)\Phi(t,x)+\sum_{j=1}^{d}\left(\sum_{j_1=1}^{d}V^{(j_1)}(x)^T V^{(j)}(x)e_{j_1}\right)\otimes\frac{\partial\Phi(t,x)}{\partial x_j}+$$
$$\left(\sum_{j=1}^{d}\mu^{(j)}(x)e_j+\sum_{j_1=1}^{d}\frac{1}{2}\sum_{j_2=1}^{d}V^{(j_1)}(x)^T V^{(j_2)}(x)e_{j_1}\otimes e_{j_2}\right)\otimes\Phi(t,x)=0$$

with the initial condition

$$\Phi(0,x)=\mathbf{1},$$

where A is the infinitesimal generator of the process X, i.e.

$$A=\sum_{j=1}^{d}\mu^{(j)}(x)\frac{\partial}{\partial x_j}+\frac{1}{2}\sum_{j_1=1}^{d}\sum_{j_2=1}^{d}V^{(j_1)}(x)^T V^{(j_2)}(x)\frac{\partial^2}{\partial x_{j_1}x_{j_2}}.$$

This result follows from the Feynman-Kac theorem, which firstly established the bridge between the parabolic PDE and the expectation of a functional of the diffusion process at the terminal time. In our case, the situation is a bit more complicated, since except for the one dimensional case, the signature of the path $X_{[0,t]}$ depends on the whole history of the path, rather than its final point X_t. But this difficulty could be overcome by taking advantage of the multiplicative property of the signature (Chen's identity), i.e. the signature of the concatenation of two paths is equal to the tensor product of the signature of each path. Generally speaking we use the martingale approach to derive the PDE of the expected signature Φ. The main idea is that for every fixed $t\in\mathbb{R}^+$, let us consider the conditional expectation $M_u=\mathbb{E}^x[S(X_{[0,t]})|\mathcal{F}_u]$. $\{M_u\}_{u\in[0,t]}$ should be a martingale under some regularity conditions, which could be rewritten as

$$S(X_{[0,u]})\otimes\mathbb{E}^x[S(X_{[u,t]})|\mathcal{F}_u]=S(X_{[0,u]})\otimes\Phi(t-u,X_u)$$

as the result of the Markov property of the Itô diffusion process X.

By definition of the signature of $X_{[0,t]}$, it allows us to derive the SDE of $S(X_{[0,u]})$, while the SDE of $\Phi(t-u,X_t)$ could be calculated using Itô formula, assuming that $\Phi\in C^{1,2}(\mathbb{R}^+\times E,T((E)))$. It means that we are able to obtain the SDE of M_u, and by setting the drift term to be zero, the PDE of Φ could be derived.

It is interesting to notice that the PDE system of Φ could be solved recursively, and the n^{th} tensor component of Φ, denoted by Φ_n could be uniquely determined by Φ_{n-1} and Φ_{n-2}, the drift term and the martingale term of the underlying X. Moreover by the Feynman-Kac theorem, Φ_n could be represented as the expectation of the integral of the function of Φ_{n-1} and Φ_{n-2} explicitly.

1.2.2 The expected signature of the Brownian motion up to the first exit time from the domain Γ

The second part of our main result is that we derive the PDE of the expected signature of Brownian motion starting at the point $x \in \Gamma \subset E$, denoted by $\Phi_\Gamma(x)$, where Γ is a domain in E satisfying certain regularity conditions. We summarize it in the following theorem; the specific regularity conditions can be found in Theorem 5.5:

Theorem 1.2 *Suppose that the domain Γ satisfies some regularity conditions. Then Φ_Γ should satisfy the following PDE*

$$\triangle\Phi_\Gamma(z) = -(\sum_{i=1}^{2} e_i \otimes e_i) \otimes \Phi_\Gamma(z) - 2\sum_{i=1}^{2}\left(e_i \otimes \frac{\partial \Phi_\Gamma(z)}{\partial z_i}\right), \ \textit{if } z \in \Gamma$$

with the boundary condition:

$$\Phi_\Gamma(z) = \mathbf{1}, \ \textit{if } z \in \partial\Gamma$$

and the initial condition:

$$\begin{aligned} \rho_0(\Phi_\Gamma(z)) &= 1, \ \textit{if } z \in \bar{\Gamma} \\ \rho_1(\Phi_\Gamma(z)) &= 0, \ \textit{if } z \in \bar{\Gamma}. \end{aligned}$$

where $\mathbf{1} = (1, 0, 0, \dots)$, and the function

$$\begin{aligned} \rho_n : T((E)) &\rightarrow E^n \\ a = (a_0, a_1, \dots, a_n, \dots) &\mapsto a_n, \text{ where } a_i \in E^{\otimes i} \ \forall i \in \mathbb{N} \end{aligned}$$

is the projection function, which maps the tensor series to the tensor powers of degree n.

In the case of the Brownian motion up to the first exit time, we choose a deterministic approach to prove it, which is a bit different from the martingale approach used in cases of diffusion processes up to a fixed time. Actually in this case we can still use the martingale approach by constructing the martingale $M_t = \mathbb{E}[S(B_{0,\tau_\Gamma})|\mathcal{F}_t]$, but we need to take care of the situations when t tends to the infinity. This difficulty can be avoided if we use the deterministic approach, but the key to the deterministic approach is still to take advantage of the Markov property of Brownian motion and the multiplicativity of signatures. The Brownian motion path started at the point $z \in \Gamma$ and running until the first hitting time of the boundary of Γ can be seen as the concatenation of the two path, i.e. the path started at z and hitting the boundary of the disk centered at z with a sufficiently small radius ε and the second path restarted at the boundary of this disk $z + \varepsilon\mathbb{D}$, and run until the first hitting time of the boundary of Γ. By Chen's identity and the strong Markov property of Brownian motion, we have that

$$\Phi_\Gamma(z) = \mathbb{E}^z[S(B_{[0,\tau_{z+\varepsilon\mathbb{D}}]}) \otimes \Phi(B_{\tau_{x+\varepsilon\mathbb{D}}})].$$

Actually the dimension of the Brownian motion is not important in this case. For simplicity, we illustrate the idea of the proof in the two dimensional case. In this case, $\Phi_\Gamma(z)$ could be rewritten in the integral form:

$$\begin{aligned} \Phi_\Gamma(z) &= \int_0^{2\pi} \mathbb{E}^x[S(B_{[0,\tau_{z+\varepsilon\mathbb{D}}]})|B_{\tau_{x+\varepsilon\mathbb{D}}} = z + \varepsilon] \otimes \Phi_\Gamma(z + \varepsilon e^{i\theta})\frac{d\theta}{2\pi} \\ &= \int_0^{2\pi} \Psi(\varepsilon e^{i\theta}) \otimes \Phi_\Gamma(z + \varepsilon e^{i\theta})\frac{d\theta}{2\pi} \end{aligned} \tag{1.2}$$

where $\Psi(z) = \mathbb{E}^x[S(B_{[0,\tau_{\varepsilon\mathbb{D}}]})|B_{\tau_{\varepsilon\mathbb{D}}} = z]$ and the last inequality follows from the translation invariance of Brownian motion.

Suppose that the function Φ_Γ is sufficiently smooth and could be written as a Taylor expansion. The right hand side(RHS) of the above equation could be written as a power series in ε, while the left hand side (LHS) of the above equation is independent of ε. This means that the coefficient of the first term in ε on the RHS should be zero, which gives us the PDE to characterize the function Φ. In the thesis we present this approach in a slightly different way by introducing the auxiliary function φ:

$$\begin{aligned} \varphi_\Gamma : z + \varepsilon\mathbb{D} &\to T((E)); \\ y &\mapsto \Psi(y) \otimes \Phi_\Gamma(z + y). \end{aligned}$$

Equation (1.2) is equivalent to saying that the function φ satisfies the mean value property at the origin, which implies

$$\triangle\varphi(y)_{y=0} = 0$$

and thus the PDE of Φ is derived from it immediately.

Similar to the case for Itô diffusion processes up to a fixed time, the PDE of Φ_Γ can be solved recursively, and the n^{th} term of Φ_Γ is uniquely determined by the previous two terms. By the PDE theorem, the solution can be written in integral form or admits a stochastic representation under regularity conditions.

1.3 The outline of the thesis

The outline of the thesis is as follows:

In Chapter 2, we give some preliminaries of Rough Path Theory. In particular we introduce the definition of the signature of paths of bounded variation and the algebraic properties of signatures. Then we give the definition of the expected signature of a stochastic process. We attempt to explain the motivation behind studying the expected signature of a random path.

In Chapter 3, we recall the existing result on the expected signature of Brownian motion up to a fixed time t using the PDE approach firstly. We consider the piecewise linear approximation of Brownian motion, and examine the convergence rate of the expected signature of such a discretized Brownian motion path.

In Chapter 4, following the PDE approach to calculate the expected signature of Brownian motion, we derive the PDE of the expected signature of a general Itô diffusion process up to a fixed time. Moreover we consider OU processes as examples, and give explicit solutions in those cases. Furthermore using our PDE theorem, we manage to derive the characteristic function of the Lévy area at a fixed time t on condition that the starting point of the Brownian motion is x, which is an alternative way to calculate the characteristic function of the Brownian motion coupled with its Lévy area on the condition that the Brownian motion starts at the origin.

In Chapter 5, we consider the domain Γ, which satisfies some regularity conditions. We are interested in the expected signature of the Brownian motion starting at $x \in \Gamma$ up to the first exit time from the domain Γ. We derive the PDE of the expected signature of such stopped Brownian motion. By solving the PDE recursively we obtain the expected signature. We also derive the finite difference equations for the expected signature of a simple random walk up to the first exit time from the domain. Then we consider the unit disk as a concrete example, and prove that the expected signature of the stopped Brownian motion up to the first exit time from the unit disk should be in the polynomial form. In a similar approach to that of the previous chapter, we examine the moment generating function and characteristic function of the Lévy area at the first exit time from the unit disk conditional on the position of the starting point of Brownian motion, and give the stochastic representation of those two functions.

In Chapter 6, we briefly summarize our work and point out future work.

Chapter 2

Rough Paths Preliminaries

In this thesis, we are interested in paths and in describing probability measures on these paths. Our examples will be quite specific, but need to be understood in a more general context. For the sake of precision, we start by introducing some notation, making essential definitions and stating the basic results we require. These can also be found in [14]. A reader experienced in rough path theory might prefer to go directly to Chapter 3.

2.1 Tensors products

Fix a real Banach space E, in which our paths will take their values. Consider the successive tensor powers $E^{\otimes n}$ of E (completed in some tensor norm). In the case where E is finite dimensional with basis $\{\alpha_1, \ldots, \alpha_d\}$, $E^{\otimes n}$ has the basis $\{\alpha_{i_1} \otimes \alpha_{i_2} \otimes \ldots \otimes \alpha_{i_n}\}$, where $i_1, \ldots, i_n \in \{1, \ldots, d\}$ and can be identified with the space of real homogeneous non-commuting polynomials of degree n in d variables. It is easy to see that $E^{\otimes}$ is isomorphic to the space spanned by indexes of length n in the numbers $\{1, \ldots, d\}$. Let us give the formal definition of words as below.

Definition 2.1.1 *Let us consider a finite set A, which we called an alphabet, whose elements are numbers $1, 2, \ldots, d$, and an index (a word) I in the alphabet A is either the empty word denoted by $()$ or it can be written as*

$$I = (i_1, i_2, \ldots, i_n)$$

which is a finite sequence of the numbers i_1, $i_2, \ldots, i_n$ from the alphabet A, where n is called the length of the word I. If I is an empty word, then its length is zero.

We note that $E^{\otimes 0} = \mathbb{R}$. It is necessary to constrain the norms used when considering tensor products (the injective and projective norms satisfy our constraints).

Definition 2.1.2 *Let E be a Banach space. We say that its tensor powers are endowed with an admissible norm $|.|$, if the following conditions hold:*

1. *For each $n \geq 1$, the symmetric group S_n acts by isometry on $E^{\otimes n}$, i.e.*

$$|\sigma v| = |v|, \ \forall v \in E^{\otimes n}, \forall \sigma \in S_n.$$

2. *$\forall n, m \geq 1$,*

$$|v \otimes w| \leq |v||w|, \forall v \in E^{\otimes n}, w \in E^{\otimes m}.$$

2.2 The algebra of tensor series

Definition 2.2.1 *A formal* E*-tensor series is a sequence of tensors $\left(a_n \in \mathrm{E}^{\otimes n}\right)_{n\in\mathbb{N}}$, denoted by $a = (a_0, a_1, \ldots)$. There are two binary operations on* E*-tensor series, an addition $+$ and a product $\otimes$, which are defined as follows. Let $\mathbf{a} = (a_0, a_1, ...)$ and $\mathbf{b} = (b_0, b_1, ...)$ be two* E*-tensor series. Then we define*

$$\mathbf{a} + \mathbf{b} = (a_0 + b_0, a_1 + b_1, ...) \tag{2.1}$$

and

$$\mathbf{a} \otimes \mathbf{b} = (c_0, c_1, ...), \tag{2.2}$$

where for each $n \geq 0$,

$$c_n = \sum_{k=0}^{n} a_k \otimes b_{n-k}. \tag{2.3}$$

The product $\mathbf{a}\otimes\mathbf{b}$ is also denoted by $\mathbf{ab}$. We use the notation $\mathbf{1}$ for the series $(1, 0, ...)$, and $\mathbf{0}$ for the series $(0, 0, ...)$. If $\lambda \in \mathbb{R}$, then we define $\lambda\mathbf{a}$ to be $(\lambda a_0, \lambda a_1, ...)$.

Definition 2.2.2 *The space $T((E))$ is defined to be the vector space of all formal E-tensors series.*

Remark 2.2.1 *The space $T((E))$ with $+$ and $\otimes$ and the action of $\mathbb{R}$ is an associative and unital algebra over $\mathbb{R}$. An element of $\mathbf{a} = (a_0, a_1, ...)$ of $T((E))$ is invertible if and only if $a_0 \neq 0$. Its inverse is then given by the series*

$$\mathbf{a}^{-1} = \frac{1}{a_0} \sum_{n\geq 0} (\mathbf{1} - \frac{\mathbf{a}}{a_0})^n \tag{2.4}$$

which is well defined because, for each given degree, only finitely many terms of the sum produce non-zero tensors of this degree. In particular, the subset $\{\mathbf{a} \in T((E)) \,|a_0 = 1\}$ forms a group.

Definition 2.2.3 *The dilation operator is the mapping* $\langle \varepsilon, \mathbf{a} \rangle$ *from* $\mathbb{R}^+ \times T((E))$ *to* $T((E))$ *defined by*

$$\langle \varepsilon, \mathbf{a} \rangle = (a_0, \varepsilon a_1, ..., \varepsilon^n a_n, ...), \ \forall \varepsilon \in \mathbb{R}^+, \mathbf{a} \in T((E)).$$

Let $B_n = \{\mathbf{a} = (a_0, a_1, ...)| a_0 = ... = a_n = 0\}$. Then B_n is an ideal of $T((E))$.

Definition 2.2.4 *Let* $n \geq 1$ *be an integer. The truncated tensor algebra* $T^{(n)}(E)$ *of order* n *over* E *is defined as the quotient algebra*

$$T^{(n)}(E) = T((E))/B_n. \tag{2.5}$$

The canonical homomorphism $T((E)) \longrightarrow T^{(n)}(E)$ *is denoted by* π_n.

2.3 Paths

Let us define the p-variation of a E-valued path X on an interval J in $\mathbb{R}^+$.

Definition 2.3.1 *Let* $p \geq 1$ *be a real number. Let* $X : J \to E$ *be a continuous path. The* p*-variation of* X *on the interval* J *is defined by*

$$||X||_{p,J} = \left[\sup_{\mathcal{D} \subset J} \sum_{j=0}^{r-1} |X_{t_j} - X_{t_{j+1}}|^p \right]^{\frac{1}{p}}$$

where $\mathcal{D}$ *is a subdivision of* J*, i.e.* $\mathcal{D} = \{t_0, t_1, \ldots, t_r\}$ *such that* $0 \leq t_0 < \cdots < t_r$ *and* $(t_0, t_r) \subset J$.

Paths also have algebraic properties:

Definition 2.3.2 *Let* $X : [0, s] \longrightarrow E$ *and* $Y : [s, t] \longrightarrow E$ *be two continuous paths. Their concatenation is the path* $X * Y$ *defined by*

$$(X * Y)_u = \begin{cases} X_u, & \text{if } u \in [0, s]; \\ X_s + Y_u - Y_s, & \text{if } u \in [s, t]. \end{cases}$$

Remark 2.3.1 $*$ *is an associative operation. Let* $X : [0, s] \longrightarrow E$*,* $Y : [s, t] \longrightarrow E$ *and* $Z : [t, v] \longrightarrow E$ *be three continuous paths. Then*

$$(X * Y) * Z = X * (Y * Z).$$

2.4 The signature of a path

Definition 2.4.1 *Let J denote an interval in $\mathbb{R}^+$. Let $X : J \longrightarrow E$ be a path of bounded variation or a rough path of finite $p-$variation such that the following integration makes sense. The signature $\mathbf{X}$ of X is an element $(1, X^1, ..., X^n, ...)$ of $T((E))$ defined for each $n \geq 1$ as follows*

$$X^n = \int_{u_1 < ... < u_n, \quad u_1,...,u_n \in J} \cdots \int dX_{u_1} \otimes ... \otimes dX_{u_n}.$$

The signature $\mathbf{X}$ is also denoted by $S(X)$ and the truncated signature of X of order n is denoted by $S^n(X)$, i.e $S^n(X) = \pi_n(\mathbf{X})$ for every $n \in \mathbb{Z}$.

Definition 2.4.2 *If the interval $I \subset J$ is a subinterval then we will use the notation $\mathbf{X}_I$ to denote the signature of X over the interval I. If $J = [S, T]$ then we use the notation Δ_J for $\{(s,t) \,|S \leq s \leq t \leq T\}$ and the notation $\mathbf{X}_{s,t}$,for $\mathbf{X}_{[s,t]}$ and $S^N(X)_{s,t}$ for $S^N(X)_{[s,t]}$ where $(s,t) \in \Delta_J$.*

Example 2.4.1

1. *If X_t is a continuous path with finite $p-$variation, where $1 \leq p < 2$, then its signature can be defined in the sense of the Young integral[14]. More generally, if X_t is a p-rough path ($p \geq 1$) then the integrals will exist as a result of the general theory of rough paths[extension theorem] ([14]).*

2. *For a Brownian path, its signature can be defined in the sense of either the Itô integral or the Stratonovich integral. Almost all Brownian paths, with their Lévy area processes, are p-rough paths for any $p > 2$. For all (s,t), with probability one, the Stratonovich signature is a geometric rough path ([14] Page 57).*

The following fundamental theorem, Chen's identity, asserts that the signature is a homomorphism between the path space and its rough path space([14]).

Theorem 2.1 *Let $X : [0, s] \longrightarrow \mathrm{E}$ and $Y : [s, t] \longrightarrow \mathrm{E}$ be two continuous paths with finite $p-$variation where $1 \leq p < 2$. Then*

$$S(X * Y) = S(X) \otimes S(Y). \tag{2.6}$$

The proof can be found in [14].

The following lemma presents the reason why the signature of a path can be described as the full non-commutative exponential of the path.

Lemma 2.4.1 *Let $X : [0,T] \to E$ be a path with finite $1-$variation. The signature $S(X_{0,\cdot})$ as a function from $[0,T]$ to $T((E))$ satisfies the following differential equation:*

$$dS(X_{[0,t]}) = S(X_{[0,t]}) \otimes dX_t, \;\; S(X_{[0,0]}) = 1.$$

2.5 Real valued functions from the range of the signature

Consider the linear forms on $T\left((E)\right)$(see [14]). Suppose $(e_1^*, ..., e_n^*)$ are elements of the dual space of E, denoted by E^*. We can introduce coordinate iterated integrals by setting

$$X_u^{(i)} := \langle e_i^*, X_u \rangle$$

and rewriting $\left\langle e_{i_1}^* \otimes ... \otimes e_{i_n}^*, S\left(X\right)\right\rangle$ as the scalar iterated integral of coordinate projections

$$\underset{\substack{u_1<...<u_n \\ u_1,...,u_n \in J}}{\int \cdots \int} dX_{u_1}^{(i_1)} \otimes \ldots \otimes dX_{u_n}^{(i_n)}.$$

For example, if $(e_1, ..., e_d)$ is a basis for a finite dimensional space E, and $(e_1^*, ..., e_d^*)$ is a basis for the dual space E^*. Then the elements $(e_I = e_{i_1} \otimes ... \otimes e_{i_n})_{I=(i_1,...,i_n)\in\{1,...,d\}^n}$ form a basis of $E^{\otimes n}$. The space $(E^*)^{\otimes n}$ with the basis $(e_I^* = e_{i_1}^* \otimes ... \otimes e_{i_n}^*)_{I=(i_1,...,i_n)\in\{1,...,d\}^n}$ can be canonically identified with the dual basis of $(E^{\otimes n})^*$ so that

$$\langle e_{i_1}^* \otimes ... \otimes e_{i_n}^*, e_{j_1} \otimes ... \otimes e_{j_n} \rangle = \delta_{i_1,j_1} ... \delta_{i_n,j_n},$$

where $\delta_{i,j} = \mathbf{1}_{\{i=j\}}$ and the linear action of $(E^*)^{\otimes n}$ on $E^{\otimes n}$ extends naturally to a linear mapping $(E^*)^{\otimes n} \to T\left((E)\right)^*$, defined by

$$e_I^*(\mathbf{a}) = e_I^*(a_n).$$

In this case we can also write

$$\mathbf{X}_J = \sum_{n=0}^{+\infty} \sum_{i_1,...i_n \in \{1,...d\}} \left(\underset{\substack{u_1<...<u_n \\ u_1,...,u_n \in J}}{\int \cdots \int} dX_{u_1}^{(i_1)} \otimes \ldots \otimes dX_{u_n}^{(i_n)} \right) e_{i_1} \otimes \ldots \otimes e_{i_n}.$$

Remark 2.5.1 *Re-parameterizing a path inside the interval of definition does not change its signature over the maximal interval. Translating a path does not change its signature. We may define an equivalence relation between paths of bounded variation by asking that they have the same signature.*

Definition 2.5.1 *If E is a Banach space then we define $T(E) \subset T((E))$ to be the tensor series $\mathbf{a} = (a_0, a_1, ...)$ for which there is an N depending on $\mathbf{a}$ so that $a_i = 0$ for all $i > N$. In other words, it is the space of polynomials (instead of series) in E(the multiplication is in the sense of tensor product).*

Remark 2.5.2 *There is a natural inclusion*

$$T(E^*) \to T((E))^*.$$

Returning to the finite dimensional example and the notation introduced above, the linear forms e_I^*, as I spans the set of finite words in the numbers $1, ..., d$, form a basis for $T(E^*)$.

Now let us consider two elements of $T(E^*)$, say $\mathbf{e}^*$ and $\mathbf{f}^*$. Their pointwise product as real-valued functions is a quadratic form on $T((E))$. Then there exists a unique third element of $T((E))$, denoted by $\mathbf{e}^* \sqcup\!\sqcup \mathbf{f}^*$, such that $\mathbf{e}^*\mathbf{f}^*$ and $\mathbf{e}^* \sqcup\!\sqcup \mathbf{f}^*$ coincide on the range of S.

Theorem 2.2 *The linear forms on $T((E))$ induced by $T(E^*)$, when restricted to the range $S(\mathcal{V}([0,T], E))$ of the signature, form an algebra of real valued functions, where $\mathcal{V}([0,T], E)$ is used to denote the set of all the continuous paths mapping $[0,T]$ to E of finite $1-$variation.*

The proof of this theorem could be found on Page 34 in [14]. By linearity it is sufficient to consider the case for $\mathbf{e}^* = e_I^*$ and $\mathbf{f}^* = e_J^*$, where $I = (i_1, \ldots, i_m)$ and $J = (j_1, \ldots, j_n)$ are two arbitrary words. It is shown that the linear form

$$e_I^* \sqcup\!\sqcup e_J^* = \sum_{\sigma \in S_{m,n}} e^*_{k_{\sigma^{-1}(1)}, \ldots, k_{\sigma^{-1}(m+n)}} \tag{2.7}$$

is equal to $e_I^* e_J^*$ on the subset $S(\mathcal{V}([0,T], E))$, where the set $S_{m,n}$ is defined below in Definition 2.5.2. The linear form $e_I^* \sqcup\!\sqcup e_J^*$ defined by (2.7) is called the *shuffle product* of e_I^* and e_J^*. By bilinearity, this product endows $T(E^*)$ with the structure of an algebra, known as the shuffle algebra. Now we defined a product on $T((E^*))$ and show that for every $\mathbf{e}^*, \mathbf{f}^* \in T(E^*)$, and for every $\mathbf{a} \in S(\mathcal{V}([0,T], E))$,

$$\mathbf{e}^*(\mathbf{a})\mathbf{f}^*(\mathbf{a}) = (\mathbf{e}^* \sqcup\!\sqcup \mathbf{f}^*)(\mathbf{a}).$$

Definition 2.5.2 *We define the set $S_{m,n}$ of (m,n) shuffles to be the subset of permutation in the symmetric group S_{m+n} defined by*

$$S_{m,n} = \{\sigma \in S_{m+n} : \sigma(1) < \cdots < \sigma(m), \sigma(m+1) < \cdots < \sigma(m+n)\}.$$

Remark 2.5.3 *For convenience, we introduce one useful functions on $T((E))$, i.e. π^I is defined by*

$$\begin{aligned} \pi^I : T((E)) &\to \mathbb{R}; \\ \mathbf{a} &\mapsto e_I^*(\mathbf{a}). \end{aligned}$$

In particular,

$$\pi^I(S(X)) = \underset{\substack{u_1<\ldots<u_n \\ u_1,\ldots,u_n \in J}}{\int \cdots \int} dX_{u_1}^{(i_1)} \otimes \ldots \otimes dX_{u_n}^{(i_n)}$$

where $I = (i_1, ..., i_n)$.

2.6 The expected signature of stochastic processes

Definition 2.6.1 *Given a probability space $(\Omega, P, \mathcal{F})$, X is a stochastic process which takes value in E. Suppose that for a.e. every $\omega \in \Omega$, the signature of $X(\omega)$ denoted by $S(X(\omega))$ is well defined and under the probability measure P, its expectation denoted by $\mathbb{E}[S(X)]$ is finite. We call $\mathbb{E}[S(X)]$ the expected signature of X.*

Remark 2.6.1 *The expected signature of a random path plays a similar role to the moment generating function of a random variable. The expected signature of the random process captures the important information about the measure on the underlying paths as well. At least it can be proved that the expected signature completely describe the law of the random signature in some simple cases, which is stated in the following theorem.*

Theorem 2.3 *Let $\{X_i\}_{i=1}^n$ be a finite collection of paths of finite variation, which have pairwise distinct signatures, ie. for every $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$, $S(X_i) \neq S(X_j)$. $\{\lambda_i\}_{i=1}^n$ satisfies $\lambda_i \geq 0, \forall i \in \{1, 2, .., n\}$ and $\sum_{i=1}^n \lambda_i = 1$. Then for every $i \in \{1, 2, \ldots, n\}$, there exists a linear form $\sigma_i : T((E)) \to \mathbb{R}$ such that*

$$\sigma_i(S(X_j)) = \delta_{i,j}$$

where $j = 1, 2, \ldots, n$.

The proof of this theorem could be found on Page 35 of [14], where the Stone-Weierstrass theorem is used to prove this statement. But in the following, I provided a slightly different approach to prove it by constructing the linear function σ_i explicitly, which is more straightforward. The underlying idea of my proof is mainly related to Lagrange polynomials.

Proof. We are going to construct a linear form σ_i such that

$$\sigma_i(S(X_j)) = \delta_{i,j}.$$

For any $i \in \{1, 2, ..., n\}$, for each $j \neq i$, there exists an index $I_{i,j}$ such that

$$\pi^{I_{i,j}}(S(X_i)) \neq \pi^{I_{i,j}}(S(X_j))$$

where $I_{i,j}$ is an index composed of the numbers $1, 2, \ldots, d$.
Let us define the linear forms σ_i as follows:

$$\sigma_i = \underset{j\neq i}{\sqcup\!\sqcup} (a_{i,j}\pi^{()} + b_{i,j}\pi^{I_{i,j}}) \tag{2.8}$$

where () is the empty index and the product is in the shuffle product sense, i.e. for every two indices I and J, which are indexes composed of the numbers $\{1, 2, \ldots, d\}$,

$$\begin{aligned} \pi^I\pi^J : T((E)) &\rightarrow \mathbb{R} \\ x &\mapsto (e_I^* \sqcup\!\sqcup\, e_J^*)(x) \end{aligned}$$

and

$$\begin{aligned} a_{i,j} &= \frac{\pi^{I_{i,j}}(S(X_j))}{\pi^{I_{i,j}}(S(X_j)) - \pi^{I_{i,j}}(S(X_i))}, \\ b_{i,j} &= \frac{-1}{\pi^{I_{i,j}}(S(X_j)) - \pi^{I_{i,j}}(S(X_i))}. \end{aligned}$$

By the shuffle product property of the signature, i.e. the multiplication of two linear forms on the signature is still a linear form on the signature, Equation (2.8) could be expressed as the linear form on the signature as well. Moreover for $i \neq j$, since

$$\pi^{I_{i,j}}(S(X_i)) \neq \pi^{I_{i,j}}(S(X_j)),$$

$a_{i,j}$ and $b_{i,j}$ are well defined.
One the other hand, it holds that

$$\sigma_i(S(X_l)) = \prod_{j\neq i} \frac{\pi^{I_{i,j}}(S(X_j)) - \pi^{I_{i,j}}(S(X_l))}{\pi^{I_{i,j}}(S(X_j)) - \pi^{I_{i,j}}(S(X_i))}.$$

If $i = l$, $\sigma_i(S(X_l)) = 1$ obviously; otherwise, there exists $j = l$, such that

$$\frac{\pi^{I_{i,j}}(S(X_j)) - \pi^{I_{i,j}}(S(X_l))}{\pi^{I_{i,j}}(S(X_j)) - \pi^{I_{i,j}}(S(X_i))} = 0$$

and thus $\sigma_i(S(X_l)) = 0$. The proof is complete now. ■

Remark 2.6.2 *We could restate this theorem in the probabilistic setting. Let us consider the probability space $(\Omega, P, \mathcal{F})$, where Ω is a finite collection of paths, i.e. $\{X_i\}_{i=1}^n$, and $\mathcal{F} = 2^{\Omega}$. Under the probability measure P,*

$$P(X = X_i) = \lambda_i, \forall i = 1, 2, ..., n$$

where $\lambda_i \geq 0$ and $\sum_{i=1}^n \lambda_i = 1$. Then the expected signature of X is the weighted sum of the signature of each path, i.e.

$$\mathbb{E}[S(X)] = \sum_{i=1}^{n} \lambda_i S(X_i).$$

Since $\{X_i\}_{i=1}^n$ has pairwise different signatures, from the above theorem, we know how to derive the linear form σ_i from the signatures of all paths, i.e. $\{S(X_i)\}_{i=1}^n$ and then we can compute the measure of the signatures of the paths from the expected signature; more precisely

$$\lambda_i = \sigma_i(\mathbb{E}[S(X)]) \tag{2.9}$$

where $i = 1, 2, \ldots, n$.

Chapter 3

The Expected Signature of Planar Brownian Motion up to a Fixed Time

3.1 The derivation of expected signature of Brownian motion up to a fixed time

For a fixed $T \in \mathbb{R}^+$, let us consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega = C_0([0,T], \mathbb{R}^d)$ is the space of $\mathbb{R}^d$-valued continuous functions defined on $[0,T]$ starting at 0, $\mathcal{F}$ is its Borel σ-field induced by the supremum norm and $\mathbb{P}$ is the Wiener measure. Now define the coordinate mapping process $B^{(i)}(t) = \omega^i(t)$ for $t \in [0,T], \omega \in \Omega$, where $i \in \{1, 2, \dots, d\}$. Then, under the Wiener measure, $B_t = (B_t^{(1)}, B_t^{(2)}, \dots, B_t^{(d)})$ is a standard d-dimensional Brownian motion starting at 0. We endow $\mathbb{R}^d$ with the canonical basis $(e_1, e_2, \dots, e_d)$, and the Brownian motion B_t could be written in the following form:

$$B_t = \sum_{i=1}^{d} B_t^{(i)} e_i$$

where $t \geq 0$.

Definition 3.1.1 *The signature of the path $B_{[0,t]} = \{B_s : s \in [0,t]\}$ is defined to be*

$$S\left(B_{[0,t]}\right) = (1, B^1_{[0,t]}, ..., B^n_{[0,t]}, ...) \in T\left((E)\right)$$

defined for almost every path B and each $n \geq 1$ by the Stratonovich iterated integral

$$B^n_{[0,t]} = \int\limits_{u_1 < ... < u_n,} \cdots \int\limits_{u_1, ..., u_n \in [0,t]} dB_{u_1} \otimes ... \otimes dB_{u_n}.$$

Definition 3.1.2 (Expected signature of Brownian motion up to t) *The expected signature of a Brownian motion up to time t, denoted by $\varphi(t)$, is defined as*

$$\varphi(t) = \mathbb{E}[S(B_{[0,t]})] \tag{3.1}$$

where $t \in \mathbb{R}^+$.

The following result is the explicit formula for the expected signature of $d-$dimensional Brownian motion up to a fixed time T, which was first understood by T. Fawcett [3].

Theorem 3.1 *Let B_t be a standard $d-$dimensional Brownian motion.*

$$\varphi(t) = \mathbb{E}[S(B_{[0,t]})] = \exp\left(\frac{t}{2}\sum_{i=1}^{d} e_i \otimes e_i\right).$$

Remark 3.1.1 *There are several possible approaches to prove this theorem. One way is to calculate explicitly by induction*

$$\pi^I(\varphi(t)) = \mathbb{E}\left[\int_{0<u_1<u_2<\cdots<u_n<T} \circ dB_{u_1}^{(i_1)} \circ dB_{u_1}^{(i_1)} \circ \cdots \circ dB_{u_n}^{(i_n)}\right]$$

where $I = (i_1, i_2, \ldots, i_n)$, and for every $j \in \{1, 2, \ldots, n\}$, $i_j \in \{1, 2, \ldots, d\}$.
A prettier approach uses the relationship between PDEs and the expectations of SDEs. The proof can be found in [17].

3.2 The expected signature of the equally spaced piecewise linear approximations of Brownian motion

3.2.1 Piecewise linear approximation of Brownian motion

Let us consider a standard $d-$dimensional Brownian motion

$$B_t = \sum_{i=1}^{d} B_t^{(i)} e_i.$$

Fix the time interval $[0, T]$ and a positive integer M. Let $\mathcal{D}_M = \{t_j\}_{j=0}^{M}$ denote an equally spaced partition of the time interval $[0, T]$, i.e. for every $j \in \{0, 1, \ldots, M\}$,

$$t_j = j\Delta t$$

where $\Delta t = \frac{T}{M}$.
Let us consider the piecewise linear approximation to the Brownian motion $B_{[0,T]}$

based on the partition $\mathcal{D}_M$, denoted by $\hat{B}^M_{[0,T]}$. More precisely, $\{\hat{B}^M_t\}_{t\in\mathcal{D}_M}$ has the same distribution as $\{B_t\}_{t\in\mathcal{D}_M}$. It means that at each j^{th} time step, the increment of $\hat{B}^M$ are mutually independent and identically distributed (i.i.d) over j, i.e.

$$\begin{pmatrix} \Delta\hat{B}_{t_1} \\ \Delta\hat{B}_{t_2} \\ \dots \\ \Delta\hat{B}_{t_{M-1}} \end{pmatrix} \sim \mathcal{N}(0, \Delta t \mathcal{I}_M)$$

where $\Delta\hat{B}_{t_j} := \hat{B}^M_{t_{j+1}} - \hat{B}^M_{t_j}$ for $j = 0, 1, \dots, M-1$, $\mathcal{N}(0, \Delta t\mathcal{I}_M)$ denotes the normal distribution with expected value 0 and covariance matrix $\Delta t\mathcal{I}_M$, where $\mathcal{I}_M$ is the $M \times M$ identity matrix.

Let us recall that the function $\varphi(.)$ is defined as the expected signature of Brownian motion, i.e.

$$\begin{aligned} \varphi : [0,T] &\to T((E)) \\ t &\mapsto \mathbb{E}[S(B_{[0,t]})]. \end{aligned}$$

For convenience, we introduce the function $\varphi^M(.)$ to denote the expected signature of the linear approximation process $\hat{B}^M_{[0,t]}$, i.e.

$$\begin{aligned} \varphi^M : [0,T] &\to T((E)) \\ t &\mapsto \mathbb{E}[S(\hat{B}^M_{[0,t]})]. \end{aligned}$$

Now we are interested in the convergence rate of the expected signature of $\hat{B}^M_{[0,T]}$, i.e. $\varphi^M(T)$ to $\varphi(T)$, as M tends to the infinity.

To do that, first of all, let us introduce the definition of an even word and a square word.

Definition 3.2.1 *Let I be a word composed of the numbers $1, 2, ..., d$. For every $i \in \{1, 2, \dots, d\}$, $n_i(I)$ denotes the number of occurrence of i in the word I, i.e*

$$n_i(I) = \sum_{k=1}^{n} \mathbf{1}_{\{i_k = i\}}$$

where $I = (i_1, \dots, i_n)$.
If for all $i \in \{1, 2, \dots, d\}$, $n_i(I)$ are all even, then the word I is called an even word. The set of all the even words of length $2n$ is denoted by $\mathcal{K}_{2n}$.

Definition 3.2.2 *Let I be a word of length $2n$, which is composed of the numbers $1, 2, ..., d$. If the word I could be written as*

$$I = (i_1, i_1, i_2, i_2, \dots, i_n, i_n)$$

where $i_j \in \{1, 2, \dots, d\}$ for $j = 1, 2, \dots, n$, then the word I is called a square word. The set of all the square words of length $2n$ is denoted by $\mathcal{S}_{2n}$.

Remark 3.2.1 *For every $n \in \mathbb{N}$,*

$$\mathcal{S}_{2n} \underset{\neq}{\subset} \mathcal{K}_{2n}.$$

There is an equivalent statement of Theorem 3.1 as follows.

Corollary 3.2.1 *For every index I, if there exists an integer $n \geq 0$, such that $I \in \mathcal{S}_{2n}$, then*

$$\pi^I(\varphi(T)) = \left(\frac{T}{2}\right)^n \frac{1}{n!};$$

otherwise,

$$\pi^I(\varphi(T)) = 0.$$

Proof. Recall Theorem 3.1, $\varphi(T)$ is equal to

$$\exp(\sum_{i=1}^{d} \frac{T}{2} e_i \otimes e_i) = \sum_{n \geq 0} (\sum_{i=1}^{d} \frac{T}{2} e_i \otimes e_i)^{\otimes n} \frac{1}{n!}.$$

From the above expansion, the result follows immediately. ■

3.2.2 The formula for the expected signature of the equally spaced piecewise linear approximations of Brownian motion

In this subsection, we aim to derive an explicit formula for the expected signature of $\hat{B}^M_{[0,T]}$, i.e. $\varphi^M(T)$. Before doing so, it is important to observe that the functions φ and φ^M both have the multiplicative property, which is stated in the following lemma.

Lemma 3.2.1 *For any $k_1, k_2 \in \{1, 2, \ldots, M\}$ and $k_2 \geq k_1$,*

$$\begin{aligned}
\varphi(k_2\Delta t) &= \varphi(k_1\Delta t) \otimes \varphi((k_2 - k_1)\Delta t); \\
\varphi^M(k_2\Delta t) &= \varphi^M(k_1\Delta t) \otimes \varphi^M((k_2 - k_1)\Delta t).
\end{aligned}$$

Proof. *Recall that $\varphi(t)$ is defined to be the expected signature of Brownian motion $B_{[0,t]}$, i.e.*

$$\varphi(k_2\Delta t) = \mathbb{E}[S(B_{[0,k_2\Delta t]})] = \mathbb{E}[\mathbb{E}[S(B_{[0,k_2\Delta t]})|\mathcal{F}_{k_1\Delta t}]].$$

By Chen's identity, we have that

$$S(B_{[0,k_2\Delta t]}) = S(B_{[0,k_1\Delta t]}) \otimes S(B_{[k_1\Delta t,k_2\Delta t]})$$

thus the following equalities holds.

$$\begin{aligned}
& \varphi(k_2\Delta t) \\
= \; & \mathbb{E}[\mathbb{E}[S(B_{[0,k_1\Delta t]}) \otimes S(B_{[k_1\Delta t,k_2\Delta t]})|\mathcal{F}_{k_1\Delta t}]], \text{ by the Tower property} \\
= \; & \mathbb{E}[S(B_{[0,k_1\Delta t]}) \otimes \mathbb{E}[S(B_{[k_1\Delta t,k_2\Delta t]})|\mathcal{F}_{k_1\Delta t}]], \text{ since } S(B_{[0,k\Delta t]}) \text{ is } \mathcal{F}_{k\Delta t}\text{-measurable.}
\end{aligned}$$

On the other hand, because Brownian motion B has the strong Markov property, and the expected signature of $B_{[0,t]}$ does not depend on the starting point, it is true that

$$\mathbb{E}[S(B_{[k_1\Delta t,k_2\Delta t]})|\mathcal{F}_{k_1\Delta t}] = \mathbb{E}[S(B_{[k_1\Delta t,k_2\Delta t]})|B_{k_1\Delta t}] = \varphi((k_2 - k_1)\Delta t).$$

Therefore it implies that

$$\begin{aligned}
\varphi(k_2\Delta t) &= \mathbb{E}[S(B_{[0,k_1\Delta t]}) \otimes \varphi((k_2 - k_1)\Delta t)] \\
&= \mathbb{E}[S(B_{[0,k_1\Delta t]})] \otimes \varphi((k_2 - k_1)\Delta t) \\
&= \varphi(k_1\Delta t) \otimes \varphi((k_2 - k_1)\Delta t).
\end{aligned}$$

Similarly, it is easy to prove that it is true for φ^M as well, since the proof only relies on the Markov property, the expected signature of the process up to a fixed time does not depend on the starting point and $\{B_s^M\}_{s\in[k_1\Delta t,k_2\Delta t]}$ has the same distribution as $\{B_s^M\}_{s\in[0,(k_2-k_1)\Delta t]}$. ■

Herein let us give the explicit representation of φ^M in the following theorem.

Theorem 3.2

$$\varphi^M(T) = \varphi^M(\Delta t)^{\otimes M}$$

where $\Delta t = \frac{T}{M}$*, and for every index* $I = (i_1, i_2, \dots, i_{2n}) \in \mathcal{K}_{2n}$*,*

$$C_I = \binom{n}{\frac{n_1(I)}{2}, \frac{n_2(I)}{2}, \dots \frac{n_d(I)}{2}} \Big/ \binom{2n}{n_1(I), n_2(I), \dots n_d(I)}$$

and

$$\varphi^M(\Delta t) \;\; = \;\; \sum_{n\geq 0} \sum_{I\in\mathcal{K}_{2n}} C_I \frac{1}{n!}\left(\frac{\Delta t}{2}\right)^n e_{i_1}\otimes e_{i_2}\otimes\cdots\otimes e_{i_{2n}}.$$

Proof. By Lemma 3.2.1, the following equality is true for every integer $k = 1, 2, \dots, M$,

$$\varphi^M(k\Delta t) \;\; = \;\; \varphi^M((k-1)\Delta t)\otimes\varphi(\Delta t).$$

By induction, it is easy to show that

$$\varphi^M(T) = \varphi^M(\Delta t)^{\otimes M}$$

where $\Delta t = \frac{T}{M}$.
Moreover since $\{B^M_s\}_{s\in[0,\Delta t]}$ is a linear path, the signature of $\{B^M_s\}_{s\in[0,\Delta t]}$ could be reduced as follows:

$$S(B^M_{[0,\Delta t]}) = \exp(B^M_{\Delta t}) \stackrel{d}{=} \exp(\sum_{i=1}^d B^{(i)}_{\Delta t} e_i)$$

since $B^M_{\Delta t}$ has the same distribution with $B_{\Delta t}$.
Thus it implies that

$$\varphi^M(\Delta t) = \mathbb{E}[S(B^M_{[0,\Delta t]})] = \mathbb{E}\left[\exp(\sum_{i=1}^d B^{(i)}_{\Delta t} e_i)\right] = \sum_{m=0}^{\infty}\frac{1}{m!}\mathbb{E}\left[\left(\sum_{i=1}^d B^{(i)}_{\Delta t} e_i\right)^{\otimes m}\right].$$

For every word $I = (i_1, i_2, \dots, i_m)$,

$$\begin{aligned}\pi^I\left(\mathbb{E}\left[\left(\sum_{i=1}^d B^{(i)}_{\Delta t} e_i\right)^{\otimes m}\right]\right) &= \mathbb{E}\left[\prod_{j=1}^m B^{(i_j)}_{\Delta t}\right] \\ = \;\; &\mathbb{E}\left[\prod_{l=1}^d (B^{(l)}_{\Delta t})^{n_l(I)}\right] \\ = \;\; &\prod_{l=1}^d \mathbb{E}[(B^{(l)}_{\Delta t})^{n_l(I)}], \text{ as } B^{(i)}_t \text{ are mutually independent.}\end{aligned}$$

Since $B^{(l)}_{\Delta t}$ is normally distributed with mean 0 and variance Δt, where $l = 1, 2, \dots, d$,

$$\mathbb{E}[(B^{(l)}_{\Delta t})^{n_l(I)}] = \begin{cases} 0, & \text{if } n_l(I) \text{ is odd;} \\ \left(\frac{\Delta t}{2}\right)^{\frac{n_l(I)}{2}}\left(\frac{n_l!}{\frac{n_l}{2}!}\right), & \text{otherwise.}\end{cases}$$

Therefore $\pi^I\left(\mathbb{E}\left[\left(\sum_{i=1}^d B_{\Delta t}^{(i)}e_i\right)^{\otimes|I|}\right]\right)$ could be further simplified, i.e if $m = 2n$ and $I \in \mathcal{K}_{2n}$,

$$\pi^I(\varphi^M(\Delta t)) = \pi^I\left(\mathbb{E}\left[\left(\sum_{i=1}^d B_{\Delta t}^{(i)}e_i\right)^{\otimes|I|}\right]\right) = \left(\frac{\Delta t}{2}\right)^n \prod_{l=1}^{l} \frac{n_l(I)!}{\frac{n_l(I)}{2}!}$$

and

$$\begin{aligned}
\pi^I(\varphi^M(\Delta t)) &= \frac{1}{(2n)!}\pi^I\left(\mathbb{E}\left[\left(\sum_{i=1}^d B_{\Delta t}^{(i)}e_i\right)^{\otimes 2n}\right]\right)\\
&= \frac{1}{n!}\left(\frac{\Delta t}{2}\right)^n \left(\frac{n!}{\prod_{l=1}^d\left(\frac{n_l(I)}{2}\right)!}\right)\left(\frac{(2n)!}{\prod_{l=1}^d n_l(I)!}\right)\\
&= C_I\frac{1}{n!}\left(\frac{\Delta t}{2}\right)^n.
\end{aligned}$$

Otherwise, at least there exists an $l \in \{1, 2, \ldots, d\}$, such that $n_l(I)$ is odd, and thus

$$\pi^I(\varphi^M(\Delta t)) = 0.$$

■

Remark 3.2.2 *For every integer $n \in \mathbb{N}$, and for every index $I \in \mathcal{K}_{2n}$,*

$$C_I = \binom{n}{\frac{n_1(I)}{2}, \frac{n_2(I)}{2}, \ldots \frac{n_d(I)}{2}} \Big/ \binom{2n}{n_1(I), n_2(I), \ldots n_d(I)} \leq 1.$$

3.2.3 The convergence result for the expected signature of the piecewise linear approximation of Brownian motion

In this subsection, we are interested in examining the convergence rate of the expected signature of the piecewise linear approximation of Brownian motion. In order to do so, we need a norm on the tensor products to measure the error term. Herein we choose an admissible norm, which is defined in Definition 2.1.2. An important property of an admissible norm $|\cdot|$, which will be used frequently in the following discussion is that $\forall n, m \geq 1$,

$$|v \otimes w| \leq |v||w|, \forall v \in E^{\otimes n}, w \in E^{\otimes m}.$$

There are various ways to define the norms on $E^{\otimes n}$, e.g. the l_1-norm and l_∞-norm defined in Definition 3.2.3 and 3.2.4 respectively. It is easy to verify that both of them are admissible norms on $E^{\otimes n}$.

Definition 3.2.3 (**l_∞-norm on $E^{\otimes n}$**) *Let E be a Banach space. For every positive integer n, the l_∞-norm denoted by $|\cdot|_\infty$ is defined as follows, i.e. for every $v \in E^{\otimes n}$,*

$$|v|_\infty = \sup_{I \in \mathcal{W}_n} \{|\pi^I(v)|\}$$

where $\mathcal{W}_n$ is the set of words of length n.

Definition 3.2.4 (**l_1-norm on $E^{\otimes n}$**) *Let E be a Banach space. For every positive integer n, the l_1-norm denoted by $|\cdot|_1$ is defined as follows, i.e. for every $v \in E^{\otimes n}$,*

$$|v|_1 = \sum_{I \in \mathcal{W}_n} |\pi^I(v)|$$

where $\mathcal{W}_n$ is the set of the words of the length n.

Remark 3.2.3 *The l_1-norm on $E^{\otimes n}$ is a projective norm and it has a nice property, which is given in Proposition 3.3.*

Proposition 3.3 *For every $a \in E^{\otimes n}$ and $b \in E^{\otimes m}$,*

$$|a \otimes b|_1 = |a|_1 |b|_1$$

where $m, n \in \mathbb{N}$.

Proof. Let $\{e_i\}_{i=1}^d$ be the basis of E. For every $a \in E^{\otimes n}$ and $b \in E^{\otimes m}$, a and b could be written as

$$a = \sum_{I \in \mathcal{W}_n} \pi^I(a) e_{i_1} \otimes \cdots \otimes e_{i_n}$$
$$b = \sum_{J \in \mathcal{W}_m} \pi^J(b) e_{j_1} \otimes \cdots \otimes e_{j_m}$$

where $I = (i_1, i_2, \ldots, i_n)$ and $J = (j_1, j_2, \ldots, j_m)$.
Thus by the definition of tensor product, we have that

$$a \otimes b = \sum_{\substack{(i_1,\ldots,i_n) \in \mathcal{W}_n \\ (j_1,\ldots,j_m) \in \mathcal{W}_m}} \pi^{(i_1,\ldots,i_n)}(a) \pi^{(j_1,\ldots,j_n)}(b) e_{i_1} \otimes \ldots e_{i_n} \otimes e_{j_1} \otimes \cdots \otimes e_{j_m}$$

Therefore it follows that

$$\begin{aligned}
|a \otimes b|_1 &= \sum_{\substack{(i_1,\ldots,i_n) \in \mathcal{W}_n \\ (j_1,\ldots,j_m) \in \mathcal{W}_m}} |\pi^{(i_1,\ldots,i_n)}(a) \pi^{(j_1,\ldots,j_n)}(b)| \\
&= \sum_{(i_1,\ldots,i_n) \in \mathcal{W}_n} \sum_{(j_1,\ldots,j_m) \in \mathcal{W}_m} |\pi^{(i_1,\ldots,i_n)}(a)||\pi^{(j_1,\ldots,j_n)}(b)| \\
&= \left(\sum_{(i_1,\ldots,i_n) \in \mathcal{W}_n} |\pi^{(i_1,\ldots,i_n)}(a)| \right) \left(\sum_{(j_1,\ldots,j_m) \in \mathcal{W}_m} |\pi^{(j_1,\ldots,j_n)}(b)| \right) \\
&= |a|_1 |b|_1
\end{aligned}$$

■

Lemma 3.2.2 *For every $k = 1, 2, \ldots, M$, for every integer $j \geq 2$,*

$$|\rho_{2j}(\varphi^M(k\Delta t))| \leq \frac{1}{j!}\left(\frac{k\Delta t}{2}\right)^j \tag{3.2}$$

where $|\cdot|$ is the admissible norm.
In particular,

$$|\rho_{2j}(\varphi^M(k\Delta t))|_1 = \frac{1}{j!}\left(\frac{k\Delta t}{2}\right)^j \tag{3.3}$$

Proof. For $k = 1$, it is just Theorem 3.2.
For $k > 1$. Suppose that this statement is true for $k-1$. Let us prove that it is true for k as well. By Lemma 3.2.1, we have that

$$\varphi^M(k\Delta t) = \varphi^M((k-1)\Delta t)) \otimes \varphi^M(\Delta t)$$

and thus

$$\rho_{2j}(\varphi^M(k\Delta t)) = \sum_{j_1+j_2=j} \rho_{2j_1}(\varphi^M((k-1)\Delta t)) \otimes \rho_{2j_2}(\varphi^M(\Delta t)).$$

Since $|\cdot|$ is an admissible norm, then

$$|\rho_{2j_1}(\varphi^M((k-1)\Delta t)) \otimes \rho_{2j_2}(\varphi^M(\Delta t))| \leq |\rho_{2j_1}(\varphi^M((k-1)\Delta t))| \cdot |\rho_{2j_2}(\varphi^M(\Delta t))|$$

In particular, in terms of l_1 norm, by Proposition 3.3 it holds that

$$|\rho_{2j_1}(\varphi^M((k-1)\Delta t)) \otimes \rho_{2j_2}(\varphi^M(\Delta t))|_1 = |\rho_{2j_1}(\varphi^M((k-1)\Delta t))|_1 \cdot |\rho_{2j_2}(\varphi^M(\Delta t))|_1$$

Then by induction hypothesis it follows that

$$\begin{aligned}
|\rho_{2j}(\varphi^M(k\Delta t))| &\leq \sum_{j_1+j_2=j} |\rho_{2j_1}(\varphi^M((k-1)\Delta t))||\rho_{2j_2}(\varphi^M(\Delta t))| \\
&\leq \sum_{j_1+j_2=j} \frac{1}{j_1!}\left(\frac{(k-1)\Delta t}{2}\right)^{j_1} \frac{1}{j_2!}\left(\frac{\Delta t}{2}\right)^{j_2} \\
&= \frac{1}{j!}\left(\frac{k\Delta t}{2}\right)^j.
\end{aligned}$$

If we choose l_1-norm, inequality signs could be replaced by equals signs in the above formula.
By induction, it shown that for every $k = 1, 2, \ldots, M$, for every integer $j \geq 2$,

$$|\rho_{2j}(\varphi^M(k\Delta t))| \leq \frac{1}{j!}\left(\frac{k\Delta t}{2}\right)^j$$

$$|\rho_{2j}(\varphi^M(k\Delta t))|_1 = \frac{1}{j!}\left(\frac{k\Delta t}{2}\right)^j.$$

■

Lemma 3.2.3
For every integer $j \geq 2$,

$$|\rho_{2j}(\varphi^M(\Delta t) - \varphi(\Delta t))| \leq \frac{1}{j!}\left(\frac{\Delta t}{2}\right)^j; \tag{3.4}$$

for every odd integer j or $j = 0, 2$,

$$\rho_j(\varphi^M(\Delta t) - \varphi(\Delta t)) = 0 \tag{3.5}$$

where $|\cdot|$ is the admissible norm.

Proof. Since the explicit formula for $\varphi^M(\Delta t)$ and $\varphi(\Delta t)$ is given in Theorem 3.2 and Corollary 3.2.1, it is known that for every word I of odd length or 0,

$$\pi^I(\varphi(\Delta t)) = \pi^I(\varphi^M(\Delta t)) = 0$$
$$\pi^{()}(\varphi(\Delta t)) = \pi^{()}(\varphi^M(\Delta t)) = 1$$

where () denotes the empty word;
For every word I of length $2n$,

$$\begin{aligned}\pi^I(\varphi(\Delta t)) &= \frac{1}{n!}\left(\frac{\Delta t}{2}\right)^n \mathbf{1}\{I \in \mathcal{S}_{2n}\}\\ \pi^I(\varphi^M(\Delta t)) &= C_I\frac{1}{n!}\left(\frac{\Delta t}{2}\right)^n \mathbf{1}\{I \in \mathcal{K}_{2n}\}.\end{aligned}$$

By Theorem 3.2

$$C_I = \binom{n}{\frac{n_1(I)}{2}, \frac{n_2(I)}{2}, \ldots \frac{n_d(I)}{2}} \Big/ \binom{2n}{n_1(I), n_2(I), \ldots n_d(I)} \leq 1.$$

In addition, using $\mathcal{S}_{2n} \underset{\neq}{\subset} \mathcal{K}_{2n}$, it follows

$$\begin{aligned}|\pi^I(\varphi(\Delta t)) - \pi^I(\varphi^M(\Delta t))| &= \begin{cases} (1 - C_I)\frac{1}{n!}\left(\frac{\Delta t}{2}\right)^n, & I \in \mathcal{S}_{2n};\\ C_I\frac{1}{n!}\left(\frac{\Delta t}{2}\right)^n, & I \in \mathcal{K}_{2n}\backslash\mathcal{S}_{2n};\\ 0, & \text{otherwise.}\end{cases}\\ &\leq \frac{1}{n!}\left(\frac{\Delta t}{2}\right)^n\end{aligned}$$

and by the definition of the admissible norm $|\cdot|$, it is also true that

$$|\rho_{2n}(\varphi^M(\Delta t)) - \rho_{2n}(\varphi(\Delta t))| \leq \frac{1}{n!}\left(\frac{\Delta t}{2}\right)^n$$

where $j \in \mathbb{N}$.
In particular, for the case $n = 1$, since $\mathcal{K}_2 = \mathcal{S}_2$ and $c_I = \mathbf{1}(I \in \mathcal{K}_2)$ where the length

of I is equal to 2, then $C_I = 1$, and the difference between $\rho_2(\varphi(\Delta t))$ and $\rho_2(\varphi(\Delta t))$ can be simplified to

$$1 - C_I = 0.$$

This means that $\varphi^M(\Delta t)$ coincides with $\varphi(\Delta t)$ for the first two levels. ■

Next we will give the bounds for the discretization error for the expected signature of Brownian motion and the result is summarised in the following theorem. In the following discussion, we restrict the dimension of Brownian motion to $d \geq 2$, because from the previous discussion, for the one dimensional case, the signature of Brownian motion $B_{[0,T]}$ only depends on the increment of Brownian motion at time T, and thus there is no discretization error for the approximation to Brownian motion.

Theorem 3.4 *For every integer $j \geq 2$, it holds that*

$$|\rho_{2j}(\varphi^M(T) - \varphi(T))| \leq T \frac{T^{j-1} - (T - \Delta t)^{j-1}}{2^{j-1}(j-1)!}.$$

Proof. Let us compute the difference of $\varphi^M(T)$ and $\varphi(T)$ by expanding it as the telescoping sum as follows:

$$\begin{aligned} & \varphi^M(T) - \varphi(T) \\ = & \sum_{k=0}^{M-1} \varphi(k\Delta t) \otimes \varphi^M((M-k)\Delta t) - \varphi((k+1)\Delta t) \otimes \varphi^M((M-(k+1))\Delta t). \end{aligned}$$

By Lemma 3.2.1, we have that

$$\begin{aligned} \varphi^M((M-k)\Delta t) &= \varphi^M(\Delta t) \otimes \varphi^M((M-(k+1))\Delta t) \\ \varphi((k+1)\Delta t) &= \varphi(k\Delta t) \otimes \varphi(\Delta t) \end{aligned}$$

and therefore the following two equalities follow that

$$\begin{aligned} & \varphi(k\Delta t) \otimes \varphi^M((M-k)\Delta t) - \varphi((k+1)\Delta t) \otimes \Phi^M((M-(k+1))\Delta t) \\ = & \varphi(k\Delta t) \otimes (\varphi^M(\Delta t) - \varphi(\Delta t)) \otimes \varphi^M((M-(k+1))\Delta t) \end{aligned}$$

and

$$\varphi^M(T) - \varphi(T) = \sum_{k=0}^{M-1} \varphi(k\Delta t) \otimes (\varphi^M(\Delta t) - \varphi(\Delta t)) \otimes \varphi^M((M-(k+1))\Delta t).$$

Then applying the projection function $\rho_{2j}(\cdot)$ to $\varphi^M(T) - \varphi(T)$, for every $j \geq 2$, $\rho_{2j}(\varphi^M(T) - \varphi(T))$ is equal to

$$\sum_{j_2=2}^{j} \sum_{j_1+j_3=j-j_2} \sum_{k=0}^{M-1} \rho_{2j_1}(\varphi(k\Delta t)) \otimes (\rho_{2j_2}(\varphi^M(\Delta t) - \varphi(\Delta t)) \otimes \rho_{2j_3}(\varphi^M((M-(k+1))\Delta t))$$

since $\pi^I(\varphi(k\Delta t)) = \pi^I(\varphi^M(k\Delta t)) = 0$ where I is a word of odd length or length 2.
For every fixed $j_2 \geq 2$, let us examine

$$U^M_{j_2}(T) = |\sum_{j_1+j_3=j-j_2} \sum_{k=0}^{M-1} \rho_{2j_1}(\varphi(k\Delta t)) \otimes (\rho_{2j_2}(\varphi^M(\Delta t) - \varphi(\Delta t)) \otimes \rho_{2j_3}(\varphi^M((M-(k+1))\Delta t))|$$

which is bounded above by

$$\sum_{j_1+j_3=j-j_2} \sum_{k=0}^{M-1} |\rho_{2j_1}(\varphi(k\Delta t))| \cdot |\rho_{2j_2}(\varphi^M(\Delta t) - \varphi(\Delta t)| \cdot |\rho_{2j_3}(\varphi^M((M-(k+1))\Delta t))|$$

since $|\cdot|$ is an admissible norm.
Recall that the following inequalities have been proved:

$$\begin{aligned}
|\rho_{2j_1}(\varphi(k\Delta t))| &\leq \frac{1}{j_1!}\left(\frac{k\Delta t}{2}\right)^{j_1} \\
|\rho_{2j_2}(\varphi^M(\Delta t) - \varphi(\Delta t))| &\leq \frac{1}{j_2!}\left(\frac{\Delta t}{2}\right)^{j_2} \\
|\rho_{2j_3}(\varphi^M((M-(k+1))\Delta t))| &\leq \frac{1}{j_3!}\left(\frac{(M-k-1)\Delta t}{2}\right)^{j_3}.
\end{aligned}$$

Therefore we have

$$\begin{aligned}
U^M_{j_2}(T) &\leq \sum_{k=0}^{M-1} \sum_{j_1+j_3=j-j_2} \frac{1}{j_1!}\left(\frac{k\Delta t}{2}\right)^{j_1} \frac{1}{j_2!}\left(\frac{\Delta t}{2}\right)^{j_2} \frac{1}{j_3!}\left(\frac{(M-k-1)\Delta t}{2}\right)^{j_3} \\
&= \frac{1}{j_2!}\left(\frac{\Delta t}{2}\right)^{j_2} \sum_{k=0}^{M-1} \frac{1}{(j-j_2)!}\left(\frac{T-\Delta t}{2}\right)^{j-j_2} \\
&= \frac{T}{2}\frac{1}{j_2!}\left(\frac{\Delta t}{2}\right)^{j_2-1} \frac{1}{(j-j_2)!}\left(\frac{T-\Delta t}{2}\right)^{j-j_2} \\
&\leq \frac{T}{2}\frac{1}{(j_2-1)!}\left(\frac{\Delta t}{2}\right)^{j_2-1} \frac{1}{(j-j_2)!}\left(\frac{T-\Delta t}{2}\right)^{j-j_2}.
\end{aligned}$$

From the above inequality, we know that

$$
\begin{aligned}
& |\rho_{2j}(\varphi^M(T)-\varphi(T))| \leq \sum_{j_2=2}^{j} U_{j_2}^M(T) \\
\leq\ & \sum_{j_2=2}^{j} \frac{T}{2}\frac{1}{(j_2-1)!}\left(\frac{\Delta t}{2}\right)^{j_2-1}\frac{1}{(j-j_2)!}\left(\frac{T-\Delta t}{2}\right)^{j-j_2} \\
=\ & \frac{T}{2}\frac{1}{(j-1)!}\left(\left(\frac{T}{2}\right)^{j-1}-\left(\frac{T-\Delta t}{2}\right)^{j-1}\right) \\
=\ & \frac{T}{2}\left(\frac{T^{j-1}-(T-\Delta t)^{j-1}}{2^{j-1}(j-1)!}\right).
\end{aligned}
$$

■

Remark 3.2.4 *It is easier to see that the convergence rate of* $|\rho_{2j}(\varphi^M(T)-\varphi(T))|$ *is no more than the order* Δt *from the following estimate:*

$$
\begin{aligned}
& |\rho_{2j}(\varphi^M(T)-\varphi(T))| \\
\leq\ & \frac{T}{2}\left(\frac{T^{j-1}-(T-\Delta t)^{j-1}}{2^{j-1}(j-1)!}\right) \\
=\ & \frac{T}{2}\sum_{j_2=1}^{j-1}\frac{1}{(j_2)!}\left(\frac{\Delta t}{2}\right)^{j_2}\frac{1}{(j-1-j_2)!}\left(\frac{T-\Delta t}{2}\right)^{j-1-j_2} \\
\leq\ & \frac{T}{2}\Delta t\sum_{j_2=0}^{j-2}\frac{1}{j_2!}\left(\frac{\Delta t}{2}\right)^{j_2}\frac{1}{(j-2-j_2)!}\left(\frac{T-\Delta t}{2}\right)^{j-2-j_2} \\
\leq\ & \frac{T}{2}\Delta t\frac{1}{(j-2)!}\left(\frac{T-\Delta t}{2}\right)^{j-2} \leq \Delta t\frac{1}{(j-2)!}\left(\frac{T}{2}\right)^{j-1}.
\end{aligned}
$$

After this discussion of the upper bound of the difference of the expected signature of Brownian motion and its piecewise linear approximation, we derive the lower bound for it as well.

Theorem 3.5 [1] *For every integer* $j \geq 0$*, consider the word* $I_j = (1,2,1,2,\underbrace{1,\dots,1}_{2j \textit{ copies of } 1})$,

[1]Thanks to the discussion with Weijun Xu, the proof has been simplified to this version. In the original work to estimate the lower bound, we considered the index J_n in the form

$$
J_n = \begin{cases} (\underbrace{1,1,2,2,\dots,1,1,2,2}_{i \text{ copies of } 1122}), & \text{if } n = 2i; \\ (\underbrace{1,1,2,2,\dots,1,1,2,2}_{i \text{ copies of } 1122},1,1), & \text{if } n = 2i+1. \end{cases}
$$

where $n \in \mathbb{N}$. It can be proved that $|\pi^{J_n}(\varphi(T)-\varphi^M(T))|$ converges with order Δt too, but it requires a bit more effort to do so.

then

$$|\pi^{I_j}(\varphi(T)-\varphi^M(T))| \geq \frac{1}{12(j+1)!}\left(\frac{T-\Delta t}{2}\right)^{j+1}\Delta t.$$

Proof. As $I_j \notin \bigcup_{n\in\mathbb{N}} \mathcal{S}_{2n}$, it follows immediately that

$$\pi^I(\varphi(T)) = 0.$$

By induction on k, we have shown that for every integer $k \in \{1,2,\ldots,M\}$, and every index I,

$$\pi^I(\varphi^M(k\Delta t)) \geq 0$$

and then

$$\begin{aligned}
& |\pi^{I_j}(\varphi^M(T)-\varphi(T))| = |\pi^{I_j}(\varphi^M(T))| = \pi^{I_j}(\varphi^M(T)) \\
\geq & \sum_{k=1}^{M-1} \pi^{(1,2,1,2)}(\varphi^M(\Delta t))\pi^{()}(\varphi^M((M-k-1)\Delta t))\pi^{J_j}(\varphi^M(k\Delta t)) \\
= & \sum_{k=1}^{M-1} \pi^{(1,2,1,2)}(\varphi^M(\Delta t))\pi^{J_j}(\varphi^M(k\Delta t)). \qquad (3.6)
\end{aligned}$$

where $J_j = \underbrace{(1,1,\ldots,1,1)}_{2j \text{ copies of } 1}$.

It is also noted that for each J_j, we have that

$$\begin{aligned}
\pi^{J_j}(\varphi^M(k\Delta t)) &= \pi^{J_j}(\varphi(k\Delta t)) = \frac{1}{j!}\left(\frac{k\Delta t}{2}\right)^j \\
\pi^{(1,2,1,2)}(\varphi^M(\Delta t)) &= C_{(1,2,1,2)}\left(\frac{\Delta t}{2}\right)^2 = \frac{1}{6}\left(\frac{\Delta t}{2}\right)^2.
\end{aligned}$$

Substituting those two equations into the inequality (3.6), we have that

$$\begin{aligned}
& \pi^{I_j}(\varphi^M(T)) \\
\geq & \sum_{k=1}^{M-1} \frac{1}{6}\left(\frac{\Delta t}{2}\right)^2 \frac{1}{j!}\left(\frac{k\Delta t}{2}\right)^j \\
= & \frac{1}{6j!}\left(\frac{\Delta t}{2}\right)^{j+2}\sum_{k=1}^{M-1} k^j.
\end{aligned}$$

Since for every $k = 1,2,\ldots,M-1$,

$$k^j = \int_{k-1}^{k} k^j dx \geq \int_{k-1}^{k} x^j dx$$

by summing up both sides of the above equation over $k = 1, \dots, M-1$, we have that

$$\sum_{k=1}^{M-1} k^j \geq \sum_{k=1}^{M-1} \int_{k-1}^{k} x^j dx = \int_{0}^{M-1} x^j dx = \frac{(M-1)^{j+1}}{j+1}.$$

Therefore it follows immediately that

$$\begin{aligned} \pi^{I_j}(\varphi^M(T)) &\geq \frac{1}{6j!}\left(\frac{\Delta t}{2}\right)^{j+2}\frac{(M-1)^{j+1}}{j+1} \\ &= \frac{1}{12(j+1)!}\left(\frac{T-\Delta t}{2}\right)^{j+1}\Delta t. \end{aligned}$$

■

Remark 3.2.5 *By Theorem 3.4 and Lemma 3.5, it is easy to see that for every $n \geq 2$,*

$$|\rho_{2n}(\varphi(T) - \varphi^M(T))|_\infty$$

is $O(\frac{1}{M})$ when M tends to the infinity.

3.2.4 The expected signature of Brownian motion under the empirical measure

Let us consider a sequence of i.i.d random paths of size K, denoted by $\{\hat{B}_n\}_{n=1}^K$, which have the same distribution as the linear approximation of Brownian motion $\hat{B}_{[0,T]}$ we discussed in the previous subsection. The average of the signature of $\{\hat{B}_n\}_{n=1}^K$ can be regarded as the expected signature of the piecewise linear approximation of Brownian motion $\hat{B}_{[0,T]}$ under the empirical measure. Herein we are interested in understanding how fast the average of the signatures of the samples of piecewise linear approximation $\{\hat{B}_n\}_{n=1}^K$ converges to the expected signature of the original Brownian motion. More precisely we focus on the convergence rate of the error term, which is defined in Definition 3.2.5.

Definition 3.2.5 *Let us define the error term of the average signature of the linear approximation of Brownian motion of size K as follows:*

$$\epsilon_{K,M} = \frac{1}{K}\sum_{n=1}^{K} S(\hat{B}_n) - \mathbb{E}[S(B_{[0,T]})].$$

Next we are interested in studying the L_2 norm of the error term $\epsilon_{K,M}$, i.e for every word I,

$$|\pi^I(\epsilon_{K,M})|_{L_2} = \mathbb{E}[\pi^I(\epsilon_{K,M})^2]$$

and the convergence rate of $|\pi^I(\epsilon_{K,M})|_{L_2}$ when K and M tend to the infinity.

Remark 3.2.6 *It is easy to see that for every index I,*

$$\mathbb{E}[\pi^I(\epsilon_{K,M})^2] = \frac{1}{K}Var(\pi^I(S(\hat{B}_{[0,T]}))) + \pi^I(\varphi(T) - \varphi^M(T))^2,$$

which indicated that the error term comes from the variance of the expected signature of approximated Brownian motion samples and the discretization error of the approximated Brownian motion. The discretization error $\pi^I(\varphi(T)-\varphi^M(T))^2$ has been studied carefully in Subsection 3.2.3. The only thing left is to understand $Var(S(\hat{B}_{[0,T]}))$.

Lemma 3.2.4 *For every index I,*

$$Var(\pi^I(S(B_{[0,T]}))) \leq \binom{2|I|}{|I|}\frac{1}{|I|!}\left(\frac{T}{2}\right)^{|I|}$$

Proof.

$$\begin{aligned} & Var(\pi^I(S(B_{[0,T]}))) = (\pi^I \sqcup\!\sqcup\, \pi^I)(\varphi(T)) - (\pi^I(\varphi(T)))^2 \\ \leq\ & (\pi^I \sqcup\!\sqcup\, \pi^I)(\varphi(T)) \\ \leq\ & \binom{2|I|}{|I|}\frac{1}{|I|!}\left(\frac{T}{2}\right)^{|I|} \end{aligned}$$

■

Lemma 3.2.5 *For every index I of length greater than or equal to 2,*

$$Var(\pi^I(S(\hat{B}_{[0,T]}))) \leq \binom{2|I|}{|I|}\left(\frac{1}{|I|!}\left(\frac{T}{2}\right)^{|I|} + \Delta t\left(\frac{T}{2}\right)^{|I|-1}\frac{1}{(|I|-2)!}\right).$$

Proof. By the shuffle property of the signature, we have that

$$Var(\pi^I(S(\hat{B}_{[0,T]}))) \leq \mathbb{E}[\pi^I(S(\hat{B}_{[0,T]}))^2] = \mathbb{E}[(\pi^I \sqcup\!\sqcup\, \pi^I)(S(\hat{B}_{[0,T]}))] = (\pi^I \sqcup\!\sqcup\, \pi^I)(\varphi^M(T)) \quad (3.7)$$

Since the number of shuffles in $I \sqcup\!\sqcup\, I$ is $\binom{2|I|}{|I|}$, it follows that

$$\mathbb{E}[(\pi^I \sqcup\!\sqcup\, \pi^I)(S(\hat{B}_{[0,T]}))] \ \leq\ \binom{2|I|}{|I|}|\rho_{2|I|}(\varphi^M(T))|_\infty.$$

By Remark 3.2.4, it is shown that for every index of length $2n$, the following estimates holds:

$$|\rho_{2n}(\varphi(T) - \varphi^M(T))|_\infty \leq \Delta t\frac{1}{(n-2)!}\left(\frac{T}{2}\right)^{n-1}$$

which implies that

$$
\begin{aligned}
|\rho_{2n}(\varphi^M(T))|_\infty &\leq |\rho_{2n}(\varphi(T)-\varphi^M(T))|_\infty + |\rho_{2n}(\varphi(T))|_\infty \\
&\leq \Delta t \frac{1}{(n-2)!}\left(\frac{T}{2}\right)^{n-1} + \left(\frac{T}{2}\right)^{n}\frac{1}{n!}.
\end{aligned}
$$

Using the above inequality in inequality (3.7), our proof is complete. ∎

Theorem 3.6 *Fix an arbitrary index I of length greater than or equal to 2. For any fixed $K \in \mathbb{N}$, $|\pi^I(\epsilon_{K,M})|_{L_2}$ converges to 0 with the speed $\frac{1}{M}$ as M tends to the infinity. For any fixed $M \in \mathbb{N}$, $|\pi^I(\epsilon_{K,M})|_{L_2}$ converges to 0 with the speed $\frac{1}{K}$ as K tends to the infinity.*

Proof. For every fixed index I, since $\hat{B}_n$ is a sequence of i.i.d sample paths with the same distribution as $\hat{B}_{[0,T]}$, we have that

$$
\mathbb{E}[\pi^I(\epsilon_{K,M})^2] = \frac{1}{K}Var(\pi^I(S(\hat{B}_{[0,T]}))) + \pi^I(\varphi(T)-\varphi^M(T))^2.
$$

By Lemma 3.2.5, for every index I of length greater than or equal to 2,

$$
Var(\pi^I(S(\hat{B}_{[0,T]}))) \leq \binom{2|I|}{|I|}\left(\frac{1}{|I|!}\left(\frac{T}{2}\right)^{|I|} + \Delta t\left(\frac{T}{2}\right)^{|I|-1}\frac{1}{(|I|-2)!}\right).
$$

Moreover, by Remark 3.2.4, it is shown that for every index of length $2n$, it holds:

$$
|\rho_{2n}(\varphi(T)-\varphi^M(T))|_\infty \leq \Delta t \frac{1}{(n-2)!}\left(\frac{T}{2}\right)^{n-1}.
$$

For every index of odd length,

$$
|\rho_{|I|}(\varphi(T)-\varphi^M(T))|_\infty = 0.
$$

Therefore we have that

$$
\begin{aligned}
\mathbb{E}[\pi^I(\epsilon_{K,M})^2] &\leq \frac{1}{K}\binom{2|I|}{|I|}\left(\frac{1}{|I|!}\left(\frac{T}{2}\right)^{|I|} + \Delta t\left(\frac{T}{2}\right)^{|I|-1}\frac{1}{(|I|-2)!}\right) \\
&\quad + \mathbf{1}_{\{|I| \text{ is even.}\}}\left(\Delta t\frac{1}{(|I|-2)!}\left(\frac{T}{2}\right)^{\frac{|I|}{2}-1}\right)^2.
\end{aligned}
$$

where $\Delta t = \frac{T}{M}$.

Now we can conclude that for any fixed K, $\frac{1}{K}\sum_{n=1}^{K} S(X_n)$ converges to the expected signature of Brownian motion with the speed of $\frac{1}{M}$ in the L_2 sense while for any fixed M, $\frac{1}{K}\sum_{n=1}^{K} S(X_n)$ converges to the expected signature of Brownian motion with the speed of $\frac{1}{K}$ in the L_2 sense as well. ∎

Chapter 4

The Expected Signature of a time homogenous Itô diffusion Process up to a Fixed Time

4.1 The setting of the problem

Fix $E = \mathbb{R}^d$, where $d \in \mathbb{N}$. Let $(\Omega, \mathcal{F}, \mathbb{P}^x, \{\mathcal{F}_t\}_{t\in[0,+\infty]})$ be a probability space, where the stochastic process X is adapted to the filtration $\{\mathcal{F}_t\}_{t\in[0,\infty]}$ and the filtration $\{\mathcal{F}_t\}_{t\in[0,\infty]}$ is complete. Moreover, let $W = \{W_t\}_{t\in[0,+\infty)}$ be a $k-$dimensional Brownian motion defined on $\mathbb{R}^k$.

Let us consider a time homogeneous Itô diffusion process X satisfying the following SDE with drift vector $\{\mu^{(j)}(x)\}_{j=1}^d$ and the dispersion matrix $\{V_i^{(j)}(x)\}_{j,i=1}^{d,k}$, where $\mu, V_1, \dots, V_k$ are functions from E to E. This is written component-wise as,

$$dX_t^{(j)} = \mu^{(j)}(X_t)dt + \sum_{i=1}^{k} V_i^{(j)}(X_t) \cdot dW_t^{(i)} \tag{4.1}$$

where $j = 1, 2, \dots, d$, and $\cdot$ is used to emphasize that the integral is taken in the Itô integral sense.

Or it can be also written in the matrix form, i.e.

$$dX_t = \mu(X_t)dt + V(X_t) \cdot dW_t$$

where $t \in [0, T]$,

$$X_t = \begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ \dots \\ X_t^{(d)} \end{pmatrix}, \mu(x) = \begin{pmatrix} \mu^{(1)}(x) \\ \mu^{(2)}(x) \\ \dots \\ \mu^{(d)}(x) \end{pmatrix}, W_t = \begin{pmatrix} W_t^{(1)}(x) \\ W_t^{(2)}(x) \\ \dots \\ W_t^{(k)}(x) \end{pmatrix},$$

and

$$V(x) = \begin{pmatrix} V_1^{(1)}(x) & V_2^{(1)}(x) & \dots & V_k^{(1)}(x) \\ V_1^{(2)}(x) & V_2^{(2)}(x) & \dots & V_k^{(2)}(x) \\ \dots & \dots & \dots & \dots \\ V_1^{(d)}(x) & V_2^{(d)}(x) & \dots & V_k^{(d)}(x) \end{pmatrix}$$

For ease of notation, let $b(x)$ denote the $(d \times d)$ matrix, which is defined as $b(x) \triangleq V(x)V^T(x)$, with elements

$$b_{j_1,j_2}(x) \triangleq \sum_{i=1}^{k} V_i^{(j_1)}(x)V_i^{(j_2)}(x); j_1, j_2 \in \{1, 2, \dots, d\}$$

and will be called the diffusion matrix.
Under the probability measure $\mathbb{P}^x$ the process X satisfies

$$\mathbb{P}^x(X_0 = x) = 1$$

Moreover it is assumed that the law of the pair (W, X) is uniquely determined in this section.

In order to discuss the regularity of the drift and volatility further, let us introduce the concept of the polynomial growth condition as follows:

Definition 4.1.1 *Let us fix a positive integer n. Suppose that f is a function mapping from $\mathbb{R}^+ \times E$ to $\mathbb{R}^n$. We say that the function f satisfies the polynomial growth condition of degree r where $r \geq 1$ if and only if for every fixed $T > 0$, there exists constants L and r such that*

$$\sup_{0 \leq t \leq T} |f(t, x)| \leq L(1 + |x|^r)$$

In the rest of this section, we assume that the drift term μ and the dispersion matrix $\{V_i\}_{i=1}^k$ satisfy the following condition.

Condition 4.1 *The drift term μ and the dispersion matrix $\{V_i\}_{i=1}^k$ satisfy the globally Lipschitz condition, i.e. there exist a constant C, such that for every $x, y \in E$,*

$$|\mu(x) - \mu(y)| + \sum_{i=1}^{k} |V_i(x) - V_i(y)| \leq C|y - x|.$$

Remark 4.1.1 *If the drift vector μ and the dispersion matrix $\{V_i\}_{i=1}^k$ satisfy Condition 4.1, then it implies that μ and $\{V_i\}_{i=1}^k$ satisfy the linear growth condition as well, i.e. there exists a constant K, such that for every $x \in E$,*

$$|\mu(x)| + \sum_{i=1}^{k} |V_i(x)| \leq K(1 + |x|).$$

By imposing the linear growth type condition and the Lipschitz type condition, it can be proved that the SDE admits a unique strong solution, as was first done by K. Itô([8]).

Now let us introduce the definition of the signature of a time homogeneous Itô diffusion process.

Definition 4.1.2 *Let X_t be a Itô diffusion process, which satisfies the SDE (4.1). Suppose that the following iterated integrals are well defined. The signature of $X_{[0,t]}$ is defined to be*

$$S(X_{[0,t]}) = (1, X^1_{[0,t]}, \ldots, X^n_{[0,t]}, \ldots) \in T((E))$$

for almost every path X and for every $n \geq 1$ by the Stratonovich iterated integral

$$X^n_{[0,t]} = \int \cdots \int_{0<u_1<u_2<\cdots<u_n<t} dX_{t_1} \otimes \cdots \otimes dX_{t_n}.$$

Remark 4.1.2 *In [5], it is shown that any continuous, d-dimensional semi-martingale, say $X = M + N$ where M is a continuous local martingale and N is a continuous process of bounded variation on any compact interval, admits an enhancement with $p-$variation rough path regularity, $p \in (2,3)$. By rough path theory([14]), the signature of $X_{[0,t]}$ is not uniquely determined until the choice of definition of the area is fixed. There is more than one choice of the integral of the semi-martingale, e.g. Itô integral or Stratonovich integral. In this thesis we choose the Stratonovich integral to define the iterated integral of the diffusion process X_t.*

Remark 4.1.3 *In the same way as Lemma 2.4.1, the signature of an Itô diffusion process $X_{[0,t]}$, i.e. $S(X_{[0,t]})$ defined in Definition 4.1.2, satisfies the following RDE:*

$$dS(X_{[0,t]}) = S(X_{[0,t]}) \otimes (\circ dX_t),\ S(X_{[0,0]}) = \mathbf{1}$$

where $\circ$ is used to emphasize that the SDE is in the Stratonovich sense.

Definition 4.1.3 *Let X be an Itô diffusion process defined by SDE (4.1). Suppose that the signature of $X_{[0,t]}$ is well defined. The expected signature of $X_{[0,t]}$ starting at point x, denoted by $\Phi(t,x)$ is defined as follows:*

$$\Phi(t,x) = \mathbb{E}^x[S(X_{[0,t]})]$$

where $t \in \mathbb{R}^+$.
For every $n \in \mathbb{N}$, the n^{th} term of the expected signature of an Itô diffusion process $X_{[0,t]}$ starting at x is denoted by $\Phi_n(t,x)$. Moreover, $\Phi_0(t,x) = 1$.

Remark 4.1.4 *By the definition of the signature, it is known that the signature of any path at time 0 should be trivial, and thus its expectation is trivial as well, i.e.*

$$\Phi(0,x) = \mathbf{1}$$

We are interested in the derivation of a PDE satisfied by $\Phi(t,x)$. By solving this PDE, subject to certain constraints, we will derive the whole expected signature $\Phi(t,x)$, for every $(t,x) \in \mathbb{R}^+ \times E$. In order to do so, the diffusion process is required to satisfy the following smoothness condition on the function Φ:

Condition 4.2 $\Phi(t,x)$ *is finite and it has continuous partial derivative with respect to t and continuous second order partial derivatives with respect to x, i.e. for any index I, $\pi^I(\Phi(t,x)) \in C^{1,2}(\mathbb{R}^+ \times E)$, where $C^{1,2}(\mathbb{R}^+ \times E)$ is the space of continuous functions mapping from $\mathbb{R}^+ \times E$ to $\mathbb{R}$, with continuous partial derivatives $\left(\frac{\partial f}{\partial t}\right)$, $\left(\frac{\partial f}{\partial x_i}\right)_{i=1}^d$, $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1}^d$.*

Remark 4.1.5 *Since our goal is to derive the PDE for the expected signature function $\Phi(t,x)$, and it will be proved later in the chapter that it is a PDE of the second order, Condition 4.2 is natural and necessary to ensure the PDE admits a unique classical solution.*

4.2 Auxiliary estimates

In this section, let us give some useful estimates in the case where the drift vector and the dispersion matrix satisfy Condition 4.1. Lemma 4.2.1 will be used in the main theorem of (Theorem 4.9), while Lemma 4.2.2 is useful in the calculation of the infinitesimal generator of $\Phi(t,x)$ (Theorem 4.6).

To show Lemma 4.2.1 and Lemma 4.2.2, we need the following proposition, which is actually Corollary 12 on page 86 in [11].

Proposition 4.3 *Let X denote a time homogeneous Itô diffusion process, which is defined by the SDE (4.1) and satisfies Condition 4.1, i.e. μ and $\{V_i\}_{i=1}^k$ all satisfy a global Lipschitz condition. Moreover let K be a constant, such that for every $x \in E$,*

$$|\mu(x)| + \sum_{i=1}^{k} |V_i(x)| \leq K(1 + |x|).$$

Then for every sufficiently small $T > 0$, there exists a constant N depending on p and K, such that for all $p \geq 0$ and $t \in [0, T]$,

$$\begin{aligned} \mathbb{E}^x[\sup_{t\in[0,T]} |X_t - x|^p] &\leq NT^{\frac{p}{2}} e^{NT}(1 + |x|)^p, \\ \mathbb{E}^x[\sup_{t\in[0,T]} |X_t|^p] &\leq Ne^{NT}(1 + |x|)^p. \end{aligned}$$

Lemma 4.2.1 *Let X denote a time homogeneous Itô diffusion process, which is defined by the SDE (4.1) and satisfies Condition 4.1, i.e. μ and $\{V_i\}_{i=1}^k$ all satisfy a global Lipschitz condition. Let f be a function from $\mathbb{R}^+ \times E$, and satisfy the polynomial growth condition, i.e. for every fixed $T > 0$, there exists constants K and $r \geq 1$, such that*

$$\sup_{t\in[0,T]} |f(t, x)| \leq K(1 + |x|^r).$$

Then for every fixed $T > 0$, $f(t, X_t)$ is uniformly L_p-integrable, i.e.

$$\sup_{t\in[0,T]} \mathbb{E}[|f(t, X_t)|^p] < +\infty.$$

where $p \geq 1$.

Proof. By the polynomial growth condition of f, there exists constants K and $r \geq 1$, such that

$$|f(t, x)|^p \quad \leq \quad K^p(1 + |x|^r)^p \leq K^p 2^{p-1}(1 + |x|^{rp}).$$

where $p \geq 1$. The last inequality comes from the inequality $(a + b)^p \leq 2^{p-1}(a^p + b^p)$, where $a, b \in \mathbb{R}^+$. Therefore, for every fixed $T > 0$,

$$\sup_{t\in[0,T]} \mathbb{E}[|f(t, X_t)|^p] \leq K^p 2^{p-1}(1 + \sup_{t\in[0,T]} \mathbb{E}^x[|X_t|^{rp}]) \leq K^p 2^{p-1}(1 + \mathbb{E}^x[\sup_{t\in[0,T]} |X_t|^{rp}])$$

By Proposition 4.3,for every $p \geq 1$, there exists a constant C depending on T, x, K and rp, such that

$$\mathbb{E}^x[\sup_{t\in[0,T]} |X_t|^{rp}] \leq C$$

and thus

$$\sup_{t\in[0,T]} \mathbb{E}[|f(t,X_t)|^p] \leq K^p 2^{p-1}(1+C) < +\infty.$$

■

Lemma 4.2.2 *Let f be a function mapping from E to $\mathbb{R}$. Suppose that for every fixed $x \in E$, there exists constants $C > 0$ and $p \geq 1$ depending only on x, such that for every $y \in E$,*

$$|f(y) - f(x)| \leq C|y-x|^p.$$

Then the following equality holds:

$$\lim_{\varepsilon\to 0} \frac{\mathbb{E}^x[\int_0^\varepsilon f(X_t)dt]}{\varepsilon} = f(x).$$

Proof. It is sufficient to show that

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\mathbb{E}^x\left[\int_0^\varepsilon |f(X_t)-f(x)|\,dt\right] = 0,$$

since

$$0 \leq \lim_{\varepsilon\to 0}\left|\frac{\mathbb{E}^x[\int_0^\varepsilon f(X_t)dt]}{\varepsilon} - f(x)\right| \leq \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathbb{E}^x\left[\int_0^\varepsilon |f(X_t)-f(x)|\,dt\right].$$

Let us give the estimate for $\mathbb{E}^x\left[\int_0^\varepsilon |f(X_t)-f(x)|\,dt\right]$. Since for every fixed $x \in E$, there exist constants $C > 0$ and $p \geq 1$ depending only on x, such that for every $y \in E$,

$$|f(y) - f(x)| \leq C|y-x|^p,$$

then we have that

$$\mathbb{E}^x\left[\int_0^\varepsilon |f(X_t)-f(x)|\,dt\right] \leq C\mathbb{E}^x\left[\int_0^\varepsilon |X_t - x|^p\,dt\right].$$

Furthermore, by Proposition 4.2.1, $|X_t - x|^p$ is uniformly L_2-integrable within the interval $[0,T]$. Then according to Fubini's theorem, for every $\varepsilon \in [0,T]$

$$\mathbb{E}^x\left[\int_0^\varepsilon |X_t - x|^p dt\right] = \int_0^\varepsilon \mathbb{E}^x\left[|X_t - x|^p\right] dt.$$

By Proposition 4.3, it follows that for every $p \geq 1$, there exists constants K and N, such that

$$\mathbb{E}^x \left[\sup_{s \in [0,t]} |X_s - x|^p \right] \leq K e^{Nt} t^{\frac{p}{2}}. \tag{4.2}$$

It follows that

$$\begin{aligned} & \frac{1}{\varepsilon} \mathbb{E}^x \left[\int_0^\varepsilon |f(X_t) - f(x)| dt \right] \\ \leq \; & C \frac{1}{\varepsilon} \int_0^\varepsilon \mathbb{E}^x \left[|X_t - x|^p \right] dt \\ \leq \; & C K \frac{1}{\varepsilon} \int_0^\varepsilon e^{Nt} t^{\frac{p}{2}} dt \end{aligned}$$

It is easy to see that

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^\varepsilon e^{Nt} t^{\frac{p}{2}} dt = 0.$$

Thus finally we have that

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathbb{E}^x \left[\int_0^\varepsilon |f(X_t) - f(x)| dt \right] = 0$$

Now the proof is complete. ■

Remark 4.2.1 *It is not difficult to see that the requirement of f in Lemma 4.2.2 could be weakened to that $f : E \to \mathbb{R}$ satisfies that for every fixed $x \in E$, there exists $2m$ constants, i.e. $\{C_i, p_i\}_{i=1}^m$ only depending on x, such that for every $y \in E$,*

$$|f(y) - f(x)| \leq \sum_{i=1}^m C_i |y - x|^{p_i}$$

where $C_i > 0, p_i > 0, \ \forall i \in \{1, 2, \ldots, m\}$.

4.3 Properties of the expected signature of a time homogeneous Itô diffusion processes

4.3.1 The mean value property of the expected signature of diffusion processes

An essential property of the expected signature $\Phi(t, x)$ is called the mean value property, which is an immediate consequence of the Markov property of the underlying process X_t. Let us state the mean value property of the expected signature of a diffusion process for a fixed time and give its brief proof.

Lemma 4.3.1 (Mean value property of the expected signature) *For any $(t,x) \in \mathbb{R}^+ \times E$,*

$$\Phi(t,x) \quad = \quad \mathbb{E}^x[S(X_{[0,\varepsilon]}) \otimes \Phi(t-\varepsilon, X_\varepsilon)] \tag{4.3}$$

where $\varepsilon \in [0,t]$.

Proof. Let $\varepsilon \in [0,t]$. By the definition of the expected signature of the process X_t

$$\begin{aligned} \Phi(t,x) &= \mathbb{E}^x[S(X_{[0,t]})] \\ &= \mathbb{E}^x[\mathbb{E}^x[S(X_{[0,t]})|\mathcal{F}_\varepsilon]], \text{ by the tower property.} \end{aligned} \tag{4.4}$$

By Chen's identity we have that

$$S(X_{[0,t]}) = S(X_{[0,\varepsilon]}) \otimes S(X_{[\varepsilon,t]}).$$

Applying it in (4.4), we obtain that

$$\begin{aligned} \Phi(t,x) &= \mathbb{E}^x[\mathbb{E}^x[S(X_{[0,\varepsilon]}) \otimes S(X_{[\varepsilon,t]})|\mathcal{F}_\varepsilon]] \\ &= \mathbb{E}^x[S(X_{[0,\varepsilon]}) \otimes \mathbb{E}^x[S(X_{[\varepsilon,t]})|\mathcal{F}_\varepsilon]]. \end{aligned}$$

Since X_s is a Markov process and it is time homogenous, we have that

$$\mathbb{E}^x[S(X_{[\varepsilon,t]})|\mathcal{F}_\varepsilon] = \mathbb{E}^x[S(X_{[\varepsilon,t]})|X_\varepsilon] = \Phi(t-\varepsilon, X_\varepsilon).$$

Thus the result follows immediately:

$$\Phi(t,x) \quad = \quad \mathbb{E}^x[S(X_{[0,\varepsilon]}) \otimes \Phi(t-\varepsilon, X_\varepsilon)].$$

$\blacksquare$

Remark 4.3.1 *The mean value property of the expected signature of an Itô diffusion process actually provides us with a way to prove the main PDE theorem, i.e. expanding the RHS of Equation (4.3) as a power series in ε. Since the LHS of Equation (4.3) is independent of ε, by setting the coefficient of ε to zero, we obtain the same PDE in Theorem 4.7. But the crucial step is to calculate the quantity $\frac{\partial \Phi(0,x)}{\partial t}$, denoted by $L(x)$, which is defined in Notation 4.5 later on, and is explicitly given in Theorem 4.6. But in this thesis, we only present the alternative approach, i.e. the martingale approach to prove the PDE result.*

4.3.2 One important quantity related to the expected signature of diffusion processes

First of all, let us recall the definition of the infinitesimal generator of a Itô diffusion process.

Definition 4.3.1 *Let $\{X_t\}$ be a time-homogeneous Itô diffusion in E. The infinitesimal generator A of X_t is defined by*

$$Af(x) = \lim_{t \to 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t}; x \in E$$

The set of functions $f : E \to \mathbb{R}$ such that the limit exists at x is denoted by $\mathcal{D}_A(x)$.

It is fundamental to associate a second order partial differential operator A to an Itô diffusion process X_t. The basic connection between A and X_t is that A is the generator of the process X_t (More details could be found in [18]).

Theorem 4.4 *Let X_t be the Itô diffusion, i.e.*

$$dX_t = \mu(X_t)dt + V(X_t) \cdot dW_t.$$

If $f \in C_0^2(E)$ then $f \in \mathcal{D}_A$ and

$$Af(x) = \sum_{j=1}^{d} \mu^{(i)}(x) \frac{\partial f(x)}{\partial x_j} + \sum_{j_1=1}^{d} \sum_{j_2=1}^{d} \frac{1}{2} b_{j_1,j_2}(x) \frac{\partial^2 f(x)}{\partial x_{j_1} \partial x_{j_2}}.$$

where $C_0^2(E)$ is the set of functions with compact support mapping from E to $\mathbb{R}$ with continuous twice partial derivatives.

Let us introduce introduce the important quantity related to the expected signature of a diffusion process, denoted by $L(x)$ as follows.

Notation 4.5 *Let X denote a time homogeneous Itô diffusion process. Let L denote a function mapping from E to $T((E))$, and defined by*

$$L : x \mapsto \lim_{t \to 0} \frac{\mathbb{E}^x[S(X_{[0,t]})] - \mathbf{1}}{t}$$

supposing that the above limit exists.

Remark 4.3.2 *Roughly speaking, $L(x)$ could be seen as the infinitesimal generator A acting on $S(X_{[0,t]})$, since $S(X_{[0,0]}) = \mathbf{1}$. Moreover, let $\Phi(t, x)$ denote the expected signature of $X_{[0,t]}$ starting at the point x(see Definition 4.1.3). Suppose that $\Phi(t, x)$ is partially differentiable with respect to t when $t = 0$. Then the function L satisfies*

$$L(x) = \frac{\partial \Phi(0, x)}{\partial t}, \ \forall x \in E.$$

In the following, Theorem 4.6 gives an explicit representation for $L(x)$.

Theorem 4.6 *Let X be an Itô diffusion process defined by SDE (4.1) and satisfy Condition 4.1. Let $\Phi(t,x)$ denote the expected signature of $X_{[0,t]}$ as before(Definition 4.1.3) and satisfying condition 4.2. Then*

$$\begin{aligned} L(x) &= \lim_{t\to 0}\frac{\mathbb{E}^x[S(X_{[0,t]})]-\mathbf{1}}{t} \\ &= \sum_{j=1}^{d}\mu^{(j)}(x)e_j + \sum_{j_1=1}^{d}\sum_{j_2=1}^{d}\frac{1}{2}b_{j_1,j_2}(x)e_{j_1}\otimes e_{j_2}. \end{aligned}$$

The following lemma will be used to complete the proof of the above theorem.

Lemma 4.3.2 *Let X be an Itô diffusion process, which is defined by SDE* (4.1) *and satisfies* Condition 4.1. *Let $\Phi(t,x)$ denote its expected signature as before* (Definition 4.1.3) *and assume that it satisfies* Condition 4.2. *Let $f:\mathbb{R}^+\times E\to\mathbb{R}$ satisfying the polynomial growth condition. Then for every index I of length greater than or equal to* 1,

$$\lim_{t\downarrow 0^+}\frac{1}{t}\mathbb{E}^x\left[\int_0^t \pi^I(S(X_{[0,s]}))f(s,X_s)ds\right]=0.$$

where $p\geq 1$.

Proof. Fix $T>0$. Let $t\in[0,T]$. By the Cauchy-Schwarz inequality, we have the following inequality

$$\mathbb{E}^x\left[|\pi^I(S(X_{[0,s]}))f(s,X_s)|\right]\leq\left(\mathbb{E}^x\left[\pi^I(S(X_{[0,s]}))^2\right]\right)^{\frac{1}{2}}\left(\mathbb{E}^x\left[f(s,X_s)^2\right]\right)^{\frac{1}{2}}. \tag{4.5}$$

Using the shuffle product property of the signature, we have that

$$\mathbb{E}^x[\pi^I(S(X_{[0,s]}))^2] = (\pi^I\sqcup\pi^I)(\mathbb{E}^x[S(X_{[0,s]})]) = (\pi^I\sqcup\pi^I)(\Phi(s,x)) \tag{4.6}$$

Since the integrand is positive, by Fubini's theorem, it holds that

$$\mathbb{E}^x\left[\int_0^t|\pi^I(S(X_{[0,s]}))f(s,X_s)|ds\right] \leq \int_0^t\mathbb{E}^x[|\pi^I(S(X_{[0,s]}))f(s,X_s)|]ds$$

Substituting (4.5) and (4.6) into the above inequality, we have that

$$\mathbb{E}^x\left[\int_0^t|\pi^I(S(X_{[0,s]}))f(s,X_s)|ds\right] \leq \int_0^t\left((\pi^I\sqcup\pi^I)(\Phi(s,x))\right)^{\frac{1}{2}}\left(\mathbb{E}^x\left[f(s,X_s)^2\right]\right)^{\frac{1}{2}}ds$$

Since f satisfies the polynomial growth condition, so does f^2. By Lemma 4.2.1, there exists a constant C depending on x, T and K, such that

$$\sup_{s\in[0,T]} \mathbb{E}^x\left[f(s,X_s)^2\right] \leq C$$

and thus for $t \leq T$,

$$\begin{aligned}
& \left|\mathbb{E}^x\left[\int_0^t \pi^I(S(X_{[0,s]}))f(s,X_s)ds\right]\right| \\
\leq\ & \mathbb{E}^x\left[\int_0^t |\pi^I(S(X_{[0,s]}))f(s,X_s)|ds\right] \\
\leq\ & C^{\frac{1}{2}}\int_0^t \left((\pi^I \sqcup\!\sqcup\, \pi^I)(\Phi(s,x))\right)^{\frac{1}{2}} ds.
\end{aligned}$$

The function Φ satisfies Condition 4.2, which implies that it has a continuous partial derivative with respect to t, which guarantees that

$$\lim_{t\downarrow 0^+}\frac{1}{t}\int_0^t \left((\pi^I \sqcup\!\sqcup\, \pi^I)(\Phi(s,x))\right)^{\frac{1}{2}} ds = \left((\pi^I \sqcup\!\sqcup\, \pi^I)(\Phi(0,x))\right)^{\frac{1}{2}} = 0.$$

Therefore it follows

$$\begin{aligned}
& \lim_{t\to 0}\frac{1}{t}\left|\mathbb{E}^x\left[\int_0^t \pi^I(S(X_{[0,s]}))f(s,X_s)ds\right]\right| \\
\leq\ & \lim_{t\to 0}\frac{1}{t}\mathbb{E}^x\left[\int_0^t |\pi^I(S(X_{[0,s]}))f(s,X_s)|ds\right] \\
\leq\ & C^{\frac{1}{2}}\lim_{t\to 0}\frac{1}{t}\int_0^t ((\pi^I \sqcup\!\sqcup\, \pi^I)(\Phi(s,x)))^{\frac{1}{2}} ds = 0,
\end{aligned}$$

where $|I| \geq 1$. Finally we have that

$$\lim_{t\to 0}\frac{1}{t}\mathbb{E}^x\left[\int_0^t \pi^I(S(X_{[0,s]}))f(s,X_s)ds\right] = 0$$

■

Now let us complete proof of Theorem 4.6 using the above lemma.

Proof. Let us recall that X satisfies the following SDE in the Itô form, i.e.

$$dX_t^{(j)} = \mu^{(j)}(X_t)dt + \sum_{i=1}^{k} V_i^{(j)}(X_t)dW_t^{(i)},$$

where $j = 1, 2, \ldots, d$.

The first step is to calculate the first term of $L(x)$, i.e.

$$\lim_{t\to 0}\frac{\rho_1(\mathbb{E}^x[S(X_{[0,t]})] - 1)}{t} = \lim_{t\to 0}\frac{\mathbb{E}^x[X_t] - x}{t} = \sum_{i=1}^{d}\mu^{(i)}(x)e_i$$

This can be easily proved by applying Theorem 4.4 for the function $f(x) := x^{(i)}$, since

$$\lim_{t\to 0}\frac{\mathbb{E}^x[\pi^{(i)}(X_t - x)])}{t} = Af(x) = \mu^{(i)}(x).$$

Next we are going to calculate the second term of $L(x)$, i.e.

$$\lim_{t\to 0}\frac{\pi^{(j_1,j_2)}(\mathbb{E}^x[S(X_{[0,t]})] - 1)}{t} = \lim_{t\to 0}\frac{\mathbb{E}^x[\int_0^t (X_s^{(j_1)} - x^{(j_1)}) \circ dX_s^{(j_2)}]}{t}$$

By the relation between Stratonovich integral and Itô integral, we have the following equation, i.e.

$$\begin{aligned}
\int_0^t (X_s^{(j_1)} - x^{(j_1)}) \circ dX_s^{(j_2)} &= \int_0^t (X_s^{(j_1)} - x) dX_s^{(j_2)} + \frac{1}{2}\int_0^t \langle dX_s^{(j_1)}, dX_s^{(j_2)}\rangle \\
&= \int_0^t \left((X_s^{(j_1)} - x^{(j_1)})\mu^{(j_2)}(X_s) + \frac{1}{2} b_{j_1,j_2}(X_s)\right) ds \\
&\quad +\text{martingale part}
\end{aligned}$$

Since μ satisfies a global Lipschitz condition, by Lemma 4.3.2 it holds that

$$\lim_{t\to 0}\frac{1}{t}\mathbb{E}\left[\int_0^t (X_s^{(j_1)} - x^{(j_1)})\mu^{(j_2)}(X_s) ds\right] = \lim_{t\to 0}\frac{1}{t}\mathbb{E}\left[\int_0^t \pi^{(j_1)}(X_s - x)\mu^{(j_2)}(X_s) ds\right] = 0$$

where $j_1, j_2 \in \{1, 2, \ldots, d\}$.

Next let us prove that

$$\lim_{t\to 0}\frac{1}{t}\mathbb{E}^x\left[\int_0^t b_{j_1,j_2}(X_s) ds\right] = b_{j_1,j_2}(x).$$

Due to the Lipschitz condition for V, and $b(x) = V(x)V^T(x)$, there exists a constant $K > 0$, such that

$$\sum_{i=1}^k |V_i(x) - V_i(y)| \leq K|x - y|$$

and therefore it further implies that

$$\begin{aligned}
|b_{j_1,j_2}(y) - b_{j_1,j_2}(x)| &= \left|\sum_{i=1}^k V_i^{(j_1)}(y)V_i^{(j_2)}(y) - V_i^{(j_1)}(x)V_i^{(j_2)}(x)\right| \\
\leq\ & \sum_{i=1}^k |V_i^{(j_1)}(y) - V_i^{(j_1)}(x)||V_i^{(j_2)}(X_s) - V_i^{(j_2)}(x)| \\
& + \sum_{i=1}^k |V_i^{(j_1)}(y) - V_i^{(j_1)}(x)|(|V_i^{(j_1)}(x)| + |V_i^{(j_2)}(x)|) \\
\leq\ & K^2|y - x|^2 + K|y - x| \max_{i=1}^k (|V_i^{(j_1)}(x)| + |V_i^{(j_2)}(x)|)
\end{aligned}$$

By Lemma 4.2.2 and Remark 4.2.1, it follows

$$\lim_{t\to 0}\frac{1}{t}\mathbb{E}^x\left[\int_0^t b_{j_1,j_2}(X_s)ds\right] = b_{j_1,j_2}(x)$$

Then we will have that

$$\lim_{t\to 0}\frac{\pi^{(j_1,j_2)}(\mathbb{E}^x[S(X_{[0,t]})]-1)}{t} = \frac{1}{2}b_{j_1,j_2}(x).$$

Finally we need to show that for every $n \geq 3$, and every index I of length n,

$$\lim_{t\to 0}\frac{\pi^I(\mathbb{E}^x[S(X_{[0,t]})])}{t} = 0.$$

By the definition of the Stratonovich integral, we have

$$\begin{aligned}\pi^I(S(X_{[0,t]})) &= \int_{0<t_1<\cdots<t_n<t} \circ dX_{t_1}^{(i_1)} \circ dX_{t_2}^{(i_2)} \cdots \circ dX_{t_n}^{(i_n)} \\ &= \int_0^t \pi^{I_1}(S(X_{[0,s]}))\cdot dX_s^{(i_n)} + \frac{1}{2}\int_0^t \pi^{I_2}(S(X_{[0,s]}))d\langle X_s^{(i_{n-1})}, X_s^{(i_n)}\rangle,\end{aligned}$$

where $I_1 = (i_1, i_2, \ldots, i_{n-1})$ and $I_2 = (i_1, i_2, \ldots, i_{n-2})$.
So the drift term of $\pi^I(S(X_{[0,t]}))$ is given as follows

$$\pi^{I_1}(S(X_{[0,t]}))\mu^{(i_n)}(X_t) + \frac{1}{2}\pi^{I_2}(S(X_{[0,t]}))b_{i_{n-1},i_n}(X_t)$$

and it holds that

$$\begin{aligned}&\pi^I(\mathbb{E}^x[S(X_{[0,t]})]) \\ = \ &\mathbb{E}^x\left[\int_0^t \pi^{I_1}(S(X_{[0,s]}))\mu^{(i_n)}(X_s) + \frac{1}{2}\pi^{I_2}(S(X_{[0,s]}))b_{i_{n-1},i_n}(X_s)]ds\right].\end{aligned}$$

It is noted that $|I_1|, |I_2| \geq 1$. Moreover, since μ and V_i satisfy a global Lipschitz condition, it implies that μ and b satisfy the polynomial growth condition. Then applying Lemma 4.3.2 to the above equality, we have that

$$\begin{aligned}\lim_{t\to 0}\frac{1}{t}\mathbb{E}^x\left[\int_0^t \pi^{I_1}(S(X_{[0,s]}))\mu^{(i_n)}(X_2)ds\right] &= 0 \\ \lim_{t\to 0}\frac{1}{t}\mathbb{E}^x\left[\int_0^t \pi^{I_2}(S(X_{[0,s]}))b_{i_{n-1},i_n}(X_s)ds\right] &= 0,\end{aligned}$$

Therefore we have that

$$\lim_{t\to 0}\frac{1}{t}\pi^I(\mathbb{E}^x[S(X_{[0,t]})]) = 0.$$

Now our proof is complete. ■

4.4 The PDE of the expected signature of an Itô diffusion process up to a fixed time

The main result of this section is that the expected signature of an Itô diffusion process X up to a fixed time should satisfy a parabolic PDE.

Theorem 4.7 [1] *Let X be an Itô diffusion process, which is defined by the SDE (4.1) and satisfies Condition 4.1. Let $\Phi(t,x)$ be its expected signature as before (Definition 4.1.3) and assume it satisfies Condition 4.2. Then Φ is a function mapping $\mathbb{R}^+ \times E$ to $T((E))$, which satisfies the following PDE:*

$$\left(-\frac{\partial}{\partial t} + A\right)\Phi(t,x) + \sum_{j=1}^{d}\left(\sum_{j_1=1}^{d} b_{j_1,j}(x)e_{j_1}\right) \otimes \frac{\partial \Phi(t,x)}{\partial x_j} +$$
$$\left(\sum_{j=1}^{d}\mu^{(j)}(x)e_j + \frac{1}{2}\sum_{j_1=1}^{d}\sum_{j_2=1}^{d} b_{j_1,j_2}(x)e_{j_1}\otimes e_{j_2}\right) \otimes \Phi(t,x) = 0$$

subject to the initial condition:

$$\Phi(0,x) = \mathbf{1},$$

and the condition:

$$\Phi_0(t,x) = 1,\ \forall (t,x) \in \mathbb{R}^+ \times E$$

where A is the infinitesimal generator of X, i.e.

$$A = \sum_{j=1}^{d}\mu^{(j)}(x)\frac{\partial}{\partial x_j} + \frac{1}{2}\sum_{j_1=1}^{d}\sum_{j_2=1}^{d} b_{j_1,j_2}(x)\frac{\partial^2}{\partial x_{j_1} x_{j_2}}.$$

Proof. Fixing time t, let us define M_u as the conditional expectation of the signature of the stochastic process $X_{[0,t]}$ under the filtration $\mathcal{F}_u$,

$$M_u = \mathbb{E}^x[S(X_{[0,t]})|\mathcal{F}_u]$$

where $u \in [0,t]$.
By Condition 4.2, $\Phi(t,x)$ is finite for every $(t,x) \in \mathbb{R}^+ \times E$, and thus for every index I,

$$\mathbb{E}^x\left[(\pi^I(S(X_{[0,t]}))^2)\right] = (\pi^I \sqcup \pi^I)(\Phi(t,x)) < +\infty,$$

[1] I warmly thank Horatio Boedihardjo for valuable discussions on the result in this chapter, which led me to consider using the martingale approach instead of expanding the expected signature function directly to prove the results, and it did simplify the proof.

i.e. $(\pi^I(S(X_{[0,t]}))$ is L^2 integrable. Therefore M_u is actually a martingale.
By the Markov property of the diffusion process $X_{[0,t]}$,

$$\begin{aligned} M_u &= S(X_{[0,u]}) \otimes \mathbb{E}^x[S(X_{[u,t]})|\mathcal{F}_u] \\ &= S(X_{[0,u]}) \otimes \Phi(t-u, X_u). \end{aligned}$$

We apply Itô's formula to $\Phi(t-u, X_u)$, and for convenience let us denote $\Phi(t-u, x)$ by $f(u, x)$:

$$df(u, X_u) = \frac{\partial f(u, X_u)}{\partial u} du + \sum_{j=1}^{d} \frac{\partial f(u, X_u)}{\partial x_j} dX_u^{(j)} + \frac{1}{2} \sum_{j_1, j_2 = 1}^{d} \frac{\partial^2 f(u, X_u)}{\partial x_{j_1} \partial x_{j_2}} d\langle X^{(j_1)}, X^{(j_2)} \rangle_u.$$

Since the quadratic variation of $X^{(j_1)}$ and $X^{(j_2)}$ is given as follows:

$$d\langle X^{(j_1)}, X^{(j_2)} \rangle_u = b_{j_1, j_2}(X_u) du; \ \forall j_1, j_2 \in \{1, 2, \ldots, d\},$$

we have that

$$\begin{aligned} & df(u, X_u) \\ = \ & \frac{\partial f(u, X_u)}{\partial u} du + \frac{1}{2} \sum_{j_1, j_2 = 1}^{d} \frac{\partial^2 f(u, X_u)}{\partial x_{j_1} \partial x_{j_1}} b_{j_1, j_2}(X_u) du + \sum_{j=1}^{d} \frac{\partial f(u, X_u)}{\partial x_j} dX_u^{(j)} \\ = \ & \left[\frac{\partial f(u, X_u)}{\partial u} + Af(u, X_u) \right] du + \sum_{j=1}^{d} \frac{\partial f(u, X_u)}{\partial x_j} \sum_{i=1}^{k} V_i^{(j)}(X_u) dW_u^{(i)} \\ = \ & \left[\frac{\partial f(u, X_u)}{\partial u} + Af(u, X_u) \right] du + \sum_{i=1}^{k} \left[\sum_{j=1}^{d} \frac{\partial f(u, X_u)}{\partial x_j} V_i^{(j)}(X_u) \right] dW_u^{(i)}. \end{aligned}$$

Recall the definition of the signature of the process $X_{[0,t]}$ and we have the following SDE, where we use $\circ$ to emphasize that it is in the sense of the Stratonovich integral,

$$dS(X_{[0,u]}) = S(X_{[0,u]}) \otimes (\circ dX_u).$$

Using the relation between the Itô integral and Stratonovich integral, we have that

$$
\begin{aligned}
& dS(X_{[0,u]}) \\
= \; & \sum_{j=1}^{d} S(X_{[0,u]}) \otimes (\cdot dX_u^{(j)} e_j) + \sum_{j_1=1}^{d} \sum_{j_2=1}^{d} \frac{1}{2} S(X_{[0,u]}) \otimes (d\langle X_u^{(j_1)}, X_u^{(j_2)} \rangle e_{j_1} \otimes e_{j_2}) \\
= \; & \left(\sum_{j=1}^{d} S(X_{[0,u]}) \otimes \mu^{(j)}(X_u) e_j + \sum_{j_1=1}^{d} \sum_{j_2=1}^{d} \frac{1}{2} S(X_{[0,u]}) \otimes b_{j_1,j_2}(X_u) e_{j_1} \otimes e_{j_2} \right) du \\
& + \sum_{j=1}^{d} S(X_{[0,u]}) \otimes \left(\sum_{i=1}^{k} V_i^{(j)}(X_u) \cdot dW_u^{(i)} e_j \right) \\
= \; & S(X_{[0,u]}) \otimes \left(\sum_{j=1}^{d} \mu^{(j)}(X_u) e_j + \sum_{j_1=1}^{d} \sum_{j_2=1}^{d} \frac{1}{2} b_{j_1,j_2}(X_u) e_{j_1} \otimes e_{j_2} \right) du \\
& + \sum_{i=1}^{k} S(X_{[0,u]}) \otimes \left(\sum_{j=1}^{d} V_j^{(i)}(X_u) e_j \right) \cdot dW_u^{(i)}.
\end{aligned}
$$

By the product rule of Itô integral, the SDE of M_u is given as follows:

$$
\begin{aligned}
dM_u \;=\; & S(X_{[0,u]}) \otimes df(u, X_u) + dS(X_{[0,u]}) \otimes f(u, X_u) \\
& + \sum_{j_2=1}^{d} \sum_{j_1=1}^{d} S(X_{[0,u]}) \otimes e_{j_1} \otimes \frac{\partial f(u, X_u)}{\partial x_{j_2}} d\langle X^{(j_1)}, X^{(j_2)} \rangle_u.
\end{aligned}
$$

Therefore the drift term of M_u should be

$$
\begin{aligned}
& S(X_{[0,u]}) \otimes \left(\frac{\partial f(u, X_u)}{\partial u} + Af(u, X_u) \right) \\
+ \; & S(X_{[0,u]}) \otimes \left(\sum_{j=1}^{d} \mu^{(j)}(X_u) e_j \right) \otimes f(u, X_u) \\
+ \; & S(X_{[0,u]}) \otimes \left(\sum_{j_1=1}^{d} \sum_{j_2=1}^{d} \frac{1}{2} b_{j_1,j_2}(X_u) e_{j_1} \otimes e_{j_2} \right) \otimes f(u, X_u) \\
+ \; & S(X_{[0,u]}) \otimes \sum_{i=1}^{k} \left(\sum_{j=1}^{d} V_i^{(j)}(X_u) e_j \otimes \left(\sum_{j_1=1}^{d} \frac{\partial f(u, X_u)}{\partial x_{j_1}} V_i^{(j_1)}(X_u) \right) \right)
\end{aligned}
$$

which could be simplified further to

$$
S(X_{[0,u]}) \otimes \mathcal{R}(u, X_u)
$$

where

$$\mathcal{R}(u, X_u) = \left(\frac{\partial f(u, X_u)}{\partial u} + Af(u, X_u)\right)$$
$$+ \left(\sum_{j=1}^{d} \mu^{(j)}(X_u)e_i + \sum_{j_1=1}^{d}\sum_{j_2=1}^{d} \frac{1}{2} b_{j_1,j_2}(X_u) e_{j_1} \otimes e_{j_2}\right) \otimes f(u, X_u)$$
$$+ \sum_{j_1=1}^{d} \left(\sum_{j=1}^{d} b_{j,j_1}(X_u)e_j\right) \otimes \frac{\partial f(u, X_u)}{\partial x_{j_1}}.$$

Recall that M_u is a martingale, and M_u is L_2 integrable. By the Doob-Meyer Decomposition Theorem, the drift term should be equal to zero, i.e.

$$S(X_{[0,u]}) \otimes \mathcal{R}(u, X_u) = 0, \ \forall u \in [0, t].$$

Since $S(X_{[0,u]})$ is invertible according to Remark 2.2.1,

$$\mathcal{R}(u, X_u) = (S(X_{[0,u]}))^{-1} \otimes S(X_{[0,u]}) \otimes \mathcal{R}(u, X_u) = (S(X_{[0,u]}))^{-1} \otimes 0 = 0.$$

It implies that

$$\mathbb{E}^x \left[\mathcal{R}(0, X_0)\right] = \mathcal{R}(0, x) = 0.$$

Recall that the definition of $\mathcal{R}(0, x)$,

$$\mathcal{R}(0, x) = \left(\frac{\partial f(0, x)}{\partial u} + Af(0, x)\right)$$
$$+ \left(\sum_{j=1}^{d} \mu^{(j)}(x)e_i + \sum_{j_1=1}^{d}\sum_{j_2=1}^{d} \frac{1}{2} b_{j_1,j_2}(x) e_{j_1} \otimes e_{j_2}\right) \otimes f(0, x)$$
$$+ \sum_{j_1=1}^{d} \left(\sum_{j=1}^{d} b_{j,j_1}(x)e_j\right) \otimes \frac{\partial f(0, x)}{\partial x_{j_1}}. \tag{4.7}$$

Since $f(u, x) = \Phi(t - u, x)$, we have the following

$$\begin{aligned}
f(0, x) &= \Phi(t, x) \\
\frac{\partial f(0, x)}{\partial u} &= -\frac{\partial \Phi(t, x)}{\partial t} \\
\frac{\partial f(0, x)}{\partial x_j} &= \frac{\partial \Phi(t, x)}{\partial x_j}, \forall j = 1, 2, \ldots, d \\
\frac{\partial^2 f(0, x)}{\partial x_{j_1} \partial x_{j_2}} &= \frac{\partial^2 \Phi(t, x)}{\partial x_{j_1} \partial x_{j_2}}, \forall j_1, j_2 = 1, 2, \ldots, d.
\end{aligned}$$

Substituting those equalities into Equation (4.7) completes the proof. ∎

Corollary 4.4.1 is an immediate consequence of Theorem 4.7, but it could be clearer to observe that the Φ_n is determined by Φ_{n-1} and Φ_{n-2} from Corollary 4.4.1.

Corollary 4.4.1 *Let X be an Itô diffusion process, which is defined by SDE (4.1) and satisfies the conditions 4.1. Let $\Phi(t,x)$ be its expected signature as before (Definition 4.1.3) and assume it satisfies Condition 4.2. For every integer $n \geq 0$, let $\Phi_n(t,x)$ denote the n^{th} term of the expected signature of $X_{[0,t]}$, which starts at x. For every $n \in \{2,3,\dots\}$, $\Phi_n(t,x)$ should satisfy the following PDE:*

$$\begin{aligned}\left(-\frac{\partial}{\partial t}+A\right)\Phi_n(t,x) &= -\left(\sum_{j=1}^{d}\mu^{(j)}(x)e_j\right)\otimes\Phi_{n-1}(t,x)\\ &-\sum_{j=1}^{d}\left(\sum_{j_1=1}^{d}b_{j_1,j}(x)e_{j_1}\right)\otimes\frac{\partial\Phi_{n-1}(t,x)}{\partial x_j}\\ &-\left(\frac{1}{2}\sum_{j_1=1}^{d}\sum_{j_2=1}^{d}b_{j_1,j_2}(x)e_{j_1}\otimes e_{j_2}\right)\otimes\Phi_{n-2}(t,x)\end{aligned}\tag{4.8}$$

Moreover, $\Phi_1(t,x)$ should satisfy the following PDE:

$$\left(-\frac{\partial}{\partial t}+A\right)\Phi_1(t,x) = -\sum_{j=1}^{d}\mu^{(j)}(x)e_j$$

where A is the infinitesimal generator of X_t, i.e.

$$A = \sum_{j=1}^{d}\mu^{(j)}(x)\frac{\partial}{\partial x_j}+\frac{1}{2}\sum_{j_1=1}^{d}\sum_{j_2=1}^{d}b_{j_1,j_2}(x)\frac{\partial^2}{\partial x_{j_1}x_{j_2}}$$

Furthermore, for every integer $n \geq 1$, $\Phi_n(t,x)$ should satisfy the initial condition:

$$\Phi_n(0,x)=0,\tag{4.9}$$

and the condition

$$\Phi_0(t,x)=1,\ \forall(t,x)\in\mathbb{R}^+\times E.\tag{4.10}$$

Next, we are interested in understanding what a sufficient condition is, which guarantees that a solution to the PDE (4.11) subject to the initial condition (Equation (4.12)) and Equation (4.13), is indeed the expected signature of $X_{[0,t]}$ starting at a point x. We give the answer in the following theorem.

Theorem 4.8 *Let X be an Itô diffusion process, which is defined by SDE (4.1) and satisfies Condition 4.1. Let $\Phi(t,x)$ be its expected signature as before (Definition*

4.1.3). Suppose that the function $\xi : \mathbb{R}^+ \times E \to T((E))$ is a classical solution to the following PDE:

$$\left(-\frac{\partial}{\partial t} + A\right)\xi(t,x) + \sum_{j=1}^{d}\left(\sum_{j_1=1}^{d} b_{j_1,j}(x)e_{j_1}\right)\otimes\frac{\partial\xi(t,x)}{\partial x_j} +$$
$$\left(\sum_{j=1}^{d}\mu^{(j)}(x)e_j + \frac{1}{2}\sum_{j_1=1}^{d}\sum_{j_2=1}^{d} b_{j_1,j_2}(x)e_{j_1}\otimes e_{j_2}\right)\otimes\xi(t,x) = 0 \tag{4.11}$$

subject to the initial condition:

$$\xi(0,x) = \mathbf{1} \tag{4.12}$$

and the condition:

$$\xi_0(t,x) = 1,\ \forall(t,x)\in\mathbb{R}^+\times E \tag{4.13}$$

where A is the infinitesimal generator of X.
Moreover suppose that $\xi(t,x)$ and the first derivative of $\xi(t,x)$ satisfy the polynomial growth condition. Then $\xi(t,x)$ is the expected signature of $X_{[0,t]}$ with the starting point x, i.e.

$$\xi(t,x) = \Phi(t,x)$$

where $(t,x)\in\mathbb{R}^+\times E$.

Proof. By induction on the tensor degree of the expected signature, it could be proved using Theorem 4.9. ■

In order to show Theorem 4.9, we need Lemma 4.4.1.

Lemma 4.4.1 *Let X be an Itô diffusion process, which is defined by SDE (4.1) and satisfies the condition 4.1. Let $\Phi(t,x)$ be its expected signature as before (Definition 4.1.3) and assume that it satisfies Condition 4.2. For every index I, and every function $f : \mathbb{R}^+\times E\to\mathbb{R}$, satisfying the polynomial growth condition, then $\pi^I(S(X_{[0,t]}))f(t,X_t)$ is uniformly L_2-integrable, i.e. for every fixed $T\in\mathbb{R}^+$,*

$$\sup_{t\in[0,T]}\mathbb{E}^x\left[\left(\pi^I(S(X_{[0,t]}))f(t,X_t)\right)^2\right] < +\infty$$

Proof. Let us estimate the upper bound of $\mathbb{E}^x\left[\left(\pi^I(S(X_{[0,t]}))f(t,X_t))^2\right)\right]$. First of all, we have that

$$\begin{aligned}&\mathbb{E}^x\left[\left(\pi^I(S(X_{[0,t]}))f(t,X_t))^2\right)\right]\\ \leq\ &\frac{1}{2}\left(\mathbb{E}^x\left[\pi^I(S(X_{[0,t]}))^4\right]+\mathbb{E}^x\left[f(t,X_t)^4\right]\right)\end{aligned}$$

By the shuffle product of the signature and the linearity of the expectation,

$$\mathbb{E}^x\left[\pi^I(S(X_{[0,t]}))^4\right]=(\pi^I\sqcup\pi^I\sqcup\pi^I\sqcup\pi^I)(\Phi(t,x))$$

Since $\Phi(t,x)$ is finite, and the projection $(\pi^I\sqcup\pi^I\sqcup\pi^I\sqcup\pi^I)$ is only the linear combination of finite many coefficients of tensors, thus $\mathbb{E}^x\left[\pi^I(S(X_{[0,t]}))^4\right]$ is finite. Moreover since $\Phi(t,x)$ is differentiable with respect to t,

$$\sup_{t\in[0,T]}\Phi(t,x)<+\infty.$$

On the other hand, it is easy to show that for every function $f(t,x)$ satisfying the polynomial growth condition, $f(t,x)^4$ satisfy the polynomial growth condition as well. By Lemma 4.2.1, we have that

$$\sup_{t\in[0,T]}\mathbb{E}^x\left[f(t,X_t)^4\right]<+\infty.$$

Now the proof of the uniformly L_2 integrability of $\pi^I(S(X_{[0,t]}))f(t,X_t)$ for $t\in[0,T]$ is complete. ■

Now let us give the main PDE theorem for $\Phi(t,x)$ as follows and complete its proof.

Theorem 4.9 *Suppose that the expectation of the truncated signature of the diffusion process X up to the degree $N-1$, i.e. $\{\Phi_n\}_{n=0}^{N-1}$ is known,and it and its first partial derivatives all satisfy the polynomial growth condition, where for each $n=0,\dots,N-1$, $\Phi_n\in C^{1,2}(\mathbb{R}^+,E^{\otimes n})$. Suppose that for each $n\in\{1,\dots,N-1\}$, Φ_n satisfies the PDE (4.8), the initial condition (4.9) and the constraint (4.10). Let the function ξ_N mapping from $\mathbb{R}^+\times E^{\otimes N}$ be the solution to the PDE (4.8), subject to the initial condition (4.9) as well as the constraint (4.10). Moreover suppose that ξ_N and $\bigtriangledown\xi_N$ both satisfy the polynomial growth condition. Then*

$$\xi_N(t,x)=\Phi_N(t,x)$$

where $(t,x)\in\mathbb{R}^+\times E$.

Proof. Fix $t \in \mathbb{R}^+$. For every $u \in [0,t]$, let us consider the semimartingale $M_N(u)$ defined as follows, which takes values in $E^{\otimes n}$, i.e.

$$M_N(u) = \xi_N(t-u, X_u) + \sum_{n=1}^{N} \rho_n(S(X_{[0,u]})) \otimes \Phi_{N-n}(t-u, X_u) \tag{4.14}$$

Since for each $n \in \{1, \ldots, N\}$ $\Phi_{N-n} \in C^{1,2}(\mathbb{R}^+, E^{\otimes(N-n)})$, by Itô formula, it holds that

$$\begin{aligned} d\Phi_{N-n}(t-u, X_u) &= \left(-\frac{\partial \Phi_{N-n}(t-u, X_u)}{\partial u} + A\Phi_{N-n}(t-u, X_u)\right) du \\ &\quad + \sum_{i=1}^{k} \left(\sum_{j=1}^{d} \frac{\partial \Phi_{N-n}(t-u, X_u)}{\partial x_j} V_i^{(j)}(X_u)\right) dW_u^{(i)} \end{aligned}$$

Similarly, since $\xi_N \in C^{1,2}(\mathbb{R}^+, E^{\otimes N})$, applying Itô's formula we have that

$$\begin{aligned} d\xi_N(t-u, X_u) &= \left(-\frac{\partial \xi_N(t-u, X_u)}{\partial u} + A\xi_N(t-u, X_u)\right) du \\ &\quad + \sum_{i=1}^{k} \left(\sum_{j=1}^{d} \frac{\partial \xi_N(t-u, X_u)}{\partial x_j} V_i^{(j)}(X_u)\right) dW_u^{(i)} \end{aligned}$$

By the definition of the iterated Stratonovich integral, the dynamics of $\rho_n(S_{[0,u]})$ can be described by the following SDE:

$$\begin{aligned} d\rho_n(S(X_{[0,u]})) &= \rho_{n-1}(S(X_{[0,u]})) \otimes \mu(X_u) du + \rho_{n-1}(S(X_{[0,u]}) \otimes \sum_{i=1}^{k} \left(\sum_{j=1}^{d} V_i^{(j)}(X_u) e_j\right) dW_u^{(i)} \\ &+ \left(\mathbf{1}_{\{n \geq 2\}} \frac{1}{2} \sum_{j_1=1}^{d} \sum_{j_2=1}^{d} \rho_{n-2}(S(X_{[0,u]})) \otimes b_{j_1,j_2}(X_u) du\right) e_{j_1} \otimes e_{j_2} \end{aligned}$$

By the product rule of the Itô's stochastic calculus, we can derive the SDE for $M_N(u)$, i.e. the drift term of $M_N(u)$ is calculated as follows:

$$\begin{aligned} &\left(-\frac{\partial \xi_N(t-u, X_u)}{\partial u} + A\xi_N(t-u, X_u)\right) \\ + &\sum_{n=1}^{N} \rho_n(S(X_{[0,u]})) \otimes \left(-\frac{\partial \Phi_{N-n}(t-u, X_u)}{\partial u} + A\Phi_{N-n}(t-u, X_u)\right) \\ + &\sum_{n=1}^{N} \rho_{n-1}(S(X_{[0,u]})) \otimes (\mu(X_u) \otimes \Phi_{N-n}(u, X_u)) \\ + &\sum_{n=2}^{N} \frac{1}{2} \sum_{j_1=1}^{d} \sum_{j_2=1}^{d} \rho_{n-2}(S(X_{[0,u]})) \otimes b_{j_1,j_2}(X_u) e_{j_1} \otimes e_{j_2} \otimes \Phi_{N-n}(u, X_u) \\ + &\sum_{n=1}^{N} \rho_{n-1}(S(X_{[0,u]})) \otimes \sum_{i=1}^{k} \left(\sum_{j_1=1}^{d} V_i^{(j_1)}(X_u) e_{j_1}\right) \otimes \left(\sum_{j=1}^{d} \frac{\partial \Phi_{N-n}(t-u, X_u)}{\partial x_j} V_i^{(j)}(X_u)\right) \end{aligned}$$

The above formula could be rewritten as follows:

$$
\begin{aligned}
& \left(-\frac{\partial \xi_N(t-u,X_u)}{\partial u} + A\xi_N(t-u,X_u)\right) + \mu(X_u)\otimes\Phi_{N-1}(u,X_u) \\
+ \quad & \left(\frac{1}{2}\sum_{j_1=1}^{d}\sum_{j_2=1}^{d} b_{j_1,j_2}(X_u)e_{j_1}\otimes e_{j_2}\right)\otimes\Phi_{N-2}(u,X_u) \\
+ \quad & \sum_{n=1}^{N}\rho_n(S(X_{[0,u]}))\otimes\mathcal{R}_{N-n}(u,X_u)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{R}_i(u,x) \quad = \quad & \left(-\frac{\partial \Phi_i(t-u,x)}{\partial u} + A\Phi_i(t-u,x)\right) + \left(\sum_{j=1}^{d}\mu^{(j)}(x)e_j\right)\otimes\Phi_{i-1}(u,x) \\
+ \quad & \left(\frac{1}{2}\sum_{j_1=1}^{d}\sum_{j_2=1}^{d} b_{j_1,j_2}(x)e_{j_1}\otimes e_{j_2}\right)\otimes\Phi_{i-2}(u,x) \\
+ \quad & \sum_{j=1}^{d}\left(\sum_{j_1=1}^{d} b_{j,j_1}(x)e_j\right)\otimes\frac{\partial\Phi_{i-1}(t-u,x)}{\partial x_j}.
\end{aligned}
$$

Since $\{\Phi_i\}_{i=1}^{N-1}$ and ξ_N satisfy the recursive PDE (4.8), then the drift term of $M_N(u)$ should be equal to 0, which implies that $M_N(u)$ is a local martingale.
Next let us consider the dW_t term of $M_N(u)$ and we prove that $M_N(u)$ is a martingale. In order to do so, let us calculate the dW_t term of $M_N(u)$, i.e.

$$
\begin{aligned}
& \sum_{i=1}^{k}\left(\sum_{j=1}^{d}\frac{\partial\xi_N(t-u,X_u)}{\partial x_j}V_i^{(j)}(X_u)\right)dW_t^{(i)} \\
+ \quad & \sum_{n=1}^{N}\rho_n(S(X_{[0,u]}))\otimes\left(\sum_{i=1}^{k}\left(\sum_{j=1}^{d}\frac{\partial\Phi_{N-n}(t-u,X_u)}{\partial x_j}V_i^{(j)}(X_u)\right)dW_t^{(i)}\right) \\
+ \quad & \sum_{n=1}^{N}\rho_{n-1}(S(X_{[0,u]}))\otimes\left(\sum_{i=1}^{k}\left(\sum_{j=1}^{d}V_i^{(j)}(X_u)e_j\right)dW_u^{(i)}\otimes\Phi_{N-n}(t-u,X_u)\right).
\end{aligned}
$$

By the assumption that ξ_N and $\{\Phi_n\}_{n=1}^{N-1}$ has continuous first partial derivatives, which satisfy the polynomial growth condition, as well as that V satisfies the linear growth condition, it guarantees that the dW_t term of M_u is L_2-integrable in $u\in[0,t]$ by Lemma 4.4.1, thus M_u is a martingale. It implies that

$$
\mathbb{E}^x[M_N(t)] = M_N(0)
$$

By the construction of the martingale M_N, plugging $u = 0, t$ into the Equation (4.14) and using that $\Phi_i(0, x) = 0, \forall i \in \mathbb{N}$, $S(X_{[0,0]}) = \mathbf{1}$ and $\Phi_0(t, x) = 1$, we have that

$$\begin{aligned} M_N(t) &= \rho_N(S(X_{[0,t]})) \\ M_N(0) &= \xi_N(t, x) \end{aligned}$$

From those two equations, we have that

$$\xi_N(t, x) = M_N(0) = \mathbb{E}^x[M_N(t)] = \mathbb{E}^x[\rho_N(S(X_{[0,t]}))]$$

■

4.5 Expected signature solved by a probabilistic approach

Let us recall the celebrated Feynman−Kac formula, that establishes a link between parabolic partial differential equations and stochastic processes. It offers a method of solving certain PDEs by simulating random paths of a stochastic process.

Theorem 4.10 *Fix an arbitrary $T > 0$. Let X be an Itô diffusion process defined by SDE (4.1) and satisfying Condition 4.1. Let $f : [0, T] \times E \to \mathbb{R}$, $G : [0, T] \times E \to \mathbb{R}$ and $\eta : E \to \mathbb{R}$. f and G both satisfy the polynomial growth condition. Suppose that $F(t, x) : [0, T] \times E \to \mathbb{R}$ is continuous of class $C^{1,2}([0, T] \times E)$, and satisfies the Cauchy problem:*

$$\begin{aligned} -\frac{\partial F(t, x)}{\partial t} &= AF(t, x) - G(t, x)F(t, x) + f(t, x); \\ F(T, x) &= \eta(x);\ x \in E \end{aligned} \tag{4.15}$$

as well as the polynomial growth condition where A is the infinitesimal generator of X. Then $F(t, x)$ admits the stochastic representation

$$F(t, x) = \mathbb{E}\left[\eta(X_T)Z_T + \int_t^T f(s, X_s)Z_s ds | X_t = x\right]$$

where $t \in [0, T]$ and

$$Z(t, x) = \exp\left(-\int_t^T G(s, x)ds\right).$$

In particular, such a solution is unique.

Proof. The proof can be found in [10]. ■

Remark 4.5.1 *Theorem 4.10 is a special case of Theorem 5.7.6 (on Page 366 in [10]). In [10], the Feynman-Kac theorem is valid for a time-inhomogeneous diffusion process under a more general condition of the drift and volatility. It is easy to prove that all the conditions are satisfied if X is a time homogenous Itô diffusion process defined by SDE (4.1) and satisfying Condition 4.1.*

Remark 4.5.2 *The existence and uniqueness of the solution to the Cauchy problem in Theorem 4.10 as well as the regularity of the gradient estimates has been studied in [11] and [1]. In [11], Theorem 10 on Page 118, it is shown that, if instead of the assumption of the polynomial growth condition on the solution to PDE F, one imposes that the first partial derivative of f, G, μ, and V satisfy the polynomial growth condition, the Feynman-Kac Theorem still holds and moreover we have that F satisfies the polynomial growth condition as well.*

Applying Theorem 4.10 to the PDE for our expected signature function Φ, the following theorem is an immediate consequence and it provides a way to present Φ in terms of a probabilistic object.

Theorem 4.11 *Let X be an E-valued Itô diffusion process defined by SDE (4.1) and satisfying Condition 4.1. Let $\Phi(t,x)$ be its expected signature as before satisfying Condition 4.2. Moreover suppose that $\Phi(t,x)$ and its partial derivatives all satisfy the polynomial growth condition. Let us define the map $f_n : \mathbb{R}^+ \times E \to E^{\otimes n}$ as follows: if $n \geq 2$, then*

$$\begin{aligned} f_n(t,x) &= \left(\sum_{j=1}^{d} \mu^{(j)}(x) e_j\right) \otimes \Phi_{n-1}(t,x) \\ &\quad + \sum_{j=1}^{d} \left(\sum_{j_1=1}^{d} b_{j_1,j}(x) e_{j_1}\right) \otimes \frac{\partial \Phi_{n-1}(t,x)}{\partial x_j} \\ &\quad + \left(\frac{1}{2} \sum_{j_1=1}^{d} \sum_{j_2=1}^{d} b_{j_1,j_2}(x) e_{j_1} \otimes e_{j_2}\right) \otimes \Phi_{n-2}(t,x) \end{aligned}$$

if $n = 1$, then

$$f_1(t,x) = \sum_{j=1}^{d} \mu^{(j)}(x) e_j$$

It holds that for each $n \geq 1$,

$$\Phi_n(t,x) = \mathbb{E}^x \left[\int_0^t f_n(s, X_s) ds\right]$$

Proof. Fix an arbitrary $T > 0$. Let $F_n : E \times E \to E^{\otimes n}$, be defined as

$$F_n(t, x) = \Phi_n(T - t, x)$$

where $x \in E$ and $t \in [0, T]$.
By the Chain rule, it is easy to verify that

$$\begin{aligned}
\frac{\partial F_n(t,x)}{\partial t} &= -\frac{\partial \Phi_n(T-t,x)}{\partial t} \\
\frac{\partial F_n(t,x)}{\partial x_j} &= \frac{\partial \Phi_n(T-t,x)}{\partial x_j} \\
\frac{\partial^2 F_n(t,x)}{\partial x_{j_1} \partial x_{j_2}} &= \frac{\partial^2 \Phi_n(T-t,x)}{\partial x_{j_1} \partial x_{j_2}} \\
F(T,x) &= \Phi_n(0,x)
\end{aligned}$$

where $j, j_1, j_2 \in \{1, 2, \ldots, d\}$.
By Corollary 4.4.1, we know that the PDE and the initial condition for $\Phi_n(t, x)$, where $n \in \mathbb{N}$. Using the relation of Φ_n and F_n, the initial value problem of Φ_n can be transformed to the terminal value problem for F_n, i.e. $F_n(t, x)$ should satisfy the following PDE: for $n \geq 2$,

$$\begin{aligned}
\left(\frac{\partial}{\partial t} + A\right) F_n(t,x) = &- \left(\sum_{j=1}^{d} \mu^{(j)}(x) e_j\right) \otimes F_{n-1}(t,x) \\
&- \sum_{j=1}^{d} \left(\sum_{j_1=1}^{d} b_{j_1,j}(x) e_{j_1}\right) \otimes \frac{\partial F_{n-1}(t,x)}{\partial x_j} \\
&- \left(\frac{1}{2} \sum_{j_1=1}^{d} \sum_{j_2=1}^{d} b_{j_1,j_2}(x) e_{j_1} \otimes e_{j_2}\right) \otimes F_{n-2}(t,x);
\end{aligned}$$

for $n = 1$,

$$\left(\frac{\partial}{\partial t} + A\right) (F_1(t,x)) = -\sum_{j=1}^{d} \mu^{(j)}(x) e_j$$

subject to

$$F_n(T, x) = 0, \ \forall n \in \mathbb{N}, \ \forall x \in E.$$

By the assumption, $\Phi(t, x)$ and its partial derivatives all satisfy the polynomial growth condition. Then f_n satisfies the polynomial growth condition as well. Therefore by applying Theorem 4.10 to the function F and plugging $V(t,x) = 0$ to it, $F_n(t,x)$ admits the following stochastic representation:

$$F_n(t,x) = \mathbb{E}\left[\int_t^T f_n(s, X_s)ds | X_t = x\right]$$

where $t \in [0,T]$.
In particular, for $t = 0$,

$$\Phi_n(T,x) = F_n(0,x) = \mathbb{E}^x\left[\int_0^T f_n(s, X_s)ds\right]$$

Since T is arbitrary, the proof is complete. ∎

4.6 Concrete examples

4.6.1 Multi-dimensional Brownian motion

Let $\varphi(t)$ denote the expected signature of the $d-$dimensional Brownian motion up to time t used in Chapter 3. Note that $\varphi(t)$ does not depend on the starting point x of the Brownian motion. Plugging $\mu(x) = 0$ and $V(x) = 1$ into Theorem 4.8, we know that $\varphi(t)$ satisfies the following PDE:

$$-\frac{\partial \varphi(t)}{\partial t} + \frac{1}{2}\left(\sum_{j=1}^{d} e_j \otimes e_j\right) \otimes \varphi(t) = 0, \tag{4.16}$$

with the initial condition:

$$\varphi(0) = \mathbf{1},$$

and the constraint:

$$\rho_0(\varphi(t)) = 1.$$

Conversely, by Theorem 4.9, the expected signature of Brownian motion denoted by $\Phi(t,x)$ should satisfy the following PDE, i.e.

$$\begin{aligned}\left(-\frac{\partial}{\partial t} + \frac{1}{2}\Delta\right)\Phi(t,x) + \sum_{j=1}^{d}\left(\sum_{j_1=1}^{d} e_{j_1}\right) \otimes \frac{\Phi(t,x)}{\partial x_j} \\ + \left(\frac{1}{2}\sum_{j_1=1}^{d}\sum_{j_2=1}^{d} e_{j_1} \otimes e_{j_2}\right) \otimes \Phi(t,x) = 0\end{aligned} \tag{4.17}$$

subject to the initial condition

$$\Phi(0, x) = \mathbf{1}$$

and the condition

$$\Phi_0(t, x) = 1.$$

By induction it could be proved that the solution to the above PDE problem does not depend on x, and it is just a function of t, since the drift vector and dispersion matrix are constant. PDE (4.17) could be further reduced to PDE (4.16). It is not difficult to compute the solution to PDE (4.16), that is indeed

$$\varphi(t) = \exp\left(\frac{t}{2}\sum_{j=1}^{d} e_j \otimes e_j\right),$$

which agrees with the representation of the expected signature of Brownian motion up to time t in Theorem 3.1.

4.6.2 One-dimensional Ornstein−Uhlenbeck process

Definition 4.6.1 *A one dimensional Ornstein− Uhlenbeck process(OU process), X_t, satisfies the following stochastic differential equation:*

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t$$

where $\theta > 0, \mu$ and $\sigma > 0$ are parameters.

In this subsection, let $\Phi(t, x)$ denote the expected signature of the one dimensional OU process $\{X_s\}_{s\in[0,t]}$ with the starting point at x. Moreover let $\Phi_n(t, x)$ denote the n^{th} term of $\Phi(t, x)$, i.e.

$$\Phi_n(t, x) = \rho_n(\Phi(t, x))$$

where $n \in \mathbb{N}$.
We give an explicit formula for the expected signature of the OU process.

Lemma 4.6.1 *Let $\Phi(t, x)$ denote the expected signature of the one dimensional OU process $\{X_s\}_{s\in[0,t]}$ from the time 0 to t with the starting point at x.*

$$\begin{aligned}
\Phi_{2k+1}(t, x) &= \frac{1}{(2k+1)!}\left(\sum_{i=0}^{k} C_{2k+1}^{2i}(\mu - x)^{2k+1-2i}(1 - e^{-\theta t})^{2k+1-2i}\left(\frac{\sigma^2}{2\theta}\right)^i\left(1 - e^{-2\theta t}\right)^i (2i-1)!!\right) \\
\Phi_{2k}(t, x) &= \frac{1}{(2k)!}\left(\sum_{i=0}^{k} C_{2k}^{2i}(\mu - x)^{2k-2i}(1 - e^{-\theta t})^{2k-2i}\left(\frac{\sigma^2}{2\theta}\right)^i (1 - e^{-2\theta t})^i (2i-1)!!\right)
\end{aligned}$$

where $k \in \mathbb{N}$ and $C_m^n = \binom{m}{n}$.

Proof. Since X_t is an Ornstein−Uhlenbeck process with parameters μ, θ and σ with starting point x, then X_t could be expressed

$$X_t = xe^{-\theta t} + \mu(1 - e^{-\theta t}) + \int_0^t \sigma e^{\theta(s-t)} dW_s$$

It implies that X_t is normally distributed with mean $xe^{-\theta t} + \mu(1 - e^{-\theta t})$ and variance $\frac{\sigma^2}{2\theta}(1 - e^{-2\theta t})$. It is a well known result, that the central moments of the normal distribution with mean a and variance b^2 are given as follows:

$$\mathbb{E}[(X-a)^p] = \begin{cases} 0 & \text{if } p \text{ is odd;} \\ b^p(p-1)!! & \text{if } p \text{ is even.} \end{cases}$$

where $p \in \mathbb{N}$. On the other hand, the expected signature of $\{X_s\}_{s \in [0,t]}$ can be reduced to a linear combination of the moments of X_t, since X_t is one dimensional, i.e.

$$\Phi(t,x) = \mathbb{E}^x[S(X_{[0,t]})] = 1 + \sum_{n=1}^{+\infty} \frac{1}{n!} \mathbb{E}^x[(X_t - x)^n] e_1^{\otimes n}$$

and

$$\begin{aligned} \mathbb{E}^x[(X_t - x)^n] &= \mathbb{E}[(X_t - \mathbb{E}[X_t] + \mathbb{E}[X_t] - x)^n] \\ &= \sum_{i=0}^{n} C_n^i \mathbb{E}[(X_t - \mathbb{E}[X_t])^i](\mathbb{E}[X_t] - x)^{n-i} \end{aligned}$$

Therefore we have the following formula for $\Phi(t,x)$:

$$\begin{aligned} \Phi_{2k+1}(t,x) &= \frac{1}{(2k+1)!}\left(\sum_{i=0}^{k} C_{2k+1}^{2i}(\mu - x)^{2k+1-2i}(1-e^{-\theta t})^{2k+1-2i}\left(\frac{\sigma^2}{2\theta}\right)^i \left(1 - e^{-2\theta t}\right)^i (2i-1)!!\right) \\ \Phi_{2k}(t,x) &= \frac{1}{(2k)!}\left(\sum_{i=0}^{k} C_{2k}^{2i}(\mu - x)^{2k-2i}(1-e^{-\theta t})^{2k-2i}\left(\frac{\sigma^2}{2\theta}\right)^i (1 - e^{-2\theta t})^i (2i-1)!!\right) \end{aligned}$$

where $k \geq 0$. ■

Lemma 4.6.2 *For any arbitrary integer $n \geq 2$, the following equations hold:*

$$\frac{\partial \Phi_n(t,x)}{\partial t} = \theta e^{-\theta t}(\mu - x)\Phi_{n-1}(t,x) + \frac{\sigma^2}{2} e^{-2\theta t}\Phi_{n-2}(t,x) \tag{4.18}$$

$$\frac{\partial \Phi_n(t,x)}{\partial x} = -(1 - e^{-\theta t})\Phi_{n-1}(t,x) \tag{4.19}$$

$$\frac{\partial^2 \Phi_n(t,x)}{\partial x^2} = (1 - e^{-\theta t})^2 \Phi_{n-1}(t,x) \tag{4.20}$$

Proof. We will prove those three equalities by considering cases when n is odd or even respectively.
Let us check Equation (4.18) firstly.
Since the following equality holds:

$$\begin{aligned}
& \frac{1}{(2k-1)!}\left(\sum_{i=0}^{k-1} C_{2k-1}^{2i}(\mu-x)^{2k-2i}(1-e^{-\theta t})^{2k-2i}\left(\frac{\sigma^2}{2\theta}\right)^i(1-e^{-2\theta t})^{i-1}(2i-1)!!\right) \\
= & \frac{1}{(2k-1)!}\left(\sum_{i=0}^{k-1} C_{2k-1}^{2i}(\mu-x)^{2k-2i-2}(1-e^{-\theta t})^{2k-2i-2}\left(\frac{\sigma^2}{2\theta}\right)^{i+1}(1-e^{-2\theta t})^{i}(2i+1)!!\right) \\
= & \frac{1}{(2k-2)!}\left(\sum_{i=0}^{k-1} C_{2k-2}^{2i}(\mu-x)^{2k-2i-2}(1-e^{-\theta t})^{2k-2i-2}\left(\frac{\sigma^2}{2\theta}\right)^{i}(1-e^{-2\theta t})^{i}(2i+1)!!\right)\frac{\sigma^2}{2} \\
= & \frac{\sigma^2}{2}\Phi_{2k-2}(t,x),
\end{aligned}$$

we have that

$$\begin{aligned}
& \frac{\partial\Phi_{2k}(t,x)}{\partial t} \\
= & \frac{1}{(2k-1)!}\theta e^{-\theta t}\left(\sum_{i=0}^{k-1} C_{2k-1}^{2i}(\mu-x)^{2k-2i}(1-e^{-\theta t})^{2k-2i-1}\left(\frac{\sigma^2}{2\theta}\right)^i(1-e^{-2\theta t})^{i}(2i-1)!!\right) \\
+ & \frac{1}{(2k-1)!}\theta e^{-2\theta t}\left(\sum_{i=1}^{k} C_{2k-1}^{2i}(\mu-x)^{2k-2i}(1-e^{-\theta t})^{2k-2i}\left(\frac{\sigma^2}{2\theta}\right)^i(1-e^{-2\theta t})^{i-1}(2i-1)!!\right) \\
= & \theta e^{-\theta t}(\mu-x)\Phi_{2k-1}(t,x)+\frac{\sigma^2}{2}e^{-2\theta t}\Phi_{2k-2}(t,x).
\end{aligned}$$

Similarly we have

$$\begin{aligned}
& \frac{\partial\Phi_{2k+1}(t,x)}{\partial t} \\
= & \frac{1}{(2k)!}\theta e^{-2\theta t}\left(\sum_{i=1}^{k} C_{2k}^{2i-1}(\mu-x)^{2k+1-2i}(1-e^{-\theta t})^{2k-2i+1}\left(\frac{\sigma^2}{2\theta}\right)^i(1-e^{-2\theta t})^{i-1}(2i-1)!!\right) \\
& +\frac{1}{(2k)!}\theta e^{-\theta t}\left(\sum_{i=0}^{k} C_{2k}^{2i}(\mu-x)^{2k+1-2i}(1-e^{-\theta t})^{2k-2i}\left(\frac{\sigma^2}{2\theta}\right)^i(1-e^{-2\theta t})^{i}(2i-1)!!\right) \\
= & \theta e^{-\theta t}(\mu-x)\Phi_{2k}(t,x)+\frac{\sigma^2}{2}e^{-2\theta t}\Phi_{2k-1}(t,x).
\end{aligned}$$

Now let us check Equation (4.19). The following is easy to verify after simple calculation:

$$\begin{aligned}
&\frac{\partial \Phi_{2k+1}(t,x)}{\partial x} \\
= \;& \frac{1}{(2k)!}\left(-\sum_{i=0}^{k} C_{2k}^{2i}(\mu - x)^{2k-2i}(1-e^{-\theta t})^{2k+1-2i}\left(\frac{\sigma^2}{2\theta}\right)^i (1-e^{-2\theta t})^i (2i-1)!!\right) \\
= \;& -(1-e^{-\theta t})\Phi_{2k}(t,x);
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial \Phi_{2k}(t,x)}{\partial x} \\
= \;& \frac{1}{(2k-1)!}\left(-\sum_{i=0}^{k-1} C_{2k-1}^{2i}(\mu - x)^{2k-1-2i}(1-e^{-\theta t})^{2k-2i}\left(\frac{\sigma^2}{2\theta}\right)^i (1-e^{-2\theta t})^i (2i-1)!!\right) \\
= \;& -(1-e^{-\theta t})\Phi_{2k-1}(t,x).
\end{aligned}$$

Finally Equation (4.20) is proved, since the following equations hold.

$$\begin{aligned}
&\frac{\partial^2 \Phi_{2k+1}(t,x)}{\partial x^2} \\
= \;& \frac{1}{(2k-1)!}\left(\sum_{i=0}^{k-1} C_{2k-1}^{2i}(\mu - x)^{2k-2i-1}(1-e^{-\theta t})^{2k+1-2i}\left(\frac{\sigma^2}{2\theta}\right)^i (1-e^{-2\theta t})^i (2i-1)!!\right) \\
= \;& -(1-e^{-\theta t})^2\Phi_{2k}(t,x)
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial^2 \Phi_{2k}(t,x)}{\partial x^2} \\
= \;& \frac{1}{(2k-2)!}\left(\sum_{i=0}^{k-1} C_{2k-2}^{2i}(\mu - x)^{2k-2-2i}(1-e^{-\theta t})^{2k-2i}\left(\frac{\sigma^2}{2\theta}\right)^i (1-e^{-2\theta t})^i (2i-1)!!\right) \\
= \;& (1-e^{-\theta t})^2\Phi_{2k-1}(t,x).
\end{aligned}$$

■

By Corollary4.4.1, $\Phi(t,x)$ should satisfy the following recursive relationship, i.e.

$$\begin{aligned}
&\left(-\frac{\partial}{\partial t} + A\right)\Phi_n(t,x) \\
= \;& \begin{cases} -\theta(\mu - x)\Phi_{n-1}(t,x) - \frac{1}{2}\sigma^2\Phi_{n-2}(t,x) - \sigma^2\frac{\partial \Phi_{n-1}(t,x)}{\partial x} & \text{if } n \geq 2; \\ -\theta(\mu - x) & \text{if } n = 1. \end{cases}
\end{aligned} \tag{4.21}$$

Since we have the explicit formula for the expected signature of the OU process, it could be used to verify our PDE theorem by checking Eqn (4.21). Now let us compute

the LHS and RHS of Equation (4.21) respectively.
For $n \geq 2$,

$$\begin{aligned}
(LHS) &= (-\frac{\partial}{\partial t} + A)(\Phi_{2k+1}(t,x)) \\
&= -\theta e^{-\theta t}(\mu - x)\Phi_{2k}(t,x) - \frac{\sigma^2}{2} e^{-2\theta t}\Phi_{2k-1}(t,x) \\
&\quad -\theta(\mu - x)(1 - e^{-\theta t})(\Phi_{2k}(t,x)) \\
&\quad +\frac{1}{2}\sigma^2(1 - e^{-\theta t})^2\Phi_{2k-1}(t,x) \\
&= -\theta e^{-\theta t}(\mu - x)\Phi_{2k}(t,x) + \sigma^2(\frac{1}{2} - e^{-\theta t})\Phi_{2k-1}(t,x) \\
(RHS) &= -\theta(\mu - x)\Phi_{2k}(t,x) - \frac{1}{2}\sigma^2\Phi_{2k-1}(t,x) - \sigma^2\frac{\partial\Phi_{2k}(t,x)}{\partial x} \\
&= -\theta(\mu - x)\Phi_{2k}(t,x) - \frac{1}{2}\sigma^2\Phi_{2k-1}(t,x) + \sigma^2(1 - e^{-\theta t})\Phi_{2k-1}(t,x) \\
&= -\theta(\mu - x)\Phi_{2k}(t,x) + \sigma^2(\frac{1}{2} - e^{-\theta t})\Phi_{2k-1}(t,x).
\end{aligned}$$

Let us consider the first term of $\Phi(t,x)$, i.e.

$$\Phi_1(t,x) = \mathbb{E}^x[X_t - x] = (\mu - x)(1 - e^{-\theta t}).$$

It is easy to verify that

$$\left(-\frac{\partial}{\partial t} + A\right)(\Phi_1(t,x)) = -\theta(\mu - x).$$

which agrees with what we state in our PDE theorem 4.7.

4.6.3 Multi-dimensional Ornstein−Uhlenbeck process

A $d-$dimensional canonical OU process X driven by Brownian motion is defined as the strong solution to the following SDE:

$$dX_t = -QX_t dt + dW_t,$$

where Q is a $d \times d$ matrix, W_t is a standard $d-$dimensional Brownian motion starting at the origin. It is actually a multi-dimensional OU process with the mean reversion level 0, like the one-dimensional case in the previous subsection. In the remaining part of this subsection, we only consider the canonical OU process. For simplicity, we call it a d-dimensional OU process.

Let $E = \mathbb{R}^d$. Suppose that the starting point of this multi-dimensional OU process is $x \in E$, X_t could be expressed as

$$X_t = e^{-tQ}x + \int_0^t e^{-(t-s)Q} dW_s, \ t \in \mathbb{R}^+.$$

Therefore X_t is normally distributed with the mean $e^{-tQ}x$ and the covariance matrix $\int_0^t \exp^{-(t-s)Q}(\exp^{-(t-s)Q})^T ds$, which enables us to compute all the moments of X_t easily. Moreover, in the same way as the one dimensional OU process case, all the moments are polynomial in $x_1, x_2, \ldots, x_t$, i.e. for every non-negative integer $k_1, \ldots, k_d$, there exists a polynomial $P_{k_1,k_2,\ldots,k_d}(x_1, \ldots, x_d)$ with coefficients, which only depend on the time parameter t, such that

$$\mathbb{E}^x[(X_t^{(i_1)})^{k_1} \ldots (X_t^{(i_d)})^{k_d}] = P^t_{k_1,k_2,\ldots,k_d}(x^{(1)}, \ldots, x^{(d)}).$$

Using the explicit form of the moments of the OU-process, combined with our PDE theorem for the expected signature, we are able to prove that the expected signature $\Phi(t, x)$ is in a polynomial form as well, which is stated in the following theorem.

Theorem 4.12 *Let $\Phi(t, x)$ denote the expected signature of multi-dimensional OU-process up to the time t with the starting point $x \in E$. Then for every positive integer n, for every index $I = (i_1, i_2, \ldots, i_n)$, where each $i_j \in \{1, 2, \ldots, d\}$ for $j = 1, 2, \ldots, n$, there exists a polynomial in the variables $x^{(1)}, \ldots, x^{(d)}$ denoted by $H^I_t(.)$, such that*

$$\pi^I(\Phi(t, x)) = H^I_t(x^{(1)}, \ldots, x^{(d)})$$

where the coefficients of the polynomial of $H^I_t(.)$ are smooth functions of t.

Proof. We are going to prove that $\Phi(t, x)$ are polynomials in x with coefficients, which are smooth functions of t by induction. It is obviously true for $n = 0, 1$, since

$$\begin{aligned} \Phi_0(t, x) &= 1 \\ \Phi_1(t, x) &= (e^{-tQ} - \mathcal{I})x \end{aligned}$$

where $\mathcal{I}$ is $d \times d$ identity matrix.
Note that the coefficients of $\Phi_0(t, x)$ and $\Phi_1(t, x)$ are differentiable functions of t, and therefore, for every fixed $T > 0$, they are bounded for every interval $[0, T]$.
Suppose that for any $i \leq n - 1$, this theorem is true. Let us consider $\Phi_n(t, x)$.
For any $n \geq 2$, for any index $I = (i_1, i_2, \ldots, i_n)$ of length n, let us define $f_I(t, x)$ as follows:

$$f_I(t, x) = \mu^{(i_1)}(x)\pi^{I_1}(\Phi(t, x)) + \sum_{j=1}^{d} \mathcal{I}_{i_1,j} \frac{\partial \pi^{I_1}(\Phi(t, x))}{\partial x_j} + \frac{1}{2}\mathcal{I}_{i_1,i_2}\pi^{I_2}(\Phi(t, x))$$

where $I_1 = (i_2, \dots, i_n)$, and $I_2 = (i_3, \dots, i_n)$.
By the induction hypothesis, $\pi^{I_1}(\Phi(t,x))$ and $\pi^{I_2}(\Phi(t,x))$ are both polynomials. In addition to that, $\mu^{(i_1)}(x)$ is a linear function of x. Therefore $f_I(t,x)$ is a polynomial in x as well and its coefficients are also smooth function of t.
Let us write $f_I(t,x)$ as follows:

$$f_I(s,x) = \sum_{k_1 \dots k_d} C^I_{k_1 \dots k_d}(s)(x^{(1)})^{k_1} \dots (x^{(d)})^{k_d}$$

Note that for any fixed I, only finitely many functions $C^I_{k_1 \dots k_d}(.)$ are non-zero.
Moreover, since the coefficients of $\pi^{I_1}(\Phi(t,x))$ and $\pi^{I_2}(\Phi(t,x))$ are smooth functions of t, f_I satisfies the polynomial growth condition, too.
Let us consider the following PDE problem, i.e.

$$\begin{cases} (-\frac{\partial}{\partial t} + A)F_I(t,x) = f_I(s,x), \\ F_I(0,x) = 0. \end{cases} \tag{4.22}$$

where A is the infinitesimal generator of OU process.
By Theorem 10 on page 118 of [11](the Feynman-Kac theorem), the solution to the PDE exists, and is unique. Moreover it has a stochastic representation as follows:

$$F_I(t,x) = \mathbb{E}^x \left[\int_0^t f_I(s, X_s) ds \right]$$

It is obvious that

$$\begin{aligned} F_I(t,x) &= \mathbb{E}^x \left[\int_0^t f_I(s, X_s) ds \right] = \mathbb{E}^x \left[\int_0^t \sum_{k_1 \dots k_d} C^I_{k_1 \dots k_d}(s)(X_s^{(1)})^{k_1} \dots (X_s^{(d)})^{k_m} ds \right] \\ &= \sum_{k_1 \dots k_d} \int_0^t C^I_{k_1 \dots k_d}(s) \mathbb{E}^x \left[(X_s^{(1)})^{k_1} \dots (X_s^{(d)})^{k_d} \right] ds \\ &= \sum_{k_1 \dots k_d} \int_0^t C^I_{k_1 \dots k_d}(s) \mathbb{E}^x \left[(X_s^{(1)})^{k_1} \dots (X_s^{(d)})^{k_d} \right] ds \\ &= \sum_{k_1 \dots k_d} \int_0^t C^I_{k_1 \dots k_d}(s) P^s_{k_1, k_2, \dots, k_d}(x^{(1)}, \dots, x^{(d)}) ds \end{aligned}$$

Since all the moments of X_t, i.e. $P^t_{k_1, k_2, \dots, k_d}(x^{(1)}, \dots, x^{(d)})$ are polynomials in x with coefficients which are smooth functions of t, then the function $\pi^I(\Phi(t,x))$ should be a polynomial in the variables $x^{(1)}, \dots, x^{(d)}$ with smooth coefficients too, so does the derivatives. $f_I(t,x)$ and $F_I(t,x)$ both satisfy that the assumption of Theorem 4.11, and thus using it, finally we have

$$\pi^I(\Phi(t,x)) = F_I(t,x).$$

By induction, the proof is complete. ∎

4.6.4 A two-dimensional Brownian motion coupled with its Lévy area

We use our PDE theorem to calculate the characteristic function of the Lévy area of a Brownian motion starting at a point $x \in \mathbb{R}^2$. Let us formulate the problem properly. Let X be a three dimensional Itô diffusion process, i.e. $X_t = (B_t^{(1)}, B_t^{(2)}, L_t)^T$, where L_t is the Lévy area of a Brownian motion up to time t, i.e.

$$L_t = \frac{1}{2}\int_0^t B_s^{(2)} \cdot dB_s^{(1)} - B_s^{(1)} \cdot dB_s^{(2)}$$

and $B_t = (B_t^{(1)}, B_t^{(2)})$ is a two dimensional standard Brownian motion. Let $V(x) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -x_2 & x_1 \end{pmatrix}$. Then X satisfies the following PDE:

$$dX_t = V(X_t) \cdot dB_t$$

with the diffusion matrix

$$b(x) = \frac{1}{4}\begin{pmatrix} 1 & 0 & -x_2 \\ 0 & 1 & x_1 \\ -x_2 & x_1 & x_1^2 + x_2^2 \end{pmatrix}$$

where $x = (x_1, x_2, x_3) \in \mathbb{R}^3$.

By Theorem 4.7 for the expected signature of a diffusion process up to a fixed time, in this example we have that for every integer $n \geq 2$,

$$\begin{aligned}(\frac{\partial}{\partial t} + A)\Phi_n(t,x) &= -\frac{1}{4}\sum_{j=1}^{3}(\sum_{j_1=1}^{3} b_{j_1,j}(x)e_{j_1}) \otimes \frac{\partial \Phi_{n-1}(t,x)}{\partial x_j} \\ &\quad -\frac{1}{8}\left(\sum_{j_1=1}^{3}\sum_{j_2=1}^{3} b_{j_1,j_2}(x)e_{j_1} \otimes e_{j_2}\right) \otimes \Phi_{n-2}(t,x)\end{aligned}$$

since the drift of X is equal to zero.

We are interested in $\pi^{I_n}(\Phi(t,x))$, where $I_n = (\underbrace{3,\ldots,3}_{n \text{ copies of } 3})$. By the above theorem, we have the recurrence relationship between $\pi^{I_n}(\Phi(t,x))$, $\pi^{I_{n-1}}(\Phi(t,x))$, and $\pi^{I_{n-2}}(\Phi(t,x))$ as follows:

$$\begin{aligned}\left(-\frac{\partial}{\partial t} + \frac{1}{2}\Delta\right)\pi^{I_n}(\Phi(t,x)) &= -\frac{1}{4}\left(-x_2\frac{\partial \pi^{I_{n-1}}(\Phi(t,x))}{\partial x_1} + x_1\frac{\partial \pi^{I_{n-1}}(\Phi(t,x))}{\partial x_2}\right) \\ &\quad -\frac{1}{8}(x_1^2 + x_2^2)\pi^{I_{n-2}}(\Phi(t,x)) \qquad (4.23)\end{aligned}$$

where $n \geq 2$, $n \in \mathbb{N}$.
Note that $\Phi(t,x)$ actually depends only on x_1 and x_2, which is independent of x_3. Therefore we write the function $\Phi(t,x)$ as $\Phi(t,x_1,x_2)$.

The next step is to derive the characteristic function of the Lévy area on the condition that it is driven by a Brownian motion starting at $(x_1,x_2) \in \mathbb{R}^2$, denoted by $\mathcal{L}(t,x_1,x_2;\lambda)$, i.e.

$$\mathcal{L}(t,x_1,x_2;\lambda) = \mathbb{E}[e^{i\lambda L_t}|B_0 = x_1e_1 + x_2e_2],$$

where $\lambda \in \mathbb{R}$.

Suppose that $\sum_{i=0}^{N} \frac{(\lambda i)^n}{n!}\pi^{I_n}(\Phi(t,x_1,x_2))$ converges as N tends to infinity, we can expand $\mathcal{L}(t,x_1,x_2;\lambda)$ by Taylor's expansion, i.e.

$$\mathcal{L}(t,x_1,x_2;\lambda) = \sum_{n\geq 0} \frac{(\lambda i)^n}{n!}\pi^{I_n}(\Phi(t,x_1,x_2)).$$

Summing the LHS and RHS of (4.23) over $n \geq 2$ respectively, we have the following equations:

$$\begin{aligned}
&\left(-\frac{\partial}{\partial t} + \frac{1}{2}\Delta\right)\mathcal{L}(t,x_1,x_2;\lambda) &(4.24)\\
= \; & -\frac{1}{4}\lambda i\left(-x_2\frac{\partial\mathcal{L}(t,x_1,x_2;\lambda)}{\partial x_1} + x_1\frac{\partial\mathcal{L}(t,x_1,x_2;\lambda)}{\partial x_2}\right) + \lambda^2\frac{1}{8}(x_1^2+x_2^2)\mathcal{L}(t,x_1,x_2;\lambda)
\end{aligned}$$

and $\mathcal{L}(0,x_1,x_2;\lambda) = 1$. Moreover, since $\pi^{I_0}(\Phi(t,x_1,x_2)) = 1$, and $\pi^{I_1}(\Phi(t,x_1,x_2)) = 0$.

By the symmetry of Brownian motion and its Lévy area, the odd moments should be equal to zero, i.e. the characteristic function of the Lévy area $\mathcal{L}(t,x_1,x_2;\lambda)$ is indeed a real valued function. By setting the imaginary part of the function to zero, PDE (4.24) implies that $\mathcal{L}(t,x_1,x_2;\lambda)$ should satisfy the following PDEs:

$$-x_2\frac{\partial\mathcal{L}(t,x_1,x_2;\lambda)}{\partial x_1} + x_1\frac{\partial\mathcal{L}(t,x_1,x_2;\lambda)}{\partial x_2} = 0 \tag{4.25}$$

$$\left(-\frac{\partial}{\partial t} + \frac{1}{2}\Delta\right)\mathcal{L}(t,x_1,x_2;\lambda) = \lambda^2\frac{1}{8}(x_1^2+x_2^2)\mathcal{L}(t,x_1,x_2;\lambda) \tag{4.26}$$

subject to the following initial condition:

$$\mathcal{L}(0,x_1,x_2;\lambda) = 1. \tag{4.27}$$

In Lemma 4.6.3 we show that the solution to (4.25) and (4.26) satisfying the boundary condition (4.27) exists and it is unique. In Lemma 4.6.4, we show that without the hypothesis of the convergence of $\frac{(\lambda i)^n}{n!}\pi^{I_n}(\Phi(t,x_1,x_2))$, it still holds that $\mathcal{L}(t,x_1,x_2)$ satisfies (4.25) and (4.26) subject to boundary condition (4.27) using the fact that the solution can be represented by power series in λ for all $\lambda \in \mathbb{R}$.

Lemma 4.6.3 *Let $f(t,x_1,x_2;\lambda)$ be a solution to the following PDEs:*

$$-x_2\frac{\partial f(t,x_1,x_2;\lambda)}{\partial x_1} + x_1\frac{\partial f(t,x_1,x_2;\lambda)}{\partial x_2} = 0 \tag{4.28}$$

$$\left(-\frac{\partial}{\partial t} + \frac{1}{2}\Delta\right) f(t,x_1,x_2;\lambda) = \lambda^2\frac{1}{8}(x_1^2+x_2^2)f(t,x_1,x_2;\lambda) \tag{4.29}$$

subject to the initial condition:

$$f(0,x_1,x_2;\lambda) = 1. \tag{4.30}$$

Then this solution exists and is unique. Moreover,

$$f(t,x_1,x_2;\lambda) = \frac{1}{\cosh(\frac{1}{2}\lambda t)}\exp\left(-\frac{1}{4}\lambda(x_1^2+x_2^2)\tanh(\frac{1}{2}\lambda t)\right).$$

Proof. By the Maximum principle, PDE (4.29), with initial condition (4.30) admits an unique solution. Note that (4.29) and initial condition (4.30) only depends on t and $|x|$. It is reasonable to suggest that $f(t,x;\lambda)$ is a function of t and $|x|$, i.e.

$$f(t,x_1,x_2;\lambda) = G(t,x_1^2+x_2^2)$$

It is easy to see that $G(t,x_1^2+x_2^2)$ automatically satisfies (4.29). Then our problem is reduced to solving PDE (4.29) with initial condition (4.30). By switching from Cartesian coordinates to polar coordinates and separating the variables, the problem of solving the PDE of f can be further reduced to solving two ODEs. The details are given in A.2. ■

Lemma 4.6.4

$$\mathcal{L}(t,x_1,x_2;\lambda) = \frac{1}{\cosh(\frac{1}{2}\lambda t)}\exp\left(-\frac{1}{4}\lambda(x_1^2+x_2^2)\tanh(\frac{1}{2}\lambda t)\right)$$

Proof. By Lemma 4.6.3,

$$f(t,x_1,x_2;\lambda) := \frac{1}{\cosh(\frac{1}{2}\lambda t)}\exp\left(-\frac{1}{4}\lambda(x_1^2+x_2^2)\tanh(\frac{1}{2}\lambda t)\right)$$

is a unique solution to PDE (4.24) with initial condition (4.27).
On the other hand, it is observed that $f(t,x_1,x_2;\lambda)$ is an infinitely differentiable real function of λ for every fixed x and t. Moreover it can be represented in terms of power series in λ, i.e.

$$f(t,x;\lambda) = \sum_{n\geq 0}(-i)^n f_n(t,x)\lambda^n \tag{4.31}$$

where for every $n \geq 0$, f_n is a function mapping $\mathbb{R}^+ \times \mathbb{R}^2$ to $\mathbb{R}$, such that

$$\begin{aligned} f_{2n}(t,x) &= \frac{1}{n!}(-1)^n \frac{\partial^{2n} f(t,x_1,x_2;0)}{\partial\lambda^{2n}} \\ f_{2n+1}(t,x) &= 0. \end{aligned}$$

Since the radius of convergence for the power series $\sum_{n\geq 0}(-i)^n f_n(t,x)\lambda^n$ is infinity, (4.31) is true for all $\lambda \in \mathbb{R}$. As f satisfies (4.24), after rewriting (4.24) as the power series in λ and comparing the coefficients of both sides of(4.24), we have $f_n(t,x)$ automatically satisfying the recursive PDEs of $\pi^{I_n}(\Phi(t,x))$, i.e. PDE (4.23). Since the solution to (4.23) exists and is unique, then we can conclude that

$$\pi^{I_n}(\Phi(t,x)) = f_n(t,x)$$

and also since $\sum_{n\geq 0}^{N}(-i)^n f_n(t,x)\lambda^n$ converges to $f(t,x;\lambda)$, the partial sum

$$\sum_{n\geq 0}^{N}(-i)^n \pi^{I_n}(\Phi(t,x))$$

converges to $\mathcal{L}(t,x;\lambda)$ as N tends to the infinity. Obviously

$$\mathcal{L}(t,x;\lambda) = f(t,x;\lambda)$$

■

The following lemma describes the relationship between the characteristic function of a Brownian motion starting at the origin coupled with its Lévy area and the characteristic function of Lévy area of a Brownian motion with starting point at x.

Lemma 4.6.5

$$\mathcal{L}(t,x_1,x_2;\lambda) = \mathbb{E}[e^{i(-\frac{1}{2}x_2\lambda B_t^{(1)}+\frac{1}{2}x_1\lambda B_t^{(2)}+\lambda L_t)}|B_0=0].$$

Proof. It comes from the translation property of Brownian motion, i.e. the Brownian motion starting at the origin added by a constant x has the same distribution of Brownian motion starting at x. ■

In [7], Theorem 3 gave the characteristic function of $(B_t^{(1)}, B_t^{(2)}, L_t)$, where $B_t = (B_t^{(1)}, B_t^{(2)})$ is the standard Brown motion starting at the origin, i.e.

$$\mathbb{E}^0[e^{i\gamma_1 B_t^1 + i\gamma_2 B_t^2 + i\lambda L_t}] = \frac{1}{\cosh(t\frac{\lambda}{2})} \exp\left(-(\sum_{i=1}^{2} \gamma_i^2)\frac{1}{\lambda}\tanh(\frac{\lambda}{2}t)\right)$$

which was first observed by Paul Lévy.

Plugging $\gamma_1 = -\frac{1}{2}x_2\lambda$ and $\gamma_2 = \frac{1}{2}x_1\lambda$ into the above equation, it is easy to see that our result for $\mathcal{L}(t, x_1, x_2, \lambda)$ coincides with Lévy and Helmes's result. But our approach is different from theirs. Instead of using Girsanov theorem to change the measure, the PDE of the characteristic function of Lévy area is directly derived from the PDE of the expected signature of Lévy area coupled with the Brownian motion with respect to time t and the starting point x, and this PDE admits a solution, which has this explicit formula.

Chapter 5

The Expected Signature of Planar Brownian Motion up to the First Exit Time of an Exponentially Twice Differentiable Domain

5.1 The setting of the problem

[1] In this chapter, we endow $\mathbb{R}^2$ with the canonical basis (e_1, e_2), and let $(B_t)_{t\geq 0}$ denote a standard two-dimensional Brownian motion with filtration $(\mathcal{F}_t)_{t\geq 0}$, where the filtration $\mathcal{F}_t$ is generated by Brownian motion $(B_s)_{s\in[0,t]}$ itself, i.e.

$$\mathcal{F}_t \triangleq \sigma(B_s; 0 \leq s \leq t)$$

the smallest $\sigma-$field with respect to which B_s is measurable for every $s \in [0, t]$. The Brownian motion B_t could be also written as

$$B_t = B_t^{(1)} e_1 + B_t^{(2)} e_2.$$

Definition 5.1.1 *Let Γ be a domain in $\mathbb{R}^2$. The first exit time of Brownian motion from the domain Γ is denoted by τ_Γ and defined as follows:*

$$\tau_\Gamma = \begin{cases} \inf\{t \geq 0 : B_t \in \Gamma^c\}, & \text{if } \exists t \geq 0, \text{ s.t. } B_t \in \Gamma^c; \\ \infty, & \text{otherwise.} \end{cases}$$

Definition 5.1.2 *The signature of the path $(B_t)_{t\in[0,\tau_\Gamma]}$ is defined to be*

$$S\left(B_{[0,\tau_\Gamma]}\right) = (1, B^1_{[0,\tau_\Gamma]}, ..., B^n_{[0,\tau_\Gamma]}, ...) \in T\left((E)\right)$$

[1]This chapter is mainly from my joint paper with Terry Lyons. ([15])

where $B^n_{[0,\tau_\Gamma]}$ *for each* $n \geq 1$ *is defined by the Stratonovich iterated integral*

$$B^n_{[0,\tau_\Gamma]} = \underset{u_1<...<u_n,\quad u_1,...,u_n\in[0,\tau_\Gamma]}{\int\cdots\int} dB_{u_1} \otimes ... \otimes dB_{u_n}.$$

Definition 5.1.3 (Expected signature of Brownian motion up to the exit time) *The expected signature of a Brownian motion starting at* z *and stopping upon the first exit time from a domain* Γ*, and denoted by* $\Phi_\Gamma(z)$*, is defined as*

$$\Phi_\Gamma(z) = \mathbb{E}^z[S(B_{[0,\tau_\Gamma]})] \tag{5.1}$$

where $\tau_\Gamma = \inf\{t \geq 0 : B_t \in \Gamma^c\}$*, i.e. the first exit time from the domain* Γ*.*

Definition 5.1.4 (Exponential smoothness of the domain) *We say that the domain* Γ *is exponentially twice differentiable if and only if the function* Φ_Γ *defined on* Γ*, mapping* $z \in \Gamma$ *to its expected signature* $\Phi_\Gamma(z) \in T((\mathbb{R}^2))$*, is a well defined and twice differentiable function.*

Remark 5.1.1 *The exponential smoothness of the domain is actually an implicit condition for the domain* Γ*, since it puts the constraints on the function* Φ_Γ *directly instead of* Γ*. Roughly speaking that the domain is exponentially smooth means that the expected signature of Brownian motion up to the first exit time is smooth enough, where "exponential " infers the signature, as the signature of the path can be regarded as the exponential of the path.*

Remark 5.1.2 *Similarly we also can extend the definition of the expected signature of a Brownian motion to that of the expected signature of the function of a Brownian motion starting at* z *upon the first exit time of domain* Γ *denoted by* $\Phi^f_\Gamma(z)$*, i.e.*

$$\Phi^f_\Gamma(z) = \mathbb{E}^z[S(f(B_t)_{t\in[0,\tau_\Gamma]})]$$

where f *is a smooth function mapping* Γ *to* $\mathbb{R}^2$*.*

In order to discuss the expected signature of Brownian motion further, we introduce an auxiliary function Ψ mapping $\mathbb{R}^2$ to $T((\mathbb{R}^2))$.

Definition 5.1.5 *Let us denote the disk centered at* 0 *with radius* r *by* $r\mathbb{D}$*. Then we denote by* $\Psi(z)$ *the expected signature of a Brownian motion started at* 0 *and run until it leaves the disk* $|z|\mathbb{D}$*, conditional on that leaving point being* z*. As a formula:*

$$\Psi(z) = \mathbb{E}^0[S(B_{[0,\tau_{|z|\mathbb{D}}]})|B_{\tau_{|z|\mathbb{D}}} = z].$$

We will see that $\Psi(z)$ plays an important role in the derivation of our PDE and in the discussions of the exponentially twice differentiability of the domain. In part, its importance comes from the strong Markov property.

In the following discussion, we will first provide a sufficient condition for the exponentially twice differentiable domain, i.e. the boundedness of domain can guarantee the exponentially twice differentiable domain.

Then assuming that Γ is an exponentially twice differentiable domain, we will derive a PDE which characterizes Φ_Γ. In the case of the disk, clearly a bounded domain, the PDE is so explicit that one can use it to identify the solution as a combination of polynomials in an explicit manner.

5.2 The exponential twice differentiability of the domain

5.2.1 The homogeneous Carnot-Caratheodory norm for the truncated signature of a path

In this subsection, in order to discuss the regularity of the domain such that the expected signature Φ_Γ mapping from Γ to $T((E))$ is a C^2 function, let us consider the space of truncated signature of order N of all continuous paths of finite $1-$variation. This forms a group and we introduce Carnot-Caratheodory norm (CC norm) to this space, which has a nice geometric property: rotational invariance. We refer to [5] for a more detailed discussion.

Let us recall the definition of paths of finite $p-$variation in Chapter 2 and use $C^{1-var}([0,T],E)$ to denote the space of continuous paths from $[0,T]$ to E of finite $1-$variation.

Definition 5.2.1 *The set of all truncated signatures of continuous paths of finite length of order N is denoted by*

$$G^N(E) := \{S^N(x)_{0,1} : x \in C^{1-var}([0,1], E)\}.$$

Theorem 5.1 *For every $g \in G^N(E)$, the so-called "Carnot-Caratheodory norm"*

$$\|g\| = \inf\left\{||\gamma||_{1-var} : S^N(\gamma) = g, \gamma \in C^{1-var}([0,1], E)\right\}$$

is finite and achieved at some minimizing path γ^*, i.e.

$$\|g\| = ||\gamma^*||_{1-var} \text{ and } S^N(\gamma^*)_{0,1} = g.$$

Definition 5.2.2 *The p-variation of a geometric rough path γ defined on $[0,T]$, and written $||S^N(\gamma)||_{p-var;[0,T]}$, is defined on paths in $G^N(E)$ to be*

$$||S^N(\gamma)||_{p-var;[0,T]} = \sup_{\mathcal{D}=(0=s_1<s_2<...<s_N=T)} \left(\sum_{\mathcal{D}} \left\|S^N(\gamma)_{s_i,s_{i+1}}\right\|^p\right)^{1/p}$$

where $\|.\|$ is the Carnot-Caratheodory norm.

Definition 5.2.3 *The norm on $G^N(E)$ denoted by $\|.\|_1$ is homogeneous if and only if it is homogeneous with respect to the dilation operator, and more precisely*

$$\|\langle\varepsilon, g\rangle\|_1 = \varepsilon \|g\|_1 .$$

Theorem 5.2 (Equivalence of homogeneous norms) *All homogeneous norms on $G^N(E)$ are equivalent. More precisely, if $\|.\|_1$ and $\|.\|_2$ are two homogeneous norms, then there exists $C \geq 1$ such that for all $g \in G^N(E)$, we have*

$$\frac{1}{C} \|g\|_1 \leq \|g\|_2 \leq C \|g\|_1 .$$

Lemma 5.2.1 *The Carnot-Caratheodory norm and the norm defined as follows*

$$|||g||| := \underset{i=1,...,N}{max} |\rho_i(g)|^{1/i}, \ \forall g \in G^N(E)$$

are both homogeneous. By the equivalence of homogeneous norms we have that

$$\begin{aligned}&|\int_{\{0<s_1<...<s_N<T\}} d\gamma_{s_1} \otimes ... \otimes d\gamma_{s_N}|^{1/N} \leq |||S^N(\gamma)||| \\ &\leq C \left\|S^N(\gamma)\right\| \leq C \left\|S^N(\gamma)\right\|_{p-var;[0,T]} .\end{aligned}$$

5.2.2 The exponential twice differentiability of the domain

In this subsection we are going to discuss the exponential twice differentiability of the domain. We will show that each expected signature of Brownian motion upon its first exit time from a bounded domain is finite, firstly.

Theorem 5.3 *Let M be a continuous, $\mathbb{R}^d$-valued local martingale starting at 0 and $S^N(M)$ be the collection of its iterated Stratonovich integrals up to level N, i.e.*

$$1, M_t, ..., \int_{0<s_1<s_2<...<s_N<t} \circ dM_{s_1} \circ ... \circ dM_{s_N}$$

viewed as path in $T^{(n)}(\mathbb{R}^d)$. Then there exists $C = C(N, F, d, |.|)$ such that for all stopping times τ

$$\begin{aligned}
&\mathbb{E}\left(F\left(|\int_{0<s_1<s_2<...<s_N<\tau} \circ dM_{s_1} \circ ... \circ dM_{s_N}|^{1/N}\right)\right) \\
&\leq C\mathbb{E}\left(F\left(||S^N(M)||_{p-var;[0,\tau]}\right)\right) \\
&\leq C\mathbb{E}\left(F\ \left(|\langle M\rangle_\tau|^{\frac{1}{2}}\right)\right)
\end{aligned}$$

where $F : \mathbb{R}+ \to \mathbb{R}+$ is moderate, i.e.

1. *F is continuous and increasing;*

2. *$F(x) = 0$ if and only if $x = 0$;*

3. *for some (and thus for every $\alpha > 1$)*

$$\sup_{x>0} \frac{F(\alpha x)}{F(x)} < +\infty$$

The proof can be found in [4].

Corollary 5.2.1 *Let Γ be a bounded domain Then for every $n \in \mathbb{N}^+$, there exists a constant $C = C(n)$ such that*

$$\begin{aligned}
&\mathbb{E}^z\left[\left|\int ... \int_{0<t_1<t_2<...<t_n<\tau_\Gamma} \circ dB_{t_1} \circ \cdots \circ dB_{t_n}\right|\right] \\
&\leq C(n)\mathbb{E}^z\left[||S^n(B)||^n_{p-var;[0,\tau]}\right] \\
&\leq C(n)\mathbb{E}^z[\tau_\Gamma^{n/2}] < +\infty.
\end{aligned}$$

Proof. Applying Theorem 5.3 to the case with $M_t = B_t$ and $F(x) = x^n$ and we get

$$\mathbb{E}^z\left[\left|\left|\int_{0<t_1<t_2<\ldots<t_n<\tau_\Gamma}\ldots\int \circ\, dB_{t_1}\circ\cdots\circ dB_{t_n}\right|\right|\right] \leq C(n)\mathbb{E}^z\left[||S^n(B)||^n_{p-var;[0,\tau_\Gamma]}\right] \leq C(n)\mathbb{E}^z[\tau_\Gamma^{n/2}]$$

The only thing we need to show is that $\mathbb{E}^z[\tau_\Gamma^{n/2}] < +\infty$. It is sufficient to show that there exists a positive number $\alpha > 0$ such that

$$\mathbb{E}^0[e^{\alpha\tau_{\mathbb{D}}}] < +\infty.$$

Let $\tau_{[-1,1]}$ denote the first exit time of the one-dimensional Brownian motion from the interval $[-1, 1]$, which is started at the origin. It is easy to show the exponential integrability of $\tau_{[-1,1]}$. Moreover since

$$\mathbb{E}^0[e^{\alpha\tau_{\mathbb{D}}}] \leq \mathbb{E}^0[e^{\alpha\tau_{[-1,1]}}]$$

Therefore we can conclude that there exists $\alpha > 0$, such that

$$\mathbb{E}^0[e^{\alpha\tau_{\mathbb{D}}}] < +\infty$$

Due to the scaling property we have

$$\forall r > 0, \mathbb{E}^0[e^{\alpha\tau_{\mathbb{D}}}] = \mathbb{E}^0[e^{\frac{\alpha}{r^2}\tau_{r\mathbb{D}}}]$$

Since Γ is bounded, there exists a positive constant $r \geq 0$ such that $\Gamma \subseteq r\mathbb{D}$. Thus $0 \leq \tau_\Gamma \leq \tau_{r\mathbb{D}}$. It implies that

$$\mathbb{E}^z[e^{\frac{\alpha}{4r^2}\tau_\Gamma}] \leq \mathbb{E}^z[e^{\frac{\alpha}{4r^2}\tau_{z+2r\mathbb{D}}}] = \mathbb{E}^0[e^{\frac{\alpha}{4r^2}\tau_{2r\mathbb{D}}}] = \mathbb{E}^0[e^{\alpha\tau_{\mathbb{D}}}] < +\infty$$

Thus τ_Γ has finite moments, i.e.

$$\forall n \in \mathbb{N}, \mathbb{E}^z[\tau_\Gamma^n] < +\infty$$

and

$$\mathbb{E}^z[\tau_\Gamma^{n/2}] \leq \sqrt{\mathbb{E}^z[\tau_\Gamma^n]} < +\infty$$

Now our proof is complete. ■

Remark 5.2.1 *Note that for the upper half plane $\mathbb{H}$, the expected coordinate signature of Brownian motion is not finite, e.g. the expected signature indexed by $(1, 1)$, since the expectation of the exit time is infinite. It is a simple example to show that the expected signature of Brownian motion upon the first exit time from some domain might explode in general. Thus $\mathbb{H}$ is not an exponentially twice differentiable domain.*

Corollary 5.2.2 *Let Γ be a bounded domain. Then $\Phi_\Gamma \in L_1$.*

Proof. Obviously Φ_Γ is a measurable function. For each index I, there exists a constant C such that $|\pi^I(\Phi_\Gamma)| < C$ is bounded and $\pi^I(\Phi_\Gamma)$ has compact support. Then

$$\int_\Gamma |\Phi_\Gamma(z)|dz \leq CA(\Gamma) < +\infty$$

where $A(\Gamma)$ is the area of Γ. So $\Phi_\Gamma \in L_1$. ■

In the rest of this subsection, we are going to prove that the expected signature of Brownian motion up to the first exit time from any bounded domain is a twice differentiable function assuming that $\Psi(.)$ is a smooth function, which is proved in the latter section(Lemma 5.3.6). Thus any bounded domain is exponentially twice differentiable.

Theorem 5.4 *Every non-empty bounded domain Γ is exponentially twice differentiable.*

Proof. Let us denote the expected signature of Brownian motion starting at any $z \in \Gamma$ upon the first exit time from a bounded domain Γ by $\Phi_\Gamma(z)$. Since we have proved that $\Phi_\Gamma(z)$ is finite, the only thing left to prove is that $\Phi_\Gamma(z)$ is twice differentiable. Let us recall the definition of the function Ψ (Definition 5.1.5). Then for every $\varepsilon > 0$, there exists $\Gamma_\varepsilon = \{z \in \Gamma | dist(z, \partial\Gamma) > \varepsilon\}$. For every $z \in \Gamma_\varepsilon$, by the Markov property of the expected signature of Brownian motion, which comes from Lemma 5.4.1 in the later section, we have that for all $r < \frac{\varepsilon}{2}$,

$$\Phi_\Gamma(z) = \int_0^{2\pi} \Psi(re^{i\theta}) \otimes \Phi(z + re^{i\theta})\frac{d\theta}{2\pi}.$$

Then

$$\Phi_\Gamma(z) = \int_0^{+\infty} \int_0^{2\pi} \Psi(re^{i\theta}) \otimes \Phi_\Gamma(z + re^{i\theta})\frac{d\theta}{2\pi}K(r)dr \tag{5.2}$$

where $K(r)$ is a smooth probability density function with a compact support of $[0, \frac{\varepsilon}{2}]$. Let F be a map from $\mathbb{R}^2 \to T((\mathbb{R}^2))$ defined by:

$$F_\varepsilon(z) = \Psi(-z)K(|z|)$$

Rewriting (5.2) we have

$$\Phi_\Gamma(z) = \int_\Gamma F_\varepsilon(z - y) \otimes \Phi_\Gamma(y)dy = F_\varepsilon * \Phi_\Gamma(z)$$

where $*$ is the convolution. Since Ψ is smooth (see Section 5.3.5) and K is a smooth function with compact support, then F_ε is still a smooth function with compact support. It is easy to show that

$$||F_\varepsilon||_{L_1} + || \bigtriangledown F_\varepsilon||_{L_1} + || \bigtriangleup F_\varepsilon||_{L_1} < +\infty$$

On the other hand, $\Phi_\Gamma \in L^1$ as well. By Theorem 1.3 on Page 3 of [19], we have

$$||F_\varepsilon * \Phi_\Gamma||_{L_1} + || \bigtriangledown F_\varepsilon * \Phi_\Gamma||_{L_1} + || \bigtriangleup F_\varepsilon * \Phi_\Gamma||_{L_1} < +\infty$$

Thus $F_\varepsilon * \Phi_\Gamma$ is twice differentiable, since $F_\varepsilon * \Phi_\Gamma, (\bigtriangledown F_\varepsilon) * \Phi_\Gamma$ and $(\bigtriangleup F_\varepsilon) * \Phi_\Gamma \in L^1$ Moreover, because $\Gamma = \bigcup_{\varepsilon>0} \Gamma_\varepsilon$, Φ_Γ is smooth in Γ. ■

Remark 5.2.2 *The boundedness of the domain is a sufficient condition for that the domain is exponentially twice differentiable; however it is not necessary. Here is a simply example to illustrate it. Let Γ_* be the disjoint union of A_n, i.e.*

$$\Gamma_* = \bigcup_{n\in\mathbb{N}} A_n$$

where A_n is the unit disk centered at $2n$. Obviously Γ_* is not bounded, but Φ_{Γ_*} is well defined and twice differentiable, since for every $z \in \Gamma_*$, there exists the unique $n \in \mathbb{N}$ such that $z \in A_n$, and

$$\Phi_{\Gamma_*}(z) = \Phi_{\mathbb{D}}(z - 2n).$$

Remark 5.2.3 *Actually following this proof we can get Φ_Γ is infinitely differentiable in a non-empty bounded domain. Alternatively, since we have proved that Φ_Γ is twice differentiable so that our PDE provides the integral representation for Φ_Γ, this also implies that Φ_Γ is infinitely differentiable.*

5.3 Main properties of the expected signature of Brownian motion

The expected signature of the Brownian motion has various nice properties, which are presented in this section.

5.3.1 Translation invariance

Lemma 5.3.1 (Translation invariance)

$$\begin{aligned} \mathbb{E}^z[S(B_{[0,\tau_{z+\Gamma}]})] &= \mathbb{E}^0[S\left(B_{[0,\tau_\Gamma]}\right)] \\ \mathbb{E}^z[S(B_{[0,\tau_{z+\Gamma}]})|B_{\tau_{z+\Gamma}} = z + a] &= \mathbb{E}^0[S(B_{[0,\tau_\Gamma]})|B_{\tau_\Gamma} = a] \end{aligned}$$

where $a \in \partial\Gamma$.

5.3.2 Scaling property

Lemma 5.3.2 (Scaling property)
For every $\varepsilon \geq 0$,

$$\begin{aligned}\Phi_{\varepsilon\mathbb{D}}(0) &= \langle \varepsilon, \Phi_{\mathbb{D}}(0)\rangle \\ \rho_n(\Phi_{\varepsilon\mathbb{D}}(0)) &= \varepsilon^n \rho_n(\Phi_{\mathbb{D}}(0)) \\ \Psi(\varepsilon) &= \langle \varepsilon, \Psi(1)\rangle \\ \rho_n(\Psi(\varepsilon)) &= \varepsilon^n \rho_n(\Psi(1))\end{aligned}$$

where $n \in \mathbb{N}$ and $\varepsilon \in (0, +\infty)$ and $\langle , \rangle$ is the dilation operator defined in Chapter 2.

5.3.3 Rotation property

Lemma 5.3.3 (Rotation property) *For every $\theta \in [0, 2\pi]$ and any index $I = i_1 i_2 ... i_n$ with each $i_k \in \{1, 2\}$:*

$$\pi^I(\Psi(e^{i\theta})) = \sum_{J=j_1...j_n} (\prod_{k=1}^{n} \Theta(\theta)_{i_k, j_k}) \pi^J(\Psi(1))$$

where $\Theta(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ and $\Theta(\theta)_{i,j}$ is (i, j) entry of the matrix $\Theta(\theta)$.

Proof. Let B_t be Brownian motion starting at the origin and stopped at the first exit time from the unit disk, on condition that the exit point is 1. Then $\tilde{B}_t = \Theta(\theta) B_t$ is Brownian motion starting at the origin and stopped at the first exit time from the unit disk on condition that the exit point is $e^{i\theta}$, by the conformal invariance of Brownian motion. It is obvious that $\tilde{B}_t$ is just a linear transformation of B_t, i.e.

$$\begin{aligned}\tilde{B}_t^{(1)} &= \cos(\theta) B_t^{(1)} - \sin(\theta) B_t^{(2)} \\ \tilde{B}_t^{(2)} &= \sin(\theta) B_t^{(1)} + \cos(\theta) B_t^{(2)}\end{aligned} \tag{5.3}$$

Then

$$\begin{aligned}\Psi(e^{i\theta}) \quad &= \pi^I(\mathbb{E}^0[S(\tilde{B}_{[0,\tau_{\mathbb{D}}]}) | \tilde{B}_{\tau_{\mathbb{D}}} = e^{i\theta}]) \\ &= \mathbb{E}^0[\int \cdots \int_{0<t_1<...<t_n<\tau_{\mathbb{D}}} \circ d\tilde{B}_{t_1}^{(i_1)} \circ ... \circ d\tilde{B}_{t_n}^{(i_n)} | \tilde{B}_{\tau_{\mathbb{D}}} = e^{i\theta}] \\ &= \mathbb{E}^0[\int \cdots \int_{0<t_1<...<t_n<\tau_{\mathbb{D}}} \circ d(\Theta(\theta)B)_{t_1}^{(i_1)} \circ ... \circ d(\Theta(\theta)B)_{t_n}^{(i_n)} | B_{\tau_{\mathbb{D}}} = 1].\end{aligned} \tag{5.4}$$

Plugging (5.3) into (5.4)

$$\pi^I(\Psi(e^{i\theta})) = \sum_{J=j_1\dots j_n} (\prod_{k=1}^{n} \Theta(\theta)_{i_k,j_k})\pi^J(\Psi(1)).$$

■

5.3.4 An application of the strong Markov property

Lemma 5.3.4 (Strong Markov property) *For every $\varepsilon > 0$ such that $z+\varepsilon\mathbb{D} \subseteq \Gamma$,*

$$\Phi_\Gamma(z) = \int_0^{2\pi} \mathbb{E}^z[S(B_{[0,\tau_{z+\varepsilon\mathbb{D}}]})|B_{\tau_{z+\varepsilon\mathbb{D}}} = z+\varepsilon e^{i\theta}] \otimes \Phi_\Gamma(z+\varepsilon e^{i\theta})\frac{d\theta}{2\pi}. \tag{5.5}$$

Proof. By Chen's identity, which states that the signature of the concatenation of two paths is the product of the signatures of those paths, and $S(B_{[0,\tau_{z+\varepsilon\mathbb{D}}]})$ is $\mathcal{F}_{\tau_{z+\varepsilon\mathbb{D}}}$-measurable, thus

$$\mathbb{E}^z[S(B_{[0,\tau_\Gamma]})|\mathcal{F}_{\tau_{z+\varepsilon\mathbb{D}}}] = S(B_{[0,\tau_{z+\varepsilon\mathbb{D}}]}) \otimes \mathbb{E}[S(B_{[\tau_{z+\varepsilon\mathbb{D}},\tau_\Gamma]})|\mathcal{F}_{\tau_{z+\varepsilon\mathbb{D}}}]$$

where $z+\varepsilon\mathbb{D} \subseteq \Gamma$.
As it is known that B_t has the strong Markov property,

$$\mathbb{E}[S(B_{[\tau_{z+\varepsilon\mathbb{D}},\tau_\Gamma]})|\mathcal{F}_{\tau_{z+\varepsilon\mathbb{D}}}] = E[S(B_{[\tau_{z+\varepsilon\mathbb{D}},\tau_\Gamma]})|B_{\tau_{z+\varepsilon\mathbb{D}}}] = \Phi_\Gamma(B_{\tau_{z+\varepsilon\mathbb{D}}})$$

which implies that

$$\mathbb{E}^z[S(B_{[0,\tau_\Gamma]})|\mathcal{F}_{\tau_{z+\varepsilon\mathbb{D}}}] = S(B_{[0,\tau_{z+\varepsilon\mathbb{D}}]}) \otimes \Phi_\Gamma(B_{\tau_{z+\varepsilon\mathbb{D}}}). \tag{5.6}$$

By the tower property, we have

$$\Phi_\Gamma(z) = \mathbb{E}^z[\mathbb{E}^z[S(B_{[0,\tau_\Gamma]})|\mathcal{F}_{\tau_{z+\varepsilon\mathbb{D}}}]]. \tag{5.7}$$

Substituting Equation (5.6) into it, we obtain

$$\begin{aligned}\Phi_\Gamma(z) &= \mathbb{E}^z[S(B_{[0,\tau_{z+\varepsilon\mathbb{D}}]}) \otimes \Phi_\Gamma(B_{\tau_{z+\varepsilon\mathbb{D}}})] \\ &= \int_0^{2\pi} \mathbb{E}^z[S(B_{[0,\tau_{z+\varepsilon\mathbb{D}}]})|B_{\tau_{z+\varepsilon\mathbb{D}}} = z+\varepsilon e^{i\theta}] \otimes \Phi_\Gamma(z+\varepsilon e^{i\theta})\frac{d\theta}{2\pi}.\end{aligned}$$

■

5.3.5 The smoothness of Ψ

In this subsection, we focus on the discussion of the smoothness of Ψ. We start with the proof of finiteness of $\Psi(1)$, and then show that $\Psi(.)$ is smooth. We end with the derivation of the first two terms of $\Psi(.)$, which is important for us to derive the PDE for $\Phi(.)$ later.

Lemma 5.3.5 $\Psi(1)$ *is well defined.*

Proof. Let $\tilde{B}_t$ be Brownian motion starting at zero in the unit disk. Let $M_t = \tilde{B}_{\tau_{\mathbb{D}}\wedge t}$ be the stopped process. It is obvious from the Burkholder inequalities that the exit time $\tau_{\mathbb{D}}$ of $\tilde{B}_t$ from the disk (or any bounded set) has moments of all orders.

By the equivalence of homogeneous norms on $G^N(E)$ (Theorem 5.2), we have

$$\left|\int_{0<s_1<s_2<...<s_N<\tau} \circ dM_{s_1}\circ ...\circ dM_{s_N}\right| \leq C\left\|S^N(M)\right\|^N_{p-var;[0,\tau_{\mathbb{D}}]}.$$

Thus it holds as well

$$\begin{aligned}&\mathbb{E}^0\left[\left|\int_{0<s_1<s_2<...<s_N<\tau} \circ dM_{s_1}\circ ...\circ dM_{s_N}\right| |M_{\tau_{\mathbb{D}}}=1\right]\\ \leq\quad & C\mathbb{E}^0\left[\left\|S^N(M)\right\|^N_{p-var;[0,\tau_{\mathbb{D}}]}|M_{\tau_{\mathbb{D}}}=1\right].\end{aligned}$$

By the definition of the Carnot-Caratheodory norm, it is obvious that

$$||S^N(e^{i\theta}M)||_{p-var;[0,\tau]}$$

is invariant with respect to θ. Therefore

$$\mathbb{E}^0\left[\left\|S^N(M)\right\|^N_{p-var;[0,\tau_{\mathbb{D}}]}|M_{\tau_{\mathbb{D}}}=1\right] = \mathbb{E}^0\left[\left\|S^N(M)\right\|^N_{p-var;[0,\tau_{\mathbb{D}}]}\right]$$

because there is nothing significant in choosing the exit point to be 1 providing the norms used respect rotation. Using the Burkholder type theorem of Victoir/Friz[4] one can control the p-variation of a martingale in terms of its bracket process and so there exists a positive constant $C>0$, such that

$$\mathbb{E}^0\left[||S^N(M)||^N_{p-var;[0,\tau_{\mathbb{D}}]}\right] \leq C\mathbb{E}^0[\tau_{\mathbb{D}}^{\frac{N}{2}}].$$

Finally we will have that there exists a constant $C>0$ such that

$$\begin{aligned}|\pi_N(\Psi(1))| &= |\mathbb{E}^0\left[\int_{0<s_1<s_2<...<s_N<\tau_{\mathbb{D}}} \circ dM_{s_1}\circ ...\circ dM_{s_N}\Big| M_{\tau_{\mathbb{D}}}=1\right]|\\ &\leq \mathbb{E}^0\left[\left|\int_{0<s_1<s_2<...<s_N<\tau_{\mathbb{D}}} \circ dM_{s_1}\circ ...\circ dM_{s_N}\right| M_{\tau_{\mathbb{D}}}=1\right]\\ &\leq C\mathbb{E}^0[\tau_{\mathbb{D}}^{\frac{N}{2}}]<+\infty.\end{aligned}$$

■

Lemma 5.3.6 Ψ *is a smooth function in* $\mathbb{R}^2$.

Proof. By the scaling property and the rotation property of the expected signature conditioned on its leaving position, we have

$$\pi^I(\Psi(re^{i\theta})) = \sum_{J=j_1\cdots j_n} a_{I,J}(re^{i\theta})\pi^J(\Psi(1))$$

where $a_{I,J}(re^{i\theta}) = r^n(\prod_{k=1}^n \Theta(\theta)_{i_k,j_k})$ and $\Theta(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$. Rewriting the coefficient $a_{I,J}(re^{i\theta})$ in Cartesian coordinate, the coefficient is the homogeneous polynomial of degree no more than n in terms of z_1 and z_2. Obviously it is infinitely differentiable. ■

Remark 5.3.1 *If* Γ *is a bounded domain and* I *is an index, then*

$$z \to \pi^I(\Psi(z))\mathbf{1}_\Gamma(z)$$

is integrable. This remark will be useful in the proof of the exponentially twice differentiability property for bounded domains.

Lemma 5.3.7

$$\pi_2(\Psi(z)) = 1 + \sum_{i=1}^2 z_i e_i + \sum_{i=1}^2 \frac{1}{2} z_i^2 e_i \otimes e_i$$

$$\begin{aligned} \frac{\partial \Psi(0)}{\partial z_i} &= e_i, \ \textit{for } i = 1,2 \\ (\Delta\Psi)(0) &= \sum_{i=1}^2 e_i \otimes e_i. \end{aligned}$$

Proof. The term of tensor degree one in Ψ is the expected increment of the conditioned path, so it is not random and $\rho_1(\Psi(z)) = z$. The expectation of Levy area of the Brownian motion conditioned to leave the disk at any particular point equals zero. Thus the term $\rho_2(\Psi(z))$ only contains the symmetric "increment squared" part, thus we have

$$\pi_2(\Psi(z)) = 1 + \sum_{i=1}^2 z_i e_i + \sum_{i=1}^2 \frac{1}{2} z_i^2 e_i \otimes e_i$$

Also since the n^{th} term of Ψ is a homogeneous polynomial of degree n, immediately we will have

$$\begin{aligned}\frac{\partial\Psi(0)}{\partial z_i} &= e_i, \text{ for } i=1,2 \\ (\Delta\Psi)(0) &= \sum_{i=1}^{2} e_i \otimes e_i.\end{aligned}$$

■

5.4 The PDE for the expected signature of two dimensional Brownian motion up to the first exit time from an exponentially twice differentiable domain

In this section, our goal is to derive the PDE for the expected signature of Brownian motion up to the first exit time from an exponentially twice differentiable domain.

Lemma 5.4.1 *For every $z \in \Gamma$,*

$$\Phi_\Gamma(z) = \int_0^{2\pi} \Psi(\varepsilon e^{i\theta}) \otimes \Phi_\Gamma(z+\varepsilon e^{i\theta})\frac{d\theta}{2\pi} \tag{5.8}$$

where ε is sufficiently small such that $z + \varepsilon\mathbb{D} \subseteq \Gamma$.

Proof. From the strong Markov property, we have

$$\begin{aligned}\Phi_\Gamma(z) &= \int_0^{2\pi} \mathbb{E}^z[S(B_{[0,\tau_{z+\varepsilon\mathbb{D}}]})|B_{\tau_{z+\varepsilon\mathbb{D}}} = z+\varepsilon e^{i\theta}] \otimes \Phi_\Gamma(z+\varepsilon e^{i\theta})\frac{d\theta}{2\pi} \\ &= \int_0^{2\pi} \mathbb{E}^0[S(B_{[0,\tau_{\varepsilon\mathbb{D}}]})|B_{\tau_{\varepsilon\mathbb{D}}} = \varepsilon e^{i\theta}] \otimes \Phi_\Gamma(z+\varepsilon e^{i\theta})\frac{d\theta}{2\pi} \\ &= \int_0^{2\pi} \Psi(\varepsilon e^{i\theta}) \otimes \Phi_\Gamma(z+\varepsilon e^{i\theta})\frac{d\theta}{2\pi} \end{aligned}\tag{5.9}$$

where ε is sufficiently small such that $z + \varepsilon\mathbb{D} \subseteq \Gamma$.
The second equality comes from the translation invariance property. ■

Theorem 5.5 *[A PDE of the expected signature of Brownian motion] Let Γ be an exponentially twice differentiable domain, the expected signature of Brownian motion up to the first exit time from the domain Γ should satisfy the following PDE:*

$$\triangle\Phi_\Gamma(z) = -(\sum_{i=1}^{2} e_i\otimes e_i)\otimes\Phi_\Gamma(z) - 2\sum_{i=1}^{2}\left(e_i \otimes \frac{\partial\Phi_\Gamma(z)}{\partial z_i}\right), \text{ if } z \in \Gamma \tag{5.10}$$

with the boundary condition:

$$\Phi_\Gamma(z) = \mathbf{1}, \ \textit{if } z \in \partial\Gamma \tag{5.11}$$

and the initial condition:

$$\begin{aligned} \rho_0(\Phi_\Gamma(z)) &= 1, \ \textit{if } z \in \bar{\Gamma} &(5.12)\\ \rho_1(\Phi_\Gamma(z)) &= 0, \ \textit{if } z \in \bar{\Gamma}. &(5.13)\end{aligned}$$

Proof. For any fixed $z \in \Gamma$, let us consider a function

$$\begin{aligned} \varphi : \epsilon\mathbb{D} &\to T((E))\\ y &\mapsto \Psi(y) \otimes \Phi_\Gamma(z+y)\end{aligned}$$

where $\epsilon = dist(z, \partial\Gamma)$.
By Lemma 5.4.1, we have

$$\Phi_\Gamma(z) = \int_0^{2\pi} \Psi(\varepsilon e^{i\theta}) \otimes \Phi_\Gamma(z + \varepsilon e^{i\theta}) \frac{d\theta}{2\pi}$$

for every $z \in \Gamma$.
It implies that φ satisfies the mean value property at 0, i.e.

$$\varphi(0) = \Psi(0) \otimes \Phi_\Gamma(z) = \Phi_\Gamma(z) = \int_0^{2\pi} \varphi(\varepsilon e^{i\theta}) \frac{d\theta}{2\pi} \tag{5.14}$$

for every $\varepsilon \leq \epsilon$.
We proved the differentiability of Ψ in Lemma 5.3.6 and by the regularity of the domain Γ we have that as Φ_Γ is twice differentiable, so is φ. By the mean value property and differentiability of φ at point 0 we will have that

$$\Delta(\varphi(y))|_{y=0} = 0.$$

By the chain rule, we obtain

$$\begin{aligned} &\Delta(\varphi(y))|_{y=0}\\ =\ & (\Delta\Psi)(0) \otimes \Phi_\Gamma(z) + 2\sum_{i=1}^{2} \frac{\partial \Psi(0)}{\partial z_i} \otimes \frac{\partial \Phi(z)}{\partial z_i} + \Psi(0) \otimes \Delta\Phi_\Gamma(z) &(5.15)\\ =\ & 0.\end{aligned}$$

Recall Lemma 5.3.7 and the following equalities:

$$\begin{aligned} \Psi(0) &= \mathbf{1} &(5.16)\\ \frac{\partial \Psi(0)}{\partial z_i} &= e_i, \text{for } i = 1, 2 &(5.17)\\ (\Delta\Psi)(0) &= \sum_{i=1}^{2} e_i \otimes e_i &(5.18)\end{aligned}$$

Thus after plugging (5.16) into (5.15) and rearranging the equation, finally we get

$$\triangle \Phi_\Gamma(z) = -(\sum_{i=1}^{2} e_i \otimes e_i) \otimes \Phi_\Gamma(z) - 2\sum_{i=1}^{2} \left(e_i \otimes \frac{\partial \Phi_\Gamma(z)}{\partial z_i} \right) \tag{5.19}$$

By the definition of the expected signature we have the boundary condition,

$$\Phi_\Gamma(z) = \mathbf{1}, \text{ if } z \in \partial\Gamma$$

By the definition of the signature, it is obvious that $\rho_0(\Phi(z)) = 1$ and

$$\rho_1(\Phi(z)) = \sum_{i=1}^{2} \mathbb{E}^z(\int_0^{\tau_\Gamma} \circ dB_t^{(i)})e_i = \sum_{i=1}^{2} \mathbb{E}^z(B_{\tau_\Gamma}^{(i)} - z^{(i)})e_i = 0$$

since Brownian motion is a martingale. ■

Remark 5.4.1 *The proof of our PDE theorem for the expected signature is actually independent of the dimension of the Brownian motion. Our theorem is still valid for higher dimensional cases.*

The following corollary is an equivalent statement to Theorem 5.5, which states how to use the PDE system to solve each term of the expected signature recursively.

Corollary 5.4.1 *Let Γ be an exponentially twice differentiable domain. For every $n \in \mathbb{N}$ and $n \geq 2$, the n^{th} term of the expected signature of Brownian motion up to the first exit time of the domain Γ should satisfy the following PDE:*

$$\Delta(\rho_n(\Phi_\Gamma(z))) = -2\sum_{i=1}^{2} e_i \otimes \frac{\partial \rho_{n-1}(\Phi_\Gamma(z))}{\partial z_i} - \left(\sum_{i=1}^{2} e_i \otimes e_i\right) \otimes \rho_{n-2}(\Phi_\Gamma(z)) \tag{5.20}$$

with the boundary condition $\rho_n(\Phi_\Gamma(z)) = 0$, if $z \in \partial\Gamma$ and $n \geq 2$.
Moreover $\rho_0(\Phi_\Gamma(z)) = 1, \rho_1(\Phi_\Gamma(z)) = 0, \forall z \in \bar{\Gamma}$.

Remark 5.4.2 *From the corollary it is important to notice that if we have computed the $(n-1)^{th}$ and $(n-2)^{th}$ term of the expected signature, the right hand side of the PDE for the n^{th} term is known. There is a classical solution to this PDE problem and it implies that from the PDE and the boundary condition we can solve all the terms of the expected signature recursively. We will explain this in the next section.*

Let us recall the following important theorem in PDE theory.

Theorem 5.6 *Let Γ be a bounded domain and $\partial\Gamma$ be the boundary of Γ. Given functions $\varphi : \partial\Gamma \to \mathbb{R}, f : \Gamma \to \mathbb{R}$, which satisfy some regularity conditions, the solution of the Dirichlet problem for the Poisson equation*

$$\begin{cases} \Delta u(x) = f(x), & \text{for } x \in \Gamma; \\ u(x) = \varphi(x), & \text{for } x \in \partial\Gamma. \end{cases}$$

is given by the representation formula

$$u(y) = \int_{\partial\Gamma} \varphi(x) \frac{\partial G(x,y)}{\partial V_x} do(x) + \int_{\Gamma} f(x)G(x,y)dx$$

where $do(x)$ is the volume element of $\partial\Gamma$, V_x is the exterior normal of $\partial\Gamma$ and $G(x,y)$ is the Green function of the domain Γ, and $y \in \Gamma$.[9]
Moreover, the solution to the Dirichlet problem for the Poisson equation has a stochastic representation

$$u(y) = \mathbb{E}^y \left[\varphi(B_{\tau_\Gamma})\right] - \mathbb{E}^y \left[\int_0^{\tau_\Gamma} \frac{1}{2} f(B_s)ds\right]$$

With this theorem in mind, we give the recurrence relation between higher order and lower order terms of the expected signature, which is a direct consequence of the above theorem, by simply plugging the PDE and the boundary condition in our setting into the theorem. It also provides insights into ways to compute the expected signature recursively, given the first two terms of expected signature, which are known.

Theorem 5.7 *Let $\Phi_\Gamma(z)$ be defined as before. Then for each index $I = (i_1, i_2, \ldots, i_n)$ of length $n \geq 2$,*

$$\pi^I(\Phi_\Gamma(z)) = \int_\Gamma \varphi^I(\nu)G(\nu,z)d\nu$$

where $I_1 = (i_2, \ldots, i_n)$, $I_2 = (i_3, \ldots, i_n)$ and

$$\varphi^I(z) = \mathbf{1}_{\{i_1=i_2\}}\pi^{I_2}(\Phi_\Gamma(z)) + 2\left(\sum_{i=1}^{2} \mathbf{1}_{\{i_1=i\}} \frac{\partial \pi^{I_1}(\Phi_\Gamma(z))}{\partial z_i}\right)$$

and $G(\nu, z)$ is the Green function of the domain Γ.
Moreover, the solution has a unique stochastic representation as follows:

$$\pi^I(\Phi_\Gamma(z)) = \mathbb{E}^z \left[\int_0^{\tau_\Gamma} \varphi^I(B_s)ds\right]$$

5.5 Concrete examples

5.5.1 The expected signature of Brownian motion in the unit disk

Before considering the case with the unit disk, we look at one-dimensional Brownian motion starting at x upon the first exit time from the interval $[-1,1]$, where $x \in (-1,1)$.

After an easy computation, the probability of hitting 1 at the exit time is

$$\mathbb{P}^x(B_{\tau_{[-1,1]}} = 1) = \frac{x+1}{2}$$

and

$$\begin{aligned}
\rho_n(\Phi_{[-1,1]}) &= \frac{(1-x)^n}{n!}\mathbb{P}^x(B_{\tau_{[-1,1]}} = 1) + \frac{(-1-x)^n}{n!}\mathbb{P}^x(B_{\tau_{[-1,1]}} = -1) \\
&= \frac{1}{2n!}(1-x^2)\left((1-x)^{n-1} - (-1-x)^{n-1}\right)
\end{aligned}$$

On the other hand, using a similar approach to the one in two dimensional case, we can derive that the expected signature $\Phi_{[-1,1]}(x)$ should satisfy the following ODE:

$$\frac{d^2\Phi_{[-1,1]}(x)}{dx^2} = -\Phi_{[-1,1]}(x) - 2\frac{d\Phi_{[-1,1]}(x)}{dx}$$

with the boundary condition

$$\rho_n(\Phi_{[-1,1]}(x)) = 0, \text{ if } x = \pm 1 \text{ for } \forall n \geq 2$$

and the initial condition

$$\begin{aligned}
\rho_0(\Phi_{[-1,1]}(x)) &= 1 \\
\rho_1(\Phi_{[-1,1]}(x)) &= 0 \text{ for all } x \in [-1,1]
\end{aligned}$$

It is remarkable that this ODE has a similar form to the PDE we derived in two dimensions, and the explicit formula for the expected signature of the 1-dimensional Brownian motion satisfies such ODEs and can be uniquely determined by the ODE system, initial condition and boundary condition.

After discussing the one-dimensional case, we will illustrate how to use our PDE approach to calculate the expected signature of the Brownian motion upon the first exit time from the unit disk. To simplify our notation, let us denote this expected signature of Brownian motion starting at z upon the first exit time by $\Phi(z)$. Based on our previous discussion, $\Phi(z)$ is the solution to the following PDE problem:

$$\begin{cases} \triangle\Phi(z) = -\left(\sum\limits_{i=1}^{2} e_i \otimes e_i\right) \otimes \Phi(z) - 2\sum\limits_{i=1}^{2}\left(e_i \otimes \frac{\partial\Phi(z)}{\partial z_i}\right) & \text{if } z \in \mathbb{D}; \\ \Phi(z) = \mathbf{1}, & \text{if } |z| = 1. \end{cases}$$

Recall Corollary 5.4.1, and we have for every $n \geq 2$,

$$\triangle\rho_n(\Phi(z)) = -\left(\sum_{i=1}^{2} e_i \otimes e_i\right) \otimes \rho_{n-2}(\Phi(z)) - 2\sum_{i=1}^{2}\left(e_i \otimes \frac{\partial\rho_{n-1}(\Phi(z))}{\partial z_i}\right)$$

We can find the whole expected signature recursively with $\rho_0(\Phi(z)) = 1$ and $\rho_1(\Phi(z)) = 0$ for every $z \in \mathbb{D}$.

In the following part, we are going to prove that the expected coordinate signature of Brownian motion upon the first exit time of the unit disk is a polynomial of z_1 and z_2.

To be convenient, we introduce a definition of what it means for a function to be of the polynomial form, when the function takes value in $(\mathbb{R}^2)^{\otimes n}$.

Definition 5.5.1 *Let the function f be a map from Γ to $(\mathbb{R}^2)^{\otimes n}$, where Γ is a domain in $\mathbb{R}^2$ endowed with the canonical basis. We say that f is of polynomial form with a factor g if for every index I with length n, $\pi^I(f(.))$ is a polynomial with a common factor g, where g is a polynomial.*

Lemma 5.5.1 *For any given polynomial of degree n denoted by $f : \mathbb{D} \to \mathbb{R}$, the solution to the following PDE*

$$\begin{cases} \Delta F(z) = f(z), & \text{if } z \in \mathbb{D}; \\ F(z) = 0, & \text{if } |z| = 1. \end{cases}$$

exists and it is unique. Moreover it is a polynomial of degree no more than $n+2$ with a factor $(1 - \sum\limits_{i=1}^{2} z_i^2)$.

We will prove this lemma in two different ways. The first one can be applied to the higher dimensional case, but the second one can give an algorithm to calculate the solution F.

Proof. For any fixed $n \in \mathbb{N}$, let us denote the space of the polynomials with degree no more than n by $P[n]$. Define the linear operator L which maps $P[n]$ to $P[n]$ as follows:

$$L(f)(z) := \Delta((1 - \sum_{i=1}^{2} z_i^2) f(z)),$$

where $f \in P[n]$.
We are going to show that $dim(Im(L)) = dim(P[n])$. Let us consider $Ker(L) = \{g \in P[n] | L(g) = 0\}$. For every $g \in Ker(L)$, let G be $G(z) := (1 - \sum_{i=1}^{2} z_i^2) g(z)$. Then $G \in P[n+2]$ and satisfies the following PDE problem

$$\Delta G(z) = 0 \text{ if } z \in \mathbb{D}$$

with the boundary condition

$$G(z) = 0, \text{ if } |z| = 1.$$

By the strong maximum principle $G(z) = 0$ is the unique solution to this PDE problem, then we have $Ker(L) = \{0\}$ and $dim(Ker(L)) = 0$. Since L is a linear operator, by the rank-nullity theorem, we have

$$dim(Im(L)) = dim(P[n]) - dim(Ker(L)) = dim(P[n])$$

It is easy to show that $Im(L) \subseteq P[n]$. Then we can claim that L is a bijection and $Im(L) = P[n]$, which means that for every $f \in P[n]$, there exists a unique polynomial $g \in P[n]$ such that

$$L(g) = f$$

Equivalently for every $f \in P[n]$, there exits a unique polynomial $F \in P[n+2]$: $F(z) := (1 - \sum_{i=1}^{2} z_i^2) g(z)$ which is the unique solution to the following PDE:

$$\begin{cases} \Delta F(z) = f(z), & \text{if } z \in \mathbb{D}; \\ F(z) = 0, & \text{if } |z| = 1. \end{cases}$$

■

The following is another way to prove Lemma 5.5.1, which gives us an algorithm to solve for the F completely.

Proof. We adopt the notation of $P[n]$ and the linear operator L used in the above proof. To prove the lemma, it is equivalent to prove that for $f \in P[n]$ with $n \geq 0$, there exists a homogeneous polynomial denoted by g_n and a polynomial R_n which is a polynomial of degree strictly less than n if $n \geq 1$ and $R_0 = 0$ such that

$$L(g_n) = f^* + R_n$$

where f^* is the leading term of f.
For $n = 0$, each polynomial $f \in P[n]$ is a constant function, i.e. $f(z) = c \; \forall z = (z_1, z_2) \in \mathbb{D}$. Then $g(z) := -\frac{c}{4}$ satisfies

$$\Delta((1 - \sum_{i=1}^{2} z_i^2) g(z)) = f(z) = c \tag{5.21}$$

It is easy to show that the solution to this PDE is unique. Thus the statement is true for $n = 0$.
For $n \geq 1$, since g_n is a homogeneous polynomial of degree n, we have

$$L(g_n)(z) = -4(n+1) g_n(z) + (1 - \sum_{i=1}^{2} z_i^2) \Delta(g_n(z)).$$

The leading term of $L(g_n)(z)$ is

$$-4(n+1) g_n(z) - \sum_{i=1}^{2} z_i^2 \Delta(g_n(z))$$

which should match f^*.
Let us write g_n in the form:

$$g_n(z) = \sum_{j=0}^{n} a_j z_1^j z_2^{n-j}$$

and f^* in the form:

$$f^*(z) = \sum_{j=0}^{n} b_j z_1^j z_2^{n-j}.$$

Then

$$\begin{aligned} & -4(n+1) g_n(z) - \sum_{i=1}^{2} z_i^2 \Delta(g_n(z)) \\ = & \sum_{j=0}^{n} [-4(n+1) - j(j-1) 1_{\{2 \leq j \leq n\}} - (n-j)(n-j-1) 1_{\{0 \leq j \leq n-2\}}] a_j z_1^j z_2^{n-j} \\ & - \sum_{j=0}^{n-2} (j+2)(j+1) a_{j+2} z_1^j z_2^{n-j} - \sum_{j=2}^{n} (n-j+2)(n-j+1) a_{j-2} z_1^j z_2^{n-j} \end{aligned}$$

Whether there exists g_n such that the leading term of $\Delta((1-\sum_{i=1}^{2} z_i^2)g_n(z))$ is $f^*(z)$ depends on whether the following linear equation has a solution or not:

$$\begin{aligned}&[-4(n+1) - j(j-1)1_{\{2\leq j\leq n\}} - (n-j)(n-j-1)1_{\{0\leq j\leq n-2\}}]a_j \\ &-(j+2)(j+1)a_{j+2}1_{\{0\leq j\leq n-2\}} - a_{j-2}(n-j+2)(n-j+1)1_{\{2\leq j\leq n\}} = b_j\end{aligned}$$

where $j = 0, 1, ..., n$.
Or equivalently in the matrix form:

$$M_n\vec{a} = \vec{b} \tag{5.22}$$

where $\vec{a} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$, $\vec{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix}$ and M_n is a $(n+1)\times(n+1)$ matrix which is defined as below:

$$\begin{cases} M_n(j,j) = [-4(n+1) - (n-j)(n-j-1)], & \text{if } j < 2; \\ M_n(j,j) = [-4(n+1) - j(j-1) - (n-j)(n-j-1)], & \text{if } j \geq 2 \text{ and } j \leq n-2; \\ M_n(j,j) = [-4(n+1) - j(j-1)], & \text{if } j > n-2 \text{ and } j \leq n; \\ M_n(j,j+2) = -(j+2)(j+1), & \text{if } j \leq n-2; \\ M_n(j,j-2) = -(n-j+2)(n-j+1), & \text{if } j \geq 2; \\ M_n(i,j) = 0, & \text{otherwise.} \end{cases}$$

The final step is to prove that M_n is invertible. It is easily verified that M_n is strictly diagonally dominant, since we have

$$|M_n(j,j)| - \sum_{\substack{i=0 \\ i\neq j}}^{n} |M_n(i,j)| = 4(n+1) > 0$$

By the Gershgorin circle theorem, a strictly diagonally dominant matrix is non-singular. So M_n is invertible and there exists the unique solution $\vec{a}$ to Eqn. 5.22:

$$\vec{a} = M_n^{-1}\vec{b}$$

i.e. there exits a unique g_n such that the leading term of $\Delta((1-\sum_{i=1}^{2} z_i^2)g_n(z))$ equals f^*. By induction, we can claim that for every polynomial f of degree n, there exists the unique polynomial g of degree no more than n such that

$$L(g)(z) = f(z)$$

■

Theorem 5.8 *For each $n \in \mathbb{N}^+$, $\rho_n(\Phi(z))$ is of polynomial form of degree no more than n with the factor $(1-\sum_{i=1}^{2} z_i^2)$.*

Proof. We proved it by induction.
For $n = 1$, $\rho_1(\Phi(z)) = 0$ and it has a factor $(1-\sum_{i=1}^{2} z_i^2)$.
For $n = 2$, by Corollary 5.4.1 we have $\rho_2(\Phi(z))$ satisfies

$$\begin{cases} \Delta\rho_2(\Phi(z)) = -\sum_{i=1}^{2} e_i \otimes e_i, & \text{if } z \in \mathbb{D}; \\ \rho_2(\Phi(z)) = 0, & \text{if } |z| = 1. \end{cases}$$

By the second proof of Lemma 5.5.1, we can calculate the unique solution $\rho_2(\Phi(z)) = \frac{1}{4}(1-\sum_{i=1}^{2} z_i^2) e_i \otimes e_i$. So our statement is true for $n = 1, 2$.
Suppose that for every $n < N$, it holds. We are going to prove that it is true for $n = N$.
By Corollary 5.4.1, we obtain that for every $n \geq 2$,

$$\triangle\rho_n(\Phi(z)) = -\left(\sum_{i=1}^{2} e_i \otimes e_i\right) \otimes \rho_{n-2}(\Phi(z)) - 2\sum_{i=1}^{2}\left(e_i \otimes \frac{\partial\rho_{n-1}(\Phi(z))}{\partial z_i}\right)$$

By the induction hypothesis, it is easy to show that the right hand side should be of polynomial form of degree no more than $n-2$. Using Lemma 5.5.1, each $\pi^I(\Phi_.)$ with $|I| = n$ should be a polynomial of degree less than n with a factor $1-\sum_{i=1}^{2} z_i^2$, so $\rho_n(\Phi(.))$ is of polynomial form of degree no greater than n with common factor $1-\sum_{i=1}^{2} z_i^2$. Now our proof is complete. ■

Now we give the truncated expected signature of Brownian motion upon the first exit time of the unit disk up to degree 4.

$$\rho_2(\Phi(z)) = \frac{1}{4}(1-\sum_{i=1}^{2} z_i^2)(\sum_{i=1}^{2} e_i \otimes e_i)$$

$$\rho_3(\Phi(z)) = (1-\sum_{i=1}^{2} z_i^2)(-\sum_{i=1}^{2} \frac{1}{8} z_i e_i) \otimes (\sum_{i=1}^{2} e_i \otimes e_i)$$

$$\begin{aligned} \rho_4(\Phi(z)) = (1-\sum_{i=1}^{2} z_i^2)&\left(D_1(z) e_1 \otimes e_1 + \frac{z_1 z_2}{24}(e_1 \otimes e_2 + e_2 \otimes e_1) + D_2(z) e_2 \otimes e_2\right) \\ &\otimes \left(\sum_{i=1}^{2} e_i \otimes e_i\right) \end{aligned}$$

where

$$\begin{aligned} D_1(z) &= \frac{7}{192}z_1^2 - \frac{1}{192}z_2^2 + \frac{1}{64} \\ D_2(z) &= \frac{7}{192}z_2^2 - \frac{1}{192}z_1^2 + \frac{1}{64} \end{aligned}$$

5.5.2 The Lévy area of a Brownian motion up to the first exit time from the disk

Let us consider the unit disk domain $\mathbb{D} \subset \mathbb{R}^2$, and the Brownian motion started at $z \in \mathbb{D}$. $\tau_{\mathbb{D}}$ is used to denote the first exit time of the Brownian motion from $\mathbb{D}$. The signature of the Lévy area of the Brownian motion up to $\tau_{\mathbb{D}}$, denoted by $S(L_{[0,\tau_{\mathbb{D}}]})$ is similar to that of the Lévy area up to a fixed time by replacing the time interval by $[0, \tau_{\mathbb{D}}]$:

$$S(L_{[0,\tau_{\mathbb{D}}]}) = 1 + \sum_{n \geq 1} \left(\int_{0<t_1<\cdots<t_n<\tau_{\mathbb{D}}} dL_{t_1} \circ dL_{t_2} \cdots \circ dL_{t_n} \right) \underbrace{e_1 \otimes \cdots \otimes e_1}_{n \text{ copies of } e_1}$$

Notation 5.9 *Let us denote the expectation of $S(L_{[0,\tau_{\mathbb{D}}]})$ on condition that the starting point of the Brownian motion is z by $F(z_1, z_2)$, where $z = (z_1, z_2) \in \mathbb{D}$.*

Then F is a function mapping $\mathbb{R}^2$ to $T((\mathbb{R}^2))$.

Recall that $\tau_{\mathbb{D}}$ denotes the first exit time of Brownian motion from the unit disk $\mathbb{D}$. Let us introduce two notations to denote the characteristic function and the moment generating function of the Lévy area of the Brownian motion at $\tau_{\mathbb{D}}$ respectively as follows.

Notation 5.10 *Denote the characteristic function of the Lévy area of the Brownian motion at $\tau_{\mathbb{D}}$ by $\mathcal{L}(z; \alpha)$, i.e.*

$$\mathcal{L}(z; \alpha) = \mathbb{E}\left[e^{i\alpha L_{\tau_{\mathbb{D}}}} | B_0 = z\right]$$

Notation 5.11 *Let us denote the moment generating function of the Lévy area of the Brownian motion at $\tau_{\mathbb{D}}$ by $\mathcal{M}(z; \alpha)$, i.e.*

$$\mathcal{M}(z; \alpha) = \mathbb{E}\left[e^{\alpha L_{\tau_{\mathbb{D}}}} | B_0 = z\right]$$

In the appendix A.1 it has been proved that the n^{th} moment of the Lévy area of the Brownian motion multiplied by $\frac{1}{n!}$ agrees with the Stratonovich iterated integral of Lévy area. We are going to use our PDE theorem for the expected signature of stopped Brownian motion in order to derive the PDE of the characteristic function and the moment generating function of the Lévy area at $\tau_{\mathbb{D}}$, i.e. $\mathcal{L}(z;\alpha)$ and $\mathcal{M}(z;\alpha)$. Moreover it turns out that both $\mathcal{L}(z;\alpha)$ and $\mathcal{M}(z;\alpha)$ admit a probabilistic representation and a deterministic representation under the assumption that $\mathcal{L}(z;\alpha)$ and $\mathcal{M}(z;\alpha)$ have continuous second partial derivatives when α is in the proper range(See Theorem 5.12 and Lemma 5.5.6). To show the above theorem, we need to have the following useful lemmas to describe the recursive relationship between lower order terms and higher order terms of the expected signature of the Lévy area up to $\tau_{\mathbb{D}}$. The idea is similar to the case for a fixed time interval.

Lemma 5.5.2 *Let F be defined as before in Notation 5.9. Then it satisfies the boundary condition:*

$$F(z) = \mathbf{1}, \text{ if } z \in \partial\mathbb{D}$$

and the PDE:

$$\triangle F(z) = -\frac{1}{2}e_1 \otimes \left(-z_2\frac{\partial F(z)}{\partial z_1} + z_1\frac{\partial F(z)}{\partial z_2}\right) - \left(\frac{1}{4}(z_1^2+z_2^2)e_1\otimes e_1\right)\otimes F(z)$$

In other words, for $n \in \mathbb{N}$ and $n \geq 2$,

$$\triangle\rho_n(F(z)) = \frac{1}{2}\left(z_2\frac{\partial\rho_{n-1}(F(z))}{\partial z_1} - z_1\frac{\partial\rho_{n-1}(F(z))}{\partial z_2}\right) - \frac{1}{4}(z_1^2+z_2^2)\rho_{n-2}(F(z)) \quad (5.23)$$

where $z \in \mathbb{D}$. Moreover $\rho_0(F(z)) = 1$ and $\rho_1(F(z)) = 0$ for all $z \in \mathbb{D}$.

Proof. For every $z \in \mathbb{D}$, there exists $\varepsilon > 0$ such that $z + \varepsilon\mathbb{D} \subset \mathbb{D}$, by Chen's identity for the signature and the strong Markov property of the Lévy area of a Brownian motion,

$$\begin{aligned} F(z) &= \mathbb{E}[S(L_{[0,\tau_{\mathbb{D}}]})|B_0 = z] \\ &= \mathbb{E}^z[S(L_{[0,\tau_{z+\varepsilon\mathbb{D}}]})\otimes F(B_{\tau_{z+\varepsilon\mathbb{D}}})] \end{aligned} \quad (5.24)$$

Since the expected signature of the Brownian motion up to $\tau_{\mathbb{D}}$ is in the polynomial form and thus infinitely smooth, the expected signature of the Lévy area up to $\tau_{\mathbb{D}}$ is just a linear functional on the expected signature of the stopped Brownian motion, thus it is infinitely smooth as well. We can apply the Taylor expansion to the function F, and get

$$\begin{aligned} F(B_{\tau_{z+\varepsilon\mathbb{D}}}) &= F(z) + \sum_{i=1}^{2}\frac{\partial F(z)}{\partial z_i}\left(\int_0^{\tau_{z+\varepsilon\mathbb{D}}} dB_{t_1}^{(i)}\right) \\ &+\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\partial^2 F(z)}{\partial z_i\partial z_j}\left(\int_0^{\tau_{z+\varepsilon\mathbb{D}}} dB_{t_1}^{(i)}\int_0^{\tau_{z+\varepsilon\mathbb{D}}} dB_{t_1}^{(j)}\right) + O(\varepsilon^3) \end{aligned}$$

Substituting this in the RHS of Equation (5.24), we write F as a power series in ε as follows:

$$F(z) + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} dB_{t_1}^{(i)} \int_0^{\tau_{z+\varepsilon\mathbb{D}}} dB_{t_1}^{(j)}\right]\frac{\partial^2 F(z)}{\partial z_i \partial z_j}$$
$$+\sum_{i=1}^{2}\mathbb{E}^z\left[L_{\tau_{z+\varepsilon\mathbb{D}}}\int_0^{\tau_{z+\varepsilon\mathbb{D}}} dB_{t_1}^{(i)}\right] e_1 \otimes \frac{\partial F(z)}{\partial z_i} + \left(\mathbb{E}^z\left[\frac{1}{2}L^2_{\tau_{z+\varepsilon\mathbb{D}}}\right] e_1 \otimes e_1\right) \otimes F(z) + o(\varepsilon^2)$$

where $B_0 = x$.

Note that, unlike the expected signature of the Brownian motion up to $\tau_{\mathbb{D}}$ centered at the start point of Brownian motion z, that of the corresponding Lévy area does not have such translation invariance, and it does depend on z. However applying the property that a Brownian motion path started at $z \in \mathbb{R}^2$ has the same distribution as the Brownian motion started at the origin translated by z, it is not difficult to get the following results. We do not give the details here; interested readers could refer to the appendix(A.3.1):

$$\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} \int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(j)}\right] = \mathbf{1}(i=j)\frac{1}{2}\varepsilon^2$$
$$\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} L_{\tau_{x+\varepsilon\mathbb{D}}}\right] = \left(-\frac{1}{4}z_2\mathbf{1}(i=1) + \frac{1}{4}z_1\mathbf{1}(i=2)\right)\varepsilon^2$$
$$\mathbb{E}^z\left[L^2_{\tau_{z+\varepsilon\mathbb{D}}}\right] = \frac{1}{8}(z_1^2+z_2^2)\varepsilon^2 + o(\varepsilon^2)$$

Since the RHS of the Equation (5.24) is independent of ε, the coefficient of the term ε is equal to

$$\frac{1}{4}\bigtriangleup F(z) + \frac{1}{4}e_1 \otimes \left(-z_2\frac{\partial F(z)}{\partial z_1} + z_1\frac{\partial F(z)}{\partial z_2}\right) + \left(\frac{1}{16}(z_1^2+z_2^2)e_1\otimes e_1\right)\otimes F(z)$$

and should be zero. Thus we have the PDE of F immediately, i.e.

$$\bigtriangleup F(z) = \frac{1}{2}e_1 \otimes \left(z_2\frac{\partial F(z)}{\partial z_1} - z_1\frac{\partial F(z)}{\partial z_2}\right) - \left(\frac{1}{4}(z_1^2+z_2^2)e_1\otimes e_1\right)\otimes F(z)$$

■

Using the above lemma we can calculate the moments of the Lévy area at time $\tau_{\mathbb{D}}$, i.e. $\rho_n(F(z))$, which is given in the following lemma.

Lemma 5.5.3 *For every integer $n \geq 1$,*

$$\rho_{2n}(F(z)) = \sum_{i=0}^{n} A_{n,i}r^{4i}$$
$$\rho_{2n+1}(F(z)) = 0$$

where $A_{n,i}$ is defined recursively as follows:

$$\begin{aligned}
A_{0,0} &= 1 \\
A_{n,i} &= -\frac{1}{64i^2}A_{n-1,i-1}, \ \textit{if } i = 1, 2, \ldots, n \\
A_{n,0} &= -\sum_{i=1}^{n} A_{n,i}.
\end{aligned}$$

Proof. By Lemma 5.5.2, it is shown that the moments of Lévy area of the Brownian motion at time $\tau_{\mathbb{D}}$, i.e. $\rho_n(F(z))$ satisfies (5.23). As the expected signature of the Brownian motion up to the first exit time from the unit disk, i.e. $\Phi_{\mathbb{D}}(x)$ is in the polynomial form and the moments of the Lévy area is the linear functional on $\Phi_{\mathbb{D}}(z)$, $\rho_n(F(z))$ is a polynomial in z. Moreover due to the rotation invariance of the Brown motion, $\rho_n(F(z))$ is indeed a function of $|z|$, and thus (5.23) can be further simplified to

$$\Delta(\rho_n(F(z))) = -\frac{1}{4}|z|^2\rho_{n-2}(F(z)), z \in \mathbb{D} \tag{5.25}$$

Since $\rho_1(F(z)) = 0$, it is easy to prove that $\rho_{2n+1}(F(z)) = 0$ by induction according to the fact that $\rho_n(F(z))$ satisfies (5.25) and $f(z) = 0$ is the unique solution to $\Delta f(z) = 0$.

Next let us focus on the even moments of Lévy area. Since $\rho_{2n}(F(z))$ is a function of z, we define the function $O_n : [0,1] \to \mathbb{R}$ as follows:

$$O_n(r) = \rho_{2n}(F(z)) \text{ if } |z| = r.$$

By the chain rule, (5.25) implies that O_n satisfies the following equation:

$$\frac{1}{r}O'_n(r) + O''_n(r) = -\frac{r^2}{4}O_{n-1}(r).$$

Assume that O_n can be written in the following form

$$O_n(r) = \sum_{i=0}^{n} A_{n,i}r^{4i}$$

Thus the PDE of O_n can be rewritten as follows:

$$\sum_{i=1}^{n} 16i^2 A_{n,i}r^{4i-2} = -\frac{1}{4}A_{n-1,i-1}r^{4i-2}, \text{ if } i = 1, 2, \ldots, n.$$

By comparing the coefficients of r^{4i-2} on both sides of the above equation, we obtain that

$$A_{n,i} = -\frac{1}{64i^2}A_{n-1,i-1}.$$

where $i = 1, 2, \ldots, n$. Since $O_n(1) = 0$, we have

$$A_{n,0} = -\sum_{i=1}^{n} A_{n,i}.$$

By the definition of the signature, it is obvious that $A_{0,0} = 1$. ∎

Suppose that the partial sum

$$\sum_{n=0}^{N} (\alpha i)^n \rho_n(F(z))$$

converges as N tends to the infinity. Under this assumption, we can expand the function $\mathcal{L}(z;\alpha)$ by power series in α, i.e.

$$\mathcal{L}(z;\alpha) = \sum_{n=0}^{\infty} (\alpha i)^n \rho_n(F(z))$$

Summing up the LHS and RHS (5.23) multiplied by $(i\alpha)^n$ over all the integer $n \geq 2$, we have the following PDE and the boundary condition:

$$\begin{aligned} \triangle \mathcal{L}(z;\alpha) &= -\frac{\alpha i}{2}(-z_2 \frac{\partial \mathcal{L}(z;\alpha)}{\partial z_1} + z_1 \frac{\partial \mathcal{L}(z;\alpha)}{\partial z_2}) + \frac{1}{4}\alpha^2 (z_1^2 + z_2^2)\mathcal{L}(z;\alpha), \text{ if } z \in \mathbb{D}; \\ \mathcal{L}(z;\alpha) &= 1, \text{ if } z \in \partial\mathbb{D}. \end{aligned}$$

By the rotation invariance of the Lévy area again, it is easy to see that $\mathcal{L}(z_1, z_2;\alpha)$ is actually a function of $z_1^2 + z_2^2$ given the parameter α, i.e. there exists a function l such that $\mathcal{L}(z_1, z_2;\alpha) = l(z_1^2 + z_2^2;\alpha)$. Then by the chain rule, it is easy to verify that

$$-z_2 \frac{\partial \mathcal{L}(z;\alpha)}{\partial z_1} + z_1 \frac{\partial \mathcal{L}(z;\alpha)}{\partial z_2} = -2z_2 z_1 l'(z_1^2 + z_2^2;\alpha) + 2z_2 z_1 l'(z_1^2 + z_2^2;\alpha) = 0$$

Therefore we can conclude that the characteristic function $\mathcal{L}(.;\alpha)$ is actually a real valued function, and the PDE for $\mathcal{L}(.;\alpha)$ could be simplified as follows:

$$\triangle \mathcal{L}(z;\alpha) = \frac{1}{4}\alpha^2 (z_1^2 + z_2^2)\mathcal{L}(z;\alpha)$$

where $z \in \mathbb{D}$.

The boundary condition is

$$\mathcal{L}(z;\alpha) = 1, \text{ if } z \in \partial\mathbb{D}$$

since $\tau_{\mathbb{D}} = 0$ if the starting point of the Brownian motion z is at the boundary of the disk, i.e. $z_1^2 + z_2^2 = 1$.

Lemma 5.5.4 *Let the function $L(.;\alpha) : \mathbb{D} \to \mathbb{R}$ be a solution to the following PDE problem:*

$$\begin{aligned} \triangle L(z_1, z_2; \alpha) &= \frac{1}{4}(z_1^2 + z_2^2)\alpha^2 L(z_1, z_2; \alpha), \text{ if } (z_1, z_2) \in \mathbb{D}; && (5.26) \\ L(z_1, z_2; \alpha) &= 1, \qquad \text{if } (z_1, z_2) \in \partial\mathbb{D}. && (5.27) \end{aligned}$$

Then the solution L exists and is unique. Besides it has stochastic representation as follows:

$$L(z; \alpha) = \mathbb{E}^z \left[\exp^{-\frac{\alpha^2}{8} \int_0^{\tau_{\mathbb{D}}} |B_t|^2 ds} \right]$$

Moreover $L(z;\alpha)$ has the following representation:

$$L(z; \alpha) = \left(\sum_{k=0}^{\infty} \left(\frac{\alpha^2}{64} |z|^2 \right)^k \left(\frac{1}{(k!)^2} \right) \right) b_0(\frac{\alpha^2}{4})$$

where for every $a \in \mathbb{R}^+$,

$$b_0(a) = \frac{1}{\sum_{k=0}^{\infty} \left(\frac{a}{16} \right)^k \left(\frac{1}{(k!)^2} \right)}$$

Proof. The stochastic representation for the solution to this PDE problem is an immediate consequence of the Feynman-Kac theorem. On the other hand, note that the PDE of $F(z;\alpha)$ is the PDE in Lemma A.3.2 in the appendix for $a = \frac{\alpha^2}{4}$. The second equality comes from Lemma A.3.2 directly. ■

Theorem 5.12 *Let the function L be defined as before in Lemma 5.5.4. Then*

$$\mathcal{L}(z; \alpha) = L(z; \alpha).$$

Proof. Similar to the fixed time case, note that

$$L(z; \alpha) = \left(\sum_{k=0}^{\infty} \left(\frac{\alpha^2}{64} |z|^2 \right)^k \left(\frac{1}{(k!)^2} \right) \right) b_0(\frac{\alpha^2}{4})$$

which means that for every fixed z, $L(z;\alpha)$ can be written as power series in α, and it is an infinitely differentiable function of α, since it is the division of two power series in α and the devisor is non-zero. It is easy to see that for every non-negative integer $n \geq 0$,

$$\frac{\partial^{2n+1} L(z; 0)}{\partial \alpha} = 0.$$

Moreover

$$L(z;\alpha) = \sum_{n\geq 0} \frac{1}{n!}(-i)^n L_n(z)\alpha^n$$

where for every integer $n \geq 0$, $z \in \mathbb{D}$,

$$L_n(z) = (-i)^n \frac{1}{n!}\frac{\partial^n L(z;0)}{\partial \alpha^n}.$$

Since $L(z;\alpha)$ satisfies PDE (5.26) with the initial condition and it can be represented as the power series in α, by comparing the coefficients of α, we can conclude that L_n satisfies the recursive PDEs of $\rho_n(F(z))$, i.e. PDE (5.23). Moreover as the solution to PDE (5.23) is unique, we have that

$$L_n(z) = \rho_n(F(z)).$$

On the other hand, as

$$\sum_{n\geq 0}(-i\alpha)^n L_n(z) = L(z;\alpha),$$

the partial sum

$$\sum_{n\geq 0}(-i\alpha)^n \rho_n(F(z))$$

converges as well and the limit should be $\mathcal{L}(z;\alpha)$. Now the proof is complete. ■

Suppose that for a sufficiently small real number $\tilde{\alpha} > 0$ such that for every $\alpha \in [0,\tilde{\alpha}]$, $\mathcal{M}(z_1,z_2;\alpha)$ is well defined, finite and has continuous second partial derivatives. Assume that the partial sum

$$\sum_{n=0}^{N} \alpha^n \rho_n(F(z))$$

converges as N tends to the infinity. We sum up the LHS and RHS of (5.23) multiplied by α^n over all the integer $n \geq 2$, and setting the first derivative part equals to zero, the function $\mathcal{M}(.;\alpha) : \mathbb{D} \to T((\mathbb{R}^2))$ satisfies the following PDE and boundary condition:

$$\begin{aligned} \triangle\mathcal{M}(z_1,z_2;\alpha) &= -\frac{1}{4}(z_1^2+z_2^2)\alpha^2\mathcal{M}(z_1,z_2;\alpha), \text{ if } (z_1,z_2)\in\mathbb{D}; && (5.28)\\ \mathcal{M}(z_1,z_2;\alpha) &= 1, \qquad \text{if } (z_1,z_2)\in\partial\mathbb{D}. && (5.29)\end{aligned}$$

The boundary condition comes from the fact the Lévy area of the Brownian motion starting at the boundary of the unit disk up to $\tau_{\mathbb{D}}$ is 0.

Lemma 5.5.5 *For every $\alpha \in [0,8]$, define the function $\bar{M}(.;\alpha) : \mathbb{R}^2 \to \mathbb{R}$ as follows:*

$$\bar{M}(z;\alpha) = \left(\sum_{k=0}^{\infty} \left(-\frac{\alpha^2}{64}|z|^2 \right)^k \left(\frac{1}{(k!)^2} \right) \right) b_0(-\frac{\alpha^2}{4})$$

where for every $a \in [-16, 0]$,

$$b_0(a) = \frac{1}{\sum_{k=0}^{\infty} \left(\frac{a}{16}\right)^k \left(\frac{1}{(k!)^2}\right)}.$$

Then $\bar{M}(.;\alpha)$ is a solution to the following PDE:

$$\triangle \bar{M}(z_1, z_2; \alpha) = -\frac{1}{4}(z_1^2 + z_2^2)\alpha^2 \bar{M}(z_1, z_2; \alpha), \text{ if } (z_1, z_2) \in \mathbb{D}; \quad (5.30)$$

$$\bar{M}(z_1, z_2; \alpha) = 1, \quad \text{if } (z_1, z_2) \in \partial\mathbb{D}. \quad (5.31)$$

Proof. This PDE problem (5.30) is actually the case for $a = -\frac{\alpha^2}{4}$ in Lemma A.3.2 in the Appendix. Since $\alpha \in [0,8]$, $-\frac{\alpha^2}{4} \in [0,16]$, Lemma A.3.2 implies that the solution should be in the form as stated in the above lemma. ■

Theorem 5.13 *For every $\alpha \in [0,8]$, let $\bar{M}$ be defined as before in Lemma 5.5.5. Then*

$$\mathcal{M}(z;\alpha) = \bar{M}(z;\alpha)$$

Proof. The proof of this theorem is similar to that of Theorem 5.12. ■

Remark 5.5.1 *Note that $\bar{M}(z;\alpha)$ can only be proved to be written as the power series in α when $\alpha \in [0,8]$. This explains why that*

$$\mathcal{M}(z;\alpha) = \bar{M}(z;\alpha)$$

is true when $\alpha \in [0,8]$ instead of for any $\alpha \in \mathbb{R}^+$.

Remark 5.5.2 *The deterministic representation of the function $\mathcal{L}(.;\alpha)$ and $\mathcal{M}(.;\alpha)$ are closely related to the modified Bessel function of the first kind (Definition A.3.1). More details can be found to Remark A.3.2 in the appendix.*

Lemma 5.5.6 *Let $\hat{\alpha}$ be a strictly positive real number such that*

$$\mathbb{E}^x[e^{\hat{\alpha}\tau_{\mathbb{D}}}] < +\infty$$

Let $\alpha = \min\{2\sqrt{\hat{\alpha}}, 8\}$. Then $\mathcal{M}(z;\alpha)$ has the following stochastic representation:

$$\mathcal{M}(z;\alpha) = \mathbb{E}^z[e^{\frac{1}{8}\alpha^2 \int_0^{\tau_{\mathbb{D}}} |B_t|^2 dt}]$$

where $z = (z_1, z_2) \in \mathbb{R}^2$.

Proof. Let us consider the semimartingale M_t defined as follows:

$$M_t = \mathcal{M}(B_{t\wedge\tau_{\mathbb{D}}};\alpha)e^{\frac{1}{8}\alpha^2\int_0^{t\wedge\tau_{\mathbb{D}}}|B_s|^2ds}.$$

The first step is to prove that for every $t \geq 0$,

$$\mathbb{E}^z[M_0] = \mathbb{E}^z[M_t].$$

To do that, we express M_t in the integral form by applying Itô's formula to M_t:

$$\begin{aligned} M_t &= M_0 + \int_0^{t\wedge\tau_{\mathbb{D}}} \left(\frac{1}{2}\triangle\mathcal{M}(B_u) + \frac{1}{8}\alpha^2|B_u|^2\mathcal{M}(B_u)\right)\exp^{\frac{1}{8}\alpha^2\int_0^u|B_{u_1}|^2du_1}du \\ &+ \int_0^{t\wedge\tau_{\mathbb{D}}}\sum_{i=1}^{2}\frac{\partial\mathcal{M}(B_u)}{\partial z_i}e^{\frac{1}{8}\alpha^2\int_0^u|B_{u_1}|^2du_1}dW_u. \end{aligned}$$

Since the function $\frac{\partial\mathcal{M}(z)}{\partial z_i}$ is continuous in $\bar{\mathbb{D}}$, it is bounded. Let C_i denote its supremum, i.e.

$$C_i = \sup_{z\in\bar{\mathbb{D}}}\left|\frac{\partial\mathcal{M}(z)}{\partial z_i}\right| < \infty$$

where $i = 1, 2$.
For every $s \in [0, t]$,

$$\begin{aligned} &\sum_{i=1}^{2}\mathbb{E}^z\left[\int_{s\wedge\tau_{\mathbb{D}}}^{t\wedge\tau_{\mathbb{D}}}\left(\frac{\partial\mathcal{M}(B_u)}{\partial z_i}e^{\frac{1}{8}\alpha^2\int_0^u|B_{u_1}|^2du_1}\right)^2du\right] \\ \leq\ &\sum_{i=1}^{2}\mathbb{E}^z\left[\int_{s\wedge\tau_{\mathbb{D}}}^{t\wedge\tau_{\mathbb{D}}}\left(\frac{\partial\mathcal{M}(B_u)}{\partial z_i}e^{\frac{1}{8}\alpha^2 u}\right)^2du\right],\ \text{by } |B_{u\wedge\tau_{\mathbb{D}}}| \leq 1; \\ \leq\ &\sum_{i=1}^{2}\mathbb{E}^z\left[\int_{s\wedge\tau_{\mathbb{D}}}^{t\wedge\tau_{\mathbb{D}}}C_i^2e^{\frac{1}{4}\alpha^2u}du\right] \leq \sum_{i=1}^{2}\mathbb{E}^z\left[\int_0^t C_i^2e^{\frac{1}{4}\alpha^2\tau_{\mathbb{D}}}du\right] \\ =\ &\sum_{i=1}^{2}C_i^2t\mathbb{E}^z\left[e^{\frac{1}{4}\alpha^2\tau_{\mathbb{D}}}\right] = \left(\sum_{i=1}^{2}C_i^2t\right)\mathbb{E}^z\left[e^{\hat{\alpha}\tau_{\mathbb{D}}}\right] < +\infty. \end{aligned}$$

Therefore M_t is L_2-integrable. Then by the Itô isometry, we have that

$$\mathbb{E}^z[M_t] = \mathbb{E}^z\left[M_0 + \int_0^{t\wedge\tau_{\mathbb{D}}}\left(\frac{1}{2}\triangle\mathcal{M}(B_u) + \frac{1}{8}\alpha^2|B_u|^2\mathcal{M}(B_u)\right)e^{\frac{1}{8}\alpha^2\int_0^u|B_{u_1}|^2du_1}du\right].$$

As $\mathcal{M}(.)$ is well defined, finite and has continuous second partial derivatives, $\mathcal{M}(.)$ satisfies (5.28)

$$\triangle\mathcal{M}(z_1, z_2;\alpha) = -\frac{1}{4}(z_1^2 + z_2^2)\alpha^2\mathcal{M}(z_1, z_2;\alpha)$$

and $B_{u\wedge\tau_{\mathbb{D}}} \in \bar{\mathbb{D}}$, the drift part of M_u should be zero, and thus for every $t \geq 0$,

$$\mathbb{E}^z[M_t] = \mathbb{E}^z[M_0].$$

It implies that

$$M(0) = \mathbb{E}^z[M_t] = \lim_{t\to+\infty} \mathbb{E}^z[M_t].$$

Furthermore, for every $t \geq 0$, it holds that

$$\begin{aligned} \mathbb{E}^z[|M_t|] &\leq \sqrt{\mathbb{E}^z[M_t^2]} = \sqrt{\mathbb{E}^z[\mathcal{M}(B_{t\wedge\tau_{\mathbb{D}}})^2 e^{\frac{\alpha^2}{4}\int_0^{\tau_{\mathbb{D}}}|B_t|^2dt}]} \\ &\leq C_3\sqrt{\mathbb{E}^z[e^{\frac{1}{4}\alpha^2\tau_{\mathbb{D}}}]} = C_3\sqrt{\mathbb{E}^z[e^{\hat{\alpha}\tau_{\mathbb{D}}}]} \end{aligned}$$

where

$$C_3 = \sup_{z\in\bar{\mathbb{D}}} \mathcal{M}(z;\alpha).$$

By the boundedness of the continuous function $\mathcal{M}(z;\alpha)$ on $\bar{\mathbb{D}}$, C_3 is finite. By assumption, we have that

$$\mathbb{E}^z[e^{\hat{\alpha}\tau_{\mathbb{D}}}] < +\infty.$$

Therefore M_u is uniformly integrable, and by the martingale convergence theorem, we have that

$$\lim_{t\to+\infty} \mathbb{E}^z\left[M_t\right] = \mathbb{E}^z\left[\lim_{t\to+\infty} M_t\right].$$

By the definition of M_u, $\mathbb{P}^z(\tau_{\mathbb{D}} < +\infty) = 1$ and $\mathbb{P}^z(M_0 = z) = 1$, we have that

$$\begin{aligned} \mathcal{M}(z;\alpha) &= \mathbb{E}^z[M_0] = \lim_{t\to+\infty} \mathbb{E}^z[M_t] = \mathbb{E}^z[M_{\tau_{\mathbb{D}}}] \\ &= \mathbb{E}^z\left[e^{\frac{1}{2}\int_0^{\tau_{\mathbb{D}}}\frac{1}{8}\alpha^2|B_t|^2dt}\right] \end{aligned}$$

as by the definition of $\mathcal{M}(.;\alpha)$, the following equation holds:

$$\mathcal{M}(B_{\tau_{\mathbb{D}}};\alpha) = 1.$$

■

Remark 5.5.3 *In the proof of Corollary 5.2.1, we show that $\tau_{\mathbb{D}}$ indeed has the exponential integrability. It means that for α is sufficiently close to zero, $\mathcal{M}(z;\alpha)$ admits this stochastic representation.*

5.6 A discrete analogy: the expected signature of a simple random walk up to an exit time

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ taking values in the integer lattice $\mathbb{Z}^d$ with

$$\mathbf{P}\{X_j = e\} = \frac{1}{2d}, |e| = 1$$

A simple random walk starting at $x \in \mathbb{Z}^d$ is a stochastic process S_n indexed by the nonnegative integers, with $S_0 = x$. We denote by Γ a regular domain of the integer lattice and denote by τ_Γ the first exit time from the domain Γ, i.e.

$$\tau_\Gamma = \inf_{n \geq 0} \{n : X_{n+1} \in ext(\Gamma)\}$$

where $ext(\Gamma)$ is the exterior of Γ.

Moreover $\Phi_\Gamma(z)$ is defined as the expected signature of such a random walk started at x run until the first exit time from the domain $\Gamma(\Gamma \subseteq \mathbb{Z}^d$ is a finite set). By the multiplicative property of the signature and the strong Markov property of the simple random walk, we have the following important equations:

$$\Phi_\Gamma(x) = \sum_{|e_j|=1} \frac{1}{2d} \exp(e_j) \otimes \Phi_\Gamma(x + e_j) \tag{5.32}$$

where e_j is the unit vector in $\mathbb{Z}^d$ with j^{th} component 1.
Rewriting (5.32), we will have

$$\rho_n(\Phi_\Gamma(x)) = \sum_{|e_j|=1} \frac{1}{2d} \sum_{i=0}^{n} \frac{(e_j)^{\otimes i}}{i!} \otimes \rho_{n-i}(\Phi_\Gamma(x + e_j)).$$

It is equivalent to

$$-\Delta \rho_n(\Phi_\Gamma(x)) = \sum_{|e_j|=1} \frac{1}{2d} \sum_{i=1}^{n} \frac{(e_j)^{\otimes i}}{i!} \otimes \rho_{n-i}(\Phi_\Gamma(x + e_j)). \tag{5.33}$$

where Δ is the discrete Laplacian, i.e.

$$\Delta \rho_n(\Phi_\Gamma(x)) = \sum_{|e_j|=1} \frac{1}{2d} \rho_n(\Phi_\Gamma(x + e_j)) - \rho_n(\Phi_\Gamma(x)).$$

It is easy to verify that the boundary condition for $\Phi_\Gamma(x)$ should be

$$\Phi_\Gamma(x) = \mathbf{1} = (1, 0, 0, \ldots), \text{if } x \in \partial\Gamma. \tag{5.34}$$

We summarize our result in the following theorem.

Theorem 5.14 *Let $\Gamma \subseteq Z^d$ be a finite set. Then $\Phi_\Gamma(x)$ should satisfies that for each $n \geq 2$*

$$\begin{aligned} &(a) \quad \Delta\rho_n(\Phi_\Gamma(x)) = -\sum_{|e_j|=1} \frac{1}{2d} \sum_{i=1}^{n} \frac{(e_j)^{\otimes i}}{i!} \otimes \rho_{n-i}(\Phi_\Gamma(x+e_j)), \;\; \textit{if } x \in \Gamma \\ &(b) \quad \rho_n(\Phi_\Gamma(x)) = 0 \;\; \textit{if } x \in \partial\Gamma \end{aligned}$$

and $\rho_0(x) = 1, \rho_1(x) = 0$ for every $x \in \bar{\Gamma}$.

Remark 5.6.1 *Notice that the right hand side of (5.33) is determined totally by the truncated expected signature up to degree $n-1$. It indicates that we can solve $\Phi_\Gamma(x)$ recursively just as the Brownian motion case. The classical potential theory in Theorem 5.15(Page 25[12])guarantees that we can recursively solve a system of the finite difference problems to obtain the whole expected signature of a simple random walk up to the first exit time from Γ.*

Theorem 5.15 *Let $\Gamma \subseteq \mathbb{Z}^d$ be a finite set, $F : \partial\Gamma \to \mathbb{R}, g : \Gamma \to \mathbb{R}$. Then the unique function $f : \bar{A} \to \mathbb{R}$ satisfying*

$$\begin{aligned} &(a) \quad \Delta f(x) = -g(x), \;\; x \in A \\ &(b) \quad f(x) = F(x), x \in \partial A, \end{aligned}$$

is

$$f(x) = \mathbb{E}^x[F(S_\tau) + \sum_{j=0}^{\tau-1} g(S_j)]$$

Immediately we have the following theorem in our setting.

Theorem 5.16 *Let $\Gamma \subseteq \mathbb{Z}^d$ be a finite set. Then $\Phi_\Gamma(x)$ should satisfies that for each $n \geq 2$*

$$\rho_n(\Phi_\Gamma(x)) = \mathbb{E}^x[\sum_{j=0}^{\tau-1} g_n(S_j)]$$

where

$$g_n(x) = \sum_{|e_j|=1} \frac{1}{2d} \sum_{i=1}^{n} \frac{(e_j)^{\otimes i}}{i!} \otimes \rho_{n-i}(\Phi_\Gamma(x+e_j))$$

Moreover $\rho_0(\Phi_\Gamma(x)) = 1, \rho_1(\Phi_\Gamma(x)) = 0$ for every $x \in \bar{\Gamma}$.

Remark 5.6.2 *We identify the finite difference equations for the simple random walk to characterize its expected signature as well. It is interesting to notice that for the discrete case, the* n^{th} *term of the expected signature of a simple random walk does not only depend on* $(n-1)^{th}$ *and* $(n-2)^{th}$ *term as in the Brownian case, but also depends on all the terms of the expected signature up to the* $(n-1)^{th}$ *degree.*

Chapter 6

Conclusion

In this thesis, we derive the PDE of the expected signature of a time homogeneous Itô diffusion process up to a fixed time. Moreover, we manage to obtain the PDE of the expected signature of Brownian motion up to the first exit time from a domain. The derivation of the PDE theorem essentially depends on the Markov property of the stochastic process. This methodology for calculating the expected signature could also be applied to the case of an Itô diffusion process up to the first exit time from a fixed domain as well. Due to the limited time and some remaining technical issues, we do not include the partial results which we have obtained, in the thesis.

In future work, we will continue working on the expected signature of a diffusion process up to the first exit time from a domain and expect to get a more generalized result on it. We would like to consider the expected signature of some non-Markov process, e.g fractional Brownian motion, or Schramm-Loewner Evolution(SLE). In terms of SLE, although it does not obey the Markov property, it has the domain Markov property as well as conformal invariance. Currently we are working on the expected signature of the SLE flow, which is a Markov process, but does not have a drift satisfying the linear growth condition, and in future we hope that we could use it to understand the expected signature of SLE.

Appendix A

Lévy area of a Brownian motion with an arbitrary starting point

A.1 The Stratonovich iterated integrals and the moments of the Lévy area

Let us recall the definition of the Lévy area denoted by L_t, i.e.

$$L_t = \frac{1}{2}\int_0^t -B_s^{(2)} dB_s^{(1)} + B_s^{(1)} dB_s^{(2)}.$$

Lemma A.1.1 *For every integer $n \geq 1$, for every $t \in \mathbb{R}^+$,*

$$\int_{0<t_1<\cdots<t_n<t} dL_{t_1} \circ dL_{t_2} \cdots \circ dL_{t_n} = \frac{1}{n!} L_t^n. \tag{A.1}$$

Proof. It can be proved by simply applying the shuffle product property of the Stratonovich integral. Alternatively we can use Itô calculus and the connection between the Stratonovich integral and the Itô integral to verify this lemma by induction. ∎

A.2 The case for a fixed time

Let $\mathcal{L}(t, x_1, x_2; \lambda)$ denote the characteristic function of Lévy area at time t on condition that the starting point of the driven Brownian motion is x,

$$\mathcal{L}(t, x_1, x_2; \lambda) = \mathbb{E}[e^{i\lambda L_t} | B_0 = x_1 e_1 + x_2 e_2]$$

In order to compute $\mathcal{L}(t, x_1, x_2; \lambda)$, we consider the following PDEs:

$$-\lambda \left(-x_2 \frac{\partial \mathcal{L}(t, x_1, x_2; \lambda)}{\partial x_1} + x_1 \frac{\partial \mathcal{L}(t, x_1, x_2; \lambda)}{\partial x_2} \right) = 0 \tag{A.2}$$

$$\left(-\frac{\partial}{\partial t} + \frac{1}{2}\Delta \right) \mathcal{L}(t, x_1, x_2; \lambda) = \lambda^2 \frac{1}{8}(x_1^2 + x_2^2) \mathcal{L}(t, x_1, x_2; \lambda) \tag{A.3}$$

Solving this PDE, we have the following solution:

$$\mathcal{L}(t, x_1, x_2; \lambda) = \frac{1}{\cosh(\frac{1}{2}\lambda t)} \exp\left(-\frac{1}{4}\lambda(x_1^2 + x_2^2)\tanh(\frac{\lambda}{2}t)\right)$$

We give the sketch of solving procedure of this PDE.
By (A.2), it is reasonable to suggest that the solution should be in the form:

$$\mathcal{L}(t, x_1, x_2; \lambda) = G(t, x_1^2 + x_2^2; \lambda)$$

By the chain rule, we have that

$$\begin{aligned}
\frac{\partial \mathcal{L}(t, x_1, x_2; \lambda)}{\partial t} &= \frac{\partial G(t, x_1^2 + x_2^2)}{\partial t} \\
\frac{\partial \mathcal{L}(t, x_1, x_2; \lambda)}{\partial x_i} &= 2x_i \frac{\partial G(t, x_1^2 + x_2^2)}{\partial x} \\
\frac{\partial^2 \mathcal{L}(t, x_1, x_2; \lambda)}{\partial x_i^2} &= 2\frac{\partial G(t, x_1^2 + x_2^2)}{\partial x} + 4x_i^2 \frac{\partial^2 G(t, x_1^2 + x_2^2)}{\partial x^2}
\end{aligned}$$

where $i = 1, 2$. Then we have

$$-\frac{\partial G(t, x_1^2 + x_2^2)}{\partial t} + 2\frac{\partial G(t, x_1^2 + x_2^2)}{\partial x} + 2(x_1^2 + x_2^2)\frac{\partial^2 G(t, x_1^2 + x_2^2)}{\partial x^2} = \frac{\lambda^2}{8}(x_1^2 + x_2^2)G(t, x).$$

Then we know that $G(t, x)$ satisfies the PDE as follows:

$$\begin{aligned}
-\frac{\partial G(t, x)}{\partial t} + 2\frac{\partial G(t, x)}{\partial x} + 2x\frac{\partial^2 G(t, x)}{\partial x^2} &= \frac{\lambda^2}{8}xG(t, x) \\
G(0, x) &= 1.
\end{aligned}$$

Suppose that the solution has the form

$$\mathcal{L}(t, x_1, x_2; \lambda) = f(t)e^{g(t)(x_1^2 + x_2^2)}.$$

We will get ODEs for the function f and g as follows:

$$\begin{aligned}
-f'(t) + 2f(t)g(t) &= 0 \\
2g(t)^2 - g'(t) &= \frac{1}{8}\lambda^2.
\end{aligned}$$

The second equation be reduced to the Riccati Equation, or the second order linear ordinary differential equation. Substituting g by $-u'/2u$, we have the second order linear ordinary differential equation as follows:

$$u''(t) = \frac{\lambda^2}{4}u(t).$$

Thus the solution $u(t)$ should be in the form:

$$u(t) = C_1 e^{\frac{\lambda}{2}t} + C_2 e^{-\frac{\lambda}{2}t}$$

where C is a constant.
Since $g(0) = 0$, it implies that $u'(0) = 0$, and therefore $C_1 = C_2$. Moreover, it implies further that

$$g(t) = -\frac{1}{2}\frac{u'(t)}{u(t)} = -\frac{1}{4}\lambda\frac{\sinh(\frac{\lambda t}{2})}{\cosh(\frac{\lambda}{2}t)} = -\frac{\lambda}{4}\tanh(\frac{\lambda}{2}t).$$

Since

$$\begin{aligned} \frac{f'(t)}{f(t)} &= 2g(t) \\ f(0) &= 0, \end{aligned}$$

using the integral equality

$$\int \tanh(\frac{\lambda}{2}t)dt = \left(\frac{\lambda}{2}\right)^{-1} \ln(\cosh(\frac{\lambda}{2}t)) + C,$$

the following equality holds

$$\ln(f(t)) = \int^t 2g(s)ds = \int^t -\frac{\lambda}{2}\tanh(\frac{\lambda}{2}s)ds = -\ln(\cosh(\frac{\lambda}{2}t)) + C.$$

Since $f(0) = 1$, $C = 0$, and

$$f(t) = \frac{1}{\cosh(\frac{\lambda}{2}t)}.$$

Finally we have that

$$\begin{aligned} \mathcal{L}(t, x_1, x_2, \lambda) &= f(t)\exp\left(g(t)(x_1^2 + x_2^2)\right) \\ &= \frac{1}{\cosh(\frac{\lambda}{2}t)}\exp\left(-\frac{\lambda}{4}(x_1^2 + x_2^2)\tanh(\frac{\lambda}{2}t)\right). \end{aligned}$$

A.3 The case for the first exit time from the unit disk

A.3.1 Some related details of derivation of the PDE of the expected signature of the Lévy area up to the first exit time of Brownian motion from the unit disk

Recall that $\tau_{\mathbb{D}}$ denotes the exit time of the Brownian motion from the unit disk. To derive the PDE of the expected signature of the Lévy area up to $\tau_{\mathbb{D}}$, we need to calculate the moments of Brownian motion and Lévy area at $\tau_{\mathbb{D}}$. We summarize the result in the following lemma, and give the detailed proof later. Let B_t denote the standard $d-$dimenional Brownian motion, where d is an integer greater than or equal to 1.

Lemma A.3.1 *For every $i, j \in \{1, 2, \ldots, d\}$,*

$$\begin{aligned}
\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} \int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(j)}\right] &= \mathbf{1}(i=j)\frac{1}{2}\varepsilon^2 \\
\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} L_{\tau_{x+\varepsilon\mathbb{D}}}\right] &= \left(-\frac{1}{4}z_2\mathbf{1}(i=1) + \frac{1}{4}z_1\mathbf{1}(i=2)\right)\varepsilon^2 \\
\mathbb{E}^z\left[L^2_{\tau_{z+\varepsilon\mathbb{D}}}\right] &= \frac{1}{8}(z_1^2+z_2^2)\varepsilon^2 + o(\varepsilon^2).
\end{aligned}$$

Proof. Let us recall the definition of $\Phi_\Gamma(z)$, i.e. the expected signature of the Brownian motion starting at $z \in \mathbb{R}^2$ and run until the first exit time from the domain Γ. We will use the notation $\Phi_\Gamma(z)$ in the following proof.
By the shuffle product property of the signature,

$$\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} \int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(j)} = \int_{0<t_1<t_2<\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} \circ dB_{t_2}^{(j)} + \circ dB_{t_1}^{(j)} \circ dB_{t_2}^{(i)}.$$

Taking the expectation of both sides of the above equation, we have that

$$\begin{aligned}
&\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} \int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(j)}\right] \\
= \;& \mathbb{E}^z\left[\int_{0<t_1<t_2<\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} \circ dB_{t_2}^{(j)} + \circ dB_{t_1}^{(j)} \circ dB_{t_2}^{(i)}\right] \\
= \;& \pi^{(i,j)}(\Phi_{z+\varepsilon\mathbb{D}}(z)) + \pi^{(j,i)}(\Phi_{z+\varepsilon\mathbb{D}}(z)).
\end{aligned}$$

By the translation invariance and the scaling property of the expected signature of the Brownian motion (Lemma 5.3.1, Lemma 5.3.2), it follows that

$$\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} \int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(j)}\right] = \varepsilon^2\left(\pi^{(i,j)}(\Phi_{\mathbb{D}}(0)) + \pi^{(j,i)}(\Phi_{\mathbb{D}}(0))\right).$$

Using the fact that the second term of $\Phi_{\mathbb{D}}(0)$ is equal to

$$\frac{1}{4}\sum_{i=1}^{2} e_i \otimes e_i,$$

the expectation $\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} \int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(j)}\right]$ could be simplified to

$$\frac{1}{2}\varepsilon^2\mathbf{1}(i=j).$$

Next, let us calculate $\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} L_{\tau_{x+\varepsilon\mathbb{D}}}\right]$. By the translation property of the Brownian motion,

$$\begin{aligned}
& \mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} L_{\tau_{z+\varepsilon\mathbb{D}}}\right] \\
= \; & \mathbb{E}^0\left[\frac{1}{2}\left(\int_0^{\tau_{\varepsilon\mathbb{D}}} -(z_2+B_t^{(2)})\circ dB_t^{(1)} + (z_1+B_t^{(1)})\circ dB_t^{(2)}\right)\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right] \\
= \; & -\frac{1}{2}z_2\mathbb{E}^0\left[\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(1)}\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right] + \frac{1}{2}z_1\mathbb{E}^0\left[\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(2)}\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right] \\
& +\mathbb{E}^0\left[\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dL_t\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right].
\end{aligned}$$

By the shuffle product property of the signature, the linearity of the expectation and the scaling invariance of the expected signature of the stopped Brownian motion, it is easy to see that

$$\begin{aligned}
\mathbb{E}^0\left[\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(1)}\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right] &= \varepsilon^2\left(\pi^{(1,i)}(\Phi_{\mathbb{D}}(0)) + \pi^{(i,1)}(\Phi_{\mathbb{D}}(0))\right) \\
\mathbb{E}^0\left[\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(2)}\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right] &= \varepsilon^2\left(\pi^{(2,i)}(\Phi_{\mathbb{D}}(0)) + \pi^{(i,2)}(\Phi_{\mathbb{D}}(0))\right) \\
\mathbb{E}^0\left[\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dL_t\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right] &= \varepsilon^3\left(-(\pi^{(1,2)} \sqcup\!\sqcup\, \pi^{(i)})(\Phi_{\mathbb{D}}(0)) + (\pi^{(i)} \sqcup\!\sqcup\, \pi^{(2,1)})(\Phi_{\mathbb{D}}(0))\right).
\end{aligned}$$

Substituting the following equations in the above equations

$$\begin{aligned}
\rho_2(\Phi_{\mathbb{D}}(0)) &= \frac{1}{4}\sum_{i=1}^{2} e_i \otimes e_i \\
\rho_3(\Phi_{\mathbb{D}}(0)) &= 0
\end{aligned}$$

then we have

$$\begin{aligned}
\mathbb{E}^0\left[\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(1)}\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right] &= \frac{1}{2}\varepsilon^2\mathbf{1}(i=1) \\
\mathbb{E}^0\left[\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(2)}\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right] &= \frac{1}{2}\varepsilon^2\mathbf{1}(i=2) \\
\mathbb{E}^0\left[\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dL_t\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(i)}\right] &= 0.
\end{aligned}$$

Therefore

$$\mathbb{E}^z\left[\int_0^{\tau_{z+\varepsilon\mathbb{D}}} \circ dB_{t_1}^{(i)} L_{\tau_{x+\varepsilon\mathbb{D}}}\right] = \left(-\frac{1}{4}z_2\mathbf{1}(i=1) + \frac{1}{4}z_1\mathbf{1}(i=2)\right)\varepsilon^2.$$

Similarly we can calculate $\mathbb{E}^z\left[L^2_{\tau_{z+\varepsilon\mathbb{D}}}\right]$ as follows:

$$
\begin{aligned}
\mathbb{E}^z\left[L^2_{\tau_{z+\varepsilon\mathbb{D}}}\right] &= \frac{1}{4}\mathbb{E}^0\left[\left(\int_0^{\tau_{\varepsilon\mathbb{D}}} -(z_2+B_t^{(2)})\circ dB_t^{(1)} + (z_1+B_t^{(1)})\circ dB_t^{(2)}\right)^2\right] \\
&= \frac{1}{4}\mathbb{E}^0\left[\left(\int_0^{\tau_{\varepsilon\mathbb{D}}} -z_2\circ dB_t^{(1)} + z_1\circ dB_t^{(2)} + \circ dL_t\right)^2\right] \\
&= \mathbb{E}^0\left[z_2^2\left(\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(1)}\right)^2 + z_1^2\left(\int_0^{\tau_{\varepsilon\mathbb{D}}} \circ dB_t^{(2)}\right)^2\right] + o(\varepsilon^2) \\
&= \frac{1}{8}(z_1^2+z_2^2)\varepsilon^2 + o(\varepsilon^2).
\end{aligned}
$$

■

A.3.2 Solving explicitly the PDE of the characteristic function and the moment generating function of the Lévy area of the Brownian motion up to the first exit time from the unit disk

The main difficulty computation of the characteristic function and the moment generating function of the Lévy area of the Brownian motion up to the first exit time from the unit disk turns out to be how to solve the following PDE (A.4) and the boundary condition (A.5) in Lemma A.3.2 when $a = \frac{\alpha^2}{4}$ and $a = -\frac{\alpha^2}{4}$ respectively.

Lemma A.3.2 *For every fixed $a \in [-16, \infty)$, the following PDE problem admits a solution:*

$$
\begin{aligned}
\triangle f(z_1, z_2; a) &= a(z_1^2+z_2^2) f(z_1,z_2;a), \quad \text{if } (z_1,z_2)\in\mathbb{D}; && \text{(A.4)} \\
f(z_1,z_2;a) &= 1, \quad \text{if } (z_1,z_2)\in\partial\mathbb{D}. && \text{(A.5)}
\end{aligned}
$$

where $f : \mathbb{D} \to \mathbb{R}$, and a is a parameter. The solution $f(z;a)$ is given as follows:

$$f(z;a) = \left(\sum_{k=0}^{+\infty}\left(\frac{a|z|^2}{16}\right)^k\left(\frac{1}{(k!)^2}\right)\right) b_0(a)$$

where

$$b_0(a) = \frac{1}{\sum_{k=0}^{+\infty}\left(\frac{a}{16}\right)^k\left(\frac{1}{(k!)^2}\right)}.$$

Proof. Let us change from the cartesian coordinates to polar coordinates. Noted that the boundary condition does not depend on the angle and the PDE only depends on the radius. It is reasonable to suggest that $f(.;a)$ is a function of r where $r = \sqrt{z_1^2 + z_2^2}$. For simplicity let us still use the function $f(r;a)$ to denote the function f after changing to polar coordinates. The PDE (A.4) is rewritten in polar form as follows:

$$\begin{aligned} \frac{1}{r}f'(r;a) + f''(r;a) &= ar^2 f(r;a),\ r \in (0,1) && \text{(A.6)} \\ f(1) &= 1. && \text{(A.7)} \end{aligned}$$

Since the function f is defined in the unit disk $\mathbb{D}$, which implies that $f(0;a)$ should be finite, and $f(r;a)$ should be analytic, and represented as a power series of r as follows:

$$f(r;a) = \sum_{n\geq 0} b_n(a) r^n.$$

Then its first derivative and the second derivative can be written in terms of the power series in r as well, supposing that the derivatives exist;

$$\begin{aligned} f'(r,a) &= \sum_{n\geq 1} n b_n(a) r^{n-1} \\ f''(r,a) &= \sum_{n\geq 2} n(n-1) b_n(a) r^{n-2}. \end{aligned}$$

Therefore the LHS of the PDE of $f(r;a)$ (A.4) can be expressed in terms of the power series in r as follows:

$$\sum_{n\geq 1} n b_n(a) r^{n-2} + \sum_{n\geq 2} n(n-1) b_n(a) r^{n-2} = b_1(a) r^{-1} + \sum_{n\geq 2} n^2 r^{n-2}$$

while the RHS of the PDE of $f(r;a)$ (A.4) is equal to

$$a \sum_{n\geq 0} b_n(a) r^{n+2}.$$

By comparing the coefficients of r^n on both sides, we have that

$$\begin{aligned} b_1(a), b_2(a), b_3(a) &= 0; \\ \forall n \geq 2, (n+2)^2 b_{n+2}(a) &= a b_{n-2}(a) \end{aligned}$$

By induction, it is easy to see that

$$b_n(a) = \begin{cases} \frac{a}{(4k)^2} b_{n-4}(a) = a^k \left(\prod_{j=0}^{k} \frac{1}{(4j)^2}\right) b_0(a), & \text{if } n = 4k; \\ 0, & \text{otherwise.} \end{cases}$$

where $n \in \mathbb{N}$.

Now we can write the function $f(r;a)$ as follows:

$$\begin{aligned} f(r;a) &= \sum_{k=0}^{\infty} b_{4k}(a) r^{4k} \\ &= \left(\sum_{k=0}^{\infty} (ar^4)^k \left(1 + \prod_{j=1}^{k} \frac{1}{(4j)^2} \right) \right) b_0(a). \end{aligned}$$

Recall the boundary condition $f(1;a) = 1$, and it gives the constraint on $b_0(a)$, i.e.

$$\left(\sum_{k=0}^{\infty} a^k \left(1 + \prod_{j=1}^{k} \frac{1}{(4j)^2} \right) \right) b_0(a) = 1$$

For every fixed $a \in \mathbb{R}$, $S(a)$ is defined as below

$$S(a) = \sum_{k=0}^{\infty} a^k \left(1 + \prod_{j=1}^{k} \frac{1}{(4j)^2} \right) = \sum_{k=0}^{+\infty} \left(\frac{a}{16} \right)^k \frac{1}{(k!)^2}.$$

By Lemma A.3.3, for every $a \in [-16, +\infty)$

$$0 < S(a) < +\infty.$$

Thus for every $a \in [-16, \infty)$, $b_0(a)$ is well defined and is equal to

$$b_0(a) = \frac{1}{S(a)} = \frac{1}{\sum_{k=0}^{\infty} \left(\frac{a}{16} \right)^k \frac{1}{(k!)^2}}.$$

It is easy to show that for every $a \in [-16, \infty)$, $f(r;a)$ is indeed twice differentiable when $r \in (0,1)$ by the ratio test. ■

Remark A.3.1 *The ODE (A.6) with the boundary condition (A.7) has solution*

$$\frac{1}{K_0(\frac{\sqrt{a}}{2})} \left[K_0 \left(\frac{\sqrt{a} r^2}{2} \right) + I_0 \left(\frac{\sqrt{a} r^2}{2} \right) K_0 \left(\frac{\sqrt{a}}{2} \right) C_1 - I_0 \left(\frac{\sqrt{a}}{2} \right) K_0 \left(\frac{\sqrt{a} r^2}{2} \right) C_1 \right]$$

where $I_c(x)$ and $K_c(x)$ are the modified Bessel function of the first kind and the second kind respectively defined in Definition A.3.1and C_1 is an arbitrary constant.

Remark A.3.2 *The solution $f(z;a)$ to PDE (A.4) with the boundary condition (A.5) is indeed the normalized Bessel function of the first kind, i.e.*

$$\frac{I_0(\frac{\sqrt{a}}{2}|z|^2)}{I_0(\frac{\sqrt{a}}{2})}$$

where $I_c(x)$ is the modified Bessel function of the first kind defined as follows (Definition A.3.1). Using the constraint that $f(0,a)$ is finite, the constant C_1 can be uniquely determined.

Definition A.3.1 *The modified Bessel functions of the first kind, denoted by $I_c(x)$ and mapping from $\mathbb{R}$ to $\mathbb{R}$ are defined as follows:*

$$I_c(x) = \sum_{n=0}^{+\infty} \frac{1}{n!\Gamma(n+c+1)} \left(\frac{x}{2}\right)^{2n+c}$$

where $\Gamma(z)$ is the gamma function and $c \in \mathbb{R}$.
The modified Bessel functions of the second kind, denoted by $K_c(x)$ and mapping from $\mathbb{R}^+$ to $\mathbb{R}$ are defined as follows:

$$K_c(x) = \frac{\pi}{2} \frac{I_{-c}(x) - I_c(x)}{\sin(c\pi)}$$

where $\Gamma(z)$ is the gamma function and $c \in \mathbb{R}$.

Lemma A.3.3 *Let us define the function $S : \mathbb{R} \to \mathbb{R}$ as follows:*

$$S(a) = \sum_{k=0}^{\infty} \left(\frac{a}{16}\right)^k \frac{1}{(k!)^2}.$$

For every $a \in \mathbb{R}$,

$$S(a) < +\infty.$$

For every $a \in [-16, \infty)$,

$$S(a) > 0.$$

Proof. First of all, let us prove for every $a \in \mathbb{R}$, $S(a)$ is finite. Let $c_k(a)$ denote $(\frac{a}{16})^k \left(\frac{1}{(k!)^2}\right)$ for every $k \in \mathbb{N}$. By the ratio test, since for every fixed $a \in \mathbb{R}$, there exists an integer $K = \lfloor \frac{\sqrt{|a|}}{4} \rfloor + 1$, such that for every integer $k \geq K$,

$$\left| \frac{c_{k+1}(a)}{c_k(a)} \right| = \left| \frac{\left(\frac{a}{16}\right)^{k+1} \left(\frac{1}{((k+1)!)^2}\right)}{\left(\frac{a}{16}\right)^k \left(\frac{1}{(k!)^2}\right)} \right| = \frac{|a|}{(4(k+1))^2} < 1$$

where $\lfloor x \rfloor$ is the largest integer which is no greater than x, for every $x \in \mathbb{R}$.
It implies the series $\sum_{i=0}^{k} c_i(a)$ absolutely converges, and thus $S(a)$ is finite.
The next step is to show that for every $a \in [-16, +\infty]$, $S(a)$ is strictly positive.
For $a \geq 0$, $S(a)$ is increasing when $a \in \mathbb{R}^+$, thus

$$S(a) \geq S(0) = 1$$

For $a \in [-16, 0]$,

$$\begin{aligned} S(a) &= \left(\sum_{k=0}^{\infty} \left(\frac{|a|}{16} \right)^{2k} \left(\frac{1}{((2k)!)^2} \right) \right) - \left(\sum_{k=0}^{\infty} \left(\frac{|a|}{16} \right)^{2k+1} \left(\frac{1}{((2k+1)!)^2} \right) \right) \\ &= \sum_{k=0}^{\infty} \left(\frac{|a|}{16} \right)^{2k} \left(\frac{1}{((2k)!)^2} \right) - \left(\frac{|a|}{16} \right)^{2k+1} \left(\frac{1}{((2k+1)!)^2} \right) \end{aligned}$$

Since $a \in [-16, 0]$, then for every $k \in \mathbb{N}$,

$$\left(\frac{|a|}{16} \right)^{2k} \left(\frac{1}{(2k)!} \right)^2 \geq \left(\frac{|a|}{16} \right)^{2k+1} \left(\frac{1}{(2k+1)!} \right)^2$$

and at least for one $k \in \mathbb{N}$, the above inequality is strict, which implies that for every $a \in [-16, 0]$,

$$S(a) > 0$$

■

Bibliography

[1] Sandra Cerrai. Elliptic and parabolic equations in with coefficients having polynomial growth. *Communications in Partial Differential Equations*, 21(1-2):281–317, 1996.

[2] Kuo-tsai Chen. Integration of pathsa faithful representation of paths by non-commutative formal power. *Trans. Amer. Math. Soc.*, (156):395–407, 1971.

[3] Thomas Fawcett. Problems in stochastic analysis. Connections between Rough paths and Non-Commutative Harmonic Analysis. *PhD thesis, University of Oxford*, 2003.

[4] Peter Friz and Nicolas Victoir. The Burkholder-Davis-Gundy Inequality for Enhanced Martingales. *Lecture notes in Mathematics*, 421-438, 2008.

[5] Peter Friz and Nicolas Victoir. *Multidimensional Stochastic Processes as Rough Paths: theory and application.* Cambridge University Press, 2010.

[6] Ben Hambly and Terry Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Annals of Mathematics*, 109-167, 2010.

[7] K Helmes and A Schwane. Levy's stochastic area formula in higher dimensions. *Journal of Functional Analysis*, 54(2):177 – 192, 1983.

[8] Kiyoshi Itô. On a stochastic integral equation. *Mem. Amer. Math. Soc*, 4:1–51, 1946.

[9] Jurgen Jost. *Partial Differential Equations.* Springer, 2 edition, 1441923802 2010.

[10] Ioannis Karatzas and Steven Shreve. *Brownian Motion and Stochastic Calculus.* Number 113 in Graduate Texts in Mathematics. Springer, 2 edition, 1991.

[11] N. V. Krylov. *Controlled Diffusion Process.* Number 14 in Applications of Mathematics. Springer-Verlag, 1980.

[12] Gregory Lawler. *Intersections of random walk.* Probability and its application. Birkhauser, 1996.

[13] Yves LeJan and Zhongmin Qian. Stratonovich's Signatures of Brownian Motion Determine Brownian Sample Paths. Feb 2011. arXiv:1102.3601v1.

[14] Terry Lyons, Thierry Lévy, and Michael Caruana. *Differential Equation driven by Rough Paths.* Springer, 2006.

[15] Terry Lyons and Hao Ni. Expected signature of two dimensional Brownian Motion up to the first exit time of the domain. *arXiv:1101.5902v4*, Jan 2011.

[16] Terry Lyons and Zhongmin Qian. *System Control and Rough Paths.* Oxford University Press, 2002.

[17] Terry Lyons and Nicolas Victoir. Cubature on Wiener Space. *Proc. R. Soc. Lond. A*, 169-198, 2004.

[18] Bernt Øksendal. *Stochastic Differential Equation: An Introduction with Applications.* Springer, 6 edition, 2007.

[19] Elias Stein and Guido Weiss. *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, May 1971.