MARTINGALE REPRESENTATION THEOREM FOR DIFFUSION IN INFINITE DIMENSIONAL SPACES AND APPLICATIONS

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

MATHEMATICS

By

Uğur Aydın

March 2023

MARTINGALE REPRESENTATION THEOREM FOR DIFFUSION IN INFINITE DIMENSIONAL SPACES AND APPLICATIONS

By Uğur Aydın

March 2023

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Ali Süleyman Üstünel(Advisor)

Mine Çağlar

Laurent Decreusefond

David Nualart

Mustafa Çelebi Pınar

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

MARTINGALE REPRESENTATION THEOREM FOR DIFFUSION IN INFINITE DIMENSIONAL SPACES AND APPLICATIONS

Uğur Aydın
M.S. in Mathematics
Advisor: Ali Süleyman Üstünel
March 2023

We show that square integrable martingales adapted to the filtration generated by a weak solution of a stochastic differential equation driven by a cylindrical Wiener process on a separable real Hilbert space that has the weak uniqueness property has a martingale representation driven by the martingale part of the stochastic differential equation.

Keywords: stocastic differential equation, martingale representation, Malliavin calculus.

ÖZET

DİFÜZYON İÇİN SONSUZ UZAYDA MARTİNGAL GÖSTERİMİ VE UYGULAMALARİ

Uğur Aydın
Matematik, Yüksek Lisans
Tez Danışmanı: Ali Süleyman Üstünel
Mart 2023

Silindirik bir Wiener süreci ile sürülen sonsuz boyutta yaşayan bir stokastik diferansiyel denkleminin zayıf nadirlik özelliğini sağlaması durumunda bütün zayıf çözümleri için yarattığı karesi integrallenebilen rassal değişkenlerin bulunduğu uzay içindeki bütün martingallerin aldığı stokastik integral gösteriminin varlığını göstereceğiz.

Anahtar sözcükler: stokastik diferansiyel denklemi, martingal gösterimi, Malliavin kalkülüsü.

Acknowledgement

I would like to thank Professor Ali Süleyman Üstünel for his guidance and advice throughout the years, as well as for always optimally motivating me to push for more.

I also would like to thank Professor Mustafa Çelebi Pınar, Professor Laurent Decreusefond, and Professor David Nualart for taking the time to read the thesis and be part of my thesis committee. I am particularly grateful to Professor Mine Çağlar for her constant encouragement and feedback on the thesis.

Finally, I thank my family and friends for being wonderful. I would be remiss if I did not thank Ahmet and Keleş for their moral support.

Contents

Chapter 1

Introduction

Construction of stochastic integral gives us a characterization of processes $\xi = (\xi_s; s \in [0,1])$ such that $(\int_0^t \xi_s dX_s; s \in [0,1])$ is a martingale, where X is a square integrable process. In this thesis we prove the converse of this statement, i.e. we characterize the elements of the space $L^2(\mathcal{F}_1(X))$ where $(\mathcal{F}_s(X); s \in [0,1])$ is the filtration generated by a diffusion process X, with unique law, in some Hilbert space H that has the following form:

$$dX_t = b(t,X)dt + \sigma(t,X)dW_t,$$

where W is a cylindrical Wiener process that lives in, perhaps another Hilbert space different than H, U.

Let $(\Omega, \mathcal{F}_1(X), P)$ be a probability space. As an application of the martingale problem associated to X, we prove that the set

$$\Gamma = \left\{ \xi \in L_2(U,K) : E\int_0^1 \operatorname{tr}(\xi_s\sigma(s,X))(\xi_s\sigma(s,X))^* ds < \infty \right\}$$

is dense in $L^2(\mathcal{F}_1(X); K)$. As a consequence of this observation, we then represent the whole $L^2(\mathcal{F}_1(X); K)$ with the stochastic integrals of the form $\int_0^1 \xi_s P_s(X) dW_s$ where $\xi_s \in L_a^2(dt \times dP, \mathcal{F}(X); L_2(U,K))$, and $P(X) = (P_s(X); s \in [0,1])$ is the measruable version of the projection to $\ker \sigma^{\perp}$, which exist by a measurable selection theorem.

To accomplish this, we start with an introduction to measure theory in Hilbert spaces, and give a detailed account of Gaussian processes in infinite dimensional space to make sense of the cylindrical Wiener processes. To this end, we introduce the concept of measure "on a Banach space" and precisely discuss the structures arise from the celebrated Cameron-Martin theorem for Gaussian measures. After the construction of stochastic integrals with respect to a cylindrical Wiener process, we give heuristic background for the Gaussian space that is generated by a cylindrical Wiener process and deal with the construction of Malliavin calculus that arise from it. We pay special attention to the case when the underlying Hilbert space is the space of Hilbert-Schmidt operators, which we will use to establish certain relations later on. As a word of caution, rather than working with Wiener spaces, we will be working on arbitrary probability spaces thus, one needs to construct the derivative by analytic means rather than the Gateux derivative, which no longer makes sense in such generality.

In chapter 3, we give a proof of Yamada-Watanabe theorem for X as we will implicity use it for existence of weak solutions.

In section 4 we give the formulation of martingale problem in the subsequent chapter and establish the martingale representation theorem.

In chapter 5, we give a direct application of this martingale representation theorem in presence of innovation process.

In chapter 6, we give a sketch of the proof of Malliavin differentiability of diffusion in our setting, and create an auxiliary differential operator via conditional expectation with respect to $\mathcal{F}_1(X)$ of Malliavin derivative. We conclude the chapter with an application of the tools we have developed through the martingale representation and obtain Log-Sobolev inequality when the underlying measure is the law of our diffusion.

Chapter 2

Preliminaries

2.1 Stochastic analysis in infinite dimensional spaces

In this section we will give some elementary background for the tools that we will use later in this thesis. Details regarding stochastic analysis in infinite dimensional spaces can be found in [3] and Gaussian measures in infinite dimensional spaces can be found in [2].

2.1.1 Elements of stochastic analysis and probability theory

A probability space for us will be the 3-tuple $(\Omega, \mathcal{F}, P)$ or $(\Omega, \mathcal{F}, \mu)$. All of our measures will be Borel. An increasing family of σ-algebras that is contained in $\mathcal{F}$, will be referred as a filtration. If $\mathbb{F}$ is a filtration in our probability space such that each of its elements are completed with respect to the measure P (including the initial σ-algebra $\mathcal{F}$) we will refer to the tuple $(\Omega, \mathcal{F}, P, \mathbb{F})$ as a normal stochastic basis. For the most part we are interested in continuous time processes, hence the filtration will be indexed with the compact interval $[0,1]$ for the most part. Implicitly we will assume that our filtration $(\mathcal{F}_t)_{t\in[0,1]}$ are right continuous, i.e. $\mathcal{F}_t = \cap_{\epsilon>0}\mathcal{F}_{t+\epsilon}$. Unless stated otherwise, we will implicitly assume that we are working on such a stochastic basis with filtration that is indexed on $[0,1]$ from now on.

A random variable will be a metric space valued measurable function for us. Particular treatment for operator valued stochastic processes will be given explicitly. We shall work on Hilbert space valued random variables for the most part hence, we will work with Bochner integrals. The most frequent space that we will use with regarding to the integrability will be the space of square integrable functions L^2. By $L^2([0,1];H)$ we will mean H-valued square integrable (equivalence class of) functions on $[0,1]$ with respect to the Lebesgue integral over $[0,1]$. If the underlying probability/measure space is understood, we will use $L^2(\mu,H)$ to indicate H-valued square integrable (equivalence class of) functions with respect to the measure μ over the whole space. If there will be no confusion with regarding to the underlying measure, to indicate the measurability with respect to a particular σ-algebra, let's say $\mathcal{F}(X)$, we will write $L^2(\mathcal{F}(X),H)$ to denote H-valued (equivalence class of) functions that are square integrable with respect to the underlying measure over the whole space that are $\mathcal{F}(X)$ measurable.

For the most part we will ignore the variable $w\in\Omega$. A random variable X is said to Bochner integrable if

$$E[\|X\|] = \int_\Omega \|X\| dP < \infty.$$

Law of a random variable X in a probability space with probability measure μ will be denoted by several different notations throughout the paper

$$\mathcal{L}(X) = X(\mu)$$

and defined as

$$\mathcal{L}(X)(\cdot) \coloneqq \mu(X^{-1}(\cdot)).$$

The notation $X(\mu)$ is referred as pushforward of μ by X. We will prefer to use

$\mathcal{L}(X)$ when we will not need to mention the underlying probability measure.

For a random variable X by $\mathcal{F}(X)$ we will denote the smallest σ-algebra generated by X. If $X = (X_t)_{t\in[0,1]}$ is a stochastic process, then the filtration generated by X will be referred as the right continuous, completed version of the filtration generated by X.

A stochastic process will be a family of random variables that depend on time. For our purposes, they will usually depend on the interval $[0,1]$ as a time parameter. One can easily take $\mathbb{R}_+$ as the domain for time parameter however, for most of our purposes space of continuous functions will play crucial role for us and in which case we will have to deal with Frechet spaces. Although dealing with Frechet spaces is doable, we shall refrain from it hence we will stick to the case $[0,1]$ (or any $[0,T]$ for $T < \infty$).

2.1.1.1 Operator valued stochastic processes

Let U and H be separable Hilbert spaces. By $L(U,H)$ we will denote the space of linear bounded operators form U to H. A trace-class operator will be an operator with finite trace (or trace-class operator). By $\mathrm{tr}\cdot$ we will denote the trace of the operator. Hence, if A is a trace class operator, then $\mathrm{tr}A < \infty$.

We say that an operator A is Hilbert-Schmidt if AA^* is trace class, where A^* denotes the adjoint of A. The space of Hilbert-Schmidt operators from U to H will be denoted by $L_2(U,H)$. At certain times to indicate the structure explicitly we may write $L_2(U,H)$ as $H \otimes U$ by abusing the notation to not indicate the completed Hilbert-Schmidt product.

Since the space $L(U,H)$ is not separable, it is pathological with regarding to most of our notion of integration (and even the measurability assumptions do not work properly). However such difficulty does not appear in the space $L_2(U,H)$ as it is separable. For this reason almost all integration that we will do, let it be with respect to a measure or a Wiener process, if the integrand is some sort of operator it will be class of $L_2(U, H)$. Therefore, we can make sense of Hilbert-Schmidt valued processes by the means of the previous subsection, although the case of a general linear bounded operator valued processes / random variables needs some extra care.

We conclude this subsection with the definition of quadratic variation in separable Hilbert spaces.

Definition 2.1.1. *We say that a trace-class operator valued process $V : H \to H$ a quadratic variation of a martingale M, if for any $a, b \in H$,*

$$\langle M_t, a\rangle\langle M_t, b\rangle - \langle V_t a, b\rangle$$

is a martingale (with respect to the underlying stochastic basis) and $V_0 = 0$, increasing, and $V_t \le V_s$ if $0 \le t \le s \le 1$. We will denote V by $\langle\langle M\rangle\rangle$.

Existence of such V follows from [3, Proposition 3.12].

2.1.1.2 Gaussian processes and measures in infinite dimensional spaces

We say that a measure is Gaussian on a Banach space B if for its pushforward with respect to the elements $\ell \in B^*$ is a Gaussian measure on $\mathbb{R}$.

Definition 2.1.2. *[2, Definition 2.2.7] Given a probability measure on a separable Banach space B, say μ, its covariance operator $I_\mu : B^* \to B^{**}$ is defined by*

$$I_\mu(\ell, \ell') = \int_B \ell(x)\ell'(x)d\mu(x)$$

Definition 2.1.3. *[2, Definition 2.2.7] Mean of a measure μ on a separable Banach space B, m_μ, is defined as $m_\mu(\ell) = \int_B \ell(x)d\mu(x)$ for any $\ell \in B^*$.*

Remark 2.1.1. *For most measures such definition of mean requires the space B to be reflexive, as m_μ is inherently in B^{**}, to obtain any reasonable theory.*

However, we will mainly work on separable Hilbert spaces hence, this will not cause any problem.

Following [2, Theorem 2.3.1] we can obtain the following definition:

Definition 2.1.4. *When μ has mean m_μ and covariance I_μ, we will denote it by $\mathcal{N}(m_\mu, I_\mu)$. In case μ satisfies the formula*

$$\int_B e^{i\ell(x)}\mu d(x) = \exp\left(-\frac{1}{2}I_\mu(\ell,\ell)\right)$$

for any $\ell \in B^*$, then we call μ as a centered Gaussian measure on B.

Since we can always shift a Gaussian measure with non zero mean to a Gaussian measure with zero mean, we can without loss of generality consider centered Gaussian processes while discussing the theory.

In our generality, for a Gaussian measure μ, the covariance operator I_μ will be compact. More can be said in case $B = H$ is a Hilbert space.

Proposition 2.1.1 ([9])**.** *If $B = H$ is a Hilbert space, then for the operator $\tilde{I}_\mu$ that is defined as $\langle \tilde{I}_\mu h, k\rangle_H = I_\mu(h,k)$ is trace-class and has the identity*

$$\int_H \|x\|_H^2 \mu d(x) = \mathrm{tr}\tilde{I}_\mu.$$

For any positive, symmetric, trace-class operator $A : H \to H$, there exists a Gaussian measure on H such that $\tilde{I}_\mu = A$.

When $B = H$ is a Hilbert space, by abusing the terminology, we can associate m_μ by an element of U, say m, and I_μ with $\tilde{I}_\mu = Q$. In such case by abusing the terminology we will also say a random variable has mean m and covariance Q if its distribution has such mean and covariance. Especially in the following subsection, when we are dealing with Wiener processes, this convention will be used as we will not be working with the underlying measures explicitly.

2.1.1.3 Cameron-Martin space of Gaussian measures

Suppose that B is a separable Banach space. For a given Guassian measure μ on B, it is possible to associate μ with a subspace H_μ of B that leaves μ quasi-invariant, i.e. translated measure in the direction of the vectors in H_μ is equivalent to μ. When H_μ is infinite dimensional, it necessarily has $\mu H_\mu = 0$, yet it completely determines the Gaussian measure as we will see. This feature is completely counter-intuitive when one considers the measures in finite dimensional spaces such as Lebesque measure, which does not exist in infinite dimensional spaces. The space H_μ is called the Cameron-Martin space and defined as follows:

Definition 2.1.5 ([2])**.** *Let $\tilde{H}_\mu \subset B$ be the linear subspace*

$$\tilde{H}_\mu = \{h \in B : \exists h^* \in B^* \text{ with } I_\mu(h^*, \ell) = \ell(h) \forall \ell \in B^*\},$$

where B^* is the dual space of B. We define the Cameron-Martin space associated to μ as the completion of the subspace $\tilde{H}_\mu$ under the norm $\|h\|_\mu^2 = I_\mu(h^*, h^*) = \langle h^*, h^* \rangle$.

Thereby, Cameron-Martin space associated to a Gaussian measure is the range of the operator $\tilde{I}_\mu$.

Proposition 2.1.2. *[2, Corollary 3.2.6] If μ and ν are Gaussian measures on B, such that $H_\mu = H_\nu$ and $\|h\|_\mu = \|h\|_\nu$ for every $h \in H_\nu$, then they are equal to each other. In particular, the Cameron-Martin space completely determines the measure.*

Theorem 2.1.1 (Cameron-Martin theorem)**.** *Define a shift operator $T : B \to B$ by $T(x) = x + h$ for some $h \in B$. Then, $T_h^*\mu$ is absolutely continuous with respect to μ if and only if $h \in H_\mu$, where $T_h^*\mu$ is the pushforward of μ under $x \to x + h$.*

Theorem 2.1.2. *[9] Let μ be a Gaussian measure on a separable Banach space K with Cameron-Martin space H_μ. Let $A : H_\mu \to B$, here B is a separable Banach space too, be a bounded linear map such that there exists a Gaussian measure ν on B with covariance $I_\nu(\ell_1, \ell_2) = I_\mu(A^*\ell_1, A^*\ell_2)$, where A^* is the adjoint of A. Then the linear measurable extension of A is unique, up to null sets.*

Theorem 2.1.3. *[9] If the measure ν as in Theorem 2.1.2 is centered Gaussian, there exists a measurable map $R : K \to B$ such that $\nu = R^*\mu$, and there exists a measurable linear subspace $V \subset B$ with $\mu V = 1$ such that $R|_V$ is linear and $R|_{H_\mu} = A$ with $H_\mu \subset V$. In case A is Hilbert-Schmidt class, then ν can be assumed to be any Gaussian measure and the statement holds verbatim.*

It is easy to see that Theorem 2.1.2 and Theorem 2.1.3 complement each other. As an interpretation of these theorems, the spaces K and B are completely irrelevant when one considers a Gaussian measure! This interpretation will be crucial to the philosophy behind the cylindrical Wiener processes, which we will consider in the next subsection

2.1.2 Stochastic integral with respect to a Wiener process in separable Hilbert spaces

2.1.2.1 Wiener processes with trace-class covariance and stochastic integrals with respect to them

Fix a probability space $(\Omega, \mathcal{F}, P)$. We further fix a sequence of increasing σ-algebras, $(\mathcal{F}_t)_{t\in[0,1]}$ such that $\mathcal{F}_1 \subset \mathcal{F}$, $\mathcal{F}_t$ are right-continuous, and are completed. Let U and H be Hilbert spaces. For a bounded linear operator $Q : U \to U$ such that tr$Q < \infty$, there exists a complete orthonormal basis $\{u_k\}_k$ of U, and a bounded sequence of non negative real numbers λ_k such that $Qe_k = \lambda_k e_k$ $k = 1, 2, \cdots$ as the trace-class operators are compact.

Definition 2.1.6. *We say that a process $W = (W_t)_{t\in[0,1]}$ is a U-valued Q-Wiener process with respect to the filtration $(\mathcal{F}_t)_{t\in[0,1]}$ if*

- $W_0 = 0$,
- *W has continuous sample trajectories almost everywhere,*
- *For $1 \geq t \geq s \geq 0$ law of $W_t - W_s$, $\mathcal{L}(W_t - W_s)$, has zero mean and its covariance is $(t-s)Q$, i.e. $\mathcal{L}(W_t - W_s) = \mathcal{N}(0, (t-s)Q)$ for $1 \geq t \geq s \geq 0$.*

- *W_t is $\mathcal{F}_t$-measurable,*
- *$W_{t+h} - W_t$ is independent of $\mathcal{F}_t$ for any $h \in [0, 1-t] \subset [0,1]$.*

Whenever $\mathrm{tr}Q < \infty$, such W always exist on U that lies in our probability space $(\Omega, \mathcal{F}, P)$ that is adapted to a given filtration, and for arbitrary $t \in [0,1]$, W can be represented as

$$W_t = \sum_{n=1}^{\infty} \sqrt{\lambda_n} \beta_n(t) e_n$$

where $\beta_n(t) = \frac{1}{\sqrt{\lambda_n}} \langle W_t, e_n \rangle$ is a stantard Wiener process for all n and β_n are mutually independent from each other in $(\Omega, \mathcal{F}, P)$. We will fix a Q-Wiener process for Q such that $\mathrm{tr}Q < \infty$ throughout this section.

We say that an $L(U,H)$-valued process $\xi = (\xi_s)_{s \in [0,1]}$ is elementary, if there exist a finite partition $0 = t_1 \leq t_2 \leq \cdots \leq t_n = 1$ of $[0,1]$, and a sequence of $L(U,H)$ valued random variables $\xi_{t_1}, \cdots, \xi_{t_{n-1}}$ such that ξ_{t_i} is $\mathcal{F}_{t_i}$ measurable for $i = 1, \cdots, n-1$ and

$$\xi_s = \sum_{i=1}^{n-1} \xi_{t_i} 1_{(t_i, t_{i+1}]},$$

where $1_\bullet$ is the indicator function. We define the stochastic integral for elementary functions with respect to a Wiener process by the formula

$$\int_0^t \xi_s dW_s = \sum_{i=1}^{n-1} \xi_{t_i} (W_{t_{i+1} \wedge t} - W_{t_i \wedge t}).$$

Hence, the stochastic integral is H-valued and adapted to the filtration $(\mathcal{F}_t)_{t \in [0,1]}$.

Extension to larger class of integrands will be done via the isometric isomorphism provided by the stochastic integral to the space that is equipped with the norm $\| \cdot \|_M^Q := (E \int_0^t \mathrm{tr}(\cdot Q^{1/2})(\cdot Q^{1/2})^* ds)^{1/2}$. In case there will be no confusion, we will omit the operator Q from the notation. For this purpose we will consider predictable processes. First, we need the following proposition:

Proposition 2.1.3 (Itô isometry)**.** *Let ξ be an elementary process that is $L(U,H)$-valued such that $E \int_0^t \mathrm{tr}(\xi_s Q^{1/2})(\xi_s Q^{1/2})^* ds < \infty$, then for any $t \in [0,1]$*

we have

$$E\left\|\int_0^t \xi_s dW_t\right\|_H^2 = E\int_0^t \mathrm{tr}(\xi_s Q^{1/2})(\xi_s Q^{1/2})^* ds,$$

and $\int_0^\cdot \xi_s dW_s$ *is continuous, square integrable* H*-valued martingale.*

Furthermore, the elementary processes themselves are dense in the space of $L_2(U,H)$ valued predictable processes:

Proposition 2.1.4. *If* ξ *is* $L_2(U,H)$*-valued predictable process with* $\|\xi\|_M < \infty$*, then there exists a sequence of elementary processes* $\{\xi^n\}$ *such that* $\|\xi^n - \xi\|_M \to 0$ *as* $n \to \infty$*.*

By using the proposition above along with the Itô isometry for elementary processes, by the Riesz-Fischer theorem we have that for $L_2(U,H)$ valued predictable processes the stochastic integral is well defined under the conditions above. Furthermore, the integral is a continuous martingale [3].

2.1.2.2 Cylindrical Wiener process and stochastic integrals with respect to them

In finite dimensional spaces, N-dimensional Wiener process is defined as follows:

$$W_t^N = \sum_{i=1}^N \beta_i(t) e_i$$

where $(e_i)_i$ is the orthonormal basis of the space and β_i are mutually independent 1-dimensional standard Wiener processes. The immediate question is that can one construct infinite dimensional standard Wiener process in a similar fashion, i.e. as an infinite tuple that consist of mutually independent copies of standard Wiener processes without any other factor in the series representation, i.e. Wiener process with series representation of the form:

$$W_t = \sum_{i=1}^\infty \beta_i(t) u_i$$

where u_i is a complete orthonormal basis of U and β_i are 1-dimensional mutually independent standard Wiener processes.

The extra factors in front of the 1-dimensional standard Wiener processes in the trace-class covariance case certainly messes up the relation of the Wiener process with the heat kernel for instance (as otherwise Wiener process would not exert the same force in every direction). However, without any extra factors in front of the individual one dimensional standard Wiener processes the covariance operator Q would be the identity operator I hence, the previous construction fails, and in general such processes explode in the initial space U since the identity operator is not trace-class in infinite dimensional spaces. It turns out that, although on the initial space this construction is not possible, we can consider slightly extended spaces to compensate such structure, and extend the notion of Q-Wiener processes with non-trace-class covariance operators. Such processes are called cylindrical Wiener processes, and the most important one, which also will be the type of Wiener process that we will consider throughout the text, will be the one with the identity as the covariance operator.

To make sense of such Wiener processes in infinite dimensional spaces, we will embed the process on a bigger space than the initial one. For the existence of such embedding, we can fix a symmetric, nonnegative, bounded and linear covariance operator Q. Take a copy of U, say $U_1 = U$, and pick a strictly positive and finite sequence $(\theta_i)_{i\in\mathbb{N}}$ such that $\sum_{i=1}^{\infty}\theta_i^2 < \infty$. We define a new operator, J, that will embed $Q^{1/2}U$ to U by

$$J(u) = \sum_{i=1}^{\infty} \theta_i \langle u, u_i\rangle u_i,$$

where u_i is a complete orthonormal basis of $Q^{1/2}U$. It immediately follows that J is one to one and Hilbert-Schmidt. In general, such operators will be called as Hilbert-Schmidt embedding and as we just showed that there always exist one. Further constructions of such embedding can be found in [20]. We will fix a Hilbert-Schmidt embedding $J : Q^{1/2}U \to U_1$ for rest of this section.

Proposition 2.1.5 ([21]). *Define $Q_1 = JJ^*$, and pick an infinite sequence of real valued mutually independent Wiener processes $(\beta_i)_{i\in\mathbb{N}}$ that are not necessarily* *standard. Then the process* $W = (W_t)_{t\in[0,1]}$ *defined as*

$$W_t = \sum_{i=1}^{\infty} \beta_i(t) J u_i$$

is a Q_1*-Wiener process in* U_1*.*

Definition 2.1.7. *We call any processes* W *as in the previous proposition as the cylindrical* Q*-Wiener process.*

Remark 2.1.2. *It is not immediate, however* Q_1 *is nonnegative, symmetric, and trace-class [3, Proposition 4.11]. Therefore, after the embedding we transform cylindrical Wiener processes to the trace-class ones. In case when* β_i *are standard real valued Wiener processes, we will obtain a cylindrical* I*-Wiener process.*

Certainly, in case when $J = I$ and Q is trace-class, any cylindrical Q-Wiener process is just a Q-Wiener process.

This construction shines when one wants to consider stochastic integrals with respect to such processes. It is easy to see that $\xi \in L_2(Q^{1/2}U, H)$ if and only if $\xi \circ J^{-1} \in L_2(Q_1^{1/2}(U_1), H)$. Thus, it is natural to define the stochastic integral as

$$\int_0^{\cdot} \xi_s dW_s := \int_0^{\cdot} \xi_s J^{-1} dW_s,$$

where J^{-1} can be thought as the left inverse of J, which exists, as a partial isometry. This integral is certainly independent of U_1 and J, which is easy to check for elementary processes as they do not depend on J, then using the uniform convergence of the truncated stochastic integral driven by W^N we can extend this result to the stochastic integrals driven by W. As in the definition of cylindrical Q-Wiener processes with trace-class covariance operator Q, in case when Q is trace-class by taking $J = I$ this integral coincides with the stochastic integral for Q-Wiener processes.

Although a cylindrical Q-Wiener processes seem to force us to work on U_1 rather than $Q^{1/2}U$ (since they do not exist on the space $Q^{1/2}U$), we will not work on U_1 at all throughout the paper. The reason being is that, when one considers the stochastic integral with respect to a cylindrical Wiener process W, there is no dependence of W on U_1 whatsoever. This crucial observation will allow us to think W as a process in the initial space rather than in U_1 (even though W lives in U_1). Hence, cylindrical Wiener processes will only be a technicality for us as the stochsatic integral with respect to it will be our main concern thus, it is useful to think them as the formal sum

$$W_t = \sum_{i=1}^{\infty} \beta_i u_i.$$

In turn, cylindrical I-Wiener process will be the infinite dimensional analogue of n-dimensional standard Wiener process.

In general it happens to be the case that if μ is the Gaussian measure on U_1 with the covariance $Q_1 = JJ^*$, then μ has U as the associated Cameron-Martin space as $\tilde{H}_\mu$ is the range of J, and JJ^* will give us the closure. Thus, associated Wiener measure of the extended Wiener processes share the same Cameron-Martin space [2].

Unless stated otherwise, all of our Wiener processes from now on will be cylindrical I- Wiener process, which we will call it simply by cylindrical Wiener process.

Similar to the finite dimensional case, when the driving process of the stochastic integral is a cylindrical Wiener process, and the integrand is adapted (hence progressively measurable, as we always consider such modification) to the (completed) filtration generated by the cylindrical Wiener process, the process also will be predictable with respect to the filtration generated by the cylindrical Wiener process. Thus, we can make sense for the integrands that are adapted in such case, similar to the finite dimensional case.

The last result that we will recall regarding cylindrical Wiener process will be the Girsanov theorem, which is an extension of Cameron-Martin theorem.

Theorem 2.1.4 (Girsanov theorem)**.** *Let ξ be a U-valued $(\mathcal{F}_t)_t$ predictable process* *such that*

$$E\exp\left(\int_0^1 \langle \xi_s, dW_s\rangle_U - \frac{1}{2}\int_0^1 \|\xi_s\|_U^2 ds\right) = 1.$$

Then the process

$$\tilde{W}_t = W_t - \int_0^t \xi_s ds$$

is a cylindrical Wiener process with respect to the measure $\tilde{P}$*, where* $\tilde{P}$ *is defined as*

$$d\tilde{P} = \exp\left(\int_0^1 \langle \xi_s, dW_s\rangle_U - \frac{1}{2}\int_0^1 \|\xi_s\|_U^2 ds\right) dP.$$

2.1.2.3 Itô formula and stochastic Fubini theorem

Suppose that the equation

$$X_t = X_0 + \int_0^1 b(s,X)ds + \int_0^t \sigma(s,X)dW_s$$

is well defined (for details check section 4), where W is a cylindrical I-Wiener process, and the process $X = (X_t)_{t\in[0,1]}$ is solution to this equation. To make any sort of calculation regarding to the process X, the standard tool to utilize is the so called Itô formula.

Definition 2.1.8. *We say that function* $F : [0,1]\times H \to \mathbb{R}$ *is regular if* $F_t, F_x,$ *and* F_{xx} *are uniformly continuous on bounded subsets of* $[0,1]\times H$.

Theorem 2.1.5 (Itô formula [3])**.** *For any regular* F*,* P*-a.e. for any* $t\in[0,T]$ *we have*

$$\begin{aligned} F(t,X_t) = F(0,X_0) &+ \int_0^t \Big(F_t(t,X_t) + \langle F_x(t,X_t), b(s,X_s)\rangle \\ &- \frac{1}{2}\mathrm{tr}\big[F_{xx}(t,X_t)\sigma(t,X_t)(\sigma(t,X_t))^*\big]\Big)dt \\ &+ \int_0^t \langle F_x(s,X_s), \sigma(s,X)dW_s\rangle \end{aligned}$$

Similar to integration with regarding to measures, it is also possible to define a Fubini-type theorem for the stochastic integrals:

Theorem 2.1.6 (Stochastic Fubini Theorem [3]). *Assume that* $\xi(\cdot, x) = (\xi_s(\cdot, x)_{s\in[0,1]})$ *is* $L_2(U, H)$ *valued predictable process for any* $x \in E$*, and*

$$\int_E \left(\int_0^1 \mathrm{tr}(\xi_s(x))(\xi_s(x))^* ds \right) d\mu(x) < \infty,$$

then P*-a.e.*

$$\int_E \left(\int_0^1 \xi_s(x) dW_s \right) d\mu(x) = \int_0^1 \left(\int_E \xi_s(x) d\mu(x) \right) dW_s.$$

2.2 Malliavin calculus in infinite dimensional spaces

Our main aim will be to work on general probability spaces rather than Wiener spaces. An equivalent theory to Malliavin calculus can be constructed in these frameworks by introducing a coordinate-process that is Gaussian, i.e. an isonormal Gaussian process.

For this purpose, we shall give a quick introduction to Gaussian spaces, with particular emphasis to those arising from a Brownian motion to motivate the coordinate-process that we use in case when we are dealing with a cylindrical Wiener process. Having done that, we will shift our focus to construction of Malliavin calculus in Hilbert spaces. Details can be found in [8] [13] [15] [28].

2.2.1 Gaussian spaces

Definition 2.2.1. *A Gaussian space is a vector space that consists of (real) centered Gaussian random variables in some probability space* $(\Omega, \mathcal{F}, P)$*. We say that a Gaussian space is complete, if it is complete as a subspace of* $L^2(\Omega, \mathcal{F}, P)$*, with respect to the underlying* L^2 *inner product.*

Although we restricted ourselves to the Gaussian spaces with centered Gaussian random variables, we can shift the space by considering $\xi \to \xi - E\xi$ for non-centred Gaussian spaces. Thus, it is sufficient to only study the Gaussian spaces that consist of centered Gaussian spaces. It is straight forward to check that L^2 completion of a Gaussian space is a complete Gaussian space, and for any $p \in (0, \infty)$ L^p topologies coincide for Gaussian spaces.

Most interesting Gaussian spaces comes from the ones that is generated by particular (uncountable or countable) families of mutually independent centered Gaussian random variables. For our purpose, and the heart of the Malliavin calculus, will be the ones that is generated by the trajectory of a standard Brownian motion (regardless of the dimension), in our case a cylindrical Wiener process.

Usual Malliavin calculus is constructed on the classical Wiener space where we can define a Gateux derivative in the direction of the Cameron-Martin space associated to the coordinate process of the setup, which in turn defines a Sobolev derivative. Our aim is to obtain the same structure, but for arbitrary Wiener process. Hence the coordinates has to be shifted in the direction of the Cameron-Space of Gaussian space rather than the probability space, but for a specified coordinate process.

2.2.1.1 Gaussian space generated by a one dimensional standard Brownian motion

Let $(B_t)_{t\in[0,1]}$ be a one dimensional standard Brownian motion on $L^2(\Omega, \mathcal{F}, P)$. Let $G(B)$ be the complete Gaussian space spanned by $\{B_t\}_{t\in[0,1]}$. Since $B_t = \int_0^\infty 1_{[0,t]} dB_t$, and

$$EB_tB_r = \int_0^\infty 1_{[0,t]}1_{[0,r]}ds,$$

as the result is deterministic we see that for deterministic family $\{a_i\}$, $\sum a_i 1_{[0,t_i]} \to \sum a_i B_{t_i}$ gives an isometry from the Itô isometry of step functions onto the incomplete Gaussian space spanned by $\{B_t\}_t$. Since step functions will be dense in $L^2([0,1])$, this isometry can be extended linearly to an isometry from

$L^2([0,1])$ to $G(B)$. Thus, if $h \in L^2([0,1])$, then $\int_0^1 h_t dB_t$ will be in $G(B)$ and the whole space is actually made of such elements. Furthermore, $W(h) = \int_0^1 h_s dB_s$ (the Paley-Wiener integral) acts as a coordinate process on $G(B)$ just like the coordinate process of the classical Wiener space. Such W is called as an isonormal Gaussian process in the literature.

2.2.1.2 Gaussian space generated by a cylindrical Wiener process

A Gaussian random variable X on a Hilbert space H, can be defined by the means of $\langle X, h\rangle$ where $h \in H$ by checking whether for each $h \in H$, $\langle X, h\rangle$ is a real Gaussian random variable.

If for a moment one assumes that a cylindrical Wiener measure W with the same setup we have discussed before actually exists, then for any $t \in [0,1]$ and $h \in U$, $\langle W_t, h\rangle_U$ will be a real valued Gaussian process. Let $G(W)$ be the space spanned by $\{\langle W_t, h\rangle\}_{h\in H, t\in[0,1]}$. Everything in the one dimensional case applies to $G(W)$ line by line, with the twist that $\sum a_i \langle W_t, h\rangle = \sum \tilde{h}_i W_t$ by Riesz representation theorem where $h_i \in L^2([0,1], L_2(U,\mathbb{R}))$. Therefore by the reasoning above, the Cameron-Martin space associated to the with the measure that is associated with the Paley-Wiener integral driven by cylindrical Wiener measure W will be the space

$$\left\{\int_0^{\cdot} h(s)ds : h \in L^2([0,1], L_2(U,\mathbb{R}))\right\}.$$

Even without the existence of a cylindrical Wiener process on the desired space, the stochastic integral driven by W exists and is a centered Gaussian process. Thus, the Paley-Wiener integral driven by W generates the space $G(W)$ with the same Cameron-Martin space nevertheless. By considering the one dimensional Brownian case as a prototype, the Malliavin calculus for such processes will be constructed by the associated Paley-Wiener integral playing the role of a coordinate process.

Since not every probability space can be characterized by the means of an isonormal Gaussian process, we will construct the Malliavin derivative with an analytical formula rather than as a directional derivative in the next section. In case our setting will be the classical Wiener space, it will correspond with the Gateux derivative in the direction of the Cameron-Martin space.

Furthermore, the cylindrical Wiener process induces a (cylindrical) Gaussian measure on the initial space rather than a full-blown Gaussian measure. As a consequence of the Sazonov theorem [16, Theorem 6.7], an extension to a full-blown Wiener process of the cylindrical Wiener process can be only obtained by applying the procedure we described before by an Hilbert-Schmidt embedding, which justifies the constructions we have done so far.

2.2.2 Malliavin calculus in Hilbert spaces for Hilbert space valued functionals

For a deterministic map $h \in L^2([0,1], L_2(U_0, \mathbb{R}))$ we define the Wiener integral with respect to a cylindrical Wiener process W as, similar to the case in finite dimensional Euclidean spaces,

$$W(h) = \int_0^1 h(t) dW_t.$$

By using function which are infinitely differentiable, and at most polynomial growth (i.e. functions that belong to $C_p^\infty(\mathbb{R}^n, R)$) we shall define the set of smooth cylindrical functions in H as follows:

$$\mathcal{S}(H) = \left\{ F \middle| F = \sum_{i=1}^{n} f_i(W(h_1), \cdots, W(h_m)) H_i, \right.$$
$$\left. H_i \in H, f_i \in C_p^\infty(\mathbb{R}^m, R), h_i \in L^2([0,1], L_2(U_0, \mathbb{R}) \right\}.$$

Since $C_p^\infty(\mathbb{R}^n, \mathbb{R})$ is dense in $L^2([0,1])$, and elements of H is dense in H, as $L^2([0,1], H) \simeq L^2([0,1]) \otimes H$ our cylindrical functions (after completion) will define the largest class which we can work with.

Definition 2.2.2. *The Malliavin derivative of $F \in \mathcal{S}(H)$ will be denoted as ∇F and is defined by*

$$\nabla F(t) = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{\partial}{\partial x_i} f_i(W(h_i), \cdots, W(h_m)) H_i \otimes h_j(t),$$

for $t \in [0,1]$, where $h_j \in L_2(U_0, \mathbb{R})$.

Remark 2.2.1. *One can define the directional Malliavin derivative for $F \in \mathcal{S}(H)$ as the right action on the right hand side of the equation above. It is not hard to see that then, ∇F is the unique operator definited by the Riesz representation theorem.*

Lemma 2.2.1. *For $F \in \mathcal{S}(H) \subset L^2(H)$, $\nabla F(t) \in L^2(L_2(U_0, H))$ for almost all $t \in [0,1]$.*

Proof. Suppose $(e_n)_n$ is an orthonormal basis of U, then we have that

$$\begin{aligned} \|H_i \otimes h_j(t)\|_2^2 &= \sum_{k=1}^{\infty} \|H_i \otimes h_j(t) e_k\|^2 \\ &= \sum_{k=1}^{\infty} \|h_j(t)(e_k) H_i\|^2 = \|H_i\|^2 \|h_j(t)\|_{L_2(U_0,\mathbb{R})}^2 < \infty, \end{aligned}$$

hence for almost all $t \in [0,1]$, $H_i \otimes h_j(t) \in L_2(U_0, H)$. $\square$

Remark 2.2.2. *It is straightforward to check that ∇F actually does not depend on the representation of F [15], hence it is well-defined. Thus, from now on without loss of generality we shall chose orthonormal basis $(H_i)_i$ of H and $(h_j)_j$ of $L_2(U_0, \mathbb{R})$ to represent any $F \in \mathcal{S}(H)$, implicitly.*

Theorem 2.2.1. *The operator $\nabla : \mathcal{S}(H) \subset L^2(H) \to L^2(\mu, L^2([0,1], L_2(U_0, H)))$ is closable.*

Proof. Almost identical to the Euclidean case. Proof for this in rather general setting can be found in [22] $\square$

Definition 2.2.3. *The Sobolev space of Malliavin differentiable functionals, $\mathbb{D}_{2,1}(H)$, is the closure of $\mathcal{S} \subset L^2(H)$ with respect to the seminorm*

$$\|F\|_{\mathbb{D}_{2,1}(H)} = (E\|F\|^2 + E[\|DF\|^2_{L^2([0,1],L_2(U_0,H)))})^{1/2}.$$

We shall extend the operator ∇ to the space $D_{2,1}(H)$, and denote the extension as ∇, by abusing the notation.

Following lemmas will be useful for us when we differentiate our diffusion process later on, which can be proven identical to the Euclidean case:

Lemma 2.2.2. *Suppose $\phi : H \to H_1$ has continuous Frechet derivative with at most linear growth, where H_1 is a real separable Hilbert space. Then for any $F \in \mathbb{D}_{2,1}(H)$, $\nabla\phi(F) = \phi'(F)\nabla F$.*

Lemma 2.2.3. *Let $\phi : H \to H_1$ be a continuous Frechet-differentiable map, where H_1 is a real separable Hilbert space. If ϕ and ϕ' are Lipschitz continuous, then for any $F \in D_{2,1}(H)$, $\nabla\phi(F) = \phi'(F)\nabla F$.*

Definition 2.2.4. *The operator $\delta : dom\delta \subset L^2(\mu, L^2([0,1], L_2(U,H)))) \to L^2(\mu, H)$ defined via the relation $E[\langle DF, G\rangle_{L^2([0,1],L_2(U,H)))}] = E[\langle F, G\rangle]$ is called the Skorohod operator.*

Closedness of δ follows from that of ∇. Furthermore, proofs of the following propositions are also identical to the Euclidean ones:

Proposition 2.2.1. *Assume $F \in L^2(\mu; L^2([0,1], L_2(U,H)))$ is predictable. Then $F \in dom\delta$ and $\delta F = \int_0^1 F(s)dW_s$. In other words, $dom\delta$ contains all the stochastic integrable functions with respect to W.*

Proposition 2.2.2. *If $F \in \mathbb{D}_{2,1}(L_2(U,H))$, then $\delta F \in \mathbb{D}_{1,1}(H)$ and $D\delta u = u + \delta(W(DF))$.*

Chapter 3

Yamada-Watanabe Theorem

To obtain the martingale representation theorem, we need the weak uniqueness of the diffusion equation, which in turn will require us to use the Yamada-Watanabe theorem to obtain large class of equations that satisfy our needs. In finite dimensional case, a detailed proof of the Yamada-Watanabe theorem can be found in [21]. In this section, we will repeat the same proof in infinite dimensional spaces, we do not claim any originality.

Let U and H be seperable real Hilbert spaces. Let $L_2(U,H)$ be the space of Hilbert-Schmidt operators from U to H. By abusing the notation, we will consider cylindrical Wiener processes on U. We will consider the following SDE on H:

$$dX_t = b(t,X)dt + \sigma(t,X)dW_t, t \in [0,1] \tag{3.1}$$

for $b : [0,1] \times C([0,1],H) \to H$ and $\sigma : [0,1] \times C([0,1],H) \to L_2(U,H)$ such that

- they are $\mathcal{B}([0,1]) \times \mathcal{B}(C([0,1],H))$ (abused the notation to not to mention the target space) measurable,
- and $b(t,\cdot)$, and $\sigma(t,\cdot)$ are $\mathcal{B}_t(C([0,1],H))$ measurable for any fixed $t \in [0,1]$, where $\mathcal{B}_t(C([0,1],H))$ is denoted by the σ-algebra generated by the projections on the space up to time t.

Definition 3.0.1. *We say that (X, W), for $(X_t)_{t\in[0,1]}$ is $(\mathcal{F}_t)$-adapted process with continuous paths and H valued and W is a cylindrical Wiener process on a complete stochastic basis $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ is a weak solution if*

i) For any $t \in [0, 1]$

$$\int_0^t \|b(s, X)\|_H^2 ds + \int_0^t \|\sigma(s, X)\|_{L_2(U,H)}^2 ds < \infty$$

P-a.e.

ii) For any $t \in [0, 1]$

$$X_t = X_0 + \int_0^t b(s, X)ds + \int_0^t \sigma(s, X)dW_s$$

P-a.e.

Set $C([0, 1], H)$ and $W_0 = \{w \in C([0, 1], U)|w(0) = 0\}$ and equip them with the uniform norm and Borel σ algebras $\mathcal{B}(A)$ and $\mathcal{B}(W_0)$.

Definition 3.0.2. *We say that weak uniqueness holds for* (1) *whenever any two solutions (X, W) and (X', W') on stochastic bases $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ and $(\Omega', \mathcal{F}', P', (\mathcal{F}'_t))$ such that*

$$X(P)(\{0\}) = X'(P')(\{0\})$$

holds as measures on $(H, \mathcal{B}(H))$, we have $X(P) = X'(P')$ as measures on $(C([0, 1], H), \mathcal{B}(C([0, 1], H)))$

In the vein of [10] we define the notion of strong solution and unique strong solution for stochastic differential equations:

Definition 3.0.3. *We say that (X, W) is a strong solution to the equation* (3.1) *if there exists a function $F(x, w) : H \times W_0 \to C([0, 1], H)$ that satisfies the following:*

i) There exists a function such that for any probabiltiy measure ν on

$(H, \mathcal{B}(H))$, $\tilde{F}_\nu(x,w) : H \times W_0 \to C([0,1], H)$ *that is* $\overline{(H \times W_0)}^{\nu \times P^W}$ *measurable and for* ν*-a.e.* $x \in H$ $F(x,w) = F_\nu(x,w)$ *for* P^W*-a.e.* $w \in W_0$.

ii) For each $x \in H$, $w \to F(x,w)$ *is* $\overline{\mathcal{B}_t(W_0)}^{P^W}$ *measurable.*

iii) $X = F(X(0), W)$ *a.e.*

Definition 3.0.4. *We say that* (1) *has a unique strong solution if there exists* $F(x,w) \to H \times W_0 \to C([0,1], H)$ *with the properties above and*

i) on any stochastic basis $(\Omega, \mathcal{F}, \mu, (\mathcal{F}_t)_t)$ *there exists a Wiener process* W *that is adapted to* $\mathcal{F}_t$ *and any* H *valued random variable* ξ *that is* $\mathcal{F}_0$ *measurable, the process* $X = F_\mu(\xi, W)$ *is a solution of* (1) *on this stochastic basis with* $X_0 = \xi$ *a.e.,*

ii) any weak solution (X, W) *to* (1) *satisfies* $X = F_{X(P)}(X_0, W)$ *a.e.*

Definition 3.0.5. *We say that pathwise uniqueness holds for* (1) *whenever any two solutions* (X, W) *and* (X', W) *on a stochastic basis* $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ *are a.e. indistinguishable when* $X_0 = X'_0$.

Theorem 3.0.1. *Existence of a strong solution implies existence of a weak solution.Furthermore, if the strong solution is unique, pathwise uniqueness holds.*

Proof. Follows from the definitions immediately. □

To prove that existence of weak solution and pathwise uniqueness implies existence of a strong solution, we start by defining two copies of $C([0,1], H)$, namely W^1 and W^2 to couple the different solutions to the equation. Each copy of the space $C([0,1], H)$ will hold a solution to the SDE and we will couple $W^1 \times W^2$ with another space, that will hold the driving Wiener process, W_0. Define a measure π on the product space $(W^1 \times W^2 \times W_0, \mathcal{B}(W^1) \otimes \mathcal{B}(W^2) \otimes \mathcal{B}(W_0))$ as follows:

$$\pi(dw_1, dw_2, dw_3) = R(w_3, dw_1)R'(w_3, dw_2)\mu(dw_3)$$

where $R(w_3, dw_1)$ (resp. $R'(w_3, dw_2)$) is a conditional probability measure on $W^1 \times W_0$ (resp. $W^2 \times W_0$) such that R (resp. R') is a measure on (W^1, W_0) (resp.

$(W^2, W_0))$ that is the image of P (resp. P') under the map $w \to (X_\cdot(w), B_\cdot(w))$ from Ω into $W^1 \times W_0$ (resp. $w \to (X'_\cdot(w), B'_\cdot(w))$ from Ω' into $W^2 \times W_0$), and $R(w_3, dw_1)$ (resp. $R'(w_3, dw_1)$) is the projection on W_0, that is characterized by the following properties:

- $R(w_2, \cdot)$ is a probability measure on $W^1 \times W_0$
- For any $A \in \mathcal{B}(W_1)$ we have $R(w_2, A)$ is $\mathcal{B}(W_0)$ measurable in w_2.
- For each $A \in \mathcal{B}(W_1)$ and $B \in \mathcal{B}(W_0)$, we have that $(X, B)(P)(A \times B) = \int_B R(w_2, A)\mu(dw_2)$.

We define R' in a similar fashion and μ is the Wiener measure on W_0. Existence of R and R' follows from [10, Theorem 3.1] since $C([0, T], H)$ is a Polish space.

Lemma 3.0.1. *For any $A \in \mathcal{B}_t(W^1)$, $w \to R(w, A)$ is $\overline{\mathcal{B}_t(W_0)}^{\mu}$-measurable.*

Proof. Let $r = (r_k)_k$ be an increasing partition of $[t, 1]$ for $t \in [0, 1]$. Set $C_r = \{w \in W_0 | w_{r_i} - w_{r_{i-1}} \in \tilde{C}_i, \tilde{C}_i \in \mathcal{B}_t(U)\}$ and $C \in \mathcal{B}_t(W_0)$, then we have

$$
\begin{aligned}
&\int_{W_0} 1_{C \cap C_r} R(w_3, A)\mu(dw_3) \\
&= \int_{W^1 \times W_0} 1_A 1_{C \cap C_r}(X, W)(P)(dw_3, dw_1) \\
&= \int_\Omega 1_A(X) 1_C(W) 1_{C_r}(W) dP
\end{aligned}
$$

from the definition. Using that $1_{C_r}(W)$ is P-independent we have

$$
\begin{aligned}
&\int_\Omega 1_A(X) 1_C(W) E_P(1_{C_r} | \mathcal{F}_t) dP \\
&= E_P(1_{C_r}) \int_\Omega 1_A(X) 1_C(W) dP \\
&= E_P(1_{C_r}) \int_{W_0} 1_C R(w_3, A)\mu(dw_3) \\
&= E_P(1_{C_r}) \int_{W_0} E_\mu 1_C [R(w_3, A) | \mathcal{F}_t] \mu(dw_3)
\end{aligned}
$$

$$= \int_{W_0} 1_{C\cap C_r} E_\mu[R(w_3, A)|\mathcal{F}_t]\mu(dw_3).$$

For C_r as given above, and C above, every $C_r \cap C$,C_r,C generates $\mathcal{B}(W_0)$. Therefore, $w \to R(w, A)$ is $\overline{\mathcal{B}_t(W_0)}^\mu$-measurable. □

Proposition 3.0.1. *$P_3 : W^1 \times W^2 \times W_0 \to W_0$ defined as $P_3(w_1, w_2, w_3) = w_3$ is a $\overline{\sigma(w_1(s), w_2(s), w_3(s); s \leq t)}^\pi$-Wiener process for the measure π.*

Proof. For any $t > s$, and $A_1 \in \mathcal{B}(W^1)$, $A_2 \in \mathcal{B}(W^2)$, and $A_3 \in \mathcal{B}(W_0)$

$$\begin{aligned}
&E_\pi\left[\exp(i\langle K, P_3(t) - P_3(s)\rangle 1_{A_1\times A_2\times A_3}\right]\\
&= \int_{W_0}\int_{W^2}\int_{W^1} \exp(i\langle K, P_3(t) - P_3(s)\rangle 1_{A_1\times A_2\times A_3} R(w_3, dw_1)R'(w_3, dw_2)\mu(dw_3)\\
&= \int_{W_0} \exp(i\langle K, P_3(t) - P_3(s)\rangle 1_{A_3} R(w_3, A_1)R'(w_3, A_2)\mu(dw_3)\\
&= \exp(-(|K|^2/2)(t-s)) \int_{W_0} 1_{A_3} R(w_3, A_1)R'(w_3, A_2)\mu(dw_3)\\
&= \exp(-(|K|^2/2)(t-s))\pi(A_1 \times A_2 \times A_3)
\end{aligned}$$

by using the previous lemma on the fourth equality. Thus, we have proved that P_3 has independent increments. The other properties follows trivially hence, P_3 is a Wiener process. □

We define two new auxiliary functions $P_1 : W^1 \times W^2 \times W_0 \to W^1$, and $P_2 : W^1 \times W^2 \times W_0 \to W^2$ as the canonical projections.

Lemma 3.0.2. *$\mathcal{L}(X, W) = \mathcal{L}(P_1, P_3)$ and $\mathcal{L}(X', W') = \mathcal{L}(P_2, P_3)$.*

Proof. Pick $A_1 \times A_3 \in \mathcal{B}(W^1) \otimes \mathcal{B}(W_0)$. Then we have that

$$\begin{aligned}
(P_1, P_3)(\pi)(A_1 \times A_3) &= \pi(\{(w_1, w_2, w_3) \in W^1 \times W^2 \times W_0 : (P_1, P_3)(w_1, w_2, w_3) \in A_1 \times A_3\})\\
&= \int_{W_0}\int_{W^2}\int_{W^1} 1_{A_1\times W^2\times A_3} R(w_3, dw_1)R'(w_3, dw_2)\mu(dw_3)\\
&= \int_{A_3} R(w_3, A_1)R'(w_3, W^2)\mu(dw_3)
\end{aligned}$$

$$= \int_{A_3} R(w_3, A_1)\mu(dw_3)$$
$$= (X, W)(P)(A_1 \times A_2),$$

where we used that $R'(w_3, W^2) = 1$ for μ-almost all w_3 and the characterization of R. $\square$

Since $\mathcal{L}(X, W) = \mathcal{L}(P_1, P_3)$, and $\mathcal{L}(X', W') = \mathcal{L}(P_2, P_3)$, we have that

$$\mathcal{L}\bigg(X, W, \int_0^t \sigma(s, X_s)dW_s, \int_0^t b(s, X_s)ds\bigg)$$
$$= \mathcal{L}\bigg(P_1, P_3, \int_0^t \sigma(s, P_1)dP_3(s), \int_0^t b(s, P_1)ds\bigg),$$

and

$$\mathcal{L}\bigg(X', W', \int_0^t \sigma(s, X'_s)dW'_s, \int_0^t b(s, X'_s)ds\bigg)$$
$$= \mathcal{L}\bigg(P_2, P_3, \int_0^t \sigma(s, P_2)dP_3(s), \int_0^t b(s, P_2)ds\bigg)$$

from the exercise 5.16 of [24]. Infinite dimensional version of this statement follows in a similar vein.

Lemma 3.0.3. (P_1, P_3) *(resp.* (P_2, P_3)*) is a weak solution for the SDE* (3.1) π*-a.e. with the same initial condition as* (X, W) *(resp. with the same initial condition as* (X', W')*.)*

Proof. Define

$$S^1 = \bigg\{P_1(t) = P_1(0) + \int_0^t b(s, P_1(s))ds + \int_0^t \sigma(s, P_1)dP_3(s) \text{ for any } t \in [0, 1]\bigg\}$$

on $\mathcal{B}(W^1) \otimes \mathcal{B}(W^2) \otimes \mathcal{B}(W_0)$ and for the canonical projection processes $P'_1 : W^1 \times W_0 \to W_1$ and $P'_3 : W^1 \times W_0 \to W_0$ define

$$S = \bigg\{P'_1(t) = P'_1(0) + \int_0^t b(s, P'_1(s))ds + \int_0^t \sigma(s, P'_1(s))dP'_3(s) \text{for any } t \in [0, 1]\bigg\}$$

on $\mathcal{B}(W^1) \otimes \mathcal{B}(W^0)$. Then

$$\begin{aligned}
\pi(S^1) &= \int_{W_0}\int_{W^2}\int_{W^1} 1_{S^1}(w_1,w_2,w_3)R(w_3,dw_1)R'(w_3,dw_2)\mu(dw_3) \\
&= \int_{W_0}\int_{W^2}\int_{W^1} 1_{S\times W^2}(w_1,w_3,w_2)R(w_3,dw_1)R'(w_3,dw_2)\mu(dw_3) \\
&= P(\{(X,W)\in S\}) = (X,W)(P)(S) = 1.
\end{aligned}$$

Therefore, (P_1,P_3) solves the equation π-a.e. We can do the same for (P_2,P_3), thus the result follows. $\square$

Proposition 3.0.2. *Suppose that (X,W) and (X',W') are weak solutions to the equation (3.1) and $X_0 = X'_0$. Suppose that the equation (3.1) has the pathwise uniqueness property then we have that $\mathcal{L}(X,W) = \mathcal{L}(X',W')$.*

Proof. Follows from the previous lemma and pathwise uniqueness since $P_1(\{0\}) = P_2(\{0\})$ as $X_0 = X'_0$ implies that (P_1,P_3) and (P_2,P_3) have the same sample trajectories as they are on the very same stochastic basis. Since $\mathcal{L}(P_1,P_3) = \mathcal{L}(X,W)$, and $\mathcal{L}(P_2,P_3) = \mathcal{L}(X',W')$ we must have that $\mathcal{L}(X,W) = \mathcal{L}(X',W')$. $\square$

Proposition 3.0.3. *Existence of a weak solution and pathwise uniqueness imply that a strong solution exists.*

Proof. Suppose (X,W) and (X',W') are two weak solutions as defined before with the same initial conditions, and satisfy pathwise uniqueness. Then from the previous remark it follows that corresponding P_1 and P_2 satisfy

$$\begin{aligned}
&\pi(\{P_1 = P_2\}) = 1 \\
&= \int_{W_0}\int_{W^1}\int_{W^2} 1_{P_1=P_2} R(w_3,dw_1)R'(w_3,dw_2)\mu(dw_3) \\
&= \int_{W_0}\int_{W^1} R'(w_3,\{w_1\})R(w_3,dw_1)\mu(dw_3)
\end{aligned}$$

and this implies that

$$\int_{W^1} R'(w_3, \{w_1\}) R(w_3, dw_1) = 1$$

μ-a.e. w_3, and which in turn then implies that $R'(w_3, \{w_1\}) = 1$ for μ-a.e. w_3 and $R(w_3, \cdot)$-a.e. w_1. So there exists $F(w)$ such that

$$R(w_3, \cdot) = \delta_{F(w_3)}$$

from the definition of a regular conditional distribution. We also have that

$$\begin{aligned}
&\int_{W_0} \int_{W^2} \int_{W^1} 1_{\{P_1 = P_2\}} R(w_3, dw_1) R'(w_3, dw_2) \mu(dw_3) \\
&= \int_{W_0} \int_{W^1} \int_{W^2} 1_{\{P_1 = P_2\}} R'(w_3, dw_2) R(w_3, dw_1) \mu(dw_3)
\end{aligned}$$

hence $R(w_3, \cdot) = R'(w_3, \cdot) = \delta_{F(w_3)}$. Suppose that W'' is another cylindrical Wiener process on a stochastic basis $(\Omega'', \mathcal{F}'', P'', (\mathcal{F}''_t)_t)$. Set $X'' = F(W'')$. First observe that

$$P(\{X = F(W)\}) = \int_{W^1} \int_{W_0} 1_{\{w_1 = F(w)\}} \delta_{F(w)}(dw_1) \mu(dw_2) = 1.$$

We want to prove that (X'', W'') is a weak solution with the same initial condition as (X, W). It follows that

$$\begin{aligned}
P''(\{X_0 = X_0'\}) &= P''(\{X_0 = F(W'')(0)\}) \\
&= \int_{W^1} \int_{W_0} 1_{x = F(w)} T(dx) \otimes \mu(dw) \\
&= P(\{X_0 = F(W)\}) \\
&= 1.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
P''(\{(X'', W'') \in S\}) &= P''(\{(F(W''), W'') \in S\}) \\
&= \int_{W_0} 1_{(F(w_3), w_3) \in S} \mu(dw_3)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{W_0} 1_{(w_1,w_3)\in S}\delta_{F(w_3)}(dw_1)\mu(dw_3) \\
&= P(\{(X,W)\in S\}) = 1
\end{aligned}
$$

so (X'', W'') is a solution to the equation (3.1) with the same initial condition as (X, W). $\square$

Since the pathwise uniqueness implies that we must necessarily have a unique strong solution we have obtained the following theorem:

Theorem 3.0.2 (Yamada-Watanabe). *A unique strong solution exists to the equation* (3.1) *if and only if there exists a weak solution and the equation satisfies pathwise uniqueness.*

Chapter 4

Martingale Representation Theorem For Diffusions

We fix a probability space $(\Omega, \mathcal{F}, P)$ for the rest of the chapter. Main purpose of this chapter will be to extend the martingale representation theorem for the processes that are adapted to the filtration generated by a diffusion process X to the infinite dimensional spaces. The finite dimensional Euclidean case is done in [29]. The main components of this result include the uniqueness of the martingale problem. For this reason, we shall investigate this problem in the first section.

It is well known that when $\sigma = 0$, if we lose the condition that b is of at most linear growth or Lipschitz continuous, there might not be solution to the (ordinary) differential equation we are working with in the very same space [7]. For this reason, to make everything simpler, we shall assume that b and σ are Lipschitz continuous, and of at most linear growth. Then there exists a unique strong solution [6] [3], hence from the Yamada-Watanabe theorem it follows that there exists a weak solution and the equation has uniqueness property. It appears that it is more common to encounter weak solutions in application, for this purpose we shall assume that our solution X is a weak solution with the uniqueness in law property.

One should notice that, in the case when Lipschitz continuity or at most linear growth conditions do not hold for the coefficients, under less regular conditions solutions to the equation (3.1) can be found not on the initial space, but on a slightly expanded space, say $\tilde{H}$ [6]. However after a minor tweak in the framework the very same proofs work line by line, hence for the simplicity we will impose Lipschitz continuity and at most linear growth conditions on the coefficients of the equation (3.1). Furthermore, the proofs given here are meant to work even in the case when $\Omega = C(0,\infty)$, but for simplicity we shall keep our time horizon finite.

4.1 The Martingale problem and weak solutions

The general scheme of being able to recover a stochastic integral, with integrand that is a square root of some operator a and integrator as a Wiener process, for processes that have the quadratic variation

$$\int_0^t a(s, X_s)ds$$

will allow us to establish a density argument in the space $L^2(\mathcal{F}_1(X), \mathbb{R})$, which in turn can be generalized to any seperable Hilbert space valued, mean integrable, and adapted to the filtration generated by the diffusion X.

Although there are different methods, one way to obtain a solution for a martingale problem in terms of a solution of a stochastic differential equation in infinite dimensional setting is through the following theorem:

Theorem 4.1.1 ([3, Theorem 8.2]). *Assume that M is a continuous, square integrable martingale in H and its quadratic-variation satisfies the following:*

$$\langle\langle M_t \rangle\rangle = \int_0^t (\phi(s)Q^{1/2})(\phi(s)Q^{1/2})^* ds,$$

where ϕ is a predictable $L_2(U, H)$-valued process, and Q is a given bounded, symmetric nonnegative operator in U. Then on an extended probability space

$(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ *of* $(\Omega, \mathcal{F}, P)$

$$M_t = \int_0^t \phi(s) dW_s,$$

where W *is a* Q*-Wiener process on the extended space.*

Although the statement above restricts one to use Q-Wiener process, the proof given in [3] uses cylindrical Wiener process. With the aid of this result, we can adapt the proof given in [14] (also see [10]) to the infinite dimensional case:

Proposition 4.1.1. *The process* X *is a solution to the equation* (3.1) *if and only if for any regular* F

$$\begin{aligned} dF(t, X_t) - \Bigg(F_t(t, X_t) + &\langle F_x(t, X_t), b(s, X_s)\rangle \\ &- \frac{1}{2}\mathrm{tr}\big[F_{xx}(t, X_t)\sigma(t, X_t)(\sigma(t, X_t))^*\big]\Bigg) dt \\ &= dF(t, X_t) - OF(t, X_t)dt. \end{aligned}$$

is a local martingale.

Proof. Only if part follows from the Itô's formula, which holds since the process X is adapted so $(\sigma(s, X_s))_s$ and $(b(s, X_s))_s$ are adapted processes hence from continuity they are progressive processes, thus

$$dM_t^F = dF(t, X_t) - OF(t, X_t)dt = \langle F_x(s, X_s), \sigma(s, X_s)dW_s\rangle$$

is a local martingale.

For the "if" part, Theorem 4.1.1 gives us the infinite dimensional machinary that we need, which says that if such equation holds, there exists Wiener process such that the equation above is satisfied, perhaps after extending the probability space solves the problem as in the finite dimensional case. Let $(H_n)_n \subset H$ be a total orthonormal sequence in H. For the functions $F_{i,j} : [0, 1] \times H \to \mathbb{R}$ defined as

$$F_{i,j}(s, x) = \langle x, H_i\rangle\langle x, H_j\rangle,$$

and $F_i : [0,1] \times H \to \mathbb{R}$ defined as

$$F_i(s,x) = \langle x, H_i \rangle$$

we have that

$$\begin{aligned} M_t^{F_i} &= \int_0^t \langle H_i, \sigma(s, X_s) dW_s \rangle \\ &= \langle X_t, H_i \rangle - \langle X_0, H_i \rangle - \int_0^t \langle H_i, b(s, X_s) dW_s \rangle, \end{aligned}$$

and

$$\begin{aligned} M_t^{F_{i,j}} &= \langle X_t, H_i \rangle \langle X_t, H_j \rangle - \langle X_0, H_i \rangle \langle X_0, H_j \rangle \\ &- \int_0^t \Big(\langle X_t, H_j \rangle \langle H_i, b(s, X_s) dW_s \rangle + \langle X_t, H_i \rangle \langle H_j, b(s, X_s) dW_s \rangle \Big) \\ &- \frac{1}{2} \int_0^t \operatorname{tr}((H_i \otimes H_j + H_j \otimes H_i) \sigma(s, X_s) \sigma(s, X_s)^*) ds \\ &= \int_0^t \langle X_t, H_i \rangle \langle H_j, \sigma(s, X_s) dW_s \rangle + \langle X_t, H_j \rangle \langle X_t, \sigma(s, X_s) dW_s \rangle \end{aligned}$$

are local martingales. It follows that $M_t^{F_{i,j}} - \langle X_0, H_i \rangle M_t^{F_j} - \langle X_0, H_j \rangle M^{F^i}$ is a local martingale and that

$$\begin{aligned} & M_t^{F_{i,j}} - \langle X_0, H_i \rangle M_t^{F_j} - \langle X_0, H_j \rangle M^{F^i} \\ & \quad - \int_0^t \int_0^s \langle H_i, b(r, X_r) ds \rangle \langle H_j, dM_s^{F_i} \rangle - \int_0^t \int_0^s \langle H_j, b(r, X_r) dr \rangle \langle H_i, dM_s^{F_j} \rangle \\ & \quad = M_t^{F_i} M_t^{F_j} - \langle \int_0^t \sigma^*(s, X_s) \sigma^*(s, X_s) ds H_i, H_j \rangle \end{aligned}$$

is a local martingale. Therefore, the quadratic variation of $M_t^{F_i}$ and $M_t^{F_j}$ is $\int_0^t \sigma(s, X_s)\sigma^*(s, X_s) ds$. From Theorem 4.1.1, we obtain that X is a solution the the desired equation. □

Remark 4.1.1. *With careful inspection, the result above holds true for SDEs with coefficients that consider the whole path rather than its state at a given time.*

Corollary 4.1.1. *If under the measure P on $(C([0,1], H), \mathcal{B}(C([0,1], H)))$ for* *any regular F*

$$M_t^F = F(y(t)) - F(y(0)) - \int_0^t OF(s, y(s))ds$$

is a continuous local martingale, then there exists a weak solution (X, W) *on some stochastic basis* $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}, (\tilde{\mathcal{F}}_t)_t)$ *such that* $X(\tilde{P}) = P$.

Proof. From the previous theorem by setting $y(t) = X_t$ we can find a solution (X, W) on some stochastic basis $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}, (\tilde{\mathcal{F}}_t)_t)$. Thus, it directly follows that $P = X(\tilde{P})$. □

Corollary 4.1.2. *In case we have weak uniqueness and weak existence for the equation, the martingale problem associated to O has unique solution.*

Proof. Since there exists a weak solution, a solution to martingale problem exists. If one takes another solution to the martingale problem, the solution will generate a weak solution for the equation from the proposition above, which in turn will have the same distribution as any other weak solution to the equation from the weak uniqueness property. Thus, all the solutions of the martingale problem are identical. □

4.2 The Martingale Representation Theorem for Diffusion Processes

Suppose (X, W) is a weak solution to the equation (3.1) with a driving cylindrical Wiener process over our probability space $(\Omega, \mathcal{F}, P)$. Since we have assumed Lipschitz continuity and at most linear growth for the coefficients of (3.1), we will have a unqiue strong solution hence, from the Yamada-Watanabe theorem law of the solution (X, W) is unique. Therefore from the previous section, we have that if the measure Q solves the associated martingale problem, every other measure that solves it must be equal to Q as $\mathcal{L}(X)$ is unique.

Let $\mathcal{F}(X) = (\mathcal{F}_t(X))_{t\in[0,1]}$ be the completion of the filtration generated by X. We say that a process L belongs to the set C if L is $L_2(H,\mathbb{R})$-valued, adapted to the filtration $\mathcal{F}(X)$, and

$$E\left[\int_0^1 \operatorname{tr}(L(s,X)\sigma(s,X))(L(s,X)\sigma(s,X))^* ds\right] < \infty.$$

We will first establish the density of potential martingales over the martingale space over the martingale part of the diffusion.

Theorem 4.2.1 ([27, Theorem 1])**.** *The set*

$$\Gamma = \{N \in L^2(\mathcal{F}_1(X)) : N = E[N] + \int_0^1 L(s,X)\sigma(s,X)dW_s, L \in C\}$$

is total in $L^2(\mathcal{F}_1(X);\mathbb{R})$.

Proof. Suppose there exists a random variable M such that $M \in L^2(\mathcal{F}_1(X)) \cap \Gamma^\perp$ and define $M_t = E[M|\mathcal{F}_t(X)]$. Without loss of generality suppose $E[M] = 0$. For the process

$$\tilde{M}_t = 1 + \frac{M_{t\wedge T_n}}{n},$$

where $T_n = \inf\{|M_t| \geq n\}$, we have that $\tilde{M} - 1 \perp \Gamma$. Using the fact that all martingales of $\mathcal{F}(X)$ is continuous, $dQ = \tilde{M}_1 dP$ gives us an equivalent measure to P. Then for any regular F, $\tilde{M}\big(F(X) - OF(X)ds\big)$ is a martingale according the the proposition above from a characterization of strong orthogonality and weak orthogonality [23, Lemma IV.2]. Uniqueness in law implies that X is a solution to the equation (1) under the measures $\tilde{M}P$ and P. It follows that $M_{t\wedge T_n} = 0$ a.s. for each n and passing to the limit it follows that $M = 0$ hence, Γ is total in $L^2(\mathcal{F}_1(X),\mathbb{R})$. □

As a consequence of the Itô isometry, it is possible to obtain the full characterization of the space $L^2(\mathcal{F}_1(X);K)$:

Theorem 4.2.2. *Let* $P(X) = (P_s(X))_{s\in[0,1]}$ *be a measurable version of the orthogonal projection from* U *onto* $\ker(\sigma(s,X))^\perp$ *and let* $F \in L^2(\mathcal{F}_1(X);K)$ *be any* *random variable with zero expectation in some separable Hilbert space K. Then there exists a process $\xi(X) \in L^2(dt \times dP; L_2(U, K))$ adapted to $\mathcal{F}(X)$, such that*

$$F(X) = \int_0^1 \xi_s(X) P_s(X) dW_s.$$

Proof. To make use of the density of the set C, we will start with the case $K = \mathbb{R}$. For any $F(X) \in L^2(\mathcal{F}_1(X), \mathbb{R})$ there exists $(F_n(X)) \subset \Gamma$ that converges to $F(X)$ in $L^2(\mathcal{F}_1(X))$. Without loss of generality suppose that $EF_n(X) = 0$ for all n. Since $(F_n(X))$ is convergent in $L^2(\mathcal{F}_1(X))$ we have that integrands of the representation of $F_n(X)$, say $L^n(s, X)\sigma(s, (X)$, convergent in $dt \times dP$ due to Itô isometry. It follows that $L^n(s, X)\sigma(s, X) = L^n(s, X)\sigma(s, X)P_s(X)$. So, $L^n(s, X)\sigma(s, X)P_s(X)$ converges to some $\xi_s(X)P_s(X)$ in $L^2(dt \times dP, L_2(U, \mathbb{R}))$ as $L^n(s, X)\sigma(s, X)$ converges to $\xi_s(X)$, since $P_s(X)$ is an orthogonal projection

Let K be a separable Hilbert space. Suppose $F(X) \in L^2(\mathcal{F}_1(X), K)$. Let $\{e_n\}_n$ be an orthonormal basis of K, Then $\langle F(X), e_n\rangle$ is in $L^2(\mathcal{F}_1(X), \mathbb{R})$ for each n. However from above, each $\langle F(X), e_n\rangle$ can be represented as $\int_0^1 \xi_s^n(X)P_s(X)dW_s$ for some $L^2(dt \times dP, L_2(U, \mathbb{R}))$ function $\xi^n(X)$ that is adapted to the filtration $\mathcal{F}(X)$. Thus, in one hand we have that

$$\begin{aligned} F(X) &= \sum_{n=1}^{\infty} \langle F(X), e_n\rangle e_n \\ &= \sum_{n=1}^{\infty} \left(\int_0^1 \xi_s^n(X)P_s(X)dW_s \right) e_n, \end{aligned}$$

for $\xi^n(X)$ described as above. On the other hand, since $E\|F(X)\|_K^2 < \infty$, we have that

$$\begin{aligned} E\|F(X)\|_K^2 &= E \sum_{n=1}^{\infty} \int_0^1 \|\xi_s^n(X)P_s(X)\|_{L_2(U,\mathbb{R})}^2 ds \\ &= E \int_0^1 \| \sum_{n=1}^{\infty} e_n \otimes (\xi_s^n(X)P_s(X))\|_{L_2(U,K)}^2 ds \\ &= E \int_0^1 \|\xi_s(X)P_s(X)\|_{L_2(U,K)}^2 ds \end{aligned}$$

where we define $\xi_s(X)P_s(X) = \sum_{n=1}^{\infty} e_n \otimes \xi_s^n(X)P_s(X)$, and $\xi_s(X) = \sum_{n=1}^{\infty} e_n \otimes \xi_s^n(X)$. Consequently, it also follows that $E\int_0^1 \|\xi_s(X)P_s(X)\|^2_{L_2(U,K)}ds < \infty$, and $E\|F(X)\|_K^2 = E\int_0^1 \|\xi_s(X)P_s(X)\|^2_{L_2(U,K)}ds$ which then implies that $F(X)$ reads as

$$F(X) = \int_0^1 \xi_s(X)P_s(X)dW_s$$

from the dominated convergence theorem, which is the desired form that we want for $F(X)$. □

Remark 4.2.1. *Under the bilinear form $(\cdot,\cdot)_K : (K\hat{\otimes}U) \times U \to K$, where $\hat{\otimes}$ is the completed Hilbert-Schmidt product and $(\cdot,\cdot)_K$ defined in obvious way, the representation reads as follows: For any $F(X) \in L^2(\mathcal{F}_1(X), K)$, there exists a process $\xi(X) \in L^2(dt \times dP, L_2(U,K))$ such that*

$$F(X) = EF(X) + \int_0^1 (\xi_s(X), P_s(X)dW_s)_K,$$

from the theorems we have just established. In particular, when $K = \mathbb{R}$, this is exactly the same form given in the Euclidean case since we can associate any elements of $L_2(U,\mathbb{R})$ with appropriate elements of U via Riesz representation theorem, we can treat the process $\xi(X)$ as an element of U, hence

$$F(X) = EF(X) + \int_0^1 \langle P_s(X)\xi_s(X), dW_s\rangle_U.$$

Remark 4.2.2. *As a consequence of the Jacod-Yor theorem [12] (also see [23, Theorem IV.57]), one can actually obtain unique martingale representation for extremal solutions of the associated martingale problem [25], which is the necessary and sufficient condition for the existence and uniqueness of such martingale representation theorems.*

Corollary 4.2.1. *Suppose that driving Wiener process of X is a Q-Wiener process in U. Let $F \in L^2(\mathcal{F}_1(X), K)$. Then there exists ξ such that $\xi Q^{1/2} \in L^2(dt \times dP; L_2(U,K))$, and P-a.s. we have*

$$F = E[F] + \int_0^1 \xi_s P_s(X)dW_s.$$

Proof. W is a cylindrical process on $Q^{1/2}H$, with the inner product induced by that of H thus, the result follows from the cylindrical case. $\square$

It turns out that all the stochastic integrals of the form

$$\int_0^1 \xi_s(X)P_s(X)dW_s$$

belong to the space $L^2(\mathcal{F}_1(X);K)$:

Corollary 4.2.2. *Suppose $\xi(X)$ is $\mathcal{F}(X)$ adapted, K valued measurable process such that*

$$E\int_0^1 \operatorname{tr}\big((\xi_s(X)P_s(X))(\xi_s(X)P_s(X))^*\big)ds < \infty.$$

Then

$$T = \int_0^1 \xi_s(X)P_s(X)dW_s$$

is a $\mathcal{F}_1(X)$-measurable, K valued random variable.

Proof. Suppose not. Let $K = \mathbb{R}$. Then $T - E[T|\mathcal{F}_1(X)] \perp L^2(\mathcal{F}_1(X), \mathbb{R})$. However, there exist $\zeta(X) \in L^2(dt \times dP, L_2(H, \mathbb{R})$ such that

$$E[T|\mathcal{F}_1(X)] = \int_0^1 \zeta_s(X)P_s(X)dW_s,$$

hence setting $L_s(X) = \xi_s(X) - \eta_s(X)$, we have that

$$E[\int_0^1 L_s(X)P_s(X)dW_s \int_0^1 h_s(X)\sigma(s,X)dW_s] = 0$$

for any $h \in C$. So $L_s(X)P_s(X) = 0$ $dt \times dP$-a.e. hence $\xi_s(X)P_s(X) = \zeta_s(X)P_s(X)$ $dt \times dP$-a.e.

Going back to the general case, for a complete orthonormal basis $(e_n)_n$ of K, we can write

$$T = \sum_{n=1}^{\infty} \int_0^1 \xi_s^n(X)P_s(X)dW_s e_n$$

and

$$E[T|\mathcal{F}_1(X)] = \sum_{n=1}^{\infty} \int_0^1 \zeta_s^n(X) P_s(X) dW_s e_n.$$

Repeating the same argument as above, we see that each coordinate is zero for $T - E[T|\mathcal{F}_1(X)]$, thus we are done. □

We are set to prove the interplay between the conditional expectation with respect to $\mathcal{F}_1(X)$ and stochastic integrals with driving Wiener process W with integrands that are adapted to the filtration generated by W. This result was first shown by Elworthy-Li [5] by using pure geometric arguments when the diffusion itself is a gradient-Brownian system then extended by Ustunel [27] to arbitrary weak solution of an equation with weak uniqueness property.

Theorem 4.2.3. *Let h be an adapted process to the filtration generated by W, square integrable, and $L_2(U, K)$ valued. Then we have that*

$$E\left[\int_0^1 h_s dW_s \middle| \mathcal{F}_1(X)\right] = \int_0^1 E[h_s|\mathcal{F}_s(X)] P_s(X) dW_s.$$

Furthermore, if $h_s = \sum_i \lambda_i(s) k_i \otimes u_i$ for some basis $(k_i \otimes u_i)_i$ of $L_2(U, K)$, then

$$E\left[\int_0^1 h_s dW_s \middle| \mathcal{F}_1(X)\right] = \sum_i \int_0^1 E[\lambda_i(s)|\mathcal{F}_s(X)] \langle P_s(X) u_i, dW_s \rangle k_i.$$

Proof. Let $(e_n)_n$ be a complete orthonormal basis of H. There exists $\xi(X) \in L^2(\mathcal{F}_1(X); \mathbb{R})$ adapted to $\mathcal{F}(X)$ such that

$$E[\int_0^1 h_s dW_s|\mathcal{F}_1(X)] = \int_0^1 \xi_s(X) P_s(X) dW_s.$$

Pick $F(X) \in L^2(\mathcal{F}_1(X); K)$, then by utilizing the martingale representation theorem for $F(X)$, we have that

$$F(X) = \int_0^1 \theta_s(X) P_s(X) dW_s$$

for some $\theta \in L^2(dt \times d\mu; L_2(U, K))$ and adapted to the filtration $\mathcal{F}(X)$. First observe that from the definition of inner product on $L_2(U, K)$ we have

$$\begin{aligned}
\langle h_s, \xi_s(X)P_s(X)\rangle_{L_2(U,K)} &= \sum \langle h_s e_i, \theta_s(X)P_s(X)e_i\rangle \\
&= \sum \langle h_s P_s(X)e_i, \theta_s(X)P_s(X)e_i\rangle \\
&= \langle h_s P_s(X), \theta_s(X)P_s(X)\rangle_{L_2(U,K)}.
\end{aligned}$$

Furthermore, using Itô isometry, in one hand we get

$$\begin{aligned}
E\int_0^1 h_s dW_s F(X) &= E\int_0^1 \langle h_s, \theta_s(X)P_s(X)\rangle_{L_2(U,K)} ds \\
&= E\int_0^1 \langle E[h_s P_s(X)|\mathcal{F}_s(X)], \theta_s(X)P_s(X)\rangle_{L_2(U,K)} ds.
\end{aligned}$$

On the other hand, we have that

$$\begin{aligned}
E\int_0^1 h_s dW_s F(X) &= E\int_0^1 \xi_s(X)P_s(X) dW_s F(X) \\
&= E\int_0^1 \langle \xi_s(X)P_s(X), \theta_s(X)P_s(X)\rangle_{L_2(U,K)} ds,
\end{aligned}$$

hence $\xi_s(X)P_s(X) = E[h_s|\mathcal{F}_s(X)]P_s(X)]$ $dt \times dP$-a.e. and as a consequence of this, we see that

$$\int_0^1 E[h_s|\mathcal{F}_s(X)]P_s(X) dW_s$$

is $\mathcal{F}_1(X)$ measurable. From the uniqueness of the martingale representation, we then have that

$$E\left[\int_0^1 h_s dW_s \middle| F_1(X)\right] = \int_0^1 E[h_s|\mathcal{F}_s(X)]P_s(X) dW_s.$$

To prove the general case, let $(k_i \otimes u_i)_i$ be a complete orthonormal basis of $L_2(U, K)$ such that $h_s = \sum_i \lambda_i(s) k_i \otimes u_i$. Since

$$\int_0^1 h_s dW_s = \int_0^1 \sum_i \lambda_i(s) k_i \otimes u_i dW_s = \sum_i \int_0^1 \lambda_i(s)\langle u_i, dW_s\rangle k_i$$

and for any i,

$$E\left[\int_0^1 \lambda_i(s)\langle u_i, dW_s\rangle \middle| \mathcal{F}_1(X)\right] = \int_0^1 \langle E[\lambda_i(s)u_i|\mathcal{F}_s(X)], P_s(X)dW_s\rangle$$

as above; from the dominated convergence theorem we have that

$$E\left[\int_0^1 \sum_i \lambda_i(s)k_i \otimes u_i dW_s \middle| \mathcal{F}_1(X)\right] = \sum_i \int_0^1 \langle P_s(X)E[\lambda_i(s)|\mathcal{F}_s(X)]u_i, dW_s\rangle k_i.$$

Since λ_i is a real valued process, we can take it out the inner product, hence the result follows. □

Corollary 4.2.3. *Let h be an adapted process to the filtration generated by W such that it is square integrable, U valued. Then we have that*

$$E\left[\int_0^1 \langle h_s, dW_s\rangle \middle| \mathcal{F}_1(X)\right] = \int_0^1 \langle E[h_s|\mathcal{F}_s(X)], P_s(X)dW_s\rangle.$$

Proof. Follows from the previous theorem after realizing $\langle h_s, \cdot\rangle$ as an operator from Riesz representation theorem. □

Corollary 4.2.4. *Let h be an adapted process to the filtration generated by W such that it is square integrable, $L_2(U,U)$ valued, and for any $s \in [0,1]$, $\ker(\sigma(s,X))^\perp$ is h_s invariant. Then it follows that*

$$E\left[\int_0^1 h_s dW_s \middle| \mathcal{F}_1(X)\right] = \int_0^1 E[P_s(X)h_s|\mathcal{F}_s(X)]dW_s. \tag{4.1}$$

Proof. By the invariance assumption we have that $P_s(X)h_s e_i = h_s P_s(X)e_i$, hence the result follows from the previous theorem. □

Corollary 4.2.5. *Suppose that $M = (M_t, t \in [0,1])$ defined by*

$$M_t = \exp\left(\int_0^t \langle h_s, dW_s\rangle_U - \frac{1}{2}\int_0^t \|h_s\|_U^2 ds\right),$$

P-a.s. for $h \in L^2([0,1], U)$ that is deterministic. Then it holds that

$$E[M|\mathcal{F}_t(X)] = \exp\left(\int_0^t \langle h_s, P_s(X)dW_s\rangle_U - \frac{1}{2}\int_0^t \|P_s(X)h_s\|_U^2 ds\right).$$

Proof. After applying the Itô's formula to the process M, we obtain that

$$M = 1 + \int_0^1 M_s \langle h_s, dW_s \rangle_U.$$

By realizing $\langle h_s, \cdot \rangle$ as a linear operator $\tilde{h}_s \cdot$ via the Riesz representation theorem, we obtain that $(M_s \tilde{h}_s, s \in [0,1]) \subset L^2([0,1], L_2(U, \mathbb{R}))$, hence applying the previous theorem we obtain following relations

$$\begin{aligned} E[M|\mathcal{F}_t(X)] &= 1 + \int_0^t E[M_s \tilde{h}_s | \mathcal{F}_s(X)] P_s(X) dW_s \\ &= 1 + \int_0^t E[M_s | \mathcal{F}_s(X)] \tilde{h}_s P_s(X) dW_s \\ &= \exp\left(\int_0^t \langle h_s, P_s(X) dW_s \rangle_U - \frac{1}{2} \int_0^t \| P_s(X) h_s \|_U^2 ds \right), \end{aligned}$$

from Itô's formula. □

Remark 4.2.3. *We have also demonstrated that*

$$E[M|\mathcal{F}_t(X)] = 1 + \int_0^t E[M|\mathcal{F}_s(X)] \tilde{h}_s P_s(X) dW_s,$$

which is the corresponding martingale representation for the martingale $(E[M|\mathcal{F}_t(X)])_{t \in [0,1]}$.

Chapter 5

Calculations Related to Entropy

Suppose that $L(X) \geq 0$ is $\mathcal{F}_1(X)$ measurable, $E[L(X)] = 1$, and $E[L(X)\log L(X)] < \infty$. Suppose $U = W + u$ is an adapted perturbation of the cylindrical Wiener process W such that

$$L(X) = \frac{dX^U(P)}{dX(P)} \circ X,$$

where

$$dX^U = \sigma(s, X_s^U)d(W_s + \int_0^s \dot{u}_s ds) + b(s, X_s^U)ds,$$

and $u = (u_s)$ is defined in $U_0 = (U)$ so that $u_s = \int_0^s \langle \dot{u}_r, dr \rangle$.

Further assume that $L(X)$ is strictly positive P-a.s. If $E\rho(-\delta u) = 1$, for any real valued continuous bounded function f, we have that:

$$\begin{aligned}
E[f(X^U)] &= \int_\Omega f(X^U)dP \\
&= \int_H f dX^U(P) \\
&= \int_H f L dX(P) \\
&= \int_\Omega f(X)L(X)dP
\end{aligned}$$

$$= \int_\Omega f(X^U)L(X^U)\rho(-\delta u)dP$$
$$= E[f(X^U)L(X^U)E[\rho(-\delta u)|\mathcal{F}_1(X)]]$$

hence $L(X^U)E[\rho(-\delta u)|\mathcal{F}_1(X)] = 1$ P-a.s. Thus,

$$\log L(X^U) = -\log E[\rho(-\delta u)|\mathcal{F}_1(X)].$$

As an application of the martingale representation theorem, in the setting above we can represent $L(X)$ as

$$L(X) = \exp\left(-\int_0^1 \langle v_s(X), P_s(X)dW_s\rangle - \frac{1}{2}\int_0^1 |P_s(X)v_s(X)|^2 ds\right),$$

where $(P_s(X)v_s(X))_s$ is adapted to the filtration of X and $\int_0^1 |P_s(X)\dot{v}_s|^2 ds < \infty$ P-a.s. and v is U valued random variable, which can be obtained after using Riesz representation theorem after the martingale representation theorem, as L is real valued. Using the fact that U is a Brownian motion under $\rho(-\delta u)dP$, it follows that

$$L(X^U) = \exp\left(-\int_0^1 \langle v_s(X^U), P_s(X^U)dU\rangle - \frac{1}{2}\int_0^1 |P_s(X^U)v_s(X^U)|^2 ds\right). \quad (5.1)$$

Using the martingale representation theorem on $E[\rho(-\delta u)|\mathcal{F}_1(X)]$ and multiplying it with $L(X^U)$, since $L(X^U)E[\rho(-\delta u)|\mathcal{F}_1(X^U)] = 1$, we have that

$$\int_0^1 (E[\dot{u}_s|\mathcal{F}_s(X^U)] + \dot{v}_s(X^U))P_s(X)dU_s$$
$$= \frac{1}{2}\int_0^1 |(E[\dot{u}_s|\mathcal{F}_s(X^U)] + v_s(X^U))P_s(X)|^2)ds = 0$$

thus $(E[\dot{u}_s|\mathcal{F}_s(X^U)] + v_s(X^U))P_s(X) = 0$ $ds \times dP$ almost surely.

Theorem 5.0.1. *Under the settings above, there exists* $\dot{u} \in L^2(ds \times dP)$ *with* $E[\rho(-\delta u)] = 1$ *with*

$$\frac{dX^U(P)}{dX(P)} = L$$

if and only if

$$P_s(X)(E[\dot{u}_s|\mathcal{F}_s(X^U)] + v_s(X^U)) = 0. \tag{5.2}$$

Furthermore

$$H(X^U(P)|X(P)) = \frac{1}{2}E\int_0^1 |P_s(X^U)E[\dot{u}_s|\mathcal{F}_s(X^U)]|^2 ds = \frac{1}{2}E\int_0^1 |P_s(X^U)\dot{v}_s(X^U)|^2 ds.$$

Proof. First part directly follows from the discussion above. For the second part, we have the following calculation:

$$\begin{aligned}
H(X^U(P)|X(P)) &= \int_H \log \frac{dX^U(P)}{dX(P)} dX^U(P) \\
&= \int_\Omega \log \frac{dX^U(P)}{dX(P)} \circ X^U(P) dP \\
&= \int_\Omega \log L(X^U) dP \\
&= \frac{1}{2}E\int_0^1 |\dot{v}_s(X^U)P_s(X)|^2 ds \\
&= \frac{1}{2}E\int_0^1 |E[\dot{u}_s|\mathcal{F}_s(X)]P_s(X)|^2 ds,
\end{aligned}$$

where the last equality comes from the equation Theorem 4.2.3. $\square$

Theorem 5.0.2. *Suppose that L and u are given as above and*

$$H(X^U(P)|X(P)) = \frac{1}{2}E\int_0^1 |P_s(X^U)\dot{u}_s|_o^2 ds. \tag{5.3}$$

Then we have the following:

$$P_s(X^U)dU_s + P_s(X^U)\dot{v}_s(X^U)ds = P_s(X^U)dB_s \tag{5.4}$$

P-a.e. In particular, X^U satisfies the following SDE:

$$dX_t^U = \sigma(t, X^U)(dB_t - \dot{v}_t(X^U)dt) + b(t, X^U)dt, \tag{5.5}$$

with the same initial condition as X.

Proof. Since

$$H(X^U(P)|X(P)) = \frac{1}{2}\int_0^1 |P_s(X^U)E[\dot{u}_s|\mathcal{F}_s(X^U)]|_o^2 ds,$$

we must have that $(P_s(X^U)\dot{u}_s, s \in [0,1])$ is adapted to the filtration $(\mathcal{F}_s(X^U), s \in [0,1])$. As a result, it holds that then

$$P_t(X^U)(\dot{v}_t(X^U) + \dot{u}_t) = 0,$$

hence, equation (5) holds. Using the fact that $\sigma(t, X^U)\eta = \sigma(t, X^U)P_t(X^U)\eta$, we obtain that

$$\begin{aligned}
dX_t^U &= \sigma(t, X^U)(dB_t + \dot{u}_t dt) + b(t, X^U)dt \\
&= \sigma(t, X^U)(dB_t + P_t(X^U)\dot{u}_t dt) + b(t, X^U)dt \\
&= \sigma(t, X^U)(dB_t - P_t(X^U)\dot{v}_t(X^U)dt) + b(t, X^U)dt \\
&= \sigma(t, X^U)(dB_t - \dot{v}_t(X^U)dt) + b(t, X^U)dt.
\end{aligned}$$

□

Theorem 5.0.3. *Assume that $u \in L_a^2(dt \times dP, U_0)$ and denote by U the process $W + u$, and suppose the Lipschitz condition on the drift and diffusion coefficients hold. Then we must have that*

$$H(X^U(P)|X(P)) \leq \frac{1}{2}E\int_0^1 |P_s(X)E[\dot{u}_s|\mathcal{F}_s(X^U)]|_o^2 ds.$$

Proof. If $u \in L_a^\infty(dt \times dP, U_0)$ then the result must hold from what we have done above, with equality. In case $u \in L_a^2(dt \times dP, U_0)$, we will bound it by using the stopping times $T_n = \inf\left\{t > 0 : \int_0^t |\dot{u}_s|^2 ds > n\right\}$, and define the stopped process by

$$u^n(t) = \int_0^t 1_{[0,T_n]}(s)\dot{u}_s ds.$$

From the definition of T_n it follows that $u^n \in L_a^\infty(dt \times dP, U_0)$ hence the result holds, with inequality as in the case described initially for not X^U but X^{U^n}.

Since Lipschitz condition implies strong uniqueness, we must have that up to time T_n, X^{U_n} and X^U agree almost surely, therefore for any continuous bounded real valued function f that takes inputs from H, we must have that

$$\int f X^{U^n} dP \to \int f X^U dP.$$

Thus, $X^{U^n}(P)$ must converge weakly to $X^U(P)$ from the Portmanteau theorem. Since entropy is weak lower semi-continuous,

$$H(X^U(P)|X(P)) \leq \liminf_n H(X^{U^n}(P)|X(P)).$$

From the dominated convergence theorem we can interchange the integral and the limit, which in turn yields the desired result. $\square$

Chapter 6

Conditional Malliavin calculus for diffusions in infinite dimensional spaces

Aim of this chapter will be the construction of Malliavin-type Sobolev operators that are related to the diffusions in infinite dimensional spaces. To this end, we will narrow our focus to Itô-type diffusions, i.e. diffusions with time homogeneous drift and diffusion coefficient, and extend the domain of the coefficients to the whole space rather than path spaces. Thus, we will be working with deterministic coefficients now. We further assume that the diffusion coefficient and drift are C^1-Lipschitz, and of linear growth, which are well-known conditions for our equations to give unique strong solution, which we will use rather than weak solution this whole chapter to utilize the classical differentiation with respect to the driving Wiener process and carry out calculationg from there.

We start with the weak differentiability of diffusion processes in our case, since we could now find any reference to it. The assumptions of Lipschitz continuity and linear growth are necessary to obtain this result since the result depends heavily on the chain rule which is only available under such conditions. Then we carry out the conditional Malliavin differentiability scheme presented in [27] in infinite dimensional setting for separable Hilbert spaces, including the space of Hilbert-Schmidt operators. Then we will end this chapter by carrying out an extrinsic Malliavin differentiability scheme for operators that are linear and bounded.

6.1 Weak derivative of a diffusion process in Hilbert spaces

Weak differentiability of diffusion processes in infinite dimensions is proved under only some higher order weak differentiability conditions for random coefficients that only depend on time in [8] and [17]. In this section we will do it under the Lipschitz continuous or at most linear growth conditions for the drift and diffusion coefficients that depend on the state of the diffusion by the chain rule.

Main ingredients of this result actually stems from the very same tools and techniques used in the finite dimensional case, which can be found in [10] and [19]. We shall consider sufficiently smooth approximation of the solution, which is possible since the solution will be unique due to the restrictions on the coefficients, and apply the tools we have developed regarding to the interplay between the Malliavin derivative and Skorohod integral operators, and the chain rule, which our coefficient have sufficient conditions to apply them.

Proposition 6.1.1 (Differentiability of a diffusion processes)**.** *The unique strong solution, X, of the equation* (1) *belongs to $D_{\infty,1}(H)$.*

Proof. Since we are considering unique strong solutions, we can approximate them (uniformly) by using the Picard iteration as in the finite dimensional case. So we shall consider the Picard sequence of X :

$$X^0 = x \in H$$
$$X_t^{n+1} = x + \int_0^t b(s, X_s^n)ds + \int_0^t \sigma(s, X_s^n)dW_s$$

for $n \geq 1$.

It is clear that $X^0 \in \mathbb{D}_{\infty,1}(H)$, so suppose that $X^n \in \mathbb{D}_{\infty,1}$. From the chain rule we have established, it follows that $b(t, X_t^n)$ and $\sigma(t, X_t^n)$ belong to $D_{\infty,1}$ from the properties of b and σ. Hereby, $\int_0^t b(s, X_s^n)ds$ and $\int_0^t \sigma(s, X_s^n)dW_s$ belong to $D_{2,1}$ and from the interchangeability of Malliavin derivative with the Skorohod integral, and the Bochner integral we have the relations:

$$\nabla_h \int_0^t b(s, X_s^n)ds = \int_0^t \nabla_h b(s, X_s^n)ds,$$
$$\nabla_h \int_0^t \sigma(s, X_s^n)dW_s = \int_0^t \nabla_h \sigma(s, X_s^n)dW_s + \int_0^t \sigma(s, X_s)\dot{h}_s ds,$$

which in turn implies that $X^{n+1} \in \mathbb{D}_{\infty,1}$. Since $X^n \to X$ uniformly, if we can show that ∇X^n is uniformly bounded, from the closed graph theorem we will obtain that $\nabla X^n \to \nabla X$, which in turn will imply that ∇X exists. However, using the approximation scheme, and the chain rule gives $b(s, X_s^n)$ and $\sigma(s, X_s^n)$ have uniformly bounded Malliavin derivatives, which in turn implies that Malliavin derivatives of the corresponding integrals are uniformly bounded. Thus, the result follows. $\square$

Remark 6.1.1. *It is actually possible to lift off the Frechet differentiable condition on the drift and diffusion coefficient as in the finite dimensional case since we are working with globally Lipschitz conditions with at most linear growth.*

One loses the measurability with respect to the original filtration generated by the process after the Malliavin derivative. A natural solution to this case is to take conditional expectation with respect to the terminal σ algebra generated by such process. However, one should notice that, even if one takes conditional expectation with respect to the terminal σ-algebra generated by the process, the subsequent Malliavin derivatives will still cause the very same problem.

Theorem 6.1.1. *For any deterministic $h \in L^2([0,1], L_2(U, \mathbb{R}))$, the operator $\hat{\nabla}$ defined as $\hat{\nabla}_h X_t := E[\nabla_h X_t | \mathcal{F}_1(X)]$ satisfies the following relation:*

$$\hat{\nabla}_h X_t = \int_0^t \sigma'(X_s)\hat{\nabla}_h X_s P_s(X) dW_s + \int_0^t b'(X_s)\hat{\nabla}_h X_s ds + \int_0^t \sigma(X_s)\dot{h}_s ds.$$

Proof. Follows from the Malliavin differentiability of X after taking conditional expectation. $\square$

6.2 Conditional derivative of functionals with respect to a diffusion

By the chain rule, and the fact that diffusion processes are differentiable, we will extend the notion of Malliavin derivative for functionals of a diffusion processes in infinite dimensional setting. Main difference between our approach and the approach in [29] is that, we will consider Hilbert space valued functions rather than real valued functions.

Definition 6.2.1. *By $\mathcal{S}(X)$ we denote the set of cylindrical functions with respect to the diffusion X:*

$$\mathcal{S}(X) = \left\{ f(X) = \sum_{i=1}^{n} f_i(\langle X_{t_1}, H_1\rangle, \cdots, \langle X_{t_n}, H_n\rangle)\tilde{H}_i : 0 \le t_1 \le \cdots \le t_k, f \in \mathcal{S}(\mathbb{R}^n), k \ge 1 \right\}$$

where $\mathcal{S}(\mathbb{R}^n)$ is the space of rapidly decreasing smooth functions, and $(H_i)_i$ and $(\tilde{H}_i)_i$ are orthonormal bases of H.

Similar to the original Malliavin derivative, by using the space of cylindrical functions, with the twist of considering projections of the diffusion process as input, we define the conditional Malliavin derivative:

Definition 6.2.2. *We call the operator*

$$\hat{\nabla} : \mathcal{S}(X) \subset L^p(\mu, H) \to L^2(\mu, L^2([0,1], L_2(U, H)))$$

as the conditional Malliavin derivative with respect to the diffusion X and for any $f(X) \in \mathcal{S}(X)$ and $h \in L^2([0,1], L_2(U, \mathbb{R}))$ it is defined as

$$\hat{\nabla}_h f(X) = \sum_{i=1}^{k} \sum_{j=1}^{n} \partial_i f_j(\langle X_{t_1}, H_1\rangle, \cdots, \langle X_{t_n}, H_n\rangle)\tilde{H}_i \otimes \hat{\nabla}_h \langle X_{t_i}, H_i\rangle$$

via the chain rule. Thereby, for any function $f(X) \in \mathcal{S}(X)$*, action of* $\hat{\nabla}$ *on* $f(X)$*,* $\hat{\nabla} f(X)$*, is defined as the vector that satisfies the relation*

$$\hat{\nabla}_h f(X) = \langle \hat{\nabla} f(X), h \rangle$$

for any $h \in L^2([0,1], L_2(U,\mathbb{R}))$*, which has to be understood by the means of the natural bilinear form rather than inner product. Hence,* $\hat{\nabla} f(X)$ *reads as*

$$\hat{\nabla} f(X) = \sum_{i=1}^{\infty} \hat{\nabla}_{e_i} f(X) \otimes e_i,$$

where $(e_i)_i \subset L^2([0,1], L_2(U,\mathbb{R}))$*. In turn, we define*

$$\hat{\nabla}_t f(X) = \hat{\nabla} f(X)(t) = \sum_{i=1}^{\infty} \hat{\nabla}_{e_i} f(X) \otimes e_i(t).$$

Remark 6.2.1. *From the well-definedness of the Malliavin derivative, it follows that* $\hat{\nabla}$ *is also well defined, i.e. the conditional Malliavin derivative does not depend on the representation of the functional.*

From the duality between ∇ and $\hat{\nabla}$, and from the closability of ∇, it follows that $\hat{\nabla}$ is also closable:

Corollary 6.2.1. *Let* $h \in L^2([0,1], L_2(U,\mathbb{R}))$*, and suppose we have a sequence* $(F_n(X))_n \subset \mathcal{S}(X)$ *that converges to zero in* $L^2(\mathcal{F}_1(X))$ *and that* $(\hat{\nabla}_h F_n(X))_n$ *Cauchy in* $L^2(\mathcal{F}_1(X))$*, then we have that*

$$\lim_{n\to\infty} \hat{\nabla}_h F_n(X) = 0$$

μ*-a.s., i.e.* $\hat{\nabla}_h$ *is a closable operator in* $L^2(X(\mu))$*.*

Proof. Suppose $G(X) \in \mathcal{S}(X)$ and let $\lim_n \hat{\nabla}_h F_n(X) = K(X)$, then we have that

$$\begin{aligned} E[K(X)G(X)] &= E[\lim_n \hat{\nabla}_h F_n(X) G(X)] \\ &= \lim_n E[\hat{\nabla}_h F_n(X) G(X)] \end{aligned}$$

$$
\begin{aligned}
&= \lim_n E[\nabla_h F_n(X)G(X)] \\
&= \lim_n E[F_n(X)(-\nabla_h G(X) + G(X)\delta(h))] \\
&= \lim_n E[F_n(X)(-\hat{\nabla}_h G(X) + G(X)\delta(\mathbb{P}(X)h))] \\
&= 0.
\end{aligned}
$$

□

It follows that $\hat{\nabla}$ is also a closable operator, and by abusing the notation we shall denote the closure as $\hat{\nabla}$ too. By $\mathbb{M}_{2,1}$ we will denote the completion of the space $\mathcal{S}(X)$ with respect to the norm

$$\|F(X)\|_{2,1} = \|\hat{\nabla} F(X)\|_{L^2(\mu, L_2(U,H))} + \|F(X)\|_{L^2(\mu)}.$$

Definition 6.2.3. *For a separable Hilbert space K, we define the space $\mathcal{S}(X, K)$ as the functions that are $\mathcal{F}_1(X)$ measurable and of the form*

$$\sum_{i=1}^{n} f_i(\langle X_{t_1}, H_1\rangle, \cdots, \langle X_{t_n}, H_n\rangle)K_i$$

where $f_i \in \mathcal{S}(\mathbb{R}^n)$ and $K_i \in K$ for finite n.

With the notation above $\mathcal{S}(X)$ reads as $\mathcal{S}(X, H)$, and it is possible to define everything with respect to an arbitrary separable Hilbert space K on $\mathcal{S}(X, K)$. We further remark that, due to its nature, what we can establish in the special case $\mathcal{S}(X, \mathbb{R})$ for finite dimensional setting carries over in infinite dimensional setting, mutatis mutandis, as in [29]. For the brevity, we will keep our focus on $\mathcal{S}(X)$.

Definition 6.2.4. *The adjoint of $\hat{\nabla}$ with respect to the measure $X(\mu)$ will be denoted as*

$$\hat{\delta} : dom\hat{\delta} \subset L^2(\mu, L^2([0,1], L_2(U,H))) \to L^2(\mu, H).$$

We shall call this as the conditional Skorohod operator with respect to the diffusion X. For any $G(X) \in \mathcal{S}(X, L_2(U, \mathbb{R}))$ $\mathcal{F}_1(X)$ measurable functions and $K(X) \in$

$\mathcal{S}(X)$ $\hat{\delta}$ *satisfies the following relation:*

$$\begin{aligned}
E[\langle \hat{\delta} G(X), K(X)\rangle_H] &= E_{X(\mu)}[\langle \hat{\delta} G, K\rangle_H] \\
&= E_{X(\mu)}[\langle G, \hat{\nabla} K\rangle_{L^2([0,1],L_2(U,H))}] \\
&= E[\langle G(X), \hat{\nabla} K(X)\rangle_{L^2([0,1],L_2(U,H))}].
\end{aligned}$$

Remark 6.2.2. *From the well-definedness and closability of $\hat{\nabla}$, it follows that $\hat{\delta}$ is well-defined and closable.*

Lemma 6.2.1. *Let $G(X) \in \mathcal{S}(X)$ and $h \in L^2([0,1]; L_2(H,\mathbb{R}))$. Then $G(X)\otimes h \in$ dom$(\hat{\delta})$ and $\hat{\delta}(G\otimes h) = W(hP(X))\otimes G - \hat{\nabla}_h G(X)$. Furthermore, if $\ker(\sigma(s,X))^{\perp}$ is h_s invariant for any $s \in [0,1]$ we have $\hat{\delta}(G\otimes h) = W(P(X)h)\otimes G - \hat{\nabla}_h G(X)$.*

Proof. For any $F(X) \in \mathcal{S}(X)$ we have

$$\begin{aligned}
&E[\langle \hat{\nabla} F(X), G(X)\otimes h\rangle_{L^2([0,1],L_2(U,H)}] \\
&\quad = E[\langle \nabla F(X), G(X)\otimes h\rangle_{L^2([0,1],L_2(U,H))}] \\
&\quad = E[\langle F(X), W(h)G(X) - \nabla_h G(X)\rangle_H] \\
&\quad = E[\langle F(X), W(hP(X))G(X) - \hat{\nabla}_h G(X)\rangle_H]
\end{aligned}$$

therefore, $\hat{\delta} G(X)\otimes h = W(hP(X))G(X) - \hat{\nabla}_h G(X)$. □

Corollary 6.2.2. *Let $G(X) \in L^2(\mu, L^2([0,1], L_2(H,H))$ be a predictable process. Then $G(X) \in$ dom$(\hat{\delta})$ and $\hat{\delta} G(X) = \delta \mathbb{P}(X) G(X)$.*

Proof. Follows from the definitions. □

Theorem 6.2.1. *Let $G(X) \in L^2(\mathcal{F}_1(X), L_2(U,H))$ be of the form*

$$G(X) = \sum_i^n K_i(X)\otimes e_i$$

where $K_i(X) \in \mathcal{S}(X)$, and $e_i \in L_2(U,\mathbb{R})$ for each $i \in \mathbb{N}$.

i) We have that

$$\hat{\delta}G(X) = \sum_{i=1}^{n} \Big[-\hat{\nabla}_{e_i} K_i(X) + K_i(X) W(\mathbb{P}(X)e_i) \Big],$$

where $\mathbb{P}(X)e_i(\cdot) = \int_0^{\cdot} \dot{e}_i(s) P(X_s) ds$. *Specifically, we have that*

$$E[\delta G(X)|\mathcal{F}_1(X)] = \hat{\delta}G(X). \tag{6.1}$$

ii) If $G(X) \in \mathcal{S}(X, L_2(H,H))$, *then*

$$\hat{\delta}G(X) = \delta \mathbb{P}(X) G \tag{6.2}$$

μ-*a.s.*

Proof. **Proof of i**) : We will proceed by a density argument. Let $H(X) \in \mathcal{S}(X)$, then we have

$$\begin{aligned}
&E[\langle G(X), \hat{\nabla} H(X) \rangle_{L^2([0,1], L_2(U,H))}] \\
&\quad = E\Big[\sum_{i=1}^{n} \langle K_i(X), \hat{\nabla}_{e_i} H(X) \rangle_H \Big] \\
&\quad = \sum_{i=1}^{n} E\Big[\langle K_i(X), \nabla_{e_i} H(X) \rangle_H \Big] \\
&\quad = \sum_{i=1}^{n} E\Big[\langle -\nabla_{e_i} K_i(X) + K_i(X)\delta(e_i), H(X) \rangle_H \Big] \\
&\quad = \star
\end{aligned}$$

from the definition of Gateaux derivative and quasi-invariance of the Wiener measure with respect to e_i. Since $E[\delta e_i | \mathcal{F}_1(X)] = \delta(\mathbb{P}(X)e_i)$ we obtain

$$\begin{aligned}
\star &= \sum_{i=1}^{n} E\Big[\langle -\nabla_{e_i} K_i(X) + K_i(X) E[\delta(e_i)|\mathcal{F}_1(X)], H(X) \rangle_{H)} \Big] \\
&= \sum_{i=1}^{n} E\Big[\langle -\nabla_{e_i} K_i(X) + K_i(X)\delta(\mathbb{P}(X)e_i), H(X) \rangle_H \Big]
\end{aligned}$$

$$= E\Big[\langle\sum_{i=1}^{n}\big[-\hat{\nabla}_{e_i}K_i(X)+K_i(X)\delta(\mathbb{P}(X)e_i)\big],H(X)\rangle_H\Big],$$

hence the result follows for the cylindrical functions. To conclude the general case, we can pass to the limit in $\mathbb{M}_{2,1}(H)$.

Proof of ii) : Let $H(X)\in\mathcal{S}(X)$, then

$$\begin{aligned}
&E[\langle\hat{\delta}G(X),H(X)\rangle_H]\\
&\quad= E[\langle G(X),\hat{\nabla}H(X)\rangle]_{L^2([0,1],L_2(U,H))}]\\
&\quad= E[\langle G(X),\nabla H(X)\rangle_{L^2([0,1],L_2(U,H))}]\\
&\quad= E[\langle\delta G(X),H(X)\rangle_H]\\
&\quad= E[\langle E[\delta G(X)|\mathcal{F}_1(X)],H(X)\rangle_H]\\
&\quad= E[\langle\delta\mathbb{P}(X)G(X),H(X)\rangle_H].
\end{aligned}$$

□

Corollary 6.2.3. *Let $F(X),G(X)\in\mathcal{S}(X)$ and $h\in L_2(U,\mathbb{R})$, then it holds that*

$$E[\hat{\nabla}F(X)G(X)] = E[F(X)(-\hat{\nabla}G(X)+G(X)W(hP(X))] \tag{6.3}$$

Proof. Follows as in the proof of the first part of theorem above. □

Lemma 6.2.2. *For any $h,k\in L_2(U,\mathbb{R})$, and $G(X)\in\mathcal{S}(X)$ we have*

$$\hat{\nabla}_k\delta(G(X)\otimes h) = G(X)\langle h,k\rangle+\hat{\nabla}_{\hat{\nabla}_h k}G(X)+\hat{\delta}(\hat{\nabla}_k G(X)\otimes h),$$

and for any $u\in U$ we have

$$\hat{\nabla}\delta(G(X)\otimes h)u = G(X)\otimes hu+\hat{\delta}(\hat{\nabla}G(X)u\otimes h),$$

and

$$\hat{\nabla}\delta(G(X)\otimes hP(X))u = G(X)\otimes hP(X)u+\hat{\delta}(\hat{\nabla}G(X)u\otimes hP(X)).$$

Proof. For the first part, we have

$$
\begin{aligned}
&\hat{\nabla}_k \delta(G(X)\otimes h)\\
&\quad = E[\nabla_k \delta(G(X)\otimes h)|\mathcal{F}_1(X)]\\
&\quad = E[\nabla_k \big(W(hP(X))G(X) - \nabla_h G(X)\big)|\mathcal{F}_1(X)]\\
&\quad = E[G(X)\langle h,k\rangle + W(h)\nabla_k G(X) - \nabla_k\nabla_h G(X)|\mathcal{F}_1(X)]\\
&\quad = E[G(X)\langle h,k\rangle + W(h)\nabla_k G(X) - \nabla_h\nabla_k G(X) + \nabla_{\nabla_h k}G(X)|\mathcal{F}_1(X)]\\
&\quad = E[G(X)\langle h,k\rangle + \nabla_{\nabla_h k}G(X) + \delta(\nabla_k G(X)\otimes h)|\mathcal{F}_1(X)]\\
&\quad = G(X)\langle h,k\rangle + \hat{\nabla}_{\hat{\nabla}_h k}G(X) + \delta(E[\nabla_k G(X)\otimes h|\mathcal{F}_1(X)]P(X))\\
&\quad = G(X)\langle h,k\rangle + \hat{\nabla}_{\hat{\nabla}_h k}G(X) + \delta\bigg(\hat{\nabla}_k G(X)\otimes h)P(X)\bigg)\\
&\quad = G(X)\langle h,k\rangle + \hat{\nabla}_{\hat{\nabla}_h k}G(X) + \hat{\delta}\bigg(\hat{\nabla}_k G(X)\otimes h)\bigg)
\end{aligned}
$$

The second part follows in a similar fashion. □

Proposition 6.2.1. *Let $G_1(X), G_2(X) \in \mathcal{S}(X)$ and $h,k \in L_2(U,\mathbb{R})$. Then, we have*

$$
\begin{aligned}
&E[\langle \hat{\delta} G_1(X)\otimes h, \hat{\delta} G_2(X)\otimes k\rangle_H]\\
&\quad = E[\langle G_1(X)\otimes h, G_2(X)\otimes k\rangle_{L_2([0,1],L_2(U,H)}]\\
&\quad + \int_0^1\int_0^1 \sum_{k,l} E[\langle \hat{\nabla} G_1(X)(r_2)e_l\otimes h(r_1)e_k, \hat{\nabla} G_2(X)(r_1)(e_k)\otimes k(r_2)e_l\rangle]dr_1dr_2.
\end{aligned}
$$

Proof. Using the relation above we obtain that

$$
\begin{aligned}
&E[\langle \hat{\delta}(G_1(X)\otimes h), \hat{\delta}(G_2(X)\otimes k)\rangle_H]\\
&\quad = E[\langle G_1(X)\otimes h, \hat{\nabla}\hat{\delta}(G_2(X)\otimes k\rangle_{L^2([0,1],L_2(U,H))}]\\
&\quad = E[\langle G_1(X)\otimes h, G_2(X)\otimes k\rangle_{L^2([0,1],L_2(U,H)}]\\
&\quad + E[\langle G_1(X)\otimes h, \hat{\delta}(\hat{\nabla} G_2(X)\otimes k\rangle_{L^2([0,1],L_2(U,H))}].
\end{aligned}
$$

Retracting back to the differential operator results:

$$
E[\langle G_1(X)\otimes h, \hat{\delta}(\hat{\nabla} G_2(X)\otimes k\rangle_{L^2([0,1],L_2(U,H)}]
$$

$$= \int_0^1 \sum_k E[\langle G_1(X) \otimes h(r_1)e_k, \hat{\delta}(\hat{\nabla} G_2(X)(r_1)(e_k) \otimes k\rangle]dr_1$$

$$= \int_0^1 \int_0^1 \sum_{k,l} E[\langle \hat{\nabla} G_1(X)(r_2)e_l \otimes h(r_1)e_k, \hat{\nabla} G_2(X)(r_1)(e_k) \otimes k(r_2)e_l\rangle]dr_1 dr_2,$$

by using the relation once again. From which the desired result follows □

6.3 Application: Explicit Representation of Martingale Representation and Functional Inequalities

In this section we will apply the conditional Malliavin derivative to obtain explicit integrand for the martingale representation for diffusion processes. Then we will apply it to functional inequalities such as Poincare and Logarithmic Sobolev inequalities.

Theorem 6.3.1. *Any $F(X) \in \mathbb{M}_{2,1}$ can be represented as*

$$F(X) = E[F(X)] + \int_0^1 E[\hat{\nabla}_s F(X) P(X_s) | \mathcal{F}_s(X)] dW_s.$$

Proof. Without loss of generality we can assume that $E[F(X)] = 0$ by considering $\tilde{F}(X) = F(X) - E[F]$ instead. In one hand we have the representation

$$F(X) = \int_0^1 \xi_s(X) P_s(X) dW_s,$$

for some $\xi(X) = (\xi_s(X))_s \in L^2(dt \times d\mu, L_2(U, H))$ that is adapted to the filtration generated by X. If for a moment one assumes that $F(X) \in \mathcal{S}(X)$, then from the explicit representation for martingales for Wiener processes, one obtains that

$$F(X) = \int_0^1 E[D_s F(X) | \mathcal{F}_s(W)] dW_s.$$

By taking conditional expectation with respect to the σ-algebra $\mathcal{F}_1(X)$ from the interplay we then obtain

$$F(X) = \int_0^1 E[D_s F(X)|\mathcal{F}_s(X)] P_s(X) dW_s.$$

From the tower property of conditional expectation it holds that $E[D_s F(X)|\mathcal{F}_s(X)] = E[E[D_s F(X)|\mathcal{F}_1(X)]|\mathcal{F}_s(X)] = E[\hat{D}_s F(X)|\mathcal{F}_s(X)]$ $dt \times d\mu$-a.e. By the density of Schwartz class functionals we can take a sequence $(F_n(X))_n \subset \mathcal{S}(X)$ which converges to $F(X)$ in $\mathbb{M}_{2,1}$, we have $(\hat{\nabla} F_n(X))_n$ converges to $\hat{\nabla} F(X)$ in $L^2(\mu, L_2(U, H))$, in particular

$$\lim_n E \int_0^1 \left\| E[\hat{D}_s F_n(X) - \hat{D}_s F(x)|\mathcal{F}_s(X)] P_s(X) \right\|^2_{L_2(U,H)} ds = 0,$$

hence the result holds for $F(X) \in \mathbb{M}_{2,1}$ by the Itô isometry. $\square$

Remark 6.3.1. *In finite dimensional Euclidean spaces several instances of this explicit representation have been known for a while. By using martingale techniques Davis have obtained this representation in [4] by using Bismut's formula [1]. Alternative formulations of the explicit integrand is also have been obtained through the PDE representations of forward-backward SDE systems [18].*

Corollary 6.3.1. *Assume that $F(X) \in \mathbb{M}_{2,1}$ such that $\hat{\nabla} F(X) = 0$ μ-a.s. Then it holds that $F(X) = 0$ μ-a.s. constant.*

Proof. Integrand in the explicit representation is zero by the assumption, hence $F(X)$ is μ-a.s. constant. $\square$

Theorem 6.3.2 (Poincare Inequality)**.** *For any $F(X) \in \mathbb{M}_{2,1}$, we have*

$$E[\|F(X) - EF(X)\|^2_H] \leq E[\|\hat{\nabla} F(X) \mathbb{P}(X)\|^2]$$

Proof. Using the explicit representation, by the Jensen's inequality we have

$$E[\|F(X) - E[F(X)]\|^2_H] = E[\| \int_0^1 E[D_s F(X)|\mathcal{F}_s(X)] P_s(X) dW_s\|^2_H$$

$$\leq E\int_0^1 \|E[D_s F(X)|\mathcal{F}_s(X)]\|^2_{L_2(Q1/2U,H)} ds$$
$$\leq E\int_0^1 \|D_s F(X)\|^2_{L_2(U,H)} ds$$

hence, we obtain the result. $\square$

Theorem 6.3.3 (Log-Sobolev inequality)**.** *For any* $F(X) \in \mathbb{M}_{2,1}$*, it holds that*

$$E[\|F(X)\|^2_H \log \|F(X)\|^2_H] \leq E[\|F\|^2_H] \log E[\|F(X)\|^2_H] + 2E[\|\hat{\nabla} F(X) P_s(X)\|^2_{L_2(U,H)}].$$

Proof. Suppose that $\|F(X)\|_H > \epsilon > 0$ almost everywhere and then set $f(X) = \|F(X)\|_H / E[\|F(X)\|_H]$. Using the representation, we obtain the following:

$$\begin{aligned} f(X) &= 1 + \int_0^1 E[\hat{\nabla}_s f(X)|\mathcal{F}_s(X)] P_s(X) dW_s \\ &= 1 + \int_0^1 \frac{E[\hat{\nabla}_s f(X)|\mathcal{F}_s(X)] P_s(X)}{E[f(X)|\mathcal{F}_s(X)]} E[f(X)|\mathcal{F}_s(X)] dW_s. \end{aligned}$$

After Itô's formula $f(X)$ reads as

$$\begin{aligned} f(X) = \exp\Bigg[&\int_0^1 \frac{E[\hat{\nabla}_s f(X)|\mathcal{F}_s(X)]}{E[f(X)|\mathcal{F}_s(X)]} P_s(X) dW_s \\ &- \frac{1}{2}\int_0^1 \left\| \frac{E[\hat{\nabla}_s f(X)|\mathcal{F}_s(X)]}{E[f(X)|\mathcal{F}_s(X)]} P_s(X) \right\|^2_{L_2(U,\mathbb{R})} ds \Bigg]. \end{aligned}$$

Define a new measure by $d\nu = f(X) d\mu$, under this measure by Itô's formula we have the following:

$$\begin{aligned} E_\nu[\log f(X)] &= \frac{1}{2} E_\nu \int_0^1 \left\| \frac{E[\hat{\nabla}_s f(X)|\mathcal{F}_s(X)]}{E[f(X)|\mathcal{F}_s(X)]} P_s(X) \right\|^2_{L_2(U,\mathbb{R})} ds \\ &= \frac{1}{2} E_\nu \int_0^1 \|E[\hat{\nabla} \log f(X) \mathcal{F}_s(X)] P_s(X)\|^2_{L_2(U,\mathbb{R})} ds \\ &\leq \frac{1}{2} E_\nu \|\hat{\nabla}_s \log(f(X))\|^2_{L_2(U,\mathbb{R})} ds \\ &= \frac{1}{2} E_\nu \left\| \frac{\hat{\nabla} f(X)}{f(X)} \mathbb{P}_s(X) \right\|^2_{L_2(U,H)}. \end{aligned}$$

Thus we have

$$E[f(X)\log f(X)] \leq \frac{1}{2}E[\|\hat{\nabla}f(X)P_s(X)\|^2_{L_2(U,\mathbb{H})}\frac{1}{f(X)}]].$$

$\square$

Bibliography

[1] J.-M. Bismut. "Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions". In: *Z. Wahrsch. Verw. Gebiete* 56.4 (1981), pp. 469–505.

[2] V. I. Bogachev. *Gaussian measures*. Vol. 62. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1998, pp. xii+433.

[3] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*. Vol. 44. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1992, pp. xviii+454.

[4] M. H. A. Davis. "The Wiener space derivative for functionals of diffusions on manifolds". In: *Nonlinearity* 1.1 (1988), pp. 241–251.

[5] K. D. Elworthy and X.-M. Li. "Itô maps and analysis on path spaces". In: *Math. Z.* 257.3 (2007), pp. 643–706.

[6] L. Gawarecki and V. Mandrekar. *Stochastic differential equations in infinite dimensions with applications to stochastic partial differential equations*. Probability and its Applications (New York). Springer, Heidelberg, 2011, pp. xvi+291.

[7] A. N. Godunov. "A counterexample to Peano's theorem in an infinite-dimensional Hilbert space". In: *Vestnik Moskov. Univ. Ser. I Mat. Meh.* 27.5 (1972), pp. 31–34.

[8] A. Grorud and É. Pardoux. "Intégrales hilbertiennes anticipantes par rapport à un processus de Wiener cylindrique et calcul stochastique associé". In: *Appl. Math. Optim.* 25.1 (1992), pp. 31–49.

[9] M. Hairer. *An Introduction to Stochastic PDEs*. 2009.

[10] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. Vol. 24. North-Holland Mathematical Library. North-Holland Publishing Co., Amsterdam-New York; Kodansha, Ltd., Tokyo, 1981, pp. xiv+464.

[11] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*. Second. Vol. 288. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 2003, pp. xx+661.

[12] J. Jacod and M. Yor. "Étude des solutions extrémales et représentation intégrale des solutions pour certains problèmes de martingales". In: *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 38.2 (1977), pp. 83–125.

[13] S. Janson. *Gaussian Hilbert spaces*. Vol. 129. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1997, pp. x+340.

[14] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*. Second. Vol. 113. Graduate Texts in Mathematics. Springer-Verlag, New York, 1991, pp. xxiv+470.

[15] R. Kruse. *Strong and weak approximation of semilinear stochastic evolution equations*. Vol. 2093. Lecture Notes in Mathematics. Springer, Cham, 2014, pp. xiv+177.

[16] H. H. Kuo. *Gaussian measures in Banach spaces*. Lecture Notes in Mathematics, Vol. 463. Springer-Verlag, Berlin-New York, 1975, pp. vi+224.

[17] J. A. León and D. Nualart. "Stochastic evolution equations with random generators". In: *Ann. Probab.* 26.1 (1998), pp. 149–186.

[18] J. Ma, P. Protter, and J. M. Yong. "Solving forward-backward stochastic differential equations explicitly—a four step scheme". In: *Probab. Theory Related Fields* 98.3 (1994), pp. 339–359.

[19] D. Nualart. *The Malliavin calculus and related topics*. Second. Probability and its Applications (New York). Springer-Verlag, Berlin, 2006, pp. xiv+382.

[20] M. Ondreját. “Brownian representations of cylindrical local martingales, martingale problem and strong Markov property of weak solutions of SPDEs in Banach spaces”. In: *Czechoslovak Math. J.* 55(130).4 (2005), pp. 1003–1039.

[21] C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*. Vol. 1905. Lecture Notes in Mathematics. Springer, Berlin, 2007, pp. vi+144.

[22] M. Pronk and M. Veraar. “Tools for Malliavin calculus in UMD Banach spaces”. In: *Potential Anal.* 40.4 (2014), pp. 307–344.

[23] P. E. Protter. *Stochastic integration and differential equations*. Vol. 21. Stochastic Modelling and Applied Probability. Second edition. Version 2.1, Corrected third printing. Springer-Verlag, Berlin, 2005, pp. xiv+419.

[24] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Third. Vol. 293. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, pp. xiv+602.

[25] D. W. Stroock and M. Yor. “On extremal solutions of martingale problems”. In: *Ann. Sci. École Norm. Sup. (4)* 13.1 (1980), pp. 95–164.

[26] A. S. Üstünel. *Analysis on Wiener Space and Applications*. 2015. arXiv: `1003.1649 [math.PR]`.

[27] A. S. Üstünel. “Martingale representation for degenerate diffusions”. In: *J. Funct. Anal.* 276.11 (2019), pp. 3468–3483.

[28] A. S. Üstünel and M. Zakai. *Transformation of measure on Wiener space*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2000, pp. xiv+296.

[29] A. S. Üstünel. *Malliavin Calculus for Degenerate Diffusions*. 2020. arXiv: `2012.07316 [math.PR]`.

[30] T. Yamada and S. Watanabe. “On the uniqueness of solutions of stochastic differential equations”. In: *J. Math. Kyoto Univ.* 11 (1971), pp. 155–167.

[31] M. Yor. “Existence et unicité de diffusions à valeurs dans un espace de Hilbert”. In: *Ann. Inst. H. Poincaré Sect. B (N.S.)* 10 (1974), pp. 55–88.