


# ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM


## Stochastic Ananlysis in Hilbert Space and Application to Interest Rate Theory


By

Maria E. Giouvanelli


A THESIS

Submitted to the Department of Statistics

of the Athens University of Economics and Business

in partial fulfilment of the requirements for

the degree of Master of Science in Statistics


Athens, Greece

January, 2013




# ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

## ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ


# Στοχαστική Αναάλυση σε χώρους Hilbert και Εφαρμογές στη Θεωρία Επιτοκίων


Μαρία Ε. Γιουβανέλλη


ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής

του Οικονομικού Πανεπιστημίου Αθηνών

ως μέρος των απαιτήσεων για την απόκτηση

Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική


Αθήνα

Ιανουάριος, 2013

# DEDICATION

This thesis is dedicated to my family…

# ACKNOWLEDGEMENTS

I am very grateful my supervisor Professor Athanasios Yannacopoulos, whose experience and knowledge led me to the completion of my thesis.

# VITA

I started my studies in 2007 at the University of Athens, Department of Mathematics. I completed it in 2011 and I started my postgraduate studies at Athens University Economics and Business in Statistical Division. These studies are completed with the preparation of my diploma thesis with supervisor Professor Athanasios Yannakopoulos

# ABSTRACT

Maria Giouvanelli

## Stochastic Analysis in Hilbert Space and Applications in Rate's Theory

January,2013

This thesis is in bond market modeling using the theory of infinite dimensional stochastic analysis. We start with a presentation of fundamental concepts such as the theory of canonical measures, the Wiener process in $\mathbb{R}$, $\mathbb{R}^n$ and infinite dimensional Hilbert space and continue with the construction of the Itô integral, the stochastic convolution and their applications in the theory of infinite dimensional S.D.E. Finally, we present applications of this theory in bond market modelling through the framework of the Heath Jarrow Morton model.

# ΠΕΡΙΛΗΨΗ

Μαρία Γιουβανέλλη

Στοχαστική Ανάλυση σε Χώρους Hilbert και Εφαρμογές στη Θεωρία Επιτοκίων

Ιανουάριος, 2013

Στην εργασία αυτή θέλουμε να παρουσιάσουμε την βασική εξίσωση στη θεωρία των επιτοκίων γνωστή ως Heath Jarrow Morton και ορμόμενοι αυτής ξεκινάμε μία μελέτη της στοχαστικής ανάλυσης σε χώρους Hilbert. Αρχικά παρουσιάζουμε τα κανονικά μέτρα, την θεωρία τελεστών και την θεωρία της διαδικασίας Wiener, τόσο στο χώρο $\mathbb{R}$, $\mathbb{R}^n$ όσο και σε απειροδιάστατους χώρους Hilbert. Ακόμα, παρουσιάζουμε το στοχαστικό ολοκλήρωμα Itô και την έννοια της στοχαστικής συνέλιξης, μαζί με την μορφή της και τις ιδιότητες της. Ολοκληρώνοντας αυτήν την εργασία εισάγουμε την έννοια των επιτοκίων και παρουσιάζουμε μοντέλα επιτοκίων τα οποία βασίζονται στην θεωρία των στοχαστικών διαφορικών εξισώσεων σε απειροδιάστατους χώρους Hilbert.

**TABLE OF CONTENTS**


**1 Intoduction** **1**

**2 Basic concepts** **3**
2.1 Hilbert spaces 3
2.2 Riesz representation theorem 4
2.3 Nuclear , Trace Class, Hilbert Schmidt and Shift Operators 6
2.4 Bochner integral 8
2.5 Sazonov's topology 10
2.6 Martingales 11
2.7 Schwarz space 12

**3 Gaussian Measures in Hilbert Spaces** **17**
3.1 Gaussian Measures and Properties 17
3.2 Covariance operator of Gaussian measures 21

**4 Wiener Process in Hilbert Space** **27**
4.1 Basic facts on Wiener Process in $\mathbb{R}^n$ 27
4.2 Properties of Wiener process 31
4.2.1 Markov property 31
4.2.2 Strong Markov property 33
4.2.3 Martingale property 38
4.2.4 Characterization of Brownian Motion 39
4.3 Properties of trajectories 41
4.4 Wiener's process construction 43
4.5 Q-Wiener process 45
4.6 Weak Wiener process 49

**5 Stochastic integral in Hilbert Space** **55**
5.1 Itô integral in $\mathbb{R}$ 55
5.2 Itô integral as a stochastic process 64
5.3 Itô processes 66
5.4 Itô processes on $\mathbb{R}^n$ 70
5.5 Stochastic integral for Q-Wiener and Weak Wiene processes 72

**6 Stochastic Convolution** **81**
6.1 $C_0$-semigroups and well-posedness of the deterministic Cauchy Problem 82

6.2 Solution to the stochastic Cauchy problem 85

**7 The Heath Jarrow Morton Model** **95**
7.1 The Bond Market 96
7.2 The HJM Evolution Equation 96
7.2.1 Function Spaces for Forward Curves 98
7.3 The Abstract HJM Model 102
7.3.1 Drift Condition and Absence of Arbitrage 104
7.3.2 Long Rates Never Fall 106

# Chapter 1

# Introduction

The aim of this thesis is an introduction to stochastic analysis in infinite dimensional Hilbert spaces, and the theory of stochastic differential equations in this setting. As an application we discuss the use of these mathematical techniques to bond market modelling, in mathematical finance. We start with a presentation of fundamental concepts such as the theory of some basic concepts of Hilbert space properties,the Riesz representation theorem,the theory of nuclear, trace class, Hilbert Schmidt and shift operators. Also we present the Bochner integral and theory about Sazonov's topology, martingales and Schwarz space.
At the first chapter we present theory of Gaussian measures, their properties and the covariance operator of Gaussian measures in Hilbert space.
The next chapter is about the Wiener process in $\mathbb{R}$, $\mathbb{R}^n$ and infinite dimensional Hilbert space, we present the properties of Wiener process like Markov property, strong Markov property and martingale property. Also, we study the characterization of Brownian motion which is based on fundamental result of Levy, properties of Wiener trajectories and a Wiener process construction using Haar functions.At next we will see the theory of Q−Wiener and weak Wiener process.
In the following chapter we present theory of the construction of the Ito integral,the lemma Ito and and its properties in $\mathbb{R}$ and $\mathbb{R}^n$ with examples and theorems.Also, we present the ito integral in Hilbert space, where we have the stochastic integral for $Q$-Wiener process.

In the next chapter we will see the theory for well posed of the equation:

$$W_A(t) := \int_0^t S(t-s)BdW(s), \;\; t \,\epsilon\, [0,T)$$

which called stochastic convolution and the theory of semigroups, its properties and their applications in the theory of infinite dimensional S.D.E.
In the end, we present applications of the theory that we present to you before in bond market modelling through the HJM equation:

$$df_t(x) = \left(\frac{\partial}{\partial x} f_t(x) + a_t(x)\right) dt + \sum_{i=1}^{\infty} \sigma_t^i(x) ds_t^i$$

of the Heath Jarrow Morton models, which are based on theory of infinite dimensional Hilbert space.

# Chapter 2

# Basic concepts

## 2.1 Hilbert spaces

A Euclidean space $\mathbb{R}^n$ is a vector space endowed with the inner product $< x, y >= x^T y$, norm $||x|| = \sqrt{x^T X} = \sqrt{< x, x >}$ and associated metric $||x - y||$, such that every Cauchy sequence obtains a limit in $\mathbb{R}^n$. This makes $\mathbb{R}^n$ a Hilbert space:

**Definition 2.1.** *A Hilbert space H is a vector endowed with an inner product and associated norm and metric, such that every Cauchy sequence in H has a limit in H.*

A Hilbert space is also a Banach space:

**Definition 2.2.** *A Banach space B is a normed space with associated metric $d(x, y) = ||x - y||$ such that every Cauchy sequence in B has a limit in B.*

**Definition 2.3.** *A norm on a vector space V is a mapping $||.|| : V \to [0, \infty)$ such that for all x and y in V and all scalars c, the following:*
*(i)* $< x, y >=< y, x >$
*(ii)*$< cx, y >= c < x, y >$
*(iii)*$< x + y, z >=< x, z > + < y, z >$
*(iv)*$< x, x >> 0$ *when* $x \neq 0$

*hold. A vector space endowed with a norm is called a normed space.*

The difference between a Banach space and a Hilbert space is the source of the norm. In the Hilbert space case the norm is defined via the inner product $||x|| = \sqrt{< x, x >}$, whereas in the Banach space case the norm is defined directly, by Definition (2.3). Thus, a Hilbert space is Banach space, but the other way around may not be true, because in some cases the norm cannot be associated with an inner product.

Example of Hilbert space are $\mathbb{R}^n,\ \mathbb{C}^n,\ L_2,\ \mathcal{L}_2$.

The space of random variables $X$ which satisfies: $\{\mathbb{E}_P[|X|^p]\}^{1/p} < \infty$ is vector space and symbolized as $L_p$.
Now, the space $L_2(\Omega,\ \mathcal{F}_0,\ P)$ or simply $L_2$ is the space of square integrable random variables.

Let now $X$ and $Y$ be Banach spaces. As usual, we denote by $\mathcal{L}(X, Y)$ the space of all bounded linear operators from $X$ into $Y$ endowed with the norm

$$||A|| := \inf\{C :\ ||Ax||_Y \leq C||x||_X,\ x \,\epsilon\, X\},\quad A \,\epsilon\, \mathcal{L}(X, Y).$$

So, the space $\mathcal{L}_N(X, Y)$ of all nuclear operators from $X$ into $Y$ endowed with the norm

$$||A||_N := \inf\{\sum_{j=1}^{\infty} ||y_j||_Y \cdot ||\phi_j||_{X^*}\ :\ Ax = \sum_{j=1}^{\infty} y_j \phi_j(x)\}$$

is a Banach space.

## 2.2 Riesz representation theorem

Let $H \neq \{0\}$ Hilbert space. We will see that $H^*$ has a lot of functionals, which representated with a specific way from the elements of $H$.

**Lemma 2.1.** *For all $a \epsilon H$, the $f_a : H \to \mathbb{R}$ with $f_a(x) = \langle x, a \rangle$ belongs to $H^*$, and $||f_a||_{H^*} = ||a||_H$.*

*Proof.* We have

$$f_a(\lambda x + \mu y) = \langle \lambda x + \mu y, a \rangle = \lambda < x, a > + \mu < y, a > =$$

$$\lambda f_a(x) + \mu f_a(y),$$

and

$$|f_a(x)| = | < x, a > | \leq ||a||\ ||x||.$$

So, $f_a \epsilon H^*$ and $||f_a|| \leq ||a||$. Finally, if $a \neq 0$,

$$||f_a|| \geq \frac{|f_a(a)}{||a||} = \frac{| < a, a > |}{||a||} = ||a||.$$

If $a = 0$, then $||f_a|| = 0$ $(f_a \equiv 0)$.
So, we finish our proof. □

Representation theorem Riesz tell us that every $f \epsilon H^*$ represented as $f = f_a$ for any $a \epsilon H$ :

**Theorem 2.1** (Representation Theorem Riesz)**.** *Let $H$ be a Hilbert space, and $f \epsilon H^*$. There exists a unique $a \epsilon H$ such that $f = f_a$.*

*Proof.* We consider $M = \text{Ker} f = \{x \epsilon H \ : \ f(x) = 0\}$. The $M$ is a linear subspace of $H$.
If $M = H$, then $f \equiv 0$ and $f = f_o$.
If $M \neq H$, then there exists $z \neq 0$, $z \epsilon H$ which is vertical to $M$. Then, for any $y \epsilon H$ we have

$$f(f(z)y - f(y)z) = f(z)f(y) - f(y)f(z) = 0.$$

So, $f(z)y - f(y)z \epsilon M$, and because $z \perp M$ we obtain

$$\langle f(z)y - f(y)z, z \rangle = 0 \ \Rightarrow \ f(z) < y, z > = f(y) < z, z >$$

$$\Rightarrow \ f(y) = \left\langle y, \frac{f(z)z}{||z||^2} \right\rangle = f_a(y),$$

where $a = f(z)z/||z||^2$. The uniqueness of $a$ is simply. If $f(y) =< y, a >=< y, a' >$ for all $y \epsilon H$, then $a - a' \perp y, \ \forall \ y \epsilon H$. So, $a = a'$. □

## 2.3 Nuclear , Trace Class, Hilbert Schmidt and Shift Operators

Consider two separable Hilbert spaces $U$ and $V$ and denote by $\mathcal{L}(U, V)$ the space of bounded linear operators $A : U \to V$. The adjoint operator $A^*$ is an element of $\mathcal{L}(U, V)$ such that

$$(Ax, y) = (x, A^*y), \quad \forall x \epsilon U, y \epsilon V$$

Two important classes of compact operators are given in the following.

**Definition 2.4.** *An operator $Q \epsilon \mathcal{L}(U, V)$ is called a nuclear operator if there exists a sequence $\{v_n\} \epsilon V$ and a sequence $\{u_n\} \epsilon U$ such that*

$$Qx = \sum_{n=1}^{\infty} v_n (u_n.x)_U \quad \forall x \epsilon U \ , and \sum_{n=1}^{\infty} \|v_n\|_V \|u_n\|_U < \infty.$$

**Definition 2.5.** *Let $U = V$. A nuclear operator $Q$ that is non-negative (i.e $(Lu, u) \geq 0$ for all $u \epsilon U$) and symmetric (i.e. $(Lu, v) = (u, Lv)$ for all $u, v \epsilon U$) is called a trace class operator.*

The following is a very usefully property of nuclear operators.

**Proposition 2.1.** *Let $Q : U \to U$ be a nuclear operator and let $\{e_n\}$ be an orthonormal basis of U. Define the trace of the operator $Q$ as the infinite series $Tr(Q) := \sum_{n=1}^{\infty} (Qe_n, e_n)$. Then $Tr(Q)$ is a well-defined finite quantity and independent of the choice of the orthonormal basis $\{e_n\}$.*

Trace class operators are interesting from the point of view of infinite-dimensional stochastic analysis since they can be considered the generalisation of the covariance matrix in infinite dimensions. The solution of the eigenvalue problem for a trace class operators provides us with an orthonormal basis for the Hilbert space $U$. An interesting subclass of nuclear operators consists of the Hilbert-Schmidt operators.

**Definition 2.6.** *A bounded linear operator $Q : U \to V$ is called a Hilbert-Schmidt operator if $\sum_{n=1}^{\infty} \|Qe_n\|^2 < \infty$, where $\{e_n\}$ is an orthonormal basis of U.We will denote the space of al Hilbert-Schmidt operators from U to V by $\mathcal{L}_2(U, V)$.*

The space of Hilbert Schmidt operators can be turned into a separable Hilbert space by defining the inner product

$$(Q_1, Q_2)_{L_2(U,V)} = \sum_{n=1}^{\infty} (Q_1 e_n, Q_2 e_n).$$

The following proposition helps us to define the "square root" of trace class operator.

**Proposition 2.2.** *If $Q : U \to U$ is a trace class operator, then there exists a unique Hilbert-Schmidt operator $R$ such that $R \circ R = Q$. We will use the notation $R = q^{\frac{1}{2}}$. Furthermore, $\|Q\|^2_{L_2(U)} = Tr(Q)$.*

The operator $Q^{\frac{1}{2}}$ has the usefully property that $L \circ Q^{\frac{1}{2}} \, \epsilon L_2(U, V)$ for any $L \, \epsilon \mathcal{L}(U, V)$.

**Definition 2.7.** *The shift operator or translation operator is an operator that takes a function $f(\cdot)$ to it translation $f(\cdot + a)$.*

Shift operators are examples of linear operators, important for their simplicity and natural occurrence. The shift operator action on functions of a real variable plays an important role in harmonic analysis.

The shift operator acting on real or complex valued functions or sequences is a linear operator which preserves most of the standard norms which appear in functional analysis.

**Example 2.1.** *Space $\ell^2(\mathbb{Z})$ is the set of functions $x : \mathbb{Z} \to \mathbb{C}$. which satisfies $\sum_{n\,\epsilon\,\mathbb{Z}} |x(n)|^2 < \infty$. It has orthonormal basis $\{e_n : n \epsilon \mathbb{Z}\}$, where $e_n(m) = \delta_{nm}(m \epsilon \mathbb{Z})$. Specially are important shift operators which defined as:*

$$Ue_n = e_{n+1} \quad \textit{right shift}$$
$$U^*e_n = e_{n+1} \quad \textit{left shift.}$$

**Definition 2.8.** *If $T,\ S \epsilon \mathcal{B}_h(H)$, we define $T \geq S$ if $\langle Tx, x\rangle \geq \langle Sx, x\rangle$ for any $x \epsilon H$, if i.e. $T - S \epsilon \mathcal{B}_+(H)$.*

## 2.4 Bochner integral

**Definition 2.9.** *A function $f : A \to E$ is $\mu$-Bochner integrable if there exists a sequence of $\mu$-simple funtions $f_n : A \to E$ such that the following two conditions are met:*
*(i)* $\lim_{n\to\infty} f_n = f$ *$\mu$-almost everywhere*
*(ii)* $\lim_{n\to\infty} \int_A ||f_n - f||d\mu = 0$.

We obtain that every $\mu$-simple function is $\mu$-Bochner integrable. For $f = \sum_{n=1}^{N} \mathbf{1}_{A_n} x_n$ we put

$$\int_A f d\mu := \sum_{n=1}^{N} \mu(A_n) x_n.$$

This definition is independent of the representation of $f$. If $f$ is $\mu$-Bochner integrable, the limit

$$\int_A f d\mu := \lim_{n\to\infty} \int_A f_n d\mu$$

exists in $E$ and is called the *Bochner integral*, of $f$ with respect to $\mu$. It is easy to check that this definition is independent of the approximating sequence $(f_n)_{n=1}^{\infty}$.

If $f$ is $\mu$-Bocner integrable and $g$ is a $\mu$-version of $f$, then $g$ is $\mu$-Bochner integrable and the Bochner integrals of $f$ anf $g$ agree. In particular, in the definition of the Bochner integral the function $f$ need not be everywhere defined: it suffices that $f$ be $\mu$-almost everywhere defined.

If $f$ is $\mu$-Bochner integrable, then for all $x^* \epsilon E^*$ we have the identity

$$\left\langle \int_A f d\mu, x^* \right\rangle = \int_A \langle f, x^* \rangle \, d\mu.$$

For $\mu$-simple functions this is trivial, and the general case follows by approximating $f$ with $\mu$ simple functions.

**Definition 2.10.** *A function $f : A \to E$ is strongly $\mathcal{A}$ -measurable if there exists a sequence of $\mathcal{A}$ -simple functions $f_n : A \to E$ such that $\lim_{n\to\infty} f_n = f$ pointwise on $A$.*

**Proposition 2.3.** *A strongly $\mu$-measurable function $f : A \to E$ is $\mu$-Bochner integrable if and only if*

$$\int_A ||f|| d\mu < \infty,$$

*and in this case we have*

$$\left|\left| \int_A f d\mu \right|\right| \leq \int_A ||f|| d\mu.$$

**Theorem 2.2** (Dominated Convergence Theorem)**.** *Let $f_n : A \to E$ be a sequence of functions, each of which is $\mu$-Bochner intagrable. Assume that there exists a function $f : A \to E$ and a $\mu$-Bochner function $g : A \to \mathbb{K}$ such that:*
*(i)* $\lim_{n\to\infty} f_n = f$ *$\mu$-almost everywhere*
*(ii)* $||f_n|| \leq |g|$ *$\mu$-almost everywhere.*

*Then $f$ is $\mu$ Bochner integrable and we have*

$$\lim_{n\to\infty}\int_A ||f_n - f||d\mu = 0.$$

*In particular we have*

$$\lim_{n\to\infty}\int_A f_n d\mu = \int_A f d\mu.$$

It is immediate from the definition of the Bochner integral that if $f : A \to E$ is $\mu$-Bochner integrable and $T$ is a bounded linear operator from $E$ into another Banach space $F$, then $Tf : A \to F$ is $\mu$-Bochner integrable and

$$T\int_A f d\mu = \int_A Tf d\mu.$$

This identity has a useful extension to a suitable class of unbounded operators. A linear operator $T$, defined on a linear subspace $\mathcal{D}(T)$ of $E$ and taking values in another Banach space $F$, is said to be *closed* if its *graph*

$$\mathcal{G}(T) := \{(x, Tx) : x \,\epsilon\, \mathcal{D}(T)\}$$

is a closed subspace of $E \times F$. If $T$ is closed, then $\mathcal{D}(T)$ is a Banach space with respect to the *graph norm*.

## 2.5 Sazonov's topology

**Definition 2.11.** *Let X locally convex random vector space and X' is dual. $\tau_S$(X',X) is called Sazanov's topology of X' and produced by the family seminorm $\{q_R, R \,\epsilon\, \mathcal{R}\}$ where $q_r(x') = \sqrt{< Rx', x' >}, x' \,\epsilon\, X'$. The random vector space $(X', \tau_S(X', X))$ is Hausdorff locally convex.*

**Theorem 2.3** (Sazonov)**.** *Let X hilbert space. A function $\mathcal{X} : H \to \mathbb{C}$ is characterized function of a normal measure probability $\mu$ on $(X, \mathcal{B}_X)$ if and only if $\mathcal{X}$ is positive defined with $\mathcal{X}(0) = 1$ and continuous for Sazonov topology $\tau_S(X)$.*

## 2.6 Martingales

In this section we will see the martingale processes which is very important for the stochastic analysis.

**Definition 2.12.** *A filtration is a family of $\sigma$-algebras $\mathcal{F}_t$ such that*

$$s \leq t \quad \Rightarrow \quad \mathcal{F}_s \subset \mathcal{F}_t$$

The $\sigma$-algebra can be consider as an information which is known until the time $t$.

**Definition 2.13.** *A family of random variables $X_t$ is called adapted to filtration $\mathcal{F}_t$, if $X_t$ is $\mathcal{F}_t$- measurable for any $t$.*

I.e. all the information of the stochastic variable $X_t$ until the time $t$ contains to $\sigma$-algebra $\mathcal{F}_t$.

**Definition 2.14.** *Let $(\Omega, \mathcal{F}, P)$ a probability space, $\mathcal{F}_t$ a filtration to $\mathcal{F}$ $(\mathcal{F}_t \subset \mathcal{F})$ and $X_t$ a family of real, integrable $(\mathbb{E}[|X_t|] < \infty)$ random variables which is adapted to filtration $\mathcal{F}_t$.*
*(i) The family $X_t$ is martingale if*

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \ \ mboxa.s. \ \ s \leq t.$$

*(ii) The family $X_t$ is supermartingale if*

$$\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s \ \ mboxa.s. \ \ s \leq t.$$

*(iii)The family $X_t$ is submartingale if*

$$\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s \ \ a.s \ \ s \leq t.$$

Martingales have also the following property:

**Theorem 2.4.** *If $X_t$ is a martingale, then*

*(i)*$\mathbb{E}[X_t] = \mathbb{E}[X_0]$

*(ii)*$\mathbb{E}[X_t - X_0] = 0$

Now we discuss about optional stopping times. The optional stopping informs us about what can be happen if we stop a martingale or a super(sub)martingale to a stopping time $T$. It is a useful method for stochastic analysis.

**Definition 2.15.** *If $T$ is a stopping time, then we can define the stopping process $X_t^T := X_{t\wedge T}$.*

**Theorem 2.5.** *(i) If $X_t$ is a martingale to filtration $\mathcal{F}_t$ and $T$ is a stopping time to the same filtration, then the stopping process $X_t^T = X_{t\wedge T}$ is also martingale to the same filtration.So, it holds that $\mathbb{E}[X_{t\wedge T}] = \mathbb{E}[X_0]$.*
*(ii)If $X_t$ is a super(sub)martingale, then the stopping process is also a super(sub)martingale. It holds also $\mathbb{E}[X_{t\wedge T}] \leq (\geq)\mathbb{E}[X_0]$.*

## 2.7 Schwarz space

Consider the Euclidean space $\mathbb{R}^n$, $n \geq 1$ with $x = (x_1, ..., x_n) \epsilon \mathbb{R}^n$ and with $|x| = \sqrt{x_1^2 + ... + x_n^2}$ and scalar product $(x, y) = \sum_{j=1}^n x_j y_j$. The open ball of radius $\delta > 0$ centered at $x \epsilon \mathbb{R}^n$ is denoted by

$$U_\delta(x) := \{y \epsilon \mathbb{R}^n : |x - y| < \delta\}.$$

Recall the Cauchy-Bunjakovsky inequality

$$|(x, y)| \leq |x||y|.$$

Following L.Schwartz we call an n-tuple $a = (a_1, ..., a_n)$, $a_j \epsilon \mathbb{N} \cup \{0\} \equiv \mathbb{N}_0$ an $n$-dimensional multi-index. Denote

$$|a| = a_1 + ... + a_n, \quad a! = a_1! \cdots a_n!$$

and $x^a = x_1^{a_1} \cdots x_n^{a_n}$, $0^0 = 1$, $0! = 1$. Moreover, multi-indices $a$ and $\beta$ can be ordered according to

$$a \leq \beta$$

if and only if $a_j \leq \beta_j$ for all $j = 1, 2, ..., n$. Let us also introduce a shorthand notation

$$\partial^a = \partial_1^{a_1} \cdots \partial_n^{a_n}, \quad \partial_j = \frac{\partial}{\partial x_j}.$$

**Definition 2.16.** *The Schwarz space $\mathcal{S}(\mathbb{R}^n)$ of rapidly decaying functions is defined as*

$$\mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) : |f|_{a,\beta} := \sup_{x \in \mathbb{R}^n} \left|x^a \partial^\beta f(x)\right| < \infty \text{ for any } a, \beta \in \mathbb{N}_0^n.\}$$

The following propetries of $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$ are readily verified.

(i) $\mathcal{S}$ is a linear space.

(ii) $\partial^a : \mathcal{S} \to \mathcal{S}$ for any $a \geq 0$.

(iii) $x^\beta : \mathcal{S} \to \mathcal{S}$ for any $\beta \geq 0$.

(iv) If $f \in \mathcal{S}(\mathbb{R}^n)$ then $|f(x)| \leq c_m(1 + |x|)^{-m}$ for any $m \in \mathbb{N}$. The converse is not true.

(v) It follows that $\mathcal{S}(\mathbb{R}^n) \subset L_p(\mathbb{R}^n)$ for any $1 \leq p \leq \infty$.

**Example 2.2.** *1) $f(x) = e^{-a|x|^2} \in \mathcal{S}$ for any $a > 0$.*

*2) $f(x) = e^{-a(1+|x|^2)^a} \in \mathcal{S}$ for any $a > 0$.*

*3) $f(x) = e^{-|x|} \notin \mathcal{S}$.*

*4)* $C_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, *where* $C_0^\infty(\mathbb{R}^n) = \{f \,\epsilon\, C^\infty(\mathbb{R}^n) : suppf \text{ is compact in } \mathbb{R}^n\}$ *and* $suppf = \overline{\{x \,\epsilon\, \mathbb{R}^n : f(x) \neq 0\}}$.

**Definition 2.17.** *The space of Schwartz test function of rapid decrease consists of those* $\phi : \mathbb{R}^n \to \mathbb{C}$ *such that for every* $a,\ \beta \,\epsilon\, \mathbb{N}_0^n$

$$\sup_{x \epsilon \mathbb{R}^n} |x^\beta D^a \phi(x)| < \infty \tag{2.1}$$

*it is the space* $\mathcal{S}(\mathbb{R}^n)$.

From (2.2) we construct norms on $\mathcal{S}(\mathbb{R}^n)$ :

$$||\phi||_k = \max_{|a|+|\beta| \leq k} \sup_{x \epsilon \mathbb{R}^n} |x^a D^\beta \phi(x)|. \tag{2.2}$$

It is straightforward to check the conditions for a norm:
(i) $||\phi||_k \geq 0,\ \ ||\phi||_k = 0 \Leftrightarrow \phi \equiv 0$
(ii) $||t\phi||_k = |t|||\phi||_k,\ t \,\epsilon\, \mathbb{C}$
(iii) $||\phi + \psi||_k \leq ||\phi||_k + ||\psi||_k \ \forall\ \phi,\ \psi \,\epsilon\, \mathcal{S}(\mathbb{R}^n)$

The space $\mathcal{S}(\mathbb{R}^n)$ is not a normed space because $|f|_{a,\beta}$ is only a semigroup for $a \geq 0$ and $\beta > 0$ i.e. the condition

$$|f|_{a,\beta} = 0 \ \text{ if and only if } \ \ f = 0$$

fails to hold for e.g. constant function $f$. But the space $(\mathcal{S}, \rho)$ is a metric space if the metric $\rho$ is defined by

$$\rho(f,g) = \sum_{a,\beta \geq 0} 2^{-|a|-|\beta|} \cdot \frac{|f-g|_{a,\beta}}{1 + |f-g|_{a,\beta}}$$

**Theorem 2.6** (Completeness)**.** *The space* $\mathcal{S}(\mathbb{R}^n)$ *is a complete space i.e. every Cauchy sequences converges.*

*Proof.* Let $\{f_k\}_{k=1}^{\infty}$, $f_k \,\epsilon\, \mathcal{S}$, be a Cauchy sequence, that is, for any $\varepsilon > 0$ there exists $n_0(\varepsilon)\,\epsilon\,\mathbb{N}$ such that

$$\rho(f_k, f_m) < \varepsilon, \quad k, m \geq n_0(\varepsilon).$$

It follows that

$$\sup_{x\epsilon K} \left|\partial^{\beta}(f_k - f_m)\right| < \varepsilon$$

for any $\beta \geq 0$ and for any compact set $K$ in $\mathbb{R}^n$. It means that $\{f_k\}_{k=1}^{\infty}$ is a Cauchy sequence in the Banach space $C^{|\beta|}(K)$. Hence there exists a function $f\,\epsilon\, C^{|\beta|}(K)$ such that

$$\lim_{k\to\infty} f_k = f.$$

That's why we may conclude that our function $f\,\epsilon\, C^{\infty}(\mathbb{R}^n)$. It only remains to prove that $f\,\epsilon\,\mathcal{S}$. It is clear that

$$\sup_{x\epsilon K} |x^a \partial^{\beta} f| \leq sup_{x\epsilon K} |x^a \partial^{\beta}(f_k - f)| + \sup_{x\epsilon K} |x^a \partial^{\beta} f_k|$$

$$\leq C_a(K) \sup_{x\epsilon K} |\partial^{\beta}(f_k - f)| + \sup_{x\epsilon K} |x^a \partial^{\beta} f_k|.$$

Taking $k \to \infty$ we obtain

$$\sup_{x\epsilon K} |x^a \partial^{\beta} f| \leq \limsup_{k\to\infty} |f_k|_{a,\beta} < \infty$$

The last inequality is valid since $\{f_k\}_{k=1}^{\infty}$ is a Cauchy sequence, so that $|f_k|_{a,\beta}$ is bounded. The last inequality doesn't depend on $K$ either. That's why we may conclude that $|f|_{a,\beta} < \infty$ or $f\,\epsilon\,\mathcal{S}$.

□

# Chapter 3

# Gaussian Measures in Hilbert Spaces

## 3.1 Gaussian Measures and Properties

**Definition 3.1.** *A measure $\mu$ on $(\mathbb{R}^n, \mathcal{B}^n)$ is called a Gaussian measure on $\mathbb{R}^n$ with parameters $a \epsilon \mathbb{R}^n$ and non negative symmetric $n \times n$ matrix $\Sigma$ if defined by the*

$$\mu(B) = \rho(T^{-1}(B)), B\epsilon\mathcal{B}^n$$

*where*

$$T(x) = a + \Sigma^{\frac{1}{2}}x, x\epsilon\mathbb{R}^n \text{ and } \rho(B) = \int_B (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}|x|^2} dx \ ,$$

$$B \epsilon \mathcal{B}^n \text{normal probability measure.}$$

It is obvious that the measure $\rho$ is Gaussian with parameters $a = 0$ and $\Sigma = I$.

When $\Sigma$ is a positive defined matrix, then the function T is $1 - 1$ and so $T^{-1}(B) = \Sigma^{-\frac{1}{2}}(B - a)$.
In this case with the transform $x = \Sigma^{-\frac{1}{2}(y-a)}$ follows that:

$$\mu(B) = \int_{T^{-1}(B)} (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}|x|^2} \, dx = \int_B (2\pi)^{-\frac{n}{2}} (det\Sigma)^{-\frac{1}{2}} e^{-\frac{1}{2}(\Sigma^{-1}(y-a), y-a)} \, dy$$

so, when $\Sigma$ is positive defined,the Gaussian measure with parameters $a$ and $\Sigma$ has density:

$$d(y) = \frac{1}{(2\pi)^{\frac{n}{2}}(det\Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}(\Sigma^{-1}(y-a),y-a)}, y\epsilon\mathbb{R}^n$$

When $\Sigma$ is degenerated (i.e. $(\Sigma x, x) = 0$ for somebody $x \neq 0$), the Gaussian measure does not have density to Lebesgue measure. (We can find set $A$ with $A \bigcap T(\mathbb{R}^n) = \emptyset$ and positive Lebesgue measure, so $\mu(A) = \rho(T^{-1}(A)) = \rho(\emptyset) = 0$)

If we consider now the function $\phi$ which defined by $\Sigma^{\frac{1}{2}}$ and with the suitable change of variable we have

$$\int_{\mathbb{R}} e^{it\cdot(\phi(z)+a)} d(\rho(z)) = \int e^{it\cdot y} d\mu(y) = \hat{\mu}(t), \quad t \,\epsilon\, \mathbb{R}^n$$

Also, we know that $t \cdot \phi(z) = z\phi^T(t)$ and now the first integral is

$$e^{it\cdot a} \int_{\mathbb{R}^n} e^{it\cdot\phi(z)} d\rho(z) = e^{it\cdot a} \int_{\mathbb{R}^n} e^{iz\cdot\phi^T(t)} d\rho(z).$$

Hence, $\hat{\mu}(t) = e^{it\cdot a}\hat{\rho}\left(\phi^T(t)\right), \quad t \,\epsilon\, \mathbb{R}^n$.
However, $\hat{\rho}\left(\phi^T(t)\right) = e^{-\frac{1}{2}|\phi^T(t)|} = e^{-\frac{1}{2}(\Sigma t,t)}, \quad t \,\epsilon\, \mathbb{R}^n$.
This form for the characteristic function of Gauss measure always applies , even when $\Sigma$ is degenerated.
Particularly, when $n = 1$ is $\Sigma = (\sigma)$ with $\sigma \geq 0$ and for any $a \,\epsilon\, \mathbb{R}$ the Gaussian measure with parameters $a, \sigma$ is

$$\mu(B) = \begin{cases} \delta_a(B) & \text{if } \sigma = 0 \\ \int_B \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma}(x-a)^2} dx & \text{if } \sigma > 0 \end{cases}$$

In this case, for $n = 1$

$$\int_{\mathbb{R}} y d\mu(y) = a \ , \quad \int_{\mathbb{R}} y^2 d\mu(y) - a^2 = \sigma.$$

**Definition 3.2.** *A measure $\mu$ on $(H, \mathcal{B}(H))$ is called a Gaussian measure if for any $h \epsilon H$, there exists $m_h \epsilon \mathbb{R}$ and $q_h > 0$ such that :*

$$\mu\{x \epsilon H : \langle h, x \rangle_H \epsilon F\} = \mathcal{N}(m_h, q_h)(F), F \epsilon \mathcal{B}(\mathbb{R})$$

If for any $h \epsilon H$, $m_h$ is equal to zero, then $\mu$ is called a *symmetric Gaussian measure.*

**Proposition 3.1.** *For a Gaussian measure $\mu$ on $(H, \mathcal{B}(H))$ , there exists $m \epsilon H$ and a symmetric non-negative bounded operator $Q \epsilon \mathcal{L}(H)$ such that*

$$\langle m, h \rangle_H = \int_H \langle h, x \rangle_H \, \mu(dx), h \epsilon H, \tag{3.1}$$

$$\langle Qh_1, h_2 \rangle_H = \int_H \langle h_1, x \rangle_H \, \mu(dx) - \langle m, h_1 \rangle_H \langle m, h_2 \rangle_H , h_1, h_2 \epsilon H. \tag{3.2}$$

*Proof.* We show that the functionals

$$h \mapsto \int_H < h, x >_H \, \mu(dx), \quad h \, \epsilon \, H,$$

$$< Qh_1, h_2 >_H = \int_H < h_1, x >_H < h_2, x >_H \, \mu(dx) \text{ on } H \times H,$$

are well defined and continuous. It follows from the definition of a Gaussian measure that for any $h \, \epsilon \, H$, the mapping $v(x) =< h, x >_h$ is a real-valued Gaussian variable on $H$ with the law $\mathcal{L}_v = \mathcal{N}(m_h, q_h)$.
In particular, it is integrable and

$$\int_h < h, x >_H \, \mu(dx) = \int_{\mathbb{R}} y d\mathcal{N}(m_h, q_h)(y) = m_h.$$

Therefore, the first functional is well defined. taking into account the inequality

$$| < h_1, x >_H < h_2, x >_H | \leq | < h_1, x >_H |^2 + | < h_2, x >_H |^2, \quad h_1, h_2, x \, \epsilon \, H,$$

and the integrability of $| < h, x >_H |^2$ for any $h \epsilon H$, we obtain that the second functional is also well defined.
Let us study properties of the introduced transformations. Let $n \epsilon \mathbb{N}$ and $k = 1, 2$. We set

$$U_n^k = \{h \epsilon H \ : \ \int_H | < h, x >_H |^k \mu(dx) \leq n\}.$$

Then $H = \bigcup_{n-1}^{\infty} U_n^k \ , \ k = 1, 2$.Since $H$ is a complete metric space, by the Baire category argument, there exists $n_0 \epsilon \mathbb{N} \ , \ h_0 \epsilon U_{n_0}^k$, and $r_0 > 0$ such that $B(h_0, r_0) \subset U_{n_0}^k$. Hence,

$$\int_H | < h_0 + y, x >_H |^k \mu(dx) \leq n_0, \ \ e \epsilon B(0, r_0),$$

and for any $y \epsilon B(0, r_0)$,

$$\int_H | < y, x >_H |^k \mu(dx) \leq 2^k \int_H | < h_0 + y, x >_H |^k \mu(dx) + 2^k \int_H | < h_0, x >_H |^k \mu(dx) \leq 2^{k+1}$$

Let $h \epsilon H$.From the last estimate for $y = r_0 h / ||h||_H$, we have that these functionals are $k$-linear symmetric continuous forms:

$$\int_H | < h, x >_H |^k \mu(dx) \leq \frac{2^{k+1}}{r_0^k} ||h||_H^k, \ \ k = 1, 2.$$

Therefore, by the Riesz representation theorem, the existence of the vector $m$ and the bounded symmetric operator $Q$ for the introduced functionals is proved. Finally, the obvious inequality

$$\left( \int_H < h, x >_H \mu(dx) \right)^2 \leq \int_H < h, x >_H^2 \mu(dx)$$

and the representation

$$< Qh, h >_H = \int_H < h, x >_H^2 \mu(dx) - \left( \int_H < h, x >_H \mu(dx) \right)^2$$

imply the non negativity of $Q$. □

## 3.2 Covariance operator of Gaussian measures

If $\mu$ is a Gaussian measure on $(H, \mathcal{B}(H))$, then $m$ defined by (2.1) is called the *mean* of $\mu$, and $Q$ defined by (2.2) is called the *covariance operator* of $\mu$. The characteristic function of a Gaussian measure $\mu$ on $(H, \mathcal{B}(H))$ is $\hat{\mu} : H \to \mathbb{R}$ and given by

$$\hat{\mu}(\lambda) = \int_H e^{i(\lambda,m)_H} e^{-\frac{1}{2}(Q\lambda,\lambda)_H} \quad , \lambda \epsilon H. \tag{3.3}$$

This formula for the characteristic function of a Gaussian measure $\mu$ formally looks like the characteristic function in the $\mathbb{R}^n$-valued case, but it turns out that not every bounded operator $Q$ can be the covariance operator of an $H$-valued Gaussian measure. In particular, we demonstrate that an $H$-valued Gaussian measure cannot have the characteristic function with $Q = I$. Let $\{\lambda_n\}$ be an orthonormal basis in $H$ and let $Q = I$; then

$$\int_H e^{i\langle \lambda_n, x\rangle_H} \mu(dx) = e^{i\langle \lambda_n, x\rangle_H} e^{-\frac{1}{2}\|\lambda_n\|_H^2}. \tag{3.4}$$

Since the Fourier coefficients $\langle \lambda_n, x\rangle_H \to 0 \ \ as \ \ n \to \infty$, in the left-hand side of (2.4) we have $\hat{\mu}(\lambda_n) \to 1$, but in the right-hand side,

$$e^{i\langle \lambda_n, x\rangle_H} e^{-\frac{1}{2}\|\lambda_n\|_H^2} \to e^{-\frac{1}{2}} \quad as \ \ n \to \infty.$$

This contradiction takes place for any $Q$ such that $\langle Q\lambda_n, \lambda_n\rangle$ does not tend to zero.
The following proposition describes the structure of covariance operator of a Gaussian measure on a Hilbert space.

**Proposition 3.2.** *The covariance operator of a symmetric Gaussian measure on $(H, \mathcal{B}(H))$ is a trace class operator.*

*Proof.* Let $\mu$ be a symmetric Gaussian measure with covariate operator Q. In this case, the characteristic function of $\mu$ has the form

$$\hat{\mu}(h) = \int_H e^{\langle h,x\rangle_H \mu(dx)} = e^{-\frac{1}{2}\langle Qh,h\rangle_H}, \quad h \,\epsilon\, H.$$

For an arbitary $c > 0$, we consider

$$1 - e^{-\frac{1}{2}\langle Qh,h\rangle_H} = \int_H (1 - cos\,\langle h,x\rangle_H\, \mu(dx) = 2\int_H sin^2 \frac{\langle h,x\rangle_H}{2}\mu(dx) \leq$$

$$\frac{1}{2}\int_{\|x\|_H \leq c} \langle h,x\rangle_H^2\, \mu(dx) + 2\mu\{x\,\epsilon\, H : \|x\|_H > c\}, \quad h\ \epsilon H.$$

The bounded linear operator $Q_c$ defined by

$$\langle Q_c h, h\rangle_H = \int_{\|x\|_H \leq c} \langle h,x\rangle_H^2\, \mu(dx), \quad h\,\epsilon\, H,$$

is a trace class operator:

$$Tr(Q_c) = \sum_{j=1}^{\infty} \langle Q_c e_j, e_j\rangle_H = \sum_{j=1}^{\infty} \int_{\|x\|_H \leq c} \langle e_j,x\rangle_H^2\, \mu(dx) = \int_{\|x\|_H \leq c} \|x\|_H^2 \mu(dx),$$

where $\{e_j\}$ an orthonormal basis in $H$.Therefore, for any $h\,\epsilon\, H$ and $c > 0$,

$$1 - e^{-\frac{1}{2}\langle Qh,h\rangle_H} \leq \frac{1}{2}\langle Q_c h, h\rangle_H + 2\mu\{x\,\epsilon\, H : \|x\|_H > c\}. \tag{3.5}$$

To prove that $Q$ is a trace class operator, we first find $\beta > 0$ such that the fulfillment of the condition

$$\langle Qh_1, h_1\rangle_H \leq \beta \tag{3.6}$$

Therefore, let $\langle Q_c h_1, h_1\rangle_H = 1$.Then we have from (2.5) that

$$e^{\frac{1}{2}\langle Qh_1,h_1\rangle_H} \leq \left(\frac{1}{2} - 2\mu\{x\,\epsilon\, H : \|x\|_H > c\}\right)^{-1} = a;$$

therefore,(2.6) holds with $\beta = 2lna$ provided that $\mu\{x\,\epsilon\, H : \|x\|_H > c\} < \frac{1}{4}$. Now, let $h$ be an arbitrary element of $H$. From (2.6) , for
$h_1 = \frac{h}{\sqrt{\langle Q_c h,h\rangle_h}} = \frac{h}{\sqrt{a_h}}$ we obtain

$$\langle Qh, h\rangle_H = a_h \langle Qh_1, h_1\rangle_H \leq a_h\beta \langle A_c h_1, h_1\rangle_H = \beta \langle Q_c h, h\rangle_H .$$

Therefore, $Q \leq \beta Q_c$ and $Q$ is a trace class operator.

□

The following theorem demonstrates a relationship between a Gaussian measure on $(H, \mathcal{B}(H))$ and an $H$-valued Gaussian random variable.

**Theorem 3.1.** *For any positive symmetric trace-class operator $Q$ in $H$ and $m \epsilon H$, there exists an $H$-valued Gaussian random variable with expectation $m$ and covariance operator $Q$. Its distribution law is a Gaussian measure on $(H, \mathcal{B}(H))$ with mean $m$ and covariance operator $Q$.*

*Proof.* Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\xi_j$ be a sequence of independent real-valued random variables distributed by $\mathcal{N}(0,1)$. Consider

$$u = m + \sum_{j-1}^{\infty} \sqrt{\lambda_j} \xi_j c_j,$$

where $\{e_j\}$ is an orthonormal basis in $H$ and $\lambda_j = \langle Qe_j, e_j \rangle_H$. This series is convergent in $L_2(\Omega, \mathcal{F}, P; H)$, since

$$\mathbb{E}\left(\sum_{j-1}^{\infty} (\sqrt{\lambda_j}\xi_j)^2\right) = \lim_{n\to\infty} \mathbb{E} \sum_{j-1}^{n} (\sqrt{\lambda_j}\xi_j)^2 = \lim_{n\to\infty} \sum_{j-1}^{n} \lambda_j \mathbb{E}(\xi_j)^2 = \sum_{j-1}^{\infty} \lambda_j = Tr(Q).$$

Let $h \epsilon H$. Consider

$$\mathbb{E}(e_H^{i\langle h,u\rangle_H}) = e^{i\langle h,m\rangle_H} \lim_{n\to\infty} \mathbb{E}\left(e^{i\sum_{j=1}^{n} \sqrt{\lambda_j}\xi_j \langle h,e_j\rangle_H}\right)$$

$$= e^{i\langle h,m\rangle_H} \lim_{n\to\infty} e^{-\frac{1}{2}\sum_{j=1}^{n} \lambda_j \langle h,e_j\rangle_H^2} = e^{i\langle h,m\rangle_H - \frac{1}{2}\langle Qh,h\rangle_H}$$

Thus, $u$ is a Gaussian random variable, $\mathbb{E}(u) = m$, and for any $h_1, h_2 \epsilon H$,

$$\langle Cov(u)h_1, h_2\rangle_H = \lim_{n\to\infty} \mathbb{E}\left(\sum_{j-1}^{n} \sqrt{\lambda_j}\xi_j \langle c_j, h_1\rangle_H \sum_{k-1}^{n} \sqrt{\lambda_k}\xi_k \langle c_k, h_2\rangle_H\right) =$$

$$\langle Qh_1, h_2\rangle_H .$$

Let us show that the distribution law of the $H$-valued Gaussian random variable $u$ is a Gaussian measure on $(H, \mathcal{B}(H))$.The distribution law of any random variable $u$ defines a measure $\mu$ on $(H, \mathcal{B}(H))$ :

$$\mu(G) = \mathcal{L}_u(G) = P\{\omega : u(\omega) \epsilon G\}, \quad G \epsilon \mathcal{B}(H).$$

Therefore,we have constructed a Gaussian random variable by a sequence of independent real-valued Gaussian random variables distributed by $\mathcal{N}(0,1)$. □

The following theorem characterizes completely the functions which can be the characteristic function for a Gaussian measure in Hilbert space .
First, we will see a useful lemma.

**Lemma 3.1.** *Let $\mu$ be a normal probability measureoin Hilbert space X, such that $\int_X \|x\|^2 d\mu x < +\infty$. Then the correlation operator S for measure $\mu$ is positive, continuous and nuclear.*

**Theorem 3.2** (E.Mourier)**.** *Let $X$ be a Hilbert space and $\mu$ probability measure on $(X, \mathcal{B}(X))$. Then the measure $\mu$ is a normal Gaussian measure if and only if the characteristic function $\hat{\mu}$ is of the form :*

$$\mathcal{X}(x) = e^{i(m,x)-\frac{1}{2}(Rx,x)}, \quad x \,\epsilon\, X \tag{3.7}$$

*where, $m \,\epsilon\, X$ and $R$ is a symmetric, positive nuclear operator on $X$.*

*Proof.* ($\Rightarrow$)Let $\mu$ be a normal Gaussian measure. The characteristic function is of the form (2.7), where $m \,\epsilon\, X$ and $R$ is its covariance operator.It suffices to prove that the operator $R$ is nuclear.
According to Sazonov theorem () the characteristic function $\hat{\mu}$ is continuous at $x = 0$ for Sazonov topology $\tau_S(X)$. So, for an arbitrary $\epsilon > 0$ there exists an operator $S \,\epsilon\, \mathbb{S}(X)$ such that:$|1 - \hat{\mu}(x)| < 1 - e^{-\frac{1}{2}}\epsilon$ when $x \,\epsilon\, X$ with $(Sx, x) < 1$.However, $1 - Re\hat{\mu}(x) \leq |1 - \hat{\mu}(x)|$ and $Re\hat{\mu}(x) \leq e^{-\frac{1}{2}(Rx,x)}$ so, $1 - e^{-\frac{1}{2}(Rx,x)} < 1 - e^{-\frac{1}{2}\epsilon}$ when $(Sx, x) < 1$ and finally

$$(Rx, x) < \epsilon, \;\; x \,\epsilon\, X \;\text{ with }\; (Sx, x) < 1 \tag{3.8}$$

Now,for an arbitrary $x_0 \,\epsilon\, X$ and an arbitrary $d > \sqrt{(Sx_0, x_0)}$ we have $\left(S\frac{x_0}{d}, \frac{x_0}{d}\right) = \frac{1}{d^2}(Sx_0, x_0) < 1$ ,so from (2.8) we have $(Rx_0, x_0) < \epsilon$ and

$$(Rx_0, x_0) < \epsilon d^2 \;\text{ for an arbitrary }\; d > \sqrt{(Sx_0, x_0)}.$$

Hence, for an arbitrary $x_0 \epsilon X$

$$(Rx_0, x_0) \leq \epsilon(Sx_0, x_0) \tag{3.9}$$

and for any orthonormal basis $\{e_j,\ j \epsilon I\}$ and for any subsystem $\{e_{j_n}, n \epsilon \mathbb{N}\}$ we have

$$\sum_n (Re_{j_n}, e_{j_n}) \leq \epsilon \sum_n (Se_{j_n}, e_{j_n}) < \infty$$

because the operator S is nuclear.
However, from Lemma 2.1 we have that $(Re_j, e_j) > 0$ up to countable number of indicators $j \epsilon I$. So, from (3.9) we have $\sum_j (Re_j, e_j) < +\infty$ (S nuclear) and $R$ also nuclear.

($\Leftarrow$) Let a function $\mathcal{X}$ be of the form (3.7). According to Theorem 3.1 it suffices to prove that exists normal measure $\mu$ for which $\hat{\mu} = \mathcal{X}$.
We set $h(x) = e^{-\frac{1}{2}(Rx,x)}$. It is obvious that $h(0) = 1$ and is positive definite. We will now show that it is continuous for Sazonov's topology $\tau_S(X)$ on $x = 0$. Actually, for an arbitrary $\epsilon > 0$ we consider the "region"
$V = \{x : (Rx, x) < -2ln(1-\epsilon)\}$ and it is easy to see that $|1 - h(x)| < \epsilon$ for all $x \epsilon V$. So as, $h$ satisfies the requirements of Sazonov's theorem and exists unique normal measure $v$ such that $\hat{v} = h$.
We define now the measure $\mu$ on $(X, \mathcal{B}(X))$ with $\mu(B) = v(\phi^{-1}(B))$ where $\phi(x) = x + m$ or else $\mu(B) = v(B - m)$. The measure $\mu$ is normal and applies:

$$\hat{\mu}(x) = \int_X e^{i(x,y)} d\mu(y) = \int_X e^{i(x,y+m)} dv(y) = e^{i(x,m)} \hat{v}(x) = \mathcal{X}(x) \ \ \forall\, x \epsilon X$$

Hence, $\mathcal{X}$ satisfies the properties of a normal Gaussian measure. □

# Chapter 4

# Wiener Process in Hilbert Space

## 4.1 Basic facts on Wiener Process in $\mathbb{R}^n$

One of more important stochastic process is the Wiener process. The Wiener process, which can somebody find as Brownian motion, has a very important role for the finance mathematics, for models in sontinuous time.

**Definition 4.1.** *The Wiener proces is an $\mathbb{R}$-valued stochastic process $W_t$ with the following properties :*

*(i)If $t_0 < t_1 < ... < t_n$ then the random variables $W_{t_0}, W_{t_1} - W_{t_0}, ..., W_{t_n} - W_{t_{n-1}}$ are independent*

*(ii)If $s,t \geq 0$ ,then*

$$P(W_{s+t} - W_s \, \epsilon \, A) = \int_A \frac{1}{(2\pi t)^{1/2}} exp\left(-\frac{|x|^2}{2t}\right),$$

*where $A$ is a Borel set , i.e. increments of Wiener process follow the Gaussian distribution.*

*(iii)Trajectories of Wiener process are continuous with probability 1, i.e. $t \to W_t$ is continuous function.*

These three properties define one unique stochastic process. Furthermore, for these properties, can we deduce the properties of measure $\mu$, i.e. the Wiener measure is a measure on path space, i.e. a measure on the space of continuous functions. The equation is the finite dimensional distribution.

$$\mu_{t_1,t_2,...,t_n}(A_1 \times A_2 \times ... \times A_n) = \int_{A_1} dx_1 \int_{A_2} dx_2 ... \int_{A_n} dx_n \prod_{i=1}^{n} p(t_i - t_{i-1}, x_{i-1}, x_i)$$

where $x_0 = x,\ t_1 = 0$, and

$$p(t,x,y) = \frac{1}{(2\pi t)^{1/2}} exp\left(-\frac{|x-y|^2}{2t}\right) \tag{4.1}$$

(3.1) help us to find the probability a stochastic process for time points $t_i$ be at subsets $A_i\ \epsilon\ \mathcal{B}(\mathbb{R})$.We can imagine these subsets as intervals in $\mathbb{R}$, so (3.1) is the probability of being Wiener process at specific intervals in $\mathbb{R}$ at specific time point $t_i$. The equation

$$\mu_{t_1,t_2,...,t_n}(A_1 \times A_2 \times ... \times A_n) = P(W_{t_1} \epsilon A_1, W_{t_2} \epsilon A_2, ..., W_{t_n} \epsilon A_n)$$

is called finite dimensional distribution and is very useful for the construction of measure $\mu$. The distribution of $W_t$ depends on the initial point at which it begins its, i.e. the point $W_0$. If $W_0 = x$, then the distribution function is defined by $P_x(W_t\ \epsilon\ A)$ for a Borel set $A$.The mean value or conditional mean value for this measure will be denoted $E_x$ or $E_x[\cdot]$ respectively.

**Definition 4.2.** *An $\mathbb{R}^n$-valued stochastic process satisfying the following conditions*

*(i)$P(\omega : W(0,\omega) = x) = 1$,that is ,process $W(t)$starts at point $x$ a.s.;*

*(ii)$W(t)$ has independent increments, that is, for any $0 \leq t_1 < ... < t_k$, the random values $W(t_1), W(t_2) - W(t_1), ..., W(t_k) - W(t_{k-1})$ are independent with respect to $P$;*

*(iii) $W(t)$ is a Gaussian process, that is, for any $0 \leq t_1 \leq ... \leq t_k$,the random variable $(W(t_1), ..., W(t_k))$is an $\mathbb{R}^{nk}$-valued Gaussian random variable. The expectation of this vector is equal to $m = (x, ..., x)\epsilon\mathbb{R}^{nk}$, and the*

*covariance matrix is equal to*

$$Q_{t_1,\ldots,t_k} = \begin{bmatrix} t_1 I_n & t_1 I_n & t_1 I_n & \cdots & t_1 I_n \\ t_1 I_n & t_2 I_n & t_2 I_n & \cdots & t_2 I_n \\ t_1 I_n & t_2 I_n & t_3 I_n & \cdots & t_3 I_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_1 I_n & t_2 I_n & t_3 I_n & \cdots & t_k I_n \end{bmatrix},$$

*where $I_n$is the identity matrix on $\mathbb{R}^n$;*

*(iv)$W(t)$ has continuous trajectories a.s., that is, the mapping $t \to W(t,\omega), \quad t \geq 0$ is continuous for almost all $\omega \in \Omega$. is called the $\mathbb{R}^n$-valued Wiener process started at point $x$.*

**Example 4.1.** ***The characteristic function of Wiener process***
*The properties of Wiener process help us to find its characteristic function and the characteristic function of its change.*
*First, we calculate the characteristic function of its changes:*

$$\phi_{W_t - W_s}(\lambda) = \mathbb{E}[exp(i\lambda(W_t - W_s))] =$$

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(t-s)}} exp\left(-\frac{x^2}{2(t-s)}\right) exp(i\lambda x) dx$$

*For calculating this interval, it suffices to write the inner quantity completely squared*

$$i\lambda x - \frac{x^2}{2(t-s)} = -\frac{(x - i\lambda(t-s))^2}{2(t-s)} - \frac{\lambda^2(t-s)}{2}$$

*After this we have:*

$$\phi_{W_t - W_s}(\lambda) = exp\left(-\frac{\lambda^2(t-s)}{2}\right) \tag{4.2}$$

*This we can calculate all polynomial moments of variables of Wiener process. This can be done by taking the higher derivatives of $\phi_{B_t - B_s}(\lambda)$ to $\lambda$ and taking $\lambda = 0$. For example:*

$$\mathbb{E}[(W_t - W_s)^4] = 3(t-s)^2$$

*So, if we take $s = 0$ to (3.2) we have:*

$$\phi_{W_t}(\lambda) = exp\left(-\frac{\lambda^2}{2}\right)$$

**Example 4.2.** *Now we want to find the joint distribution for random variables $W_t$ and $W_s$ using expectation $\mathbb{E}[W_t W_s]$.*
*We consider without loss of generality $s < t$. Using properties of Wiener process (i) and (ii) we have:*

$$f_{W_t W_s}(x,y) = p(s,0,x)p(t-s,x,y) = \frac{1}{2\pi\sqrt{s(t-s)}} exp\left(-\frac{x^2}{2s}\right) exp\left(-\frac{(x-y)^2}{2(t-s)}\right)$$

*The mean value of $W_t W_s$ can be an interval of joint distribution:*

$$\begin{aligned}\mathbb{E}[W_t W_s] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyp(s,0,x)p(t-s,x,y)dxdy \\ &= \int_{-\infty}^{\infty} xp(s,0,x)\left(\int_{-\infty}^{\infty} yp(t-s,x,y)dy\right)dx \\ &= \int_{-\infty}^{\infty} x^2 p(s,0,x)dx = s\end{aligned}$$

*So, we conclude to $\mathbb{E}[W_t W_s] = s \wedge t = min(s,t)$*

**Definition 4.3.** *We say that $u(t)$ is*
*(i)measurable if the mapping $u(\cdot,\cdot) : \mathcal{T} \times \Omega \to H$ is $\mathcal{B}(\mathcal{T}) \times \mathcal{F}$ measurable;*
*(ii)stochastically continuous at $t_0$, $t_0 \epsilon \mathcal{T}$, if for any positive numbers $\varepsilon$ and $\delta$, there exists a positive number $\rho$ such that*

$$P(||u(t) - u(t_0)||_H \geq \varepsilon) \leq \delta \quad for\ any \quad t\,\epsilon[t_0 - \rho, t_0 + \rho] \cap \mathcal{T};$$

*$u(t)$ is stochastically continuous in $\mathcal{T}$ if it is stochastically continuous at each point of $\mathcal{T}$;*
*(iii)mean square continuous at $t_0$, $t_0 \epsilon \mathcal{T}$, if*

$$\lim_{t \to t_0} \mathbb{E}[||u(t) - u(t_0)||_H^2] = 0;$$

*$u(t)$ is mean square continuous in $\mathcal{T}$ if it is mean square continuous at each point of $\mathcal{T}$;*
*(iv) continuous with probability 1 (or continuous) if its trajectories $u(\cdot,\omega)$ are continuous almost everywhere with respect to $P$.*

## 4.2 Properties of Wiener process

### 4.2.1 Markov property

The fact that Wiener process has the Markov property means that if we take an arbitrary $s \geq 0$, then $W_{t+s} - W_s$ is a Wiener process which is independent of what happened before the time point $s$. So, Wiener process forgets its past fully and what happened for the time point $s$ and after depends only from the final value of Wiener process, i.e. $W_s$. Furthermore, $W_{t+s} - W_s$ is also a Wiener process with mean value 0 and variance $(t+s) - s = t$, i.e. the process $W_{t+s} - W_t$ is a Wiener process which starts at 0 and is running for time $t$.
Let $\mathcal{F}_s = \sigma(W_u, u \leq s)$ the sigma-algebra which produced by Wiener process until the time point s. This sigma-algebra is the smallest sigma-algebra which make the random variable $W_r, r \leq s$ measurable.Then, $\mathcal{F}_s$ has all the information about what happened to Wiener process up to the time point $s$. If $f$ is bounded function, then $\forall\ x\ \epsilon\ \mathbb{R}^d$ it holds that

$$\mathbb{E}_x[f(W_{t+s} - W_s)|\mathcal{F}_s] = \mathbb{E}_x[f(W_{t+s} - W_s)]] \int_{-\infty}^{\infty} f(y) \frac{1}{\sqrt{2\pi t}} exp\left(-\frac{y^2}{2t}\right) dy$$

because of the independence of Wiener process increments. So,we obtain :

$$\mathbb{E}_x[f(W_{t+s} - W_s)|\mathcal{F}_s] = \mathbb{E}_x[f(W_{t+s} - W_s + W_s)|\mathcal{F}_s]$$

and because of independence of $W_{t+s} - W_s$ from algebra $\mathcal{F}_s$ and the $W_s$ fully known up to the time point $s$, can we consider that we have $W_s = z$, so:

$$\mathbb{E}_x[f(W_{t+s} - W_s + W_s)|\mathcal{F}_s] = \mathbb{E}_x[f(W_{t+s} - W_s + z)]$$

However,is $W_{t+s} - W_s$ also a Wiener process which starts at 0 and is running for time $t+s-s = t$. So, the process $W_{t+s} - W_s + z$ has the same distribution with a Wiener process which starts at point $z$ and is running for time $t$.So,

$$\mathbb{E}_x[f(W_{t+s} - W_s + z)] = \int_{-\infty}^{\infty} f(y+z) \frac{1}{\sqrt{2\pi t}} exp\left(-\frac{y^2}{2t}\right) dy$$

$$= \int_{-\infty}^{\infty} f(y) \frac{1}{\sqrt{2\pi t}} exp\left(-\frac{(y-z)^2}{2t}\right) dy = \mathbb{E}_x[f(W_t)] = \mathbb{E}_{W_s}[f(W_t)]$$

And now we have :

$$\mathbb{E}_x[f(W_{t+s})|\mathcal{F}_s] = \mathbb{E}_{W_s}[f(W_t)] = \int_{-\infty}^{\infty} f(y)\frac{1}{\sqrt{2t}}exp\left(-\frac{(y-W_s)^2}{2t}\right)dy \quad (4.3)$$

This property is Markov property for the Wiener process.
We can also write the Markov property with probabilities.For example:

$$P_x[W_{t+s}\epsilon\ A|\mathcal{F}_s] = P_{W_s}(B_t\ \epsilon\ A)$$

for a Borel set A.
A equivalent form of this Markov property is

$$\mathbb{E}_x[f(W_t)|\mathcal{F}_s] = \mathbb{E}_{W_s}[f(W_{t-s})], s \leq t$$

i.e. if we take the condition expectation for a function of Wiener process at the time point t and we know the history of Wiener process until time point s , it suffices to calculate the same in a new Wiener process which starts at the position where the first Wiener process is at time point $s$,i.e. $W_s$ and is running for time length $t-s$. The full history of Wiener process before the time point $s$ is not pay at all.

**Theorem 4.1** (Markov property)**.** *Let $Y$ a bounded measurable function and $\theta_s$ the shift operator which has the following property:$(\theta_s\omega)(t) = \omega(t+s)$. Then*

$$\mathbb{E}_x[Y \circ \theta_s|\mathcal{F}_s] = \mathbb{E}_{B_s}[Y].$$

We give here the Markov property with this form, because $Y$ can be a variable which depends from $W_t$ but not necessarily from $f(W_t)$.

**Example 4.3.** *We define the function $u(x,t) = \mathbb{E}_x[\int_0^t g(W_r)dr]$, where $g$ is a bounded function. Using the Markov property we can obtain that if $0 < s < t$ then*

$$\mathbb{E}_x[\int_0^t g(W_r)dr|\mathcal{F}_s] = \int_0^s g(W_r)dr + u(t-s, W_s).$$

*First, we write the left side in the form:*

$$\mathbb{E}_x[\int_0^s g(B_r)dr|\mathcal{F}_s] + \mathbb{E}_x[\int_s^t g(W_r)dr|\mathcal{F}_s] = \int_0^s g(W_r)dr + \mathbb{E}_x[\int_s^t g(W_r dr|\mathcal{F}_s)] \quad (4.4)$$

*where we use the fact that* $\int_0^s g(W_r)dr$ *is* $\mathcal{F}_t$*-measurable and also properties of conditional expectation.*
*We define now* $Y = \int_0^{t-s} g(W_r)dr$ *and we can see that*

$$\mathbb{E}_x[\int_s^t g(W_r)dr|\mathcal{F}_s] = \mathbb{E}_x[Y \circ \theta_s|\mathcal{F}_s]$$

*and using the Markov property we obtain*

$$\mathbb{E}_x[\int_s^t g(W_r)dr|\mathcal{F}_s] = \mathbb{E}_{B_s}[Y] = \mathbb{E}_{B_s}[\int_0^{t-s} g(W_r)dr] = u(t-s, W_s).$$

*Replacing this form to (4.4) we obtain the result.*

### 4.2.2 Strong Markov property

The strong Markov property is the fact that Markov property apply also for a specific class random time points,called the **stopping times** .

**Remark** :
(a) An increasing family of $\sigma$-fields $\{\mathcal{F}_t,\ t\ \epsilon\ \mathcal{T}\}$ on $\Omega$ is called filtration. A filtration $\{\mathcal{F}_t,\ t\ \epsilon\ \mathcal{T}\}$ is said to be normal if

$$\{G\ \epsilon\ \mathcal{F}\ :\ P(G) = 0\} \subset \mathcal{F}_0 \quad \text{and} \quad \mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s\ \ \forall\ t\ \epsilon\ \mathcal{T}.$$

(b)If, for any $t\ \epsilon\ \mathcal{T}$, a random variable $u(t)$ is $\mathcal{F}_t$-measurable, then the process $\{u(t),\ t\ \epsilon\ \mathcal{T}\}$ is said to be adapted (to the family $\{\mathcal{F}_t,\ t\ \epsilon\ \mathcal{T}\}$).

**Definition 4.4.** *Let* $(\mathcal{F}_t)_{t\epsilon\ I}$ *a filtration on a set* $\Omega$*, where* $I$ *is a set of indicators (not necessarily discrete). A stopping time about this filtration is a function* $T : \Omega \to I$ *such that*

$$\{T \leq t\}\epsilon\ \mathcal{F}_t,\ \forall\ t\ \epsilon\ I$$

The independence of Wiener process increments applies also for stopping times.

**Definition 4.5.** *The $\sigma$-algebra $\mathcal{F}_T$ where $\tau$ is a stopping time defined as*

$$\mathcal{F}_\tau = A\epsilon\, \mathcal{F}_\infty :\ A \cap \tau \leq t\ \epsilon \mathcal{F}_t,\ \forall t > 0$$

**Theorem 4.2.** *Let $W_t$ a Wiener process and $\tau$ a finite stopping time for $W_t$. Then it holds that $W_{t+T} - W_T$ is a Wiener process independent of algebra $\mathcal{F}_T$.*

*Proof.* We take for this proof an approximation by $T$ with a sequence of stopping times $T_n$ which defined as $T_n = \frac{k}{2^n}$, if $T(\omega)\ \epsilon\ \left[\frac{(k-1)}{2^n}, \frac{k}{2^n}\right)$. This sequence convergences monotically to stop time T,$T_n \downarrow T$.It suffices to prove that

$$\mathbb{E}_x\left[e^{iv(W_{T+t}-W_T)}e^{i\lambda W_{T-\tau}}\right] = \mathbb{E}_x\left[e^{iv(W_{T+t}-W_T)}\right]\mathbb{E}_x\left[e^{i\lambda W_{T-\tau}}\right]$$

The stopping time $\tau$ consider bounded, so we can write

$$\mathbb{E}_x\left[e^{iv(W_{T_n+t}-W_{T_n})}e^{i\lambda W_{T_n-\tau}}\right] = \sum_{k=1}^{\infty}\mathbb{E}_x\left[e^{iv(W_{T_{n+t}}-W_{T_n})}e^{i\lambda W_{T_n-\tau}}; \{T_n = \frac{k}{2^n}\}\right]$$

$$= \sum_{k=1}^{\infty}\mathbb{E}_x\left[e^{iv(W_{T_{n+t}}-W_{T_n})}\right]\mathbb{E}_x\left[e^{i\lambda W_{T_n-\tau}}\right]\mathbf{1}_{\{T_n=\frac{k}{2^n}\}} = \mathbb{E}_x\left[e^{iv(W_{T_{n+t}}-W_{T_n})}\right]\mathbb{E}_x\left[e^{i\lambda W_{T_n-\tau}}\right]$$

We use here the independence of $W_{\frac{k}{2^n}+t} - W_{\frac{k}{2^n}}$ and $W_{\frac{k}{2^n}-\tau},\ \forall\ k$. We can take the limit at $n \to \infty$ and from the Theorem Dominated Converges we can interchange with the expectation and finally have the result.
So, now we prove the independence and we have to prove that $W_{T+t} - W_T$ is a Brownian motion and have the same distribution with the $W_t$. It suffices to calculate the characteristic function of random variable $W_{T+t} - W_T$. We use the same approach process of stop time T.

$$\mathbb{E}_x[e^{i\lambda(W_{T_{n+t}}-W_{T_n})}] = \sum_{k=1}^{\infty}\mathbb{E}_x\left[e^{i\lambda(W_{T_{n+t}}-W_{T_n})\mathbf{1}_{\{T_n=\frac{k}{2^n}\}}}\right]$$

$$= \sum_{k=1}^{\infty} \mathbb{E}_x \left[ e^{i\lambda(W_{\frac{k}{2^n}+t} - W_{\frac{k}{2^n}})} \mathbf{1}_{\{T_n = \frac{k}{2^n}\}} \right]$$

$$= \sum_{k=1}^{\infty} e^{-\frac{\lambda^2 t}{2}} \mathbf{1}_{\{T_n = \frac{k}{2^n}\}} = e^{-\frac{\lambda^2 t}{2}} \mathbf{1}_{\{T < \infty\}}$$

Now, we take the limit as $n \to \infty$ and using the Theorem Dominated Converges we have

$$\mathbb{E}_x[e^{i\lambda(W_{T+t} - W_T)}] = e^{-\frac{\lambda^2 t}{2}}$$

which show us that $W_{T+t} - W_T$ is a Wiener process.

□

Also, we can prove after that, the following theorem.

**Theorem 4.3.** *Let $\theta_s$ the shift operator, $Y$ a bounded measurable function and $T$ is a stop time. Then,*

$$\mathbb{E}_x[f(Y \circ \theta_T)|\mathcal{F}_T] = \mathbb{E}_{W_T} Y$$

*Proof.* For this proof we take the approximation of stopping time $T$ by the sequence $T_n$. It suffices to prove that:

$$\mathbb{E}_x[f(W_{T+t})|\mathcal{F}_T] = \mathbb{E}_{W_T}[f(W_t)]$$

for $f$ continuous and bounded.
If A is a set which belongs to $\sigma$-algebra $\mathcal{F}_{T_n}$, we can write

$$\mathbb{E}_x[f(W_{T_n+t}); A \cap \{T_n = \frac{k}{2^n}\}] = \mathbb{E}_x[f(B_{t+\frac{k}{2^n}}); A \cap \{T_n = \frac{k}{2^n}\}]$$

$$\mathbb{E}_x[\mathbb{E}_{W_{\frac{k}{2^n}}} f(W_t); A \cap \{T_n = \frac{k}{2^n}\}] = \mathbb{E}_x[\mathbb{E}_{W_{T_n}} f(W_t); A \cap \{T_n = \frac{k}{2^n}\}]$$

where we use the Markov property for the Wiener process at time $\frac{k}{2^n}$. Based on this calculation we have

$$\mathbb{E}_x[f(W_{T_n+t}); A \cap \{T < \infty\}] = \sum_{k=1}^{\infty} \mathbb{E}_x[f(W_{T_n+t}); A \cap \{T_n = \frac{k}{2^n}\}]$$

$$= \sum_{k=1}^{\infty} \mathbb{E}_x[\mathbb{E}_{W_{T_n}} f(W_t); A \cap \{T_n = \frac{k}{2^n}\}] = \mathbb{E}_x[\mathbb{E}_{W_{T_n}} f(W_t); A \cap \{T < \infty\}]$$

Now we obtain the limit $n \to \infty$ and using the Theorem Dominated Converges we have

$$\mathbb{E}x[f(W_{T+t}); A \cap \{T < \infty\}] = \mathbb{E}_x[\mathbb{E}_{W_{T_n}}[f(W_t)]; A \cap \{T < \infty\}]$$

From the definition of conditional expectation we obtain finally

$$\mathbb{E}_x[f(W_{T+t})|\mathcal{F}_T] = \mathbb{E}_{W_T}[f(W_t)]$$

General Y can be approximated as $\prod_{i=1}^{n} f_i(W_{t_i})$ in the limit $n \to \infty$.

□

An interesting particular case is

$$\mathbb{E}_x[f(W_{t+T})|\mathcal{F}_T] = \mathbb{E}_{W_T}[W_T],$$

where $f$ is a bounded function. This property informs us that the conditional expectation of function $f$ calculated at position where the Wiener process arrived at time point t+T given the information of Wiener process until stopping time T, is the expectation of the same on a Wiener process which starts at time T and is running for time length $t$. If as $f$ we chose the indicator function from a set, then the Markov property can be written as:

$$P_x[W_{t+T} \, \epsilon \, A|\mathcal{F}_T] = P_{W_T}(W_t)$$

**Example 4.4.** *We want to find the distribution of random times $\tau_a$. First, we have that*

$$P(\tau_a \leq t) = P(S_a \geq a) = P(S_t \geq a, W_t \leq a) + P(S_t \geq a, W_t > a).$$

*We wrote this using*
*(i) the fact that the facts $\{\tau_a \leq t\}$ and $\{S_t \geq a\}$ is identical*
*and*
*(ii)the fact that $\{S_t \geq a\}$ can happen with two ways: either Wiener process came out of set $x \leq a$ before the time point t and the time point t has returned again into this set, or came out of this set before time point t and at time t*

*continue to be out of this set. These two facts exclude one another. However, because of* $\{B_t > a\} \subset \{S_t \geq a\}$ *holds that:*

$$P(S_t \geq a, B_t > a) = P(S_t > a).$$

*Also, because of* $P(S_t \geq a, B_t \leq a - y) = P(B_t \geq a + y)$, *for* $y = 0$ *we obtain:*

$$P(S_t \geq a, B_t <\leq a) = P(B_t \geq a)$$

*So,* $P(\tau_a \leq t) = P(S_t \geq a) = 2P(B_t \geq a) = \frac{2}{\sqrt{2\pi t}} \int_a^\infty \exp\left(-\frac{x^2}{2t} dx\right) = \sqrt{\frac{2}{\pi}} \int_{\frac{a}{\sqrt{t}}}^\infty \exp\left(-\frac{y^2}{2} dy\right)$.
*Now we can calculate the distribution of* $\tau_a$ *and the result is:*

$$f_{\tau_a}(t) = \frac{a}{\sqrt{2\pi}} t^{-\frac{3}{2}} \exp\left(-\frac{a^2}{2t}\right)$$

*So, we find the distribution.*

**Example 4.5.** *With this example we want to find the probability of Wiener process to have at least a zero point on set* $[t_0, t_1]$, *conditional* $W_0 = 0$.

*We can write*

$$\begin{aligned} P[min_{0 \leq u \leq t} W_u \leq 0 | W_0 = a] &= P[max_{0 \leq u \leq t} W_u \geq 0 | W_0 = -a] \\ &= P[max_{o \leq u \leq t} W_u \geq a | W_0 = 0] = P(\tau_a \leq t) \\ &= \frac{a}{\sqrt{2\pi}} \int_0^t \frac{1}{s^{3/2}} \exp\left(-\frac{a^2}{2s} ds\right), \quad a > 0. \end{aligned}$$

*We use here the symmetry of Wiener process around to 0 and the space homogeneity of Wiener process.*

*To calculate the probability of Wiener process to be zero once on set* $(t_0, t_1)$ *conditional* $W_0 = 0$ *we bind on every value* $W_{t_0}$. *If* $W_{t_0} = a$ *the probability of* $W_t$ *to be zero on set* $(t_0, t_1)$ *is equal to* $P(a)$. *So,*

$$p = \int_{-\infty}^\infty P(a) P[W_{t_0} = a | W_0 = 0] da = \sqrt{\frac{2}{\pi t_0}} \int_0^\infty P(a) \exp\left(-\frac{a^2}{2t_0} da\right)$$

*We calculate the integral and finally we have:*

$$p = \frac{2}{\pi} \arctan \sqrt{\frac{t_1 - t_0}{t_0}} = \frac{2}{\pi} \arccos \sqrt{\frac{t_0}{t_1}}$$

### 4.2.3 Martingale property

The theory of Wiener process can be developed by using the results of **martingale** theory.We will prove that Wiener process is a martingale. First, we will see the following theorem.

**Theorem 4.4.** *Let $W_t$ a Wiener proces and $\mathcal{F}_s = \sigma(W_u, u \leq s)$. The following stochastic processes are martingale for the filtration $\mathcal{F}_s$:*
*(i)$W_t$*
*(ii)$(W_t)^2 - t$*
*(iii)$M_t^\lambda = exp\left(\lambda W_t - \frac{\lambda^2 t}{2}\right)$.*

*Proof.* (i)We know that $\mathbb{E}[|W_t|] < \infty$ because the integral $\int_{-\infty}^{+\infty} |x| exp\{-x^2/2t\}dx$ converges . So,it holds that

$$\mathbb{E}[W_t|\mathcal{F}_s] = \mathbb{E}[W_t - W_s + W_s|\mathcal{F}_s] = \mathbb{E}[W_t - W_s|\mathcal{F}_s] + \mathbb{E}[W_s|\mathcal{F}_s] = \mathbb{E}[W_t - W_s] + W_s = W_s$$

We use here the independence of $W_t - W_s$ from $\mathcal{F}_s$ and that the expectation is 0.

(ii)We obtain now

$$\mathbb{E}[W_t^2 - W_s^2|\mathcal{F}_s] = \mathbb{E}[(W_t - W_s)^2|\mathcal{F}_s] + 2W_s\mathbb{E}[W_t - W_s|\mathcal{F}_s] = \mathbb{E}[W_{t-s}^2|\mathcal{F}_s] = t - s$$

So, $\mathbb{E}[W_t^2 - t|\mathcal{F}_s] = W_s^2 - s$

(iii)We now recall that if g is a standard normal random variable, we know that $\mathbb{E}(e^{\lambda g}) = \int_{-\infty}^{\infty} e^{\lambda x} e^{-x^2/2} \frac{dx}{\sqrt{2\pi}}$.
On the other hand, if $s < t$, $\mathbb{E}\left(e^{\sigma W_t - \sigma^2 t/2}|\mathcal{F}_s\right) = e^{\sigma W_s - \sigma^2 t/2}\mathbb{E}\left(e^{\sigma(W_t - W_s)}|\mathcal{F}_s\right)$ because $W_s$ is $\mathcal{F}_s$-measurable. Since $W_t - W_s$ is independent of $\mathcal{F}_s$, it turns out that

$$\mathbb{E}\left(e^{\sigma(W_t - W_s)}|\mathcal{F}_s\right) = \mathbb{E}\left(e^{\sigma(W_t - W_s)}\right) = \mathbb{E}\left(e^{\sigma W_{t-s}}\right) = \mathbb{E}\left(e^{\sigma g\sqrt{t-s}}\right) = exp\left(\frac{1}{2}\sigma^2(t-s)\right)$$

□

**Example 4.6** (Law of exit time of Wiener process)**.** *Using Theorem 4.4 we will find a different way for the distribution of* $\tau_a = \inf\{t : S_t \geq a\}$*, where* $S_t = \sup_{s \leq t} W_s$ *and* $W$ *is dimensional Wiener process which starts at 0.*

*From Theorem 4.4 we can see that for* $\lambda \geq 0$*, the stochastic process* $M_t^\lambda$ *is martingale. Let, consider the stopping martingale* $M_{t \wedge \tau_a}^\lambda$*. This is a martingale, which is bounded from* $e^{\lambda a}$ *and so using the theorem of optimal stopping we can see that:*

$$\mathbb{E}[M_{\tau_a}^\lambda] = \mathbb{E}[M_0^\lambda] = 1.$$

*Also,*

$$\mathbb{E}\left[\exp\left(\lambda a - \frac{\lambda^2 \tau_a}{2}\right)\right] = 1$$

*So,*

$$\mathbb{E}\left[\exp\left(-\frac{\lambda^2 \tau_a}{2}\right)\right] = \exp(-\lambda a).$$

*Now, we consider* $s = \frac{\lambda^2}{2}$ *and we obtain*

$$\mathbb{E}(e^{s\tau_a}) = e^{-\sqrt{2s}a}$$

*However, if* $f(t)$ *is probability density of* $\tau_a$*, then*

$$\mathbb{E}(e^{s\tau_a}) = \int_0^\infty e^{-st} f(t) dt$$

*so,* $\mathbb{E}(e^{s\tau_a})$ *is Laplace transform of probability density of* $\tau_a$*.Inverting Laplace transform and using typical methods we obtain:*

$$f(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right).$$

### 4.2.4 Characterization of Brownian Motion

We will show a way to check a stochastic process is a Wiener process or not. This way is based on a fundamental result of Paul Lévy and characterized the Wiener process using martingale property.

**Theorem 4.5** (Lévy)**.** *Let $X_t, t \geq 0$ a stochastic process and $\mathcal{G}_t = \sigma(X_s, s \leq t)$ the filtration which produced by it. $X_t$ is a Wiener process if and only if the following hold:*

*(i) $X_0 = 0$ a.s.*

*(ii)The trajectories of $X_t$ are continuous functions of time.*

*(iii) $X_t$ is a martingale to filtration $\mathcal{G}_t = \sigma(X_s, s \leq t)$.*

*(iv)$X_t^2 - t$ is a martingale with the respect to filtration $\mathcal{G}_t = \sigma(X_s, s \leq t)$.*

*Proof.* ($\Rightarrow$) This side is simple to prove and in fact is similar with the proof of theorem (3.3)
($\Leftarrow$)Here we will use the characteristic function.If $X_t$ has these four properties, then we can prove that the stochastic process $M_t = exp\left(i\lambda X_{t\wedge T} + \frac{1}{2}\lambda^2(t \wedge T)\right)$ is martingale $\forall\ \lambda$, where $T$ is a (bounded) stop time and the arrangement $i^2 = -1$. For a random $A\ \epsilon\ \mathcal{F}_s, s < t < T$ we have

$$\mathbb{E}[\mathbf{1}_A M_t | \mathcal{F}_s] = \mathbf{1}_A M_s$$

where with the suitable organise we finally take

$$\mathbb{E}[\mathbf{1}_A e^{i\lambda(X_t - X_s)}] = P(A)e^{-\frac{\lambda^2}{2}(t-s)}$$

From this we obtain that $W_t - W_s$ is independent of $\mathcal{F}_s$ and follows the distribution of normal $N(0, t-s)$. So, $X_t$ is a Wiener process.

□

**Example 4.7.** *Using Lévy theorem we prove that stochastic process $X_t = W_{t+T} - W_T$, where $W_t$ is Wiener process and $T$ a deterministic time, is also a Wiener process.*

*The filtration $\mathcal{G}_t$ which produces the stochastic process $X_t$ associated with filtration $\mathcal{F}_t$ which produce the Wiener process. Actually,*

$$\mathcal{G}_t = \sigma(X_s, s \leq t) = \sigma(B_{s+T} - B_T, s \leq t) = \sigma(B_s, s \leq t + T) = \mathcal{F}_{t+T}$$

*The properties (i) and (ii) of Lévy theorem is obvious from the definition of Wiener process.*

*We check now the (iii) property*

$$\mathbb{E}[X_t|\mathcal{G}_s] = \mathbb{E}[X_t - X_s + X_s|\mathcal{G}_s] = X_s + \mathbb{E}[X_t - X_s|\mathcal{G}_s]$$

$$= \mathbb{E}[W_{t+T} - W_{s+T}|\mathcal{F}_{s+T}] + X_s = \mathbb{E}[W_{t+T} - W_{s+T}] + X_s = X_s$$

*And now we check the (iv) property*

$$\mathbb{E}[X_t^2 - t|\mathcal{G}_s] = \mathbb{E}[(X_t - X_s + X_s)^2 - t|\mathcal{G}_s]$$

$$= \mathbb{E}[(X_t - X_s)^2|\mathcal{G}_s] - t + X_s^2$$

$$= \mathbb{E}[(W_{t+T} - W_{s+T})^2|\mathcal{F}_{s+T}] - t + X_s^2$$

$$= (t + T - (s + T)) - t + X_s^2 = X_s^2 - s$$

*where we use the property* $\mathbb{E}[X_t - X_s|\mathcal{G}_s] = 0$.

*So, the stochastic process* $X_t$ *is a Wiener process.*

## 4.3 Properties of trajectories

The trajectories of Wiener process have some characteristic properties, which we will see in this section.

**Theorem 4.6.** *Let* $0 < \epsilon < \frac{1}{2}$ *and* $0 < T < \infty$. *Then, there is a random variable* $N$ *, which depends from* $\epsilon$ *and* $T$*, such that* $\mathbb{E}[N^p] < \infty \; \forall \; p, \; 0 < p < \infty$ *and for this we have also:*

$$|W_t(\omega) - W_s(\omega)| \leq N(\omega)|t - s|^{\frac{1}{2}-\epsilon}, \forall \; s, t \; \epsilon \; [0, T]$$

Kintchin has prove on another result of Wiener process trajectories , which gives us a better bound.

**Theorem 4.7** (Law of the Iterated Logarithm)**.** *For a Wiener process holds the followings with probability 1:*

$$\limsup_{t\downarrow 0} \frac{W_t}{\sqrt{lnln\left(\frac{1}{n}\right)}} = \limsup_{t\uparrow\infty} \frac{W_t}{\sqrt{lnln(t)}} = 1$$

$$\liminf_{t\downarrow 0} \frac{W_t}{\sqrt{lnln\left(\frac{1}{t}\right)}} = \liminf_{t\uparrow\infty} \frac{W_t}{\sqrt{lnln(t)}} = -1$$

With this property we have in fact that the limit $\frac{W_t}{\sqrt{lnln\left(\frac{1}{t}\right)}}$ for $t = 0$ doesn't exist. This idea can be proved with the following theorem.

**Theorem 4.8.** *Let $W_t$ is a Wiener process. Then the stochastic processes*

*(i)*$X_t = \frac{1}{c}W_{c^2t}$

*(ii)*$Y_t = tW_{\frac{1}{t}}$
*is also Wiener process.*

Finally, on more characteristic property of Wiener process is that t is not differentiable nowhere a.s. I.e. is a continuous function but not differentiable.This result had be proved by Paley, Wiener and Zygmund.

**Theorem 4.9.** *The Wiener process $W_t$ is not differentiable a.s. at point t,* $\forall\ t \geq 0$.

The quadratic variation of Wiener process is equal to $t$. For the other side, Wiener process has infinite variation.

**Theorem 4.10.** *The Wiener process has the following properties:*

*(i) The quadratic variation in [0,t] is equal to t.*

*(ii)Wiener process has infinite variation.*

## 4.4 Wiener's process construction

We show now a representation of Wiener process using suitable functions of bases. We first, insert a base of the space of the continuous functions, the base of **Haar** functions.

**Definition 4.6.** *Functions* $\phi_{ij}(t) : [0,1] \to \mathbb{R}$ *which defined as*

$$\phi_{ij}(t) = \begin{cases} 2^{(i-1)/2}, & \frac{2j-2}{2^i} \leq t < \frac{2j-1}{2^i} \\ -2^{(i-1)/2}, & \frac{2j-1}{2^i} \leq t < \frac{2j}{2^i} \\ 0, & \textit{for any else } t \end{cases}$$

*for* $i = 1, 2, ...$ *and* $j = 1, 2, ..., 2^{i-1}$ *is called* **Haar functions**.

The Haar functions have some useful properties:

(i) The Haar functions is rectangular on [0,1] to $< f, g >= \int_0^1 \bar{g} f dt$

(ii)The Haar functions is a fully orthonormal system on
$\mathcal{L}^2[0,1] = \{f :< f, f >< \infty\}$.

(iii)Linear combinations of Haar functions is thick to set of continuous functions.So, because the set of continuous functions is thick on $\mathcal{L}^2[0,1]$, linear combinations of Haar functions are thick on $\mathcal{L}^2[0,1]$.

So, we can see that every continuous function on [0,1] and every function which belong on $\mathcal{L}^2[0,1]$ can be approximated as linear combination of Haar function. So, Wiener process is also a case of this and can be represented with these functions.

**Approach to Wiener Process**

We consider

$$\psi_{ij}(t) = \int_0^t \phi_{ij}(s) ds$$

and also a sequence from independent with same distribution normal variables $Y_{ij}$ with $\mathbb{E}[Y_{ij}] = 0$ and $\mathbb{E}[Y_{ij}^2] = 1$.
We write the sums:

$$V_0(t) = Y_{00}\psi_{00}(t)$$

$$V_i(t) = \sum_{j=1}^{2^{i-1}} Y_{ij}\psi_{ij}(t), \quad i \geq 1.$$

The sum

$$X_t = \sum_{i=0}^{\infty} V_i(t) = \sum_{i=0}^{\infty}\sum_{j=1}^{2^{i-1}} Y_{ij}\psi_{ij}(t), ; i \geq 1$$

is an approach of Wiener process on [0,1].

For the approach of Wiener process on any set, it suffices to use the property of Wiener process, that if $W_t$ is a Wiener process, then $\hat{B}_t = tB_{1/t}$ is a Wiener process. This property extend the Wiener process for $t \in [0, 1]$ to any $t$.

The well defined for this construction of Wiener process proved with the following theorem:

**Theorem 4.11.** *The series $\sum_{i=0}^{\infty} V_i(t)$ converges uniformly at t, a.s. If $X_t = \sum_{i=0}^{\infty} V_i(t)$, then the $X_t$ is Wiener process which satisfies $X_0 = 0$.*

**Remark**: The domination of Wiener process using the Haar functions is not the only one domination. One famous domination is when Wiener process is as a Fourier series with arbitrary coefficients. This domination is called "spread Karhunen-Loéve". According to this domination:

$$X_t(\omega) = \sum_{n=0}^{\infty} Z_n(\omega)\Phi_n(t), \quad 0 \leq t \leq T$$

where

$$\Phi_n(t) = \frac{2\sqrt{2T}}{(2n+1)\pi} \sin\left(\frac{(2n+1)\pi t}{2T}\right), \quad n = 0, 1, ...$$

and $\{Z_i\}$, $i = 0, 1, ...$ is independent normal random variables with $\mathbb{E}[Z_i] = 0$ and $\mathbb{E}[Z_i^2] = 1$, converges on $\mathcal{L}^2$ to a Wiener process on [0,T].

## 4.5 Q-Wiener process

Now we want to generalize a real-valued Wiener process to a Hilbert space valued process.

**Definition 4.7.** *Let $U$ a Hilbert space. A $U$-valued Stochastic process $\{u(t),\ \ t > 0\}$ is called Gaussian process if for any $n \in \mathbb{N}$ and arbitrary positive numbers $t_1, t_2, ..., t_n$, the $U^n$-valued random variable $(u(t_1), ..., u(t_n))$ is Gaussian.*

I.e. $\{u(t),\ \ t > 0\}$ is a Gaussian process if and only if for any $n \in \mathbb{N}$, the real-valued random variable

$$\langle (u(t_1), ..., u(t_n)), (h_1, ..., h_n) \rangle_{U^n} = \sum_{i=1}^{n} < u(t_i), h_i >_U$$

is Gaussian for any choice of positive numbers $t_1, t_2, ..., t_n$ and elements $h_1, ..., h_n \in U$.Equivalently, the vector $(< u(t_1), h_1 >_U, ..., < u(t_n), h_n >_U)$ is an $\mathbb{R}^n$-valued Gaussian random variable.
A Gaussian variable cannot have the covariance operator equal to the identity operator on a Hilbert space. Hence, we cannot use the definition of a Wiener process for $U$-valued stochastic process. Taking into account Definition 4.2 of the $\mathbb{R}^n$-valued Wiener process, we consider an important class of stochastic process, the $Q$-Wiener processes, with $Q$ being a trace class operator on $U$.

**Definition 4.8.** *Let $Q$ be a symmetric nonnegative trace class operator in $U$. A $U$-valued stochastic process $\{W(t),\ \ t \geq 0\}$ is called a $Q$-Wiener process, if*

*(i)$W(0) = 0$ $P$ a.s.*

*(ii)$W(t)$ has independent increments*

*(iii)$\mathcal{L}_{|W(t)-W(s)|} = \mathcal{N}(0, (t-s)Q),\ a \leq s \leq t$*

*(iv)$W(t)$ has continuous trajectories $P$ a.s.*

Since $Q$ is a non negative symmetric trace-class operator, it is compact and

there exists an orthonormal basis $\{e_j\}$ of eigenvectors of $Q$ and $\lambda_j$ eigenvalues of the space $U$. Then

$$\mathrm{Tr}Q := \sum_{j=1}^{\infty} < Qe_j, e_j >_U = \sum_{j=1}^{\infty} \lambda_j.$$

We now consider some properties of the introduced process, in particular, the connection of properties (i)-(iv) of the $Q$ Wiener process with properties (i)-(iv) of the Wiener process.

**Proposition 4.1.** *Let $\{W(t),\ t \geq 0\}$ be a Q-Wiener process on U. Then it is a Gaussian process and*

$$\mathbb{E}(W(t)) = 0 \quad , \quad Cov(W(t)) = tQ, \ \ t \geq 0 \tag{4.5}$$

*Furthermore, for any $t \geq 0$, the random variable W(t) has the following expansion in $\mathcal{L}_2(\Omega, \mathcal{F}, P; U)$:*

$$W(t) = \sum_{j=1}^{\infty} \sqrt{\lambda_j} w_j(t) e_j, \tag{4.6}$$

*where*

$$w_j(t) = \frac{1}{\sqrt{\lambda_j}} < W(t), e_j >_U, \quad t \geq 0, \ \ j = 1, 2, ... \tag{4.7}$$

*are independent real-valued Wiener processes on $(\Omega, \mathcal{F}, P)$.*

*Proof.* Property (iii) of Definition 4.7 implies that for $s = 0$ and any fixed $t \geq 0, W(t)$ is a Gaussian $U$-valued random variable with the law $\mathcal{N}(0, tQ)$; therefore,(4.4) holds.
To show that $\{W(t),\ t \geq 0\}$ is a Gaussian process we take arbitrary $0 \leq t_1 < ... < t_n$ and $h_1, ..., h_n \, \epsilon \, U$ and consider the real-valued random variable

$$Z = \sum_{j=1}^{n} \langle W(t_j), h_j \rangle_U ,$$

which can be written in the form

$$Z = \left\langle W(t_1), \sum_{j=1}^{n} h_j \right\rangle_U + \left\langle W(t_2) - W(t_1), \sum_{j=2}^{n} h_j \right\rangle_U + ...$$

$$+ \langle W(t_n) - W(t_{n-1}), h_n \rangle_U .$$

Since the increments of $W(t),\ t \geq 0$ are independent random variables distributed according to the Gaussian law (properties (ii) and (iii) of Definition 4.7) the products on the right-hand side are independent real-valued Gaussian variables. Consequently, their sum $Z$ is Gaussian and the process $W(t),\ t \geq 0$ is Gaussian by definition.
For any $t \geq 0$ and $j\ \epsilon\ \mathbb{N}$, the random variable $w_j(t)$ defined by (4.6) is Gaussian and satisfies:

$$\mathbb{E}(w_j(t)) = \mathbb{E}\left(\frac{1}{\sqrt{\lambda_j}} \langle W(t), e_j \rangle_U\right) = \frac{1}{\sqrt{\lambda_j}} \langle \mathbb{E}(W(t)), e_j \rangle_U = 0,$$

$$\mathrm{Cov}(w_j(t)) = \mathbb{E}(w_j^2(t)) = \frac{1}{\lambda_j}\mathbb{E}(\langle W(t), e_j \rangle_U^2) = \frac{1}{\lambda_j} \langle \mathrm{Cov}(W(t)e_j, e_j) \rangle_U = t;$$

therefore, $w_j(t),\ j\ \epsilon\ \mathbb{N}$, are Gaussian with the law $\mathcal{N}(0,t)$. We now show that they are independent. For $0 \leq s \leq t$, let us consider

$$\mathbb{E}(w_i(t)w_j(s)) = \frac{1}{\sqrt{\lambda_i\lambda_j}}\mathbb{E}\left(\langle W(t), e_i \rangle_U \langle W(s), e_j \rangle_U\right)$$

$$= \frac{1}{\sqrt{\lambda_i\lambda_j}} \left[\mathbb{E}\left(\langle W(t) - W(s), e_i \rangle_U \langle W(s), e_j \rangle_U\right) + \mathbb{E}\left(\langle W(s), e_i \rangle_U \langle W(s), e_j \rangle_U\right)\right].$$

Taking into account the independence of the random variables $W(t) - W(s)$ and $W(s)$ for $t > s$ , we have also

$$\mathbb{E}[w_i(t)w_j(s)] = \frac{s}{\sqrt{\lambda_i\lambda_j}} \langle Qe_i, e_j \rangle_U = s\delta_{ij}, \quad i, j\ \epsilon\ \mathbb{N}.$$

Hence, $w_i(t),\ i\ \epsilon\ \mathbb{N}$, are uncorrelated. Since, $w_i(t)$ are Gaussian, this implies that they are independent.
To prove representation (4.5), we show that the series is convergent in $\mathcal{L}_2(\Omega, \mathcal{F}, P; U)$. It is enough to prove that is partial sums form a fundamental sequence in the complete space $\mathcal{L}_2(\Omega, \mathcal{F}, P; U)$. For $1 \leq n \leq m$, we have

$$\mathbb{E}\left|\left|\sum_{j=n}^{m} \sqrt{\lambda_j} w_j(t) e_j\right|\right|_U^2 = t\sum_{j=n}^{m} \lambda_j.$$

To complete the proof, we note that

$$\sum_{j=1}^{\infty} \lambda_j = \mathrm{Tr}Q < \infty.$$

□

Thus, we have proved that for any $Q$-Wiener process, there exists a sequence of independent real-valued Wiener processes. The converse follows from the proof of the next theorem.

**Theorem 4.12.** *For an arbitrary trace class, symmetric, non negative operator $Q$ on a Hilbert space $U$, there exists a $U$-valued $Q$-Wiener process.*

*Proof.* Let $\{w_j(t),\ t \geq 0\},\ j \in \mathbb{N}$, be a sequence of independent real-valued Wiener processes on $(\Omega, \mathcal{F}, P)$. Consider series (4.5) with $\lambda_j$ being the eigenvalues of $Q$. Since

$$\mathbb{E}\left( \| \sum_{j=n}^{m} \sqrt{\lambda_j} w_j(t) e_j \|_U^2 \right) = t \sum_{j=n}^{m} \lambda_j ||e_j||_U^2 = t \sum_{j=n}^{m} \lambda_j$$

for any $1 \leq n \leq m$, series (4.5) is convergent in $\mathcal{L}_2(\Omega, \mathcal{F}, P; U)$. The obtained process is a $Q$-Wiener process. Indeed, without loss of generality, we may assume that $w_j(0) = 0,\ j \in \mathbb{N}$. then $W(0) = 0$, that is, property (i) holds. We note for any $0 \leq s \leq t,\ W(t) - W(s)$ is a Gaussian random variable, since it is the mean square limit of the sequence of Gaussian random variables $W_n(t) - W_n(s)$, where

$$W_n(t) = \sum_{j=1}^{n} \sqrt{\lambda_j} w_j(t) e_j.$$

Therefore, the increments of $W(t)$ are independent if and only if they are uncorrelated. Let $0 \leq t_1 < t_2 < ... < t_n$. Set

$$\Delta W_i = W(t_{i+1}) - W(t_i),\ \ i = 1, 2, ..., n-1.$$

Obviously, $\mathbb{E}(\Delta W_i) = 0$, and hence

$$\langle \mathrm{Cor}(\Delta W_i, \Delta W_k) h_1, h_2 \rangle_U = \mathbb{E}\left[ \langle (\Delta W_i \otimes \Delta W_k) h_1, h_2 \rangle_U \right] = \mathbb{E}\left[ \langle \Delta W_i, h_1 \rangle_U \langle \Delta W_k, h_2 \rangle_U \right]$$

$$= \mathbb{E}\left[\left\langle \sum_{j=1}^{\infty} \sqrt{\lambda_j}[w_j(t_{j+1}) - w_j(t_i)]e_j, h_1 \right\rangle_U \left\langle \sum_{l=1}^{\infty} \sqrt{\lambda_l}[w_l(t_{k+1}) - w_l(t_k)]e_l, h_2 \right\rangle_U\right]$$

$$= \sum_{j=1}^{\infty}\sum_{l=1}^{\infty} \sqrt{\lambda_j}\sqrt{\lambda_l}\mathbb{E}[w_j(t_{i+1}) - w_j(t_i)][w_l(t_{k+1}) - w_l(t_k)] \langle e_j, h_1\rangle_U \langle e_l, h_2\rangle_U .$$

The independence of increments of the Wiener process and the independence of $w_j(t)$ and $w_l(t)$ for $j \neq l$ imply

$$\langle \mathrm{Cor}(\Delta W_i, \Delta W_k)h_1, h_2\rangle_U = 0$$

for $i \neq k$. Therefore, (ii) holds. For any $0 < s \leq t$, the covariance operator of $W(t) - W(s)$ is defined by

$$\langle \mathrm{Cov}(W(t) - W(s))h, g\rangle_U = \mathbb{E}(\langle h, W(t) - W(s)\rangle_U , \langle g, W(t) - W(s)\rangle_U)$$

$$= \mathbb{E}\left(\left\langle h, \sum_{j=1}^{\infty} \sqrt{\lambda_j}(w_j(t) - w_i(s))e_j \right\rangle_U \left\langle g, \sum_{j=1}^{\infty} \sqrt{\lambda_j}(w_j(t) - w_j(s))e_j \right\rangle_U\right)$$

$$= (t-s)\sum_{j=1}^{\infty} \lambda_j \langle h, e_j\rangle_U \langle g, e_j\rangle_U = (t-s)\sum_{j=1}^{\infty} \langle Qh, g\rangle_U ,$$

that is, (iii) holds. Since $w_j(t),\ j \,\epsilon\, \mathbb{N}$ are continuous, (iv) also holds.

□

**Remark** : Let $H$ be a Hilbert space and $h_1,\ h_2 \,\epsilon\, H$, the linear operator $h_1 \otimes h_2$ is defined by

$$(h_1 \otimes h_2)h := h_1 \langle h_2, h\rangle_H ,\quad h \,\epsilon\, H.$$

## 4.6 Weak Wiener process

We now consider the general case of $Q$ being a bounded operator in $U$ ($\mathrm{Tr}Q \leq \infty$). If $Q$ is not a trace class operator. then series (4.5) can be divergent in $U$. We dive two ways of avoiding the difficulties that arise and of generalizing the Wiener process to a Hilbert space. The first way is to construct a $Q_1$ Wiener process in an appropriate space $U \subset U_1$. The second way is to construct a

Wiener process in $U$ in a weak sense. We begin with the construction of the space $U_1$. To do this, we need the following definition 2.6 of Hilbert Schmidt operator.
Recall the space $\mathcal{L}_{HS}(U,H)$ of Hilbert Schmidt operators with norm

$$||A||_{HS} = \left(\sum_{j=1}^{\infty} ||Ae_j||_H^2\right)^{1/2}$$

is a Hilbert orthonormal generated by:

$$< A, B >_{HS} = \sum_{j=1}^{\infty} < Ae_j, Be_j >_H .$$

The following connection between nuclear and Hilbert Schmidt operators holds.

**Proposition 4.2.** *Let $U$, $H$ and $E$ be Hilbert spaces. For any $A \epsilon \mathcal{L}_{HS}(U,H)$ and $B \epsilon \mathcal{L}_{HS}(H,E)$, the operator $BA \epsilon \mathcal{L}_N(U,E)$ and*

$$||BA||_N \leq ||A||_{HS} \cdot ||B||_{HS}.$$

*Proof.* If $\{e_j\}$ is a basis of $H$, then

$$Ax = \sum_{j=1}^{\infty} \langle Ax, e_j \rangle_H e_j, \quad x \epsilon U,$$

and, therefore, we have the representation

$$BAx = \sum_{j=1}^{\infty} \langle Ax, e_j \rangle_H Be_j, \quad x \epsilon U.$$

Hence, the operator $BA$ is nuclear from $U$ to $E$. It also follows from the definition of nuclear operators that

$$||BA||_N \leq \sum_{j=1}^{\infty} ||A^* e_j||_U \cdot ||Be_j||_E \leq \left(\sum_{j=1}^{\infty} ||A^* e_j||_U^*\right)^{1/2} \left(\sum_{j=1}^{\infty} ||Be_j||_E^2\right)^{1/2}.$$

□

We now consider the $Q$ where $\mathrm{Tr}Q = \infty$

**Definition 4.9.** *Define the space $U_0 := Q^{1/2}(U)$ equipped with the inner product $\langle u, v\rangle_{U_0} := \left\langle Q^{-1/2}u, Q^{-1/2}n\right\rangle_U$. Define the Hilbert space $U_1$ such that the embedding $J$ of $U$ in $U_1$ is a continuous operator and the embedding $J_0$ of $U_0$ in $U_1$ is a Hilbert-Schmidt operator.*

Obviously, $U_0$ is a subspace of $U$.

**Theorem 4.13.** *Let $Q$ be a bounded self-adjoint positive operator on $U$ with $TrQ = \infty$. Let $\{w_j(t),\ t > 0,\ j \in \mathbb{N}\}$ be a family of independent real-valued Wiener processes and $\{g_j\}$ be an orthonormal basis on $U_0$. Then the series*

$$W(t) = \sum_{j=1}^{\infty} w_j(t)g_j, \quad t \geq 0, \tag{4.8}$$

*defines a $U_1$-valued $Q_1$-Wiener process with zero expectation and the covariation operator $Q_1 := J_0 J_0^*$, where $J_0$ is the embedding $J_0 : U_0 \to U_1$. Moreover,*

$$Q_1^{1/2}(U_1) = U_0 \quad \textit{and} \quad ||h||_{U_0} = ||Q_1^{-1/2}h||_{U_1}.$$

*Proof.* To show that series (4.7) is convergent in $\mathcal{L}_2(\Omega, \mathcal{F}, P; U_1)$, we note that for any $1 \leq n \leq m$,

$$\mathbb{E}\left[\left|\left|\sum_{j=n}^{m} w_j(t)g_j\right|\right|_{U_1}^2\right] = \sum_{j=n}^{m} ||g_j||_{U_1}^2 = \sum_{j=n}^{m} ||J_0 g_j||_{U_1}^2.$$

Since the embedding $J_0$ is a Hilbert Schmidt operator, we have $\sum_{j=1}^{\infty} ||J_0 g_j||_{U_1}^2 < \infty$. The fact that obtained process is a $Q_1$-Wiener process can be proved by a similar argument as in the previous theorem. We define the covariance operator $Q_1$ of $\{W(t),\ t \geq 0\}$. Let $0 \leq s \leq t$. We have

$$\langle \mathrm{Cov}(W(t) - W(s))h, g\rangle_{U_1} = \mathbb{E}\left(\langle h, W(t) - W(s)\rangle_{U_1} \langle g, W(t) - W(s)\rangle_{U_1}\right)$$

$$= \mathbb{E}\left(\left\langle h, \sum_{j=1}^{\infty}(w_j(t) - w_j(s))g_j\right\rangle_{U_1} \left\langle g, \sum_{j=1}^{\infty}(w_j(t) - w_j(s))g_j\right\rangle_{U_1}\right)$$

$$= (t-s)\sum_{j=1}^{\infty} \langle h, g_j\rangle_{U_1} \langle g, g_j\rangle_{U_1} = (t-s)\sum_{j=1}^{\infty} \langle h, J_0 g_j\rangle_{U_1} \langle g, J_0 g_j\rangle_{U_1}$$

$$= (t-s)\sum_{j=1}^{\infty} \langle J_0^* h, g_j\rangle_{U_0} \langle J_0^* g, g_j\rangle_{U_0} = (t-s)\langle J_0^* h, J_0^* g\rangle_{U_0}$$

$$= (t-s)\langle J_0 J_0^* h, g\rangle_{U_1} \quad \text{for any } \ h, g \,\epsilon\, U_1.$$

It follows from Definition 4.7 that $Q_1 = J_0 J_0^*$. The definition of $Q_1$ implies that it is a self-adjoint and positive operator. It follows from Proposition 4.2 that is nuclear. Next, $\mathrm{Im}Q_1^{1/2} = \mathrm{Im}J_0^* = U_0$, and the operator $Q_1^{-1/2} J_0$ from $U_0$ onto $U_1$ is bounded. For $t - s = 1$ and $g = h$, we obtain from the last equality that

$$||Q_1^{1/2} h||_{U_1}^2 = \langle J_0 J_0^* h, h\rangle_{U_1} = ||J_0^* h||_{U_0}^2, \quad h \,\epsilon\, U_1. \tag{4.9}$$

Since, $J_0^* Q_1^{-1/2}$ is an isometry, $Q_1^{-1/2} J_0$ is also an isometry. Therefore,

$$||Q_1^{-1/2} h||_{U_1} = ||Q_1^{-1/2} J_0 h|| = ||h||_{U_0}.$$

□

We now consider the second approach to a Wiener process in a Hilbert space $U$ in the case where $Q$ is not a trace-class operator.

**Proposition 4.3.** *Let $\{w_j(t),\ t \geq 0\}$ $j \,\epsilon\, \mathbb{N}$, be a family of independent real-valued Wiener processes. For an arbitrary $h \,\epsilon\, U$, the process defined by*

$$\langle h, W(t)\rangle_U = \sum_{j=1}^{\infty} w_j(t) \langle h, g_j\rangle_U, \quad t \geq 0, \tag{4.10}$$

*is a real-valued process in $\mathbb{R}$ with zero expectation, $\{g_j\}$ be an orthonormal basis and the covariance $tQ_h = t < Qh, h >_U$ . Furthermore,*

$$\mathbb{E}\left(\langle h, W(t)\rangle_U \langle g, W(s)\rangle_U\right) = (t \wedge s)\langle Qh, g\rangle_U, \quad h, g \,\epsilon\, U$$

*Proof.* For arbitrary $h \,\epsilon\, U$ and $1 \leq n \leq m$, consider

$$\mathbb{E}\left(|\sum_{j=n}^{m} w_j(t) \langle h, g_j\rangle_U|^2\right) = \sum_{j=n}^{m} |\langle h, g_j\rangle_U|^2$$

$$\leq ||h||_U^2 \sum_{j=n}^{m} ||g_j||_U^2 \leq C||J||_{\mathcal{L}(U,U_1)}^2 ||h||_U^2 \sum_{j=n}^{m} ||g_j||_{U_1}^2.$$

Since, $J_0$ is a Hilbert-Schmidt operator, we have

$$\sum_{j=1}^{\infty} ||J_0 g_j||_{U_1}^2 = \sum_{j=1}^{\infty} ||g_j||_{U_1}^2 < \infty.$$

Hence, the series defining the random variable $\langle h, W(t)\rangle_U$ is convergent in $\mathcal{L}_2(\Omega, \mathcal{F}, P)$. If $0 \leq s \leq t$, then

$$\mathbb{E}(\langle h, W(t)\rangle_U \langle g, W(s)\rangle_U) = \mathbb{E}(\langle h, W(s)\rangle_U \langle g, W(s)\rangle_U)$$

$$= s \sum_{j=1}^{\infty} \langle h, g_j\rangle_U \langle g, g_j\rangle_U = s \sum_{j=1}^{\infty} \left\langle q^{1/2}h, Q^{-1/2}g_j\right\rangle_U \left\langle Q^{1/2}g, Q^{-1/2}g_j\right\rangle_U$$

$$s \sum_{j=1}^{\infty} \langle Qh, g_j\rangle_{U_0} \langle Qg, g_j\rangle_{U_0} = s \langle Qh, Qg\rangle_{U_0} = s \langle Qh, g\rangle_{U_0}$$

for any $h, g \in U$. □

Proposition 4.3 loads the following definition:

**Definition 4.10.** *For any $h \in U$, let a real-valued process $\{\langle h, W(t)\rangle_U,\ t \geq 0\}$ be Gaussian with independent increments and a continuous version. Let*

$$\mathbb{E}(\langle h, W(t)\rangle) = 0 \quad \textit{and} \quad Cov(\langle h, W(t)\rangle) = t \langle Qh, h\rangle_U$$

*for all $t \geq 0,\ h \in U$. Then we say that $\{W(t),\ t \geq 0\}$ is a weak Wiener process in $U$.*

Also, because we consider that $U \subset H \subset U^*$ we can say that $\{W(t), t \geq 0\}$ is also a weak Wiener process in $U^*$.

# Chapter 5

# Stochastic integral in Hilbert Space

## 5.1 *Itô* integral in $\mathbb{R}$

We consider also that $(\Omega, \mathcal{F}, P)$ is a probability space.

**Definition 5.1.** *We consider the partition $a = t_0 < t_1 < ... < t_n = b$ for the interval $[a, b]$ and that we approximate the function $f(t, \omega)$ as*

$$f(t, \omega) \simeq \sum_{i=0}^{n-1} f(t_i, \omega) \mathbf{1}_{[t_i, t_{i+1})}(t).$$

*The Itô integral can be defined as a limit in $L_2$, because it is adapted. In particular*

$$\int_a^b f(t, \omega) dW_t(\omega) = \lim_{n \to \infty} \sum_{i=1}^{n} f(t_i, \omega)[W_{t_{i+1}} - W_{t_i}](\omega)$$

*which is measurable with respect to $\{\mathcal{F}_t, \ \mathcal{F}_t = \sigma(W(t), \ t \geq 0)\}$*

The Wiener process is a function which is nowhere differentiable, but never can be used as an integrator. These integrators are very useful in stochastic

analysis.In this chapter we will present the construction of the *Itô* integral. We consider a random function $f$ which depends on a Wiener process and we want to define the integral of increments of the Wiener process. I.e. we want to define

$$\int_a^b f(t,\omega)dW_t(\omega)$$

where $W_t$ is one-dimensional Wiener process which starts at 0, while $f$ is a function $f : (0,\infty) \times \Omega \to \mathbb{R}$.

**Example 5.1.** *This example, illustrates the usefulness of a possible application. We consider that an investor can invest some money on a title where rate of return $S_t$, is a stochastic process. Assume the $S_t$ is a model of the Wiener process i.e. $W_t = S_t$. At each time the investor decides on how her position $h_t$ on the esser. This is allowed of change over time, and is modelled as an adapted stochastic process; $h(t,\omega)$ is the position of the investor held in $[t, t+dt]$ on the esser, given the realisation $\omega$ of the ..... . The gain or loss of the investor form this position on the interval $[t, t+dt]$ is $h_t(S_{t+1} - S_t) = h_t(W(t+1) - W(t))$. The all wealth, $V_t$, at time point $t$, or after $N$ time periods, will be the sum of these changes, i.e.*

$$V_t = V_0 + \sum_{i=1}^{N} h_{t_i}(W_{t_{i+1}} - W_{t_i})$$

*where $t_N = t$.*

Now we want to define step processes.

**Definition 5.2.** *A stochastic process f which can be written in the form*

$$f(t) = \sum_{j=1}^{n-1} \eta_j \mathbf{1}_{[t_j, t_{j+1})}(t) \tag{5.1}$$

*for any partition $a = t_0 < t_1 < ... < t_n = b$ of the space $[a,b]$ where $\eta_j$ are random variables which are $\mathcal{F}_{t_j}$ measurable and $\mathbb{E}[\eta_j^2] < \infty$ is called step function. We will symbolize with $M_{step}([a,b])$ the set of step functions in space $[a,b]$.*

So, the function $I : M^2_{step} \to L_2(\Omega, \mathcal{F}, P)$ which satisfies the Proposition 5.1 is also continuous. This function is the *stepfunction* which we use in *Itô* integral.

**Definition 5.3.** *If the stochastic process f is a step function of the form (5.1) then its stochastic integral of Wiener process defined as*

$$\int_a^b f(t)dW_t = \sum_{j=0}^{n-1} \eta_j (W_{t_{j+1}} - W_{t_j})$$

**Proposition 5.1.** *The function $I : M^2_{step} \to L_2$ is an isometry, hence is continuous. This property called Itô isometry. I.e.*

$$||I(f_{step})||^2_{L_2} = \mathbb{E}[|I(f_{step})|^2] := \mathbb{E}\left[\left|\int_a^b f_{step}(t)dW_t\right|^2\right] = \mathbb{E}\left[\int_a^b |f_{step}(t)|^2 dt\right]$$

*Proof.* We consider the step function

$$f_{step}(t) = \sum_{j=0}^{n-1} \eta_j \mathbf{1}_{[t_j, t_{j+1})}(t)$$

which has the stochastic integral

$$I(f_{step}) := \int_a^b f_{step}(t)dt \equiv \sum_{j=0}^{n-1} \eta_j (W_{t_{j+1}} - W_{t_j}).$$

We can calculate the expectation of the square of this form. So, we obtain

$$|I(f_{step})|^2 = \sum_{j=0}^{n-1} \eta_j^2 (W_{t_{j+1}} - W_{t_j})^2 + 2\sum_{k<j} \eta_j (W_{t_{j+1}} - W_{t_j})\eta_k (W_{t_{k+1}} - W_{t_k})$$

The increments of Wiener process $(W_{t_{j+1}} - W_{t_j})$ are independent from the facts before the time $t_j$. So, because of the random variable $\eta_j$ is $\mathcal{F}_{t_j}$ measurable, the random variables $(W_{t_{j+1}} - W_{t_j})$ and the $\eta_j$ are independent. The same is true and for any function of these variables. Hence,

$$\mathbb{E}[\eta_j^2 (W_{t_{j+1}} - W_{t_j})^2] = \mathbb{E}[\eta_j^2]\mathbb{E}[(W_{t_{j+1}} - W_{t_j})^2] = \mathbb{E}[\eta_j^2](t_{j+1} - t_j)$$

because of $\mathbb{E}[(W_{t_{j+1}} - W_{t_j})^2] = (t_{j+1} - t_j)$ by the Wiener's process properties. Furthermore, because $k < j$ the arbitrary variables $\eta_k,\ W_{t_{k+1}} - W_{t_k},\ \eta_j,\ W_{t_{j+1}} - W_{t_j}$ are independent of each and so,

$$\mathbb{E}[\eta_j(W_{t_{j+1}} - W_{t_j})\eta_k(W_{t_{k+1}} - W_{t_k})] = \mathbb{E}[\eta_j(W_{t_{k+1}} - W_{t_k})\eta_k]\mathbb{E}[W_{t_{j+1}} - W_{t_j})] = 0$$

because $\mathbb{E}[W_{t_{j+1}} - W_{t_j})] = 0$ from the Wiener's process properties. Using all of these, we have that

$$||I(f_{step})||^2_{L_2} = \mathbb{E}[I(f_{step})^2] = \sum_{j=0}^{n-1} \mathbb{E}[\eta_j^2](t_{j+1} - t_j).$$

From this we can obtain

$$\mathbb{E}\left[\int_a^b |f_{step}(t)|^2 dt\right] = \mathbb{E}\left[\int_a^b \left|\sum_{j=0}^{n-1} \eta_j \mathbf{1}_{[t_j, t_{j+1})}\right|^2 dt\right].$$

Actually, we have

$$|f_{step}|^2 = \sum_{j=0}^{n-1} \eta_j^2 \mathbf{1}_{[t_j, t_{j+1})} + \sum_{k<j} \eta_k \eta_j \mathbf{1}_{[t_j, t_{j+1})} \mathbf{1}_{[t_k, t_{k+1})} = \sum_{j=0}^{n-1} \eta_j^2 \mathbf{1}_{[t_j, t_{j+1})}$$

because $\mathbf{1}_{[t_j, t_{j+1})} \mathbf{1}_{[t_k, t_{k+1})} = 0$ for $k < j^2$, and so we obtain

$$\mathbb{E}\left[\int_a^b |f_{step}(t)|^2 dt\right] = \mathbb{E}\left[\int_a^b \left|\sum_{j=0}^{n-1} \eta_j \mathbf{1}_{[t_j, t_{j+1})}\right|^2 dt\right] = \sum_{j=0}^{n-1} \mathbb{E}[\eta_j^2](t_{j+1} - t_j).$$

So, we prove that step functions have the property of *Itô isometry*.

□

Now, we define a general doss of stochastic processes for which we can define the stochastic integral.

**Definition 5.4.** *A stochastic process f, which is continuous, belongs to space* $M^2([a,b])$ *if is adapted to filtration* $\mathcal{F}_t = \sigma(B_s, s \leq t)$ *and satisfies the following:*

$$||f||_{M^2([a,b])} := \mathbb{E}\left[\int_a^b |f|^2 dt\right] < \infty$$

Often we can write $M^2$ instead to $M^2([a,b])$.

The stochastic processes which belong to space $M^2$ can be approximated from step functions according to the following theorem.

**Remark** : $I : M^2_{step} \to L_2$ is a continuous mapping hence it can be uniquely extended to $\overline{M^2_{step}}$ as $M^2$. This extension is the $Itô$ integral. The following theorem characterizes $\overline{M^2_{step}}$ as $M^2$ given by Definition 5.4.

**Theorem 5.1.** *For all* $a \in M^2$ *exists a sequence of step functions* $f_{step,n}$, *such that*

$$\lim_{n\to\infty} ||f(t) - f_{step,n}(t)||_{M^2([a,b])} = \lim_{n\to\infty} \mathbb{E}\left[\int_a^b |f(t) - f_{step,n}(t)|^2 dt\right] = 0$$

*Proof.* First, we consider a function $f \in M^2([a,b])$. We define the sequence of stochastic processes $\phi_n(t) = [-n \vee f(t)] \wedge n$. We can easily prove that the sequence $\phi_n(t)$ is bounded and that $\phi_n(t) \in M^2([a,b])$ for any $n$. Furthermore, $\lim_{n\to\infty} \phi_n(t) = f(t)$. So, from the Theorem of Dominated Convergence we have

$$\lim_{n\to\infty} \mathbb{E}\left[\int_a^b |f(t) - \phi_n(t)|^2 dt\right] = 0.$$

Now, we consider that $\phi(t) \in M^2([a,b])$ is bounded. We can make the sequence $\psi_n(t)$ in this way:

For any $n$ we consider $\rho_n : \mathbb{R} \to \mathbb{R}^+$ a continuous function such that $\rho_n(t) = 0$ for $t \leq -\frac{1}{n}$ and $t \geq 0$ and $\int_{-\infty}^{\infty} \rho_n(t)dt = 1$. We define

$$\psi_n(t) = \int_a^b \rho_n(s-t)\phi(s)ds$$

The sequence $\psi_n(t)$ is a sequence of stochastic processes because $\phi(s)$ is also a stochastic process. The integral which defined $\psi_n$ is a typical Riemann integral. From Riemann's integral properties and because of $\psi(t)$ is bounded, we obtain that the process $\psi_n(t)$ consists from continuous functions and so is bounded. Also, belongs to $M^2([a,b])$. From the Theorem of Bounded Convergence we have that:

$$\lim_{n\to\infty} \mathbb{E}\left[\int_a^b |f(t)-\psi_n(t)|^2 dt\right] = 0$$

Finally, if $\psi(t)\ \epsilon\ M^2([a,b])$ and is bounded and continuous we can make a sequence of step functions $f_{step,n}$ in this way:

$$f_{step,n}(t) = \psi(a)\mathbf{1}_{[a,a+\frac{b-a}{n}]}(t) + \sum_{i=1}^{n-1} \psi\left(a+i\frac{b-a}{n}\right)\mathbf{1}_{(a+i\frac{b-a}{n},a+(i+1)\frac{b-a}{n}]}(t)$$

This sequence is bounded, so from Theorem of Bounded Convergence we can obtain that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_a^b |\psi(t)-f_{step,n}(t)|^2 dt\right] = 0.$$

So, using all of these facts and triangle inequality we have:

$$\lim_{n\to\infty} \mathbb{E}\left[\int_a^b |f(t)-f_{step,n}(t)|^2 dt\right] = 0$$

With this we have finish the proof.

□

Now, we have approximate any stochastic process $f\ \epsilon\ M^2([a,b])$ with a sequence of step functions $f_{step,n}$ and because the stochastic integral is well defined for a stochastic step function we can define the *Itô* stochastic integral with the following way.

**Definition 5.5.** *Let $f\ \epsilon\ M^2([a,b])$. The Itô integral of $f$ is defined as the following limit*

$$I(f) := \int_a^b f(t)dt = \lim_{n\to\infty} I(f_{step,n}) := \lim_{n\to\infty} \int_a^b f_{step,n}(t)dW_t$$

*where $f_{step,n}$ is a sequence of step functions which approximates in $L_2$ the function $f$ and the integrals $\int_a^b f_{step,n}(t)dW_t := I(f_{step,n})$ defined according to Definition 5.2.*

This definition makes sense only if the sequence $I(f_{step,n}) := \lim_{n\to\infty} f_{step,n}(t)dB_t$ converges. This is the main result of the following theorem.

**Theorem 5.2.** *Let $f \epsilon M^2([a,b])$ and $f_{step,n}$ is a sequence of step functions which approximated in $L_2$ the function $f$. Then, the sequence $I(f_{step,n}) = \int_a^b f_{step,n}dW_t$ where the Itô stochastic integral had defined according to Definition 5.3 converges as $n \to \infty$ to a random variable which belongs in $L_2$.*

*Proof.* Because of the completeness of $L_2$ it suffices to prove that the sequence $r_n := I(f_{step,n})$ is a Cauchy sequence i.e. $||r_n - r_m||_{L_2} := \mathbb{E}[|r_n - r_m|^2] \to 0$ as $n, m \to \infty$.
To prove the Cauchy property we use the fact that the sequence $f_{step,n}$ approximate in $L_2$ the function $f$, i.e.

$$\lim_{n\to\infty} \mathbb{E}\left[\int_a^b |f(t) - f_{step,n}(t)|^2 dt\right] = 0$$

and the property of *Itô* stochastic integral for step functions,

$$\mathbb{E}\left[\left|\int_a^b f_{step,n}(t)dW_t\right|^2\right] = \mathbb{E}[|f_{step,n}(t)|^2 dt].$$

Also, we use the linear property of *Itô* integral for step functions.Using all of these we have

$$\begin{aligned}
\mathbb{E}[|r_n - r_m|^2] &= \mathbb{E}\left[\left|\int_a^b f_{step,n}(t)dt - \int_a^b f_{step,m}(t)dt\right|^2\right] \\
&= \mathbb{E}\left[\left|\int_a^b (f_{step,n}(t) - f_{step,m}(t))\right|^2\right] \\
&= \mathbb{E}\left[\int_a^b |f_{step,n}(t) - f_{step,m}(t)|^2 dt\right]
\end{aligned}$$

$$= \mathbb{E}\left[\int_a^b |f_{step,n}(t) - f(t) + f(t) - f_{step,m}(t)|^2 dt\right]$$

$$\leq \mathbb{E}\left[\int_a^b |f_{step,n}(t) - f(t)|^2 dt\right] + \mathbb{E}\left[\int_a^b |f_{step,m}(t) - f(t)|^2 dt\right]$$

Because the $f_{step,n}$ approximate well in $L_2$ the function $f(t)$ we can see that the right side tends to 0 as $m \to \infty$ and $n \to \infty$. So, we end the proof.

□

**Example 5.2.** *We want to prove that*

$$I(W_t) = \int_0^T W_t dW_t = \frac{1}{2}W_T^2 - \frac{T}{2}.$$

*The stochastic process which is under integration is $f(t,\omega) = W_t$. We take the partition $0 < t_1^n < t_2^n < ... < t_n^n = T$, $t_j^n = \frac{jT}{n}$ and the approximation of the stochastic process*

$$f_{step,n}(t) = \sum_{j=0}^{n} W_{t_j^n} \mathbf{1}_{[t_j^n, t_{j+1}^n)}(t).$$

*We take the sequence of the random variables*

$$I(f_{step,n}) = \sum_{j=0}^{n-1} W_{t_j^n}(W_{t_{j+1}^n} - W_{t_j^n}^n),$$

*and its limit in $L_2$ is the stochastic integral which we want. To calculate this limit we use the identity*

$$a(b-a) = \frac{1}{2}(b^2 - a^2) - \frac{1}{2}(a-b)^2$$

*with $a = W_{t_j^n}$, $b = W_{t_{j+1}^n}$. So, we have*

$$W_{t_j^n}(W_{t_{j+1}^n} - W_{t_j^n}) = \frac{1}{2}(W_{t_{j+1}^n}^2 - W_{t_j^n}^2) - \frac{1}{2}(W_{t_{j+1}^n} - W_{t_j^n})^2.$$

*Summing on all $j$ from $j = 0$ until $j = n-1$ we obtain*

$$I(f_{step,n}) = \frac{1}{2}W_T^2 - \frac{1}{2}\sum_{j=0}^{n-1}(W_{t_{j+1}}^n - W_{t_j}^n)^2.$$

*Since,*

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

*as* $n \to \infty$.

With this example we can see that the *Itô* stochastic integral has some different properties of Riemann integral. This is obvious from $\int_0^T W_t dW_t = \frac{1}{2}W_T^2 - \underbrace{\frac{1}{2}T}$.

Because, a Riemann integral given by $\int_0^T f dt = \frac{f^2(T)}{2} - \frac{f^2(0)}{2}$.

Now, we will see some properties of *Itô* integral.

**Theorem 5.3.** *The Itô integral has the following important properties:*

*(i) For two stochastic processes* $f_1$ *and* $f_2$ *is true that* $I(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 I(f_1) + \lambda_2 I(f_2)$, *where* $\lambda_1, \lambda_2 \epsilon \mathbb{R}$.

*(ii)* $\mathbb{E}\left[\int_a^b f dW_t\right] = 0$

*(iii)* $\mathbb{E}\left[\left|\int_a^b f(t,\omega) dW_t\right|^2\right] = \mathbb{E}\left[\int_a^b |f(t,\omega)|^2 dt\right]$ *(Itô isometry)*

*On all of these we consider that the function in the integral belongs to the suitable space* $M^2$.

Now, we can see one more example.

**Example 5.3.** *We prove now that for any* $f, g \epsilon M^2$ *it holds that*

$$\mathbb{E}[I(f)I(g)] = \mathbb{E}\left[\int_0^T f(t)dW_t \int_0^T g(t)dW_t\right] \mathbb{E}\left[\int_0^T f(t)g(t)dt\right]$$

*For this proof we use the identity*

$$ab = \frac{1}{4}(|a+b|^2 - |a-b|^2)$$

*where $a = I(f)$ and $b = I(g)$. So, using this identity for the expected value we have*

$$\mathbb{E}[I(f)I(g)] = \frac{1}{4}(\mathbb{E}[|I(f)+I(g)|^2]-\mathbb{E}[|I(f)-I(g)|^2]) = \frac{1}{4}(\mathbb{E}[|I(f+g)|^2]-\mathbb{E}[|I(f-g)|^2])$$

*where we use also the linear property of stochastic integral. However,*

$$\mathbb{E}[|I(f+g)|^2] = \mathbb{E}\left[\int_0^T |f+g|^2 dt\right]$$

*and from the Itô isometry*

$$\mathbb{E}[|I(f-g)|^2] = \mathbb{E}\left[\int_0^T |f-g|^2 dt\right]$$

*So, using these and the linear property of expected value we have*

$$\mathbb{E}[I(f)I(g)] = \mathbb{E}\left[\int_0^T \frac{1}{4}(|f+g|^2 - |f-g|^2)dt\right]\mathbb{E}\left[\int_0^T f(t)g(t)dt\right]$$

*and now we end the proof.*

## 5.2 *Itô* integral as a stochastic process

In this section we consider that the indefinite *Itô* integral, where the lower limit is fixed but the upper limit is allowed to take values $t \in [a,b]$, i.e. we write $\int_0^t f(t)dW_t \quad$ where $\quad 0 \leq t \leq T$. All these properties are fine if $f \in M^2_{step}$, hence to prove that any $f \in M^2$, we use an approximation sequence $f_n \in M^2_{step}$, $f_n \to f$ and go to the limit. So, for any value of $t$ we obtain a random variable which is square integrable and its value is equal to integral $\int_0^t f(t)dW_t$. Therefore, we construct the stochastic process

$$I_t := \int_0^t f(t)dW_t$$

This stochastic process is called the indefinite *Itô* integral. This stochastic process can also be defined as the stochastic integral from 0 to $T$ of $f(t)\mathbf{1}_{[0,t]}$ i.e.

$$\int_0^t f(s)dW_s = \int_0^T f(s)\mathbf{1}_{[0,t]}(s)dW_s \tag{5.2}$$

The following theorem present basic properties of the indefinite *Itô* integral.

**Theorem 5.4.** *Let* $f \in M^2([0,T]),\ 0 \leq t \leq T$ *and* $I_t = \int_0^t f(s)dW_s$.
*The following are true:*
*(i)The stochastic process* $I_t$ *is a square integrable martingale.*
*(ii)The bracket process of* $I_t$ *is*

$$< I >_t = \int_0^t |f(s)|^2 ds$$

*Proof.* (i) To prove that $I_t$ is square integrable it suffices to use the *Itô* isometry and modelled $f \in M^2([0,T])$. We can also see that $I_t$ is adapted to $\mathcal{F}_t$, from the definition of the stochastic integral. So, we have to prove only that $\mathbb{E}[I_t|\mathcal{F}_s] = I_s,\ 0 \leq s \leq t \leq T$. To prove this we have to see that

$$I_t = I_s + \int_s^t f(r)dW_r$$

and remember that the cause of the independence of $\int_s^t f(r)dW_r$ from $\mathcal{F}_s$

$$\mathbb{E}\left[\int_s^t f(r)dW_r|\mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t f(r)dW_r\right] = 0$$

To obtain this result we use Theorem 5.3. We can also obtain that

$$\mathbb{E}[I_t|\mathcal{F}_s] = \mathbb{E}[I_s|\mathcal{F}_s] + \mathbb{E}\left[\int_s^t f(r)dW_r|\mathcal{F}_s\right] = I_s$$

So, $I_t$ is martingale.
(ii)To prove that the bracket process of $I_t$ is $< I >_t = \int_0^t |f(s)|^2 ds$ it suffices to prove that $M_t = I_t^2 - < I >_t$ is a continuous martingale which is equal to zero for $t = 0$. Actually,

$$\mathbb{E}[M_t|\mathcal{F}_s] = \mathbb{E}\left[I_t^2 - \int_0^t |f(r)|^2 dr|\mathcal{F}_s\right]$$

$$= \mathbb{E}\left[(I_s + \int_s^t f(r)dW_r)^2 - \int_0^s |f(r)|^2 dr - \int_s^t |f(r)|^2 dr|\mathcal{F}_s\right]$$

$$= I_s^2 - \int_0^s |f(r)|^2 dr + 2I_s \mathbb{E}\left[\int_s^t f(r)dW_r|\mathcal{F}_s\right]$$
$$+\mathbb{E}\left[\left|\int_s^t f(r)dW_r\right|^2 |\mathcal{F}_s\right] - \mathbb{E}\left[\int_s^t |f(r)|^2 dr|\mathcal{F}_s\right]$$

Using the properties of *Itô* integral we can see that

$$\mathbb{E}\left[\int_s^t f(r)dW_r|\mathcal{F}_s\right] = 0$$

and also

$$\mathbb{E}\left[\left|\int_s^t f(r)dW_r\right|^2 |\mathcal{F}_s\right] = \mathbb{E}\left[\left|\int_s^t f(r)dW_r\right|^2\right]$$
$$\mathbb{E}\left[\int_s^t |f(r)|^2 dr\right] = \mathbb{E}\left[\int_s^t |f(r)|^2 dr|\mathcal{F}_s\right]$$

So, we finally obtain

$$\mathbb{E}\left[I_t^2 - \int_0^t |f(r)|^2 dr|\mathcal{F}_s\right] = I_s^2 - \int_0^s |f(r)|^2 dr$$

and from the definition of square integrable process for a martingale we end the proof. □

## 5.3 *Itô* processes

Using the *Itô* stochastic integral we can define a new group of stochastic processes from Wiener process, the *Itô* processes.

**Definition 5.6.** *An Itô process is a stochastic process of $X_t$ such that*

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dW_s \tag{5.3}$$

*where $u, v$ satisfy the properties:*

$$\int_0^t v^2(s,\omega)ds < \infty \ \ a.s., \quad \int_0^t u(s,\omega)ds < \infty \ \ a.s.$$

*This process can be written as*

$$dX_t = udt + vdW_t$$

We can see from this that an *Itô* process can be broken in two pieces: The $M_t := \int_0^t vdW_t$ which is a martingale and the $A_t := \int_0^t uds$ which is a process of finite variation.

**Example 5.4** (A model of stock prices with *Itô* process)**.** *The stochastic process*

$$X_t = X_0 + \int_0^t \left(\mu(s) - \frac{\sigma^2(s)}{2}\right) ds + \int_0^t \sigma(s) dW_s$$

*where $\mu(t)$ and $\sigma(t)$ satisfy the properties of the Definition 5.6. In differential form we can write this process as*

$$dX_t = \left(\mu(t) - \frac{\sigma^2(t)}{2}\right) dt + \sigma(t) dW_t$$

*If we take $X_t = \ln S_t$ then $S_t$ is also an Itô process which used in finance for modelling of stock prices. If $\mu$ and $\sigma$ are fixed, $S_t$ is called geometric Wiener process.*

**Example 5.5** (A model for interest rates with *Itô* processes)**.** *The stochastic process*

$$X_t = X_0 + a\mu \int_0^t e^{as} ds + \sigma \int_0^t e^{as} dW_s$$

*is an Itô process which can be written in the differential form*

$$dX_t = a\mu e^{at} dt + \sigma e^{at} dW_t$$

*If we take $X_t = e^{at} r_t$ we will see that it is an Itô process which called Ornstein Uhlenbeck. We can see that*

$$\mathbb{E}[r_t] = e^{-at} X_0 + \mu(1 - e^{-at}) \to \mu, \quad as \quad t \to \infty$$

*The expected value depends on time, and for long time tends to $\mu$. Furthermore, using properties of Itô integral we can calculate the variation of $r_t$ which is*

$$Var(r_t) = \frac{\sigma^2}{2a}(1 - e^{-2at})$$

**Example 5.6.** *We use now (5.3) in the function* $g(t,x)=\frac{1}{2}x^2$ *to calculate the integral* $I=\int_0^t W_s dW_s$.
*Let* $Y_t=g(t,W_t)=\frac{1}{2}W_t^2$. *Using (5.3) in this function we obtain*

$$\frac{1}{2}W_t^2=\int_0^t W_s dW_s+\frac{1}{2}\int_0^t dt=\int_0^t W_s dW_s+\frac{1}{2}t$$

*so,*

$$I=\frac{1}{2}W_t^2-\frac{1}{2}t.$$

An interesting equation is which is the form of an *Itô* function and if this is also an *Itô* process, and if is true for Riemann's sums and *Itô* integrals.
So, we can see the next theorem and we have the answer.

**Theorem 5.5** (*Itô's* lemma)**.** *We consider that* $X_t$ *is an Itô process which can be written as*

$$X_t=X_0+\int_0^t u(s,\omega)ds+\int_0^t v(s,\omega)dW_s.$$

*Then, anyone function of* $X_t$ *which is of the form* $g(t,x)\,\epsilon\, C^{1,2}$ *can be written as a stochastic integral*

$$g(t,X_t)=g(0,X_0)+\int_0^t\left(\frac{\partial g}{\partial t}+u\frac{\partial g}{\partial x}+\frac{1}{2}v^2\frac{\partial^2 g}{\partial x^2}\right)dt+v\frac{\partial g}{\partial x}dW_t$$

*With* $C^{1,2}$ *we symbolize the space of* $g(t,x)$ *functions which have continuous the first derivative to the first variable and continuous second derivative to the second variable.*

*Proof.* We present here the basic steps for the proof of *Itô's* lemma.
Let consider the partition $t_j=\frac{jt}{n}$ for the interval $[0,t]$ and let write

$$g(t,X_t)-g(0,X_0)=\sum_{j=0}^{n-1}[g(t_{j+1}^n,X_{t_{j+1}^n})-g(t_j^n,X_{t_j^n})]=$$

$$\sum_{j=0}^{n-1}\underbrace{g(t_{j+1}^n,X_{t_{j+1}^n})-g(t_j^n,X_{t_{j+1}^n})}-\sum_{j=0}^{n-1}\underbrace{g(t_j^n,X_{t_{j+1}^n}-g(t_j^n,X_{t_j^n}))}$$

The first doss of underline terms consists of terms which are calculate for the same value of $Itô$ process, but for different times. In first doss terms we use Taylor analysis for $t$. The second doss consists of terms which are calculate for the same value of time but for different values of $X$. In the second doss we use also Taylor analysis for variable $X$. We use $\Delta t_j = t^n_{j+1} - t^n_j$ and $\Delta X_j = X_{t^n_{j+1}} - X_{t^n_j}$.

• Taylor for $t$: Exists $\overline{t}_j \, \epsilon \, [t^n_j, t^n_{j+1}]$ such that

$$g(t^n_{j+1}, X_{t^n_{j+1}}) - g(t^n_j, X_{t^n_{j+1}}) = \frac{\partial g}{\partial t}(\overline{t}_j, X_{t^n_{j+1}})\Delta t_j$$

Using the continuity of derivatives for $t$ of the function $g$ we can prove that

$$\lim_{n\to\infty} \sum_{j=0}^{n-1} \frac{\partial g}{\partial t}(\overline{t}_j, X_{t^n_{j+1}})\Delta t_j = \int_0^t \frac{\partial g}{\partial s}(s, X_s)ds \;\; a.s.$$

• Taylor analysis for $x$: Exists $\overline{X}_j \, \epsilon \, [X_{t^n_j}, X_{t^n_{j+1}}]$, such that

$$g(t^n_j, X_{t^n_{j+1}}) - g(t^n_j, X_{t^n_j}) = \frac{\partial g}{\partial x}(t^n_j, X_{t^n_j})\Delta X_j + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t^n_j, \overline{X}_j)\Delta X_j^2$$

This second doss needs more attention. First we obtain the approximation

$$\Delta X_j = u(\overline{t}_j, \omega)\Delta t_j + v(\overline{t}_j, \omega)\Delta B_j$$

where $\Delta B_j = B_{t^n_{j+1} - B^n_t}$ and as $\overline{t}_j$ we can choose $\overline{t}_j = t^n_j$. Using this approximation in Taylor for $x$ we obtain

$$g(t^n_j, X_{t^n_{j+1}}) - g(t^n_j, X_{t^n_j}) = \frac{\partial g}{\partial x}(t^n_j, X_{t^n_j})u(\overline{t}_j, \omega)\Delta t_j + \frac{\partial g}{\partial x}(t^n_j, X_{t^n_j})u(\overline{t}_j, \omega)\Delta B_j$$

$$+\frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t^n_j, \overline{X}_j)v(\overline{t}_j, \omega)^2\Delta B_j^2 + \underline{\frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t^n_j, \overline{X}_j)u(\overline{t}_j, \omega)\Delta t_j^2}$$

$$+\underline{\frac{\partial^2 g}{\partial x^2}(t^n_j, \overline{X}_j)u(\overline{t}_j, \omega)v(\overline{t}_j, \omega)\Delta t_j, \Delta B_j}$$

The sum of the two underlined terms can be proved that tends to zero in $L_2$ as $n \to \infty$. From the other terms we can see that

$$\sum_{j=0}^{n-1} \frac{\partial g}{\partial x}(t^n_j, X_{t^n_j})u(\overline{t}_j, \omega)\Delta t_j \to \int_0^t \frac{\partial g}{\partial x}(s, X_s)u(s, X_s)ds, \;\; n \to \infty$$

and

$$\sum_{j=0}^{n-1} \frac{\partial g}{\partial x}(t_j^n, X_{t_j^n}) u(\overline{t}_j, \omega) \Delta B_j \to \int_0^t \frac{\partial g}{\partial x}(s, X_s) v(s, X_s) dB_s, \quad n \to \infty$$

where the limit is always in $L_2$. The final term must be write as

$$\frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t_j^n, \overline{X}_j) v(\overline{t}_j, \omega)^2 \Delta B_j^2 = \frac{1}{2}\frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t_j^n, \overline{X}_j) v(\overline{t}_j, \omega)^2 \Delta t_j$$

$$+\frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t_j^n, \overline{X}_j) v(\overline{t}_j, \omega)^2 (\Delta B_j^2 - \Delta t_j)$$

$$+\frac{1}{2}\left(\frac{\partial^2 g}{\partial x^2}(t_j^n, \overline{X}_j) - \frac{\partial^2 g}{\partial x^2}(t_j^n, X_{t_j^n}) v^2 \Delta B_j^n\right)$$

The sum of the semifinal term is zero because in $L_2$ the limit of $\Delta B_j^2 - \Delta t_j$ is zero and the sum of final term is also zero because of the continuity of the second derivatives of the function g as x. The sum of the first term is

$$\sum_{j=0}^{n-1} \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t_j^n, X_{t_j^n}) v(t_j^n, \omega)^2 \Delta t_j \to \frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(s, X_s) v(s, \omega)^2 dt$$

With this interval we and the proof.

□

*Itô's* lemma is a very important result for stochastic analysis and is very useful in many applications. Also, this Lemma applies in case when $t$ is a stopping time and is bounded. Finally, there are other forms of this Lemma in which the function can satisfy weaker conditions than $C^{1,2}$.

## 5.4 *Itô* processes on $\mathbb{R}^n$

In this section we want to see *Itô* integral and processes on $\mathbb{R}^n$, i.e. when Wiener process is on $\mathbb{R}^n$ and the elements in integral are possible a vector function.

First, we will see the *Itô* integral on a multidimensional Wiener process.

**Definition 5.7.** *Let consider* $W_t = (W_{1,t}, ..., W_{d,t})^T$ *a d-dimensional Wiener process and* $f \in \mathbb{R}^{n\times d}$ *a family of adapted stochastic processes on filtration* $\mathcal{F}_t = \sigma(W_s,\ s \leq t)$ *which take values on space of matrices* $n \times d$*. Equivalent we can consider that* $f = \{f_{ij}\},\ i = 1, ...n,\ j = 1, ..., d$ *where* $f_{ij} \in \mathbb{R}$ *are dimensional stochastic processes which are adapted on filtration* $\mathcal{F}_t = \sigma(W_s,\ s \leq t)$*. The Itô integral* $\int_0^t f dW_s$ *is a stochastic process* $I_t \in \mathbb{R}^{n\times 1}$ *which has the form* $I_t = (I_{1,t}, ..., I_{n,t})^T$ *where*

$$I_{i,t} = \sum_{j=1}^{d} \int_0^t f_{i,j} dW_{j,s},\ \ i = 1, ..., n$$

*and* $\int_0^t f_{ij} dW_{j,s}$ *is the dimensional integral of Itô.*

So, we can write using matrices

$$\int_0^t f dW_s = \int_0^t \begin{pmatrix} f_{11} & \cdots & f_{1d} \\ \vdots & & \vdots \\ f_{n1} & \cdots & f_{nd} \end{pmatrix} \begin{pmatrix} dW_{1,s} \\ \vdots \\ dW_{d,s} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^d \int_0^t f_{1j} dW_{j,s} \\ \vdots \\ \sum_{j=1}^d \int_0^t f_{nj} dW_{j,s} \end{pmatrix}$$

The multidimensional *Itô* integral satisfies the properties of dimensional *Itô* integral and for the proof use the properties of the dimensional *Itô* integral.

We want also see the *Itô* processes on $\mathbb{R}^n$. Using the multidimensional *Itô* integral we can define these.

**Definition 5.8.** $X_t$ *is a dimensional Itô process if*

$$X_t = X_0 + \int_0^t u(s, \omega) ds + \sum_{i=1}^{d} \int_0^t v_i(s, \omega) dW_{i,s}$$

*where $u$ and $v_i$ are adapted on $\mathcal{F}_t$ for all $i$ and satisfies the next:*

$$\int_0^T |u(s,\omega)|ds < \infty, \quad \int_0^T |v_i(s,\omega)|^2 ds < \infty \ \ a.s. \ \ 1 \leq i \leq d.$$

*In differential form we can write*

$$dX_t = u(t,\omega)dt + \sum_{i=1}^{d} v_i(s,\omega)dW_{i,t}$$

Now we can define the n-dimensional *Itô* process.

**Definition 5.9.** *A n-dimensional Itô process is a stochastic process $X_t = (X_{1,t}, ..., X_{n,t})$ where any $X_{i,t}$, $i = 1, ..., n$ is an Itô process. More specifically, a n-dimensional Itô process is a stochastic process in the form $X_t = (X_{1,t}, ..., X_{n,t})^T$ where*

$$X_{i,t} = X_{i,0} + \int_0^t u_i(s,\omega)ds + \sum_{i=1}^{d} \int_0^t v_{ij}(s,\omega)dW_{j,s}$$

The most simple multidimensional *Itô* process is a n-dimensional Wiener process $W_t = (W_{1,t}, ..., W_{n,t})^T$.

## 5.5 Stochastic integral for $Q$-Wiener and Weak Wiener processes

In this section, we define the integral

$$\int_0^t \Psi(s)dW_s, \ \ t \, \epsilon \, [0,T],$$

of an operator-valued stochastic process $\{\Psi(t), \ t \geq 0\}$ with respect to a $Q$-Wiener process $W_t, \ t \geq 0 \ \ (\mathrm{Tr}Q < \infty)$, which is called the *stochastic integral*.

Integrals of such form will be used for constructing solutions to stochastic Cauchy problems in Hilbert spaces.
Let $(\Omega, \mathcal{F}, P)$ be a probability space and $U$ and $H$ be Hilbert spaces. Let us discuss some properties of stochastic processes, which are essential for constructing the stochastic integral.

In this section, we assume that $\mathcal{T}$ is equal to $[0, \infty)$ or $[0, T]$. Let us introduce the notation

$$\Omega_\infty = [0, \infty) \times \Omega, \quad \Omega_T = [0, T] \times \Omega.$$

On $\Omega_\infty$, we introduce the $\sigma$-field $\mathcal{B}_\infty$ generated by sets of the form

$$(s, t] \times F, \quad F \, \epsilon \, \mathcal{F}_s, \; 0 \leq s < t < \infty,$$

$$\{0\} \times F, \quad F \, \epsilon \mathcal{F}_0.$$

Denote by $P_\infty$ the product of the Lebesgue measure on $[0, \infty)$ with the probability measure $P$ on $\Omega$. Denote by $\mathcal{B}_T$ the $\sigma$-field that is a restriction of $\mathcal{B}_\infty$ to $\Omega_T$. It is easy to see that $\mathcal{B}_T$ is generated by sets of the form

$$(s, t] \times F, \quad F \, \epsilon \, \mathcal{F}_s, \; 0 \leq s < t < T, \quad \{0\} \times F, \quad F \, \epsilon \mathcal{F}_0. \tag{5.4}$$

Denote by $\mathcal{P}_T$ the product of the Lebesgue measure on $[0, T]$ by the probability measure $P$ on $\Omega$.

**Definition 5.10.** *(a) A measurable mapping from $(\Omega_\infty, \mathcal{P}_\infty)$ or $(\Omega_T, \mathcal{P}_T)$ into $U, \mathcal{B}(U)$ is said to be predictable.*
*(b)If, for any $t \, \epsilon \, \mathcal{T}$, the mapping $u : (\cdot, \cdot) : [0, t] \times \Omega \to U$ is $\mathcal{B}([0, t]) \times \mathcal{F}_t$-measurable, then the process $\{u(t), \; t \, \epsilon \, \mathcal{T}\}$ is said to be progressively measurable.*

**Proposition 5.2.** *An adapted and stochastically continuous process on an interval $[0, T]$ has a progressively measurable and predictable version.*

We assume that $(\Omega, \mathcal{F}, P)$ is a probability space with a given normal filtration $\{\mathcal{F}_t,\ t \geq 0\}$.

**Definition 5.11.** *An $\mathcal{L}(U,H)$-valued process $\{\Phi(t),\ t \in [0,T]\}$ is called step if there exists a set $0 = t_0 < t_1 < ... < t_k = T$ and $\mathcal{L}(U,H)$-valued random variables $\Phi_0, \Phi_1, ..., \Phi_{k-1}$ such that $\Phi_m$ is $\mathcal{F}_{t_m}$-measurable and*

$$\Phi(t) = \Phi_m, \quad t \in (t_m, t_{m+1}],$$

*for any $m = 0, ..., k-1$.*

Here and in what follows, the measurability of an $\mathcal{L}(U,H)$-valued random variables means the strong measurability.

**Definition 5.12.** *A function $\Phi : \Omega \to \mathcal{L}(U,H)$ is called to be strong measurable if for any $u \in U$, $\Phi(\cdot)u$ is measurable as a mapping from $(\Omega, \mathcal{F})$ into $(H, \mathcal{B}(H))$.*

Let $\{W_t,\ t \geq 0\}$ be a $Q$-Wiener process with values in $U$. We assume that $W_t,\ t \geq 0$ is a process with respect to $\{\mathcal{F}_t\},\ t \geq 0$, i.e.
(i) the random variable $W_t$ is $\mathcal{F}_t$-measurable for any $t \geq 0$;
(ii) the increment $W_{t+h} - W_t$ is independent of $\mathcal{F}_t$ for all $t,\ h \geq 0$.
For an elementary process $\{\Phi(t),\ t \in [0,T]\}$, the stochastic integral with respect to $W(\cdot)$, denoted by $\Phi \cdot W(\cdot)$, is defined as follows:

$$\int_0^t \Phi(s)dW_s := \sum_{m=0}^{k-1} \Phi_m(W_{t_{m+1}\wedge t} - W_{t_m \wedge t}), \quad t \in [0,T] \tag{5.5}$$

Further, we indicate the class of $\mathcal{L}(U,H)$-valued process for which the stochastic integral can be defined as the mean-square limit(Definition 4.3) of sums of the form (5.5).
Since $Q$ is the covariance operator of $\{W_t,\ t \geq 0\}$, it is a symmetric nonnegative trace class operator, and there exists an orthonormal basis of eigenvectors of $Q$ $\{e_j\}$ in $U$. In the previous section, we introduce the Hilbert space

$U_0 = Q^{1/2}U$ with the norm $||u||_{U_0} = ||Q^{-1/2}u||_U$ (Definition 4.9). It follows from the definition that the system $\{g_j\} = \{\sqrt{\lambda_j}e_j\}$, where $\lambda_j e_j = Qe_j$, forms an orthonormal basis in $U_0$. Consider $\mathcal{L}_{GS}(U_0, H)$-the space of Hilbert-Schmidt operators mapping $U_0$ into $H$.

**Proposition 5.3.** *For any $\Psi \epsilon \mathcal{L}_{GS}(U_0, H)$, the operators $\Psi\Psi^*$ and $\Psi Q^{1/2}(\Psi Q^{1/2})^*$ act on H and*

$$Tr[\Psi\Psi^*] = ||\Psi||^2_{GS} = Tr[\Psi Q\Psi^*].$$

*Proof.* Let $\{f_k\}_{k=1}^{\infty}$ be an orthonormal basis in H. Using the definition of adjoint operator, we obtain

$$Tr[\Psi\Psi^*] = \sum_{k=1}^{\infty} \langle \Psi\Psi^* f_k, f_k\rangle_H = \sum_{k=1}^{\infty} \langle \Psi^* f_k, \Psi^* f_k\rangle_{U_0}$$

$$= \sum_{k=1}^{\infty} \left\langle \sum_{j=1}^{\infty} \langle \Psi^* f_k, g_j\rangle_{U_0 g_j}, \sum_{j=1}^{\infty} \langle \Psi^* f_k, g_j\rangle_{U_0 g_j} \right\rangle_{U_0}$$

$$= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \langle \Psi^* f_k, g_j\rangle_{U_0} \langle \Psi^* f_k, g_j\rangle_{U_0} = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \langle f_k, \Psi g_j\rangle^2_H$$

$$= \sum_{k=1}^{\infty}\sum_{k=1}^{\infty} \langle f_k, \Psi g_j\rangle^2_H = \sum_{j=1}^{\infty} ||\Psi g_j||^2_H = ||\Psi||^2_{GS}.$$

By the definition of the inner product in $U_0$, for all $h_1, h_2 \epsilon U_0$, we have

$$\langle h_1, h_2\rangle_{U_0} = \left\langle Q^{1/2}h_1, Q^{1/2}h_2\right\rangle_U.$$

Taking into account that $Q^{1/2}$ is self-adjoint in $U_0$, we obtain

$$||\Psi||^2_{GS} = \sum_{k=1}^{\infty} \langle \Psi^* f_k, \Psi^* f_k\rangle_{U_0}$$

$$= \sum_{k=1}^{\infty} \left\langle Q^{1/2}\Psi^* f_k, Q^{1/2}\Psi^* f_k\right\rangle_U$$

$$= \sum_{k=1}^{\infty} \langle \Psi Q \Psi^* f_k, f_k \rangle_H = Tr[\Psi Q \Psi^*].$$

□

The following proposition shows that an operator from $\mathcal{L}(U, H)$ can be regarded as an element from $\mathcal{L}_{GS}(U_0, H)$.

**Proposition 5.4.** *If $\Psi \in \mathcal{L}(U, H)$, then $\Psi_0 := \Psi_{|U_0}$ belongs to $\mathcal{L}_{GS}(U_0, H)$ and*

$$Tr[\Psi Q \Psi^*] = ||\Psi_0||^2_{GS} = Tr[\Psi_0 \Psi_0^*]. \tag{5.6}$$

*Proof.* First, we show that $\Psi_0 \in \mathcal{L}_{GS}(U_0, H)$. We have

$$||\Psi_0||^2_{GS} = \sum_{j=1}^{\infty} ||\Psi_0 g_j||^2_H = \sum_{j=1}^{\infty} \left|\left| \Psi \left( \sqrt{\lambda_j} e_j \right) \right|\right|^2_H =$$

$$\sum_{j=1}^{\infty} \lambda_j ||\Psi e_j||^2_H \leq ||\Psi||^2 \sum_{j=1}^{\infty} \lambda_j < \infty.$$

The second equality in (5.6) is valid by the previous proposition. We show the first one. For the orthonormal basis $\{f_k\}_{k=1}^{\infty}$ in $H$, using the definiton of adjoint operator, we obtain

$$Tr[\Psi Q \Psi^*] = \sum_{k=1}^{\infty} \langle \Psi Q \Psi^* f_k, f_k \rangle_H = \sum_{k=1}^{\infty} \langle Q \Psi^* f_k, \Psi^* f_k \rangle_U$$

$$= \sum_{k=1}^{\infty} \left\langle \sum_{j=1}^{\infty} Q \langle \Psi^* f_k, e_j \rangle_U e_j, \sum_{j=1}^{\infty} \langle \Psi^* f_k, e_j \rangle_U e_j \right\rangle_U$$

$$= \sum_{k=1}^{\infty} \left\langle \sum_{j=1}^{\infty} \lambda_j \langle \Psi^* f_k, e_j \rangle_U e_j, \sum_{j=1}^{\infty} \langle \Psi^* f_k, e_j \rangle_U e_j \right\rangle_U$$

$$= \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j \langle \Psi^* f_k, e_j \rangle^2_U = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \left\langle \Psi^* f_k, \sqrt{\lambda_j} e_j \right\rangle^2_U$$

$$= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \langle \Psi^* f_k, g_j\rangle_U^2 = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \langle f_k, \Psi g_j\rangle_H^2$$

$$= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \langle f_k, \Psi_0 g_j\rangle_H^2 = \sum_{j=1}^{\infty} ||\Psi_0 g_j||_H^2 = ||\Psi_0||_{GS}^2.$$

□

The following theorem states a fundamental relation, which will be used in the definition of the stochastic integral and in the description of the class of integrable process. First, we will see another useful proposition.

**Proposition 5.5.** *Let $(H, \mathcal{B}(H))$ and $(U, \mathcal{B}(U))$ be measurable spaces and $\psi : U \times H \to \mathbb{R}$ be a bounded measurable function. Let $u$ and $w$ be two random variables on $(\Omega, \mathcal{F}, P)$ assuming values in U and H, respectively, and let $\mathcal{G}$ be a $\sigma$-field contained in $\mathcal{F}$. Assume that $w$ is $\mathcal{G}$-measurable and $u$ is independent of $\mathcal{G}$. Then*

$$\mathbb{E}(\psi(u, w)|\mathcal{G}) = \mathbb{E}(\psi(u, w)) \ \ P\ a.s.$$

**Theorem 5.6** (*Itô* isometry)**.** *If $\{\Phi(t),\ t \geq 0\}$ is an elementary process with values in $\mathcal{L}_{GS}(U_0, H)$, and for some $T \leq \infty$,*

$$\left\{\mathbb{E}\left[\int_0^T ||\Phi(s)||_{GS}^2 ds\right]\right\}^{1/2} < \infty, \tag{5.7}$$

*then*

$$\mathbb{E}[||\Phi \cdot W_t||_H^2] = \mathbb{E}\left[\int_0^t ||\Phi(s)||_{GS}^2 ds\right], \quad 0 \leq t \leq T. \tag{5.8}$$

*Proof.* Let $t\ \epsilon\ [0, T]$. For definiteness, we assume that $t\ \epsilon\ (t_m, t_{m+1}]$. Denote, $\Delta W_{j,t} = W_{t_{j+1}\wedge t} - W_{t_j \wedge t}, j = 1, ..., m$. Then

$$\mathbb{E}[||\Phi \cdot W_t||_H^2] = \mathbb{E}\left[\left|\left|\sum_{j=0}^{m} \Phi_j \Delta W_{j,t}\right|\right|_H^2\right]$$

$$= \mathbb{E}\left[\sum_{j=0}^{m} ||\Phi_j \Delta W_{j,t}||_H^2\right] + 2\mathbb{E}\left[\sum_{j=0}^{m}\sum_{i=0}^{j-1} \langle \Phi_i \Delta W_{i,t}, \Phi_j \Delta W_{j,t}\rangle_H\right].$$

For any $0 \leq j \leq m$,

$$\mathbb{E}[||\Phi_j \Delta W_{j,t}||_H^2] = \mathbb{E}\left[\sum_{k=1}^{\infty} \langle \Phi_j \Delta W_{j,t}, f_k \rangle_H^2\right] = \sum_{k=1}^{\infty} \mathbb{E}[\left\langle \Delta W_{j,t}, \Phi_j^* f_k \right\rangle_H^2].$$

By the definition of the elementary process, for any $u \in U$, the random variable $\Phi_j u$ is $\mathcal{F}_{t_j}$-measurable with respect to $\mathcal{B}(H)$. Therefore, $\langle \Phi_j u, h \rangle_H$ is $\mathcal{F}_{t_j}$ measurable with respect to $\mathcal{B}(\mathbb{R})$ for all $h \in H$. Hence $\left\langle u, \Phi_j^* h \right\rangle_U$ is $\mathcal{F}_{t_j}$ measurable with respect to $\mathcal{B}(\mathbb{R})$ for all $u \in U$; this implies that $\Phi_j^* h$ is $\mathcal{F}_{t_j}$ measurable with respect to $\mathcal{B}(U)$. This and the fact that $\Delta W_{j,t}$ is independent of $\mathcal{F}_{t_j}$ imply

$$\mathbb{E}[||\Phi_j \Delta W_{j,t}||_H^2] = \sum_{k=1}^{\infty} \mathbb{E}[\mathbb{E}[\left\langle \Delta W_{j,t}, \Phi_j^* f_k \right\rangle_U^2 | \mathcal{F}_{t_j}]],$$

where, by Proposition 5.4 and the definition of the covariance operator,

$$\mathbb{E}[\left\langle \Delta W_{j,t}, \Phi_j^* f_k \right\rangle_U^2 | \mathcal{F}_{t_j}] = \mathbb{E}[\left\langle \Delta W_{j,t}, \Phi_j^* f_k \right\rangle_U^2]$$

$$= \left\langle Cov[\Delta W_{j,t}] \Phi_j^* f_k, \Phi_j^* f_k \right\rangle_U^2 = (t_{j+1} - t_j) \left\langle Q \Phi_j^* f_k, \Phi_j^* f_k \right\rangle_U.$$

Hence,

$$\sum_{k=1}^{\infty} \mathbb{E}[\mathbb{E}[\left\langle \Delta W_{j,t}, \Phi_j^* f_k \right\rangle_H^2 | \mathcal{F}_{t_j}]] = \sum_{k=1}^{\infty} \mathbb{E}[\left\langle (t_{j+1} - t_j) Q \Phi_j^* f_k, \Phi_j^* f_k \right\rangle_U]$$

$$= (t_{j+1} - t_j) \sum_{k=1}^{\infty} \mathbb{E}[\left\langle \Phi_j Q \Phi_j^* f_k, f_k \right\rangle_H] = (t_{j+1} - t_j) \mathbb{E}\left[\sum_{k=1}^{\infty} \left\langle \Phi_j Q \Phi_j^* f_k, f_k \right\rangle_H\right]$$

$$= (t_{j+1} - t_j) \mathbb{E}[Tr[\Phi_j Q \Phi_j^*]].$$

For $j = m$, the increment $\Delta W_{m,t} = W_t - W_{t_m}$ is independent of $\mathcal{F}_{t_m}$ since $t > t_m$. The same argument yields

$$\mathbb{E}[||\Phi_m \Delta W_{m,t}||_H^2] = (t - t_m) \mathbb{E}[Tr[\Phi_m Q \Phi_m^*]].$$

Therefore,

$$\mathbb{E}\left[\sum_{j=0}^{m} ||\Phi_j \Delta W_{j,t}||_H^2\right] = \sum_{j=0}^{m} \left((t_{j+1} \wedge t) - (t_j \wedge t)\right) \mathbb{E}[||\Phi_j||_{GS}^2]$$

$$= \mathbb{E}\left[\sum_{j=0}^{m}\left((t_{j+1}\wedge t)-(t_j\wedge t)\right)||\Phi_j||^2_{GS}\right] = \mathbb{E}\left[\int_0^t ||\Phi(s)||^2_{GS}ds\right].$$

For $j \neq i$, we have $\mathbb{E}[\langle \Phi_i \Delta W_{i,t}, \Phi_j \Delta W_{j,t}\rangle_H] = 0$.

Finally we obtain (5.8).

□

Now we are ready to define the stochastic integral with respect to a $Q$-Wiener process.However, we will see first one more proposition.

**Proposition 5.6.** *For a Hilbert space $H$ (a separable metric space $(H,\rho)$) and an $H$-valued random variable $u$, there exists a sequence of simply $H$-valued random variables $\{u_n\}$ such that $\rho(u(\omega),u_n(\omega))$ decreases monotonically to zero for any $\omega \epsilon \Omega$.*

**Theorem 5.7.** *The following statements hold:*
*(i)if a mapping of the set $\Omega_T$ into $\mathcal{L}(U,H)$ is $\mathcal{L}(U,H)$-predictable, then it is $\mathcal{L}_{GS}(U_0,H)$-predictable. In particular, elementary process are $\mathcal{L}_{GS}(U_0,H)$-predictable;*
*(ii) if $\{\Psi(t),\ t\ \epsilon\ [0,T]\}$ is an $\mathcal{L}_{GS}(U_0,H)$-predictable process with property (5.7), then there exists a sequence $\{\Psi_n(t),\ t\ \epsilon\ [0,T]\}$ of elementary processes such that*

$$\mathbb{E}\left[\int_0^T ||\Psi(s)-\Psi_n(s)||^2_{GS}ds\right] \to 0 \quad as \quad n\to\infty. \tag{5.9}$$

*Proof.* It is proved that for any set from a $\sigma$-field on $\mathcal{L}_{GS}(U_0,H)$, its inverse image $G$ is measurable in $\Omega_T$, which is due to the fact that $G$ can be approximated by sets of the form (5.4). Hence, by Proposition 5.4, $\{\Psi(t),\ t\geq 0\}$ is $\mathcal{L}_{GS}(U_0,H)$-predictable.
By Proposition 5.4, the space $\mathcal{L}(U,H)$ is densely embedded into $\mathcal{L}_{GS}(U_0,H)$. Next, by Proposition 5.6, there exists a sequence $\{\Psi_n\}$ of elementary $\mathcal{L}(U,H)$-valued predictable processes on [0,T] such that

$$||\Psi(\omega,t)-\Psi_n(\omega,t)||_{GS} \to 0 \ \ as \ \ n\to\infty$$

for all $(\omega, t)\ \epsilon\ \Omega_T$. Consequently, (5.9) holds.

□

Thus, we can introduce the class of stochastically integrable processes. It follows from Theorem 5.7 that under condition (5.7) they are $\mathcal{L}(U, H)$-predictable processes.
We note that the integral is constructed with respect to a $Q$-Wiener process. However, as was shown in the previous section, a weak Wiener process can also be regarded as a $Q_1$-Wiener process in the Hilbert space $U_1$. Therefore, the above definition is applicable to integrals with respect to $Q$-Wiener processes and also those with respect to weak Wiener processes.

# Chapter 6

# Stochastic Convolution

Let $(\Omega, \mathcal{F}, P)$ be a probability space with a given normal filtration $\{\mathcal{F}_t,\ t \geq 0\}$, and let $U$ and $H$ be a separable Hilbert spaces.
In this chapter we consider the stochastic Cauchy problem:

$$du(t) = Au(t)dt + BdW(t)\ , \quad t\ \epsilon\ [0,T),\ \ T \leq \infty, \quad u(0) = \xi, \tag{6.1}$$

with $A$ being the generator of a strongly continuous semigroup of operators $\{S(t), t \geq 0\}$ in $H$.

**Definition 6.1.** *Let $T$ be a semigroup. The generator of $T$, denoted by $A$, is given by the equation:*

$$Af = \lim_{t \to 0^+} A_t f = \lim_{t \to 0+} \frac{T(t) - f}{t}$$

*where the limit is evaluated in terms of the norm on $H$ and $f$ is in the domain of $A$ if and only if this limit exists.*

**Definition 6.2.** *A one parameter family of bounded linear operators $\{S(t),\ t \geq 0\}$ on $H$ is called a strongly continuous semigroup (or a semigroup of class $C_0$) if the following conditions hold:*
*(i) the semigroup property:$S(t+s) = S(t) + S(s), t, s \geq 0$*
*(ii) $S(0) = I$*

*(iii) the operator function $S(t)$ is strongly continuous with respect to $t \geq 0$.*

We assume that $\{W(t),\ t \geq 0\}$ is a $Q$-Wiener process with respect to $\{\mathcal{F}_t,\ t \geq 0\}$ with values in $U, B$ is a bounded linear operator from $U$ into $H$, and $\xi$ is an $\mathcal{F}_0$-measurable $H$-valued random variable.

## 6.1 $C_0$-semigroups and well-posedness of the deterministic Cauchy Problem

Before defining solutions of the stochastic Cauchy problem (6.1), we turn to the deterministic problems

$$u'(t) = Au(t) + f(t), \quad t\epsilon\ [0,t),\ \ T \leq \infty,\ \ u(0) = x\ \epsilon\ D(A) \subset H, \tag{6.2}$$

$$u'(t) = Au(t),\ \ t\ \epsilon\ [0,T),\ \ T \leq \infty,\ \ u(0) = x\ \epsilon\ D(A) \subset H, \tag{6.3}$$

and recall the conditions for their well-posedness. The results on the well-posedness of problems (6.2) and (6.3) given here are valid in an arbitrary Hilbert space $H$.

**Remark** : The IVP given by

$$\frac{d}{dt}u(t) = a[u(t)],\ \ t \leq 0,\ \ u(0) = f$$

with $A$ being linear, is well posed if and only if $A$ is the generator of a semigroup $T$.

This turns out to be quite important as it provides both necessary and sufficient conditions to determine if a problem is well-posed.

**Definition 6.3.** *A solution of the deterministic Cauchy problem (6.2) is a function* $u(\cdot) \epsilon C([0,T); D(A)) \cap C^1([0,T); H^*)$, *satisfying the initial condition and Eq. (6.2)*

**Definition 6.4.** *The homogeneous deterministic Cauchy problem (6.3) is said to be uniformly well posed if for any* $x \epsilon D(A)$, *if there exists a unique solution to (6.3) such that for all* $t \epsilon [0,T)$,

$$||u(t)||_H \leq C||x||_H,$$

*for some constant* $C > 0$.

Owing to the differential operator structure of the equation, the well-posedness of (6.2), (6.3) is closely to the theory of semigroups of operators.

The operator $A$ in Definition 6.2 defined by

$$Ax := S'(0)x = \lim_{t \to 0} t^{-1}(S(t) - I)x,$$

with the domain

$$D(A) = \{x \epsilon H \ : \ \lim_{t \to 0} t^{-1}(S(t) - I)x \ exists\},$$

is called the *generator* of the semigroup $\{S(t), t \geq 0\}$.

The generator of a strongly continuous semigroup is a closed densely defined operator commuting with the semigroup on its own domain, and

$$S(t)Ax = AS(t)x = S'(t)x, \ \ t \geq 0, x \epsilon D(A). \tag{6.4}$$

A strongly continuous semigroup is exponential bounded:

$$||S(t)||_H \leq Ce^{at}, \ \ t \geq 0,$$

for some constants $C > 0$ and $a \epsilon \mathbb{R}$. Owing to this property, the Laplace transform of the semigroup is well defined and gives an equivalent definition

of the generator as an operator $A$ satisfying the relation

$$(\lambda I - A)^{-1}x = \int_0^\infty e^{-\lambda t}S(t)xdt, \;\; x \,\epsilon\, H, \;\; \mathrm{Re}\lambda > a.$$

This relation obviously implies that the domain of the generator coincides with the range of its resolvent:$D(A) = \mathrm{ran}(\lambda I - A)^{-1}$.

**Remark** :[Hille-Yosida Theorem]
A linear unbounded operator $A$ is the generator of a $C_0$ semigroup if and only if:
(i) $A$ is a closed operator,
(ii) $A$ has dense domain $(D(A))$,
(iii) for each $\lambda > 0, \; \lambda \,\epsilon\, \rho(A)$, and
(iv) $||R(\lambda, A)|| \leq \frac{1}{\lambda}$.

The Hille-Yosida Theorem is very powerful as it gives us both necessary and sufficient conditions.

One of the main well-posedness results for the deterministic Cauchy problem can be formulated as follows.

**Theorem 6.1.** *Let $A$ be a linear operator with nonempty resolvent set $(\rho(A) \neq \emptyset)$. Then the following assertions are equivalent:*
*(i) the homogeneous Cauchy problem (6.3) is uniformly well posed;*
*(ii) the operator $A$ is the generator of the strongly continuous semigroup $\{S(t), \; t \geq 0\}$ in $H$;*
*(iii) the resolvent of $A$ satisfies the Hille-Yosida type conditions: there exists constants $C > 0$ and $a \,\epsilon\, \mathbb{R}$ such that*

$$\left|\left|\frac{d^k}{d\lambda^k}(\lambda I - A)^{-1}\right|\right|_{\mathcal{L}(H)} \leq \frac{Ck!}{(Re\lambda - a)^{k+1}}, \;\; k = 0, 1, 2, ... \qquad (6.5)$$

*for all $\lambda \,\epsilon\, C$ with $Re\lambda > a$.*
*In this case, for any $x \,\epsilon\, D(A)$, the solution to the problem (6.3) is given*

*by* $u(t) = S(t)x,\ t \geq 0$. *If, moreover,* $f \in C([0,T), D(A))$, *then for any* $x \in D(A)$,

$$u(t) = S(t)x + \int_0^t S(t-s)f(s)ds,\ \ t \in [0,T), \tag{6.6}$$

*is a solution to (6.2).*

## 6.2 Solution to the stochastic Cauchy problem

We now discuss the notion of solutions to the stochastic Cauchy problem (6.1).

**Definition 6.5.** *An H-valued predictable process* $\{u(t),\ t \in [0,T)\}$ *is said to be strong solution to (6.1) if*
*(i)*$u(t)$ *takes values in* $D(A)$ $P_T$ *a.s.*
*(ii)* $\int_0^T ||Au(t)||_H dt < \infty$ $P$ *a.s.*
*(iii) for any* $t \in [0,T)$,

$$u(t) = \xi + \int_0^t Au(s)ds + BW(t),\ \ P \text{ a.s.}$$

In other words, a strong solution is a predictable process $\{u(t),\ \in [0,T)\}$ taking values in $D(A)$ for almost all $t \in [0,T)$ and $\omega \in \Omega$ such that the trajectories of the process $\{Au(t),\ t \in [0,T)\}$ are integrated for almost $\omega \in \Omega$, it satisfies the integrable equation (6.1) and are called *mild solution*.

Note that the definition of a strong solution implies that the process $\{BW(t),\ t \geq 0\}$ should be well defined as an $H$-valued stochastic process. Therefore, this definition has a sense for a $Q$-Wiener process $\{W(t),\ t \geq 0\}$ only if $\mathrm{Tr}BQB^* < \infty$, where $B$ is a bounded operator $U$ into $H$.

We will show that the requirement that a solution belong to $D(A)$ is a strong restriction on the class of admissible processes (Theorem 6.6). For this reason, we introduce one more definition of a solution to (6.1).

**Definition 6.6.** *An H-valued predictable process* $\{u(t),\ t \epsilon [0,T)\}$ *is said to be a weak solution to (6.1) if*
*(i)*$\int_0^T ||u(t)||_H dt < \infty$ *P a.s.*
*(ii)for any* $y \epsilon D(A^*)$ *and* $t \epsilon [0,T)$.

$$\langle u(t), y\rangle_H = \langle \xi, y\rangle_H + \int_0^t \langle u(s), A^* y\rangle_H \, ds + \langle BW(t), y\rangle_H , \quad P \ \ a.s. \tag{6.7}$$

Similarly to the solution (6.6) of the deterministic Cauchy problem, we show that any solution to the stochastic problem (6.1) is the sum of the process $\{S(t)\xi,\ t \geq 0\}$ and a special kind of the stochastic integral, called the *stochastic convolution* :

$$W_A(t) := \int_0^t S(t-s)BdW(s), \ \ t \epsilon [0,T) \tag{6.8}$$

that is, we show that the process

$$u(t) = S(t)\xi + \int_0^t S(t-s)BdW(s), \ \ t \epsilon [0,T) \tag{6.9}$$

gives a solution to (6.1). it is easy to see that $\{S(t)\xi,\ t \geq 0\}$ is the adapted process as an action of deterministic process $\{S(t),\ t \geq 0\}$ on the $\mathcal{F}_0$-measurable random variable $\xi$. For the same reason, it is stochastically continuous, and hence, by Proposition 5.2, it has a predictable version. The expectation of the process is

$$\mathbb{E}(S(t)\xi) = S(t)\mathbb{E}(\xi), \ \ t \geq 0.$$

Let us find its covariation operator. For all $h_1, h_2 \epsilon H$ and $t \geq 0$,

$$\langle Cov(S(t)\xi)h_1, h_2\rangle_H = \mathbb{E}\,\langle S(t)\xi - S(t)\mathbb{E}(\xi), h_1\rangle_H \,\langle S(t)\xi - S(t)\mathbb{E}(\xi), h_2\rangle_H$$
$$= \mathbb{E}\,\langle S(t)\xi, h_1\rangle_H \,\langle S(t)\xi, h_2\rangle_H - \mathbb{E}\,\langle S(t)\mathbb{E}(\xi), h_1\rangle_H \,\langle S(t)(\xi), h_2\rangle_H$$

$$-\mathbb{E}\left\langle S(t)\xi, h_1\right\rangle_H \left\langle S(t)\mathbb{E}(\xi), h_2\right\rangle_H + \mathbb{E}\left\langle S(t)\mathbb{E}(\xi), h_1\right\rangle_H \left\langle S(t)\mathbb{E}(\xi), h_2\right\rangle_H$$

$$= \mathbb{E}\sum_{j=1}^{\infty} \left\langle S(t)\xi, e_j\right\rangle_H^2 \left\langle h_1, e_j\right\rangle_H \left\langle h_2, e_j\right\rangle_H - \left\langle S(t)\xi, h_1\right\rangle_H \left\langle S(t)\mathbb{E}(\xi), h_2\right\rangle_H .$$

By the definition of the adjoint operator and the covariance of a real-valued random variable, we obtain

$$\left\langle Cov(S(t)\xi)h_1, h_2\right\rangle_H$$

$$= \sum_{j=1}^{\infty} \left\langle h_1, e_j\right\rangle_H \left\langle h_2, e_j\right\rangle_H \mathbb{E}\left\langle \xi, S(t)^* e_j\right\rangle_H^2 - \left\langle S(t)\mathbb{E}(\xi), h_1\right\rangle_H \left\langle S(t)\mathbb{E}(\xi), h_2\right\rangle_H$$

$$= \sum_{j=1}^{\infty} \left\langle h_1, e_j\right\rangle_H \left\langle h_2, e_j\right\rangle_H \left\langle Cov(\xi), S(t)^* e_j\right\rangle_H - \left\langle S(t)\mathbb{E}(\xi), h_1\right\rangle_H \left\langle S(t)\mathbb{E}(\xi), h_2\right\rangle_H$$

$$\sum_{j=1}^{\infty} \left\langle h_1, e_j\right\rangle_H \left\langle h_2, e_j\right\rangle_H \left\langle S(t)Cov(\xi), e_j\right\rangle_H - \left\langle S(t)\mathbb{E}(\xi), h_1\right\rangle_H \left\langle S(t), \mathbb{E}(\xi), h_2\right\rangle_H$$

$$= \left\langle S(t)Cov(\xi)h_1, h_2\right\rangle_H - \left\langle S(t)\mathbb{E}(\xi), h_1\right\rangle_H \left\langle S(t)\mathbb{E}(\xi), h_2\right\rangle_H$$

and all $h_1, h_2 \,\epsilon\, H$ and $t \geq 0$. In particular, if $\mathbb{E}(\xi) = 0$, then

$$Cov(S(t)\xi) = S(t)Cov(\xi), \quad t \geq 0.$$

Let us study the properties of the stochastic convolution.

**Theorem 6.2.** *Let $B$ be a bounded operator from $U$ into $H$ and $A$ be the generator of a strongly continuous semigroup $\{S(t),\ t \geq 0\}$ in $H$ such that $S(t)B \,\epsilon\, \mathcal{L}_{HS}(U, H)$ for all $t \geq 0$. Let*

$$\int_0^T ||S(t)B||_{GS}^2 dt < \infty \tag{6.10}$$

*for some $T \leq \infty$. Then the stochastic convolution $W_A(t),\ t \,\epsilon\, [0, T)$ is an $H$-valued Gaussian process with a predictable version and mean square continuous on $[0, T)$ with the characteristics*

$$\mathbb{E}(W_A(t)) = 0, \quad Cov(W_A(t)) = \int_0^t S(t-s)BQB^*S^*(t-s)ds, \quad t \,\epsilon\, [0, T). \tag{6.11}$$

*Proof.* We note that $S(t)B,\ t \geq 0$ is $\mathcal{L}_{HS}(U,H)$ predictable as a deterministic process, and hence

$$\mathbb{E}\int_0^t ||S(r)B||_{GS}^2 dr = \int_0^t ||S(r)B||_{GS}^2 dr, \quad t \,\epsilon\, [0,T).$$

By Theorem 5.7, it follows from (6.10) and the above relation that the stochastic convolution is well defined. By definition, the stochastic convolution is the limit of integrals of elementary processes approximating the deterministic process $S(t)B,\ t \geq 0$. Taking into account that $\{\Delta W(t),\ t \geq 0\}$ are Gaussian processes, we obtain that the process $\{W_A(t),\ t \geq 0\}$ is also Gaussian. To show its mean-square continuity, we take $0 \leq s \leq t < T$; then

$$\mathbb{E}[||W_A(t) - W_A(s)||_H^2] = \mathbb{E}\left[\left|\left|\int_0^t S(t-r)BdW(r) - \int_0^s S(s-r)BdW(r)\right|\right|_H^2\right]$$

$$= \mathbb{E}\left[\left|\left|\int_0^t S(t-r)BdW(r) - \int_{t-s}^t S(t-r)BdW(r)\right|\right|_H^2\right]$$

$$\mathbb{E}\left[\left|\left|\int_0^{t-s} S(t-r)BdW(r)\right|\right|_H^2\right].$$

*Itô's* isometry (Theorem 5.6) implies

$$\mathbb{E}[||W_A(t) - W_A(s)||_H^2] = \mathbb{E}\left[\int_0^{t-s} ||S(t-r)B||_{GS}^2 dr\right] = \int_0^{t-s} ||S(t-r)B||_{GS}^2 dr,$$

which proves the mean-square continuity. To prove (6.11), we recall that $S(t)B \,\epsilon\, \mathcal{L}_{HS}(U,H)$ for all $t \,\epsilon\, [0,T)$. Therefore, $S(t)BQ^{1/2} \,\epsilon\, \mathcal{L}_{HS}(U,H)$ and $(S(t)BQ^{1/2})^* \,\epsilon\, \mathcal{L}_{HS}(H,U)$. Then by Proposition 4.2, the operator $S(t)BQB^*S^*(t) \,\epsilon\, \mathcal{L}_N(H,U)$ and

$$||S(t)BQB^*S^*(t)||_N \leq ||S(t)BQ^{1/2}||_{HS} \cdot ||S(t)BQ^{1/2}||_{HS}, \ \ t \,\epsilon\, [0,T).$$

Hence the integral in (6.11) is well defined as a Bochner integral. Now consider the moments of stochastic convolution. We have

$$\mathbb{E}(W_A(t)) = \mathbb{E}\left(\int_0^t S(t-s)BdW(s)\right) = \int_0^t S(t-s)B\mathbb{E}(dW(s)) = 0$$

For the covariance operator, we obtain

$$\langle Cov(W_A(t))h_1, h_2\rangle_H = \mathbb{E}\,\langle W_A(t), h_1\rangle_H \langle W_A(t), h_2\rangle_H$$

$$= \mathbb{E}\left\langle \int_0^t S(t-s)BdW(s), h_1\right\rangle_H \left\langle \int_0^t S(t-s)BdW(s), h_2\right\rangle_H$$

$$= \mathbb{E}\sum_{k=1}^{n}\langle S(t-s_k)B\Delta W_k, h_1\rangle_H \sum_{k=1}^{n}\langle S(t-s_k)B\Delta W_k, h_2\rangle_H$$

$$= \mathbb{E}\sum_{0}^{n}\langle \Delta W_k, (S(t-s_k)B)^* h_1\rangle_H \sum_{k=1}^{n}\langle \Delta W_k, (S(t-s_k)B^*)h_2\rangle_H$$

$$= \mathbb{E}\sum_{k=1}^{n}\left\langle \Delta W_k, \sum_{j=1}^{\infty}\langle (S(t-s_k)B)^* h_1, e_j\rangle e_j\right\rangle_H \sum_{k=1}^{n}\left\langle \Delta W_k, \sum_{j=1}^{\infty}\langle (S(t-s_k)B)^* h_2, e_j\rangle e_j\right\rangle_H$$

$$= \mathbb{E}\sum_{k=1}^{n}\sum_{j=1}^{\infty}\langle \Delta W_k, e_j\rangle_H^2 \langle (S(t-s_k)B)^* h_1, e_j\rangle_H \langle (S(t-s_k)B)^* h_2, e_j\rangle_H$$

$$= \sum_{k=1}^{n}\sum_{j=1}^{\infty}\langle \Delta s_k Q e_j, e_j\rangle_H \langle (S(t-s_k)B)^* h_1, e_j\rangle_H \langle (S(t-s_k)B)^* h_2, e_j\rangle_H$$

$$= \sum_{k=1}^{n}\sum_{j=1}^{\infty}\lambda_j \langle (S(t-s_k)B)^* h_1, e_j\rangle_H \langle (S(t-s_k)B)^* h_2, e_j\rangle_H \Delta s_k$$

$$\sum_{k=1}^{n}\sum_{j=1}^{\infty}\left\langle (S(t-s_k)B)^* h_1, Q^{1/2}e_j\right\rangle_H \left\langle (S(t-s_k)B)^* h_2, Q^{1/2}e_j\right\rangle_H \Delta s_k$$

$$\sum_{k=1}^{n}\left\langle Q^{1/2}(S(t-s_k)B)^* h_1, Q^{1/2}(S(t-s_k)B)^* h_2\right\rangle_H \Delta s_k$$

$$= \int_0^t \langle S(t-s)BQB^* S(t-s)^* h_1, h_2\rangle_H\, ds$$

for all $h_1, h_2 \,\epsilon\, H$.

□

**Remark** :[Fubini's Theorem]

An always important tool for the computation of integral on product spaces is Fubini's theorem. We consider first the case of non-negative functions.

Let $(E, \varepsilon)$ be a measureble space and let $\Phi : (t, \omega, x) \to \Phi_t(\omega, x)$ be a measurable mapping from $(\mathbb{R}_+ \times \Omega \times E, \mathcal{P} \otimes \mathcal{E})$ into $(L_2^0(H), \mathcal{B}(L_2^0(H)))$. In addition let $\mu$ denote a finite positive measure on $(E, \mathcal{E})$.
Fix $T > 0$. By localization we get the following version of

**Theorem** (The Stochastic Fubini Theorem)**.** *Assume that*

$$\int_0^T \int_E ||\Phi_t(x)||^2_{L_2^0(H)} \mu(dx)dt < \infty, \ \ \mathbb{P} - a.s.$$

*Then there exists an $\mathbb{F}_T \otimes \mathcal{E}$ -measurable version $\xi(\omega, x)$ of the stochastic integral $\int_0^T \Phi_t(x)dW_t$ which is $\mu$-integrable $\mathbb{P}$-a.s. such that*

$$\int_E \xi(x)\mu(dx) = \int_0^T \left( \int_E \Phi_t(x)\mu(dx) \right) dW_t, \ \ \mathbb{P} - a.s.$$

**Theorem 6.3.** *Under the conditions of Theorem 6.2 , the process*

$$u(t) = S(t)\xi + \int_0^t S(t-s)BdW(s), \ \ t \, \epsilon \, [0, T)$$

*is a weak solution to*

$$du(t) = Au(t) + BdW(s), \ \ t \, \epsilon \, [0, T), \ T \leq \infty, \ \ u(0) = \xi$$

*Proof.* We note that process (6.9) is predictable with integrable trajectories as a sum of processes with these properties. To prove (6.7), we take arbitrary $t \, \epsilon \, [0, T)$ and $y \, \epsilon \, D(A^*)$ and consider the integral

$$\int_0^t \langle u(s), A^*y \rangle_H \, ds = \int_0^t \langle S(s)\xi, A^*y \rangle_H \, ds + \int_0^t \langle W_A(s), A^*y \rangle_H \, ds. \quad (6.12)$$

Since $A$ generates the strongly continuous semigroup $\{S(t), \ t \geq 0\}$ in $H$, its dual $A^*$ generates the strongly continuous semigroup $\{S^*(t), \ t \geq 0\}$ in $H^*$.

Therefore, for $y \epsilon D(A^*)$, property (6.4) of a strongly continuous semigroup implies

$$\int_0^t \langle S(s)\xi, A^*y\rangle_H \, ds = \int_0^t \langle \xi, S^*(s)A^*y\rangle_H \, ds = \left\langle \xi, \int_0^t S^*(s)A^*y ds \right\rangle_H$$

$$= \left\langle \xi, \int_0^t \frac{d}{ds}(S^*(s)y)ds \right\rangle_H = \langle \xi, (S^*(t) - S^*(0)y)\rangle_H$$

$$= \langle S(t)\xi - \xi, y\rangle_H = \langle S(t)\xi, y\rangle_H - \langle \xi, y\rangle_H$$

for the first term in (6.12). we apply the stochastic Fubini theorem to the second term in (6.12) and obtain

$$\int_0^t \langle W_A(s), A^*y\rangle_H \, ds = \int_0^t \left\langle \int_0^s S(s-r)BdW(r), A^*y \right\rangle_H ds$$

$$= \left\langle \int_0^t \int_r^t S(s-r)BdsdW(r), A^*y \right\rangle_H .$$

By the definition of the stochastic integral and using the properties of the adjoint operator and the continuity of inner product, we have

$$\int_0^t \langle W_A(s), A^*y\rangle_H \, ds = \left\langle \lim_{n\to\infty} \sum_{m=0}^{n-1} \left( \int_{r_m}^t S(s-r_m)Bds \right) (W(r_m+1) - W(r_m)), A^*y \right\rangle_H$$

$$= \lim_{n\to\infty} \sum_{m=0}^{n-1} \left\langle W(r_{m+1}) - W(r_m), \left( \int_{r_m}^t S(s-r_m)Bds \right)^* A^*y \right\rangle_H$$

$$= \lim_{n\to\infty} \sum_{m=0}^{n-1} \left\langle W(r_{m+1}) - W(r_m), \left( \int_{r_m}^t B^*S^*(s-r_m)ds \right) A^*y \right\rangle_H$$

$$= \lim_{n\to\infty} \sum_{m=0}^{n-1} \left\langle W(r_{m+1}) - W(r_m), \int_{r_m}^t B^*S^*(s-r_m)A^*yds \right\rangle_H$$

$$= \lim_{n\to\infty} \sum_{m=0}^{n-1} \left\langle W(r_{m+1}) - W(r_m), \int_{r_m}^t \frac{d}{ds}(B^*S^*(s-r_m)y)ds \right\rangle_H$$

$$= \lim_{n\to\infty} \sum_{m=0}^{n-1} \left\langle W(r_{m+1}) - W(r_m), [B^*S^*(s-r_m)y]_{r_m}^t \right\rangle_H$$

$$= \lim_{n\to\infty} \sum_{m=0}^{n-1} \langle (S(t-r_m)B - B)(W(r_{m+1}) - W(r_m)), y\rangle_H$$

$$= \left\langle \int_0^t (S(t-r)B - B)dW(r), y \right\rangle_H = \langle W_A(t), y\rangle_H - \langle BW(t), y\rangle_H .$$

Substituting the above expressions in (6.12), we obtain

$$\int_0^t \langle u(s), A^*y\rangle_h \, ds = \langle S(t)\xi, y\rangle_H - \langle \xi, y\rangle_H + \langle W_A(t), y\rangle_H - \langle BW(t), y\rangle_H$$

$$= \langle u(t), y\rangle_H - \langle \xi, y\rangle_H - \langle BW(t), y\rangle_H$$

which proves (6.7).

□

To prove that the solution (6.9) is unique we need the following result.

**Lemma 6.1.** *Let $u(t),\ t \in [0,T)$ be a weak solution to (6.1) with $\xi = 0$. Then for arbitrary $y(\cdot) \in C^1([0,T); dD(A^*))$ and $t \in [0,T)$ we have the representation*

$$\langle u(t), y(t)\rangle_H = \int_0^t \langle u(s), y'(s) + A^*y(s)\rangle_H \, ds + \int_0^t \langle BdW(s), y(s)\rangle_H . \quad (6.13)$$

*Proof.* First, let us consider the function $y(\cdot)$ of the form $y(t) = f(t)y_0,\ \ t \in [0,T)$, where $y_0 \ \in D(A^*)$ and $f(\cdot) \in C^1([0,T))$. For any $t \in [0,T)$, we have

$$\langle u(t), y_0\rangle_H = \int_0^t \langle u(s), A^*y_0\rangle_H \, ds + \langle BW(t), y_0\rangle_H .$$

For the real-valued process $\langle u(s), y_0\rangle_H f(s),\ \ s\,\epsilon\,[0,T)$, $It\hat{o}'s$ formula gives

$$d(\langle u(s), y_0\rangle_H f(s)) = f(s)d\langle u(s), y_0\rangle_H + f'(s)\langle u(s), y_0\rangle_H ds.$$

Hence,

$$\langle u(t), y_0\rangle_H f(t) = \int_0^t f(s)\langle u(s), A^*y_0\rangle_H ds + \int_0^t \langle BdW(s), f(s)y_0\rangle_H + \int_0^t f'(s)\langle u(s), y_0\rangle_H ds,$$

which implies (6.13) for the functions $y(\cdot)$ considered. Since these functions form a dense subset in $C^1([0,T); D(A^*))$, the proof is completed.

□

**Theorem 6.4.** *The weak solution (6.9) to the problem (6.1) is unique.*

*Proof.* By the properties of the well-posed deterministic abstract Cauchy problem, it suffices, to prove that the weak solution (6.9) corresponding to $\xi = 0$ is unique. Let $y_0\,\epsilon\,D(A^*)$ and $t\,\epsilon\,[0,T)$. We apply Lemma 6.1 to $y(s) = S^*(t-s)y_0,\ \ s\,\epsilon\,[0,t]$; using the properties of strongly continuous semigroups, we obtain

$$\langle u(t), y(t)\rangle_H = \langle u(t), S^*(0)y_0\rangle_H = \langle u(t), y_0\rangle_H$$

$$= \int_0^t \langle u(s), y'(s)\rangle_H ds + \int_0^t \langle u(s), A^*y(s)\rangle_H ds + \int_0^t \langle y(s), BdW(s)\rangle_H$$

$$= \int_0^t \left\langle u(s), \frac{d}{ds}S^*(t-s)y_0\right\rangle_H - \int_0^t \left\langle u(s), \frac{d}{ds}S^*(t-s)y_0\right\rangle_H + \int_0^t \langle y(s), BdW(s)\rangle_H$$

$$\int_0^t \langle y(s), BdW(s)\rangle_H = \int_0^t \langle S^*(t-s)y_0, BdW(s)\rangle_H = \int_0^t \langle y_0, S(t-s)BdW(s)\rangle_H .$$

Since $D(A^*)$ is defined in $H$, we conclude that $u(t) = W_A(t),\ t\,\epsilon\,[0,T)$.

□

The following result is devoted to the continuity of the obtained weak solution to (6.1).

**Theorem 6.5.** *Assume that $U = H$, $B = I$, and for some $a > 0$,*

$$\int_0^T s^{-a} ||S(s)||_{GS}^2 ds < \infty.$$

*Then the weak solution to (6.1) has a continuous version.*

Thus, we have discussed weak solutions to (6.1). For strong solutions, the following result holds.

**Theorem 6.6.** *Let $Q$ be a trace class operator in $U$, $U = H$, and $A \epsilon \mathcal{L}_{GS}(H)$. Let $\xi \epsilon D(A)$, $P$ a.s. Then (6.9) is a strong solution to (6.1).*

In accordance with this theorem, the existence of a strong solution to (6.1) is not guaranteed even for a bounded operator $A$. A strong solution exists only for Hilbert-Schmidt operators. This condition on $A$ is connected with requirements (i) and (ii) of Definition 6.4 : the stochastic convolution should take values in $D(A)$ and trajectories of the process $\{AW_A(t), t \epsilon [0, T)\}$ should be integrated almost surely. The lack of smoothness of a Wiener process implies a stronger requirement as compared with the requirements on $A$ in the deterministic case. Therefore, the existence of weak solutions is more interesting from the applied point of view. The regularity for the deterministic problem $u'(t) \epsilon Au(t) + f$, we need $A$ be a bounded generator for strongly semigroup and also $f \epsilon C^1([0, T]; H)$, where $f$ has a form $f = BdW(t)$.

# Chapter 7

# The Heath Jarrow Morton Model

We now present the stochastic model of a bond market riskless zero coupon bonds, and we take advantage of our short excursion in the world of infinite dimensional stochastic analysis to generalized the HJM model.

We introduce the time value of money by valuing the simplest possible fixed income instrument. Like for all the other financial instruments considered in this book, we define it by specifying its cash flow. In the present situation,the instrument provides a single payment of a fixed amount (the principal or nominal value X) at a given date in the future. This date is called the maturity date. If the time to maturity is exactly n years, the present value of this instrument is:

$$P(X, n) = \frac{1}{(1+r)^n} X.$$

This formula gives the present value of a nominal amount X due in n years time. Such an instrument is called a discount bond or a zero coupon bond because the only cash exchange takes place at the end of the life of the instrument, i.e. at the date of maturity. The positive number r is referred to as the annual discount rate or spot interest rate for time to maturity n since it is the interest rate which is applicable today on an n-year loan.

## 7.1 The Bond Market

Throughout this chapter we assume the existence of a frictionless market (in particular we ignore transaction costs) for riskless zero coupon bonds of all maturities. As before, we follow the convention in use in the financial mathematics literature and we denote by $P(t,T)$ the price at time $t$ of a zero coupon bond with maturity date $T$ and nominal value 1 euro. So we assume the existence of a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ and for each $T > 0$, of a non-negative adapted process $\{P(t,T); 0 \leq t \leq T\}$ which satisfies $P(T,T) = 1$. We shall specify the dynamics of the bond prices in an indirect way, namely through prescriptions for the instantaneous forward rates, but as explained earlier, this is quite all right. We assume that our bond prices $P(t,T)$ are differentiable functions of the maturity date $T$, so we define the instantaneous forward rates as:

$$f(t,T) = -\frac{\partial \log P(t,T)}{\partial T}$$

in such a way that:

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right).$$

We shall use also Musiela's notation:

$$P_t(x) = P(t,t+x) \quad \text{and} \quad f_t(x) = f(t,t+x), \quad t,x \geq 0.$$

## 7.2 The HJM Evolution Equation

One of the goals of this chapter is to analyse the HJM equation:

$$df_t(x) = \left(\frac{\partial}{\partial x} f_t(x) + a_t(x)\right) dt + \sum_{i=1}^{\infty} \sigma_t^i(x) dw_t^i \tag{7.1}$$

where $\{w^{(i)}\}_i$ are independent scalar Wiener processes, and where the drift is given by the famous HJM no-arbitrage condition

$$a_t(x) = \sum_{i=1}^{\infty} \sigma_t^i(x) \left(\int_0^x \sigma_t^i(u)du + \lambda_t^i\right). \tag{7.2}$$

The theory of financial mathematics is based upon the assumption of (long term) arbitrage opportunities. This Assumption is compatible with the theory of generator function in financial mathematics.
An arbitrage opportunity is a portfolio that generates a certain process.

We would like to think of the forward rate curve $x \mapsto f_t(x)$ as an $\mathcal{F}_t$ measurable random vector taking values in a function space $F$. Once we choose an appropriate space $F$, we will interpret Eq. (7.1) by rewriting it as a stochastic evolution equation in $F$.

We consider $w = \{w^{(i)}\}$ a cylindrical Wiener process. The cylindrical Wiener process is a generalization of the Q-Wiener process.

**Definition 7.1.** *A cylindrical Wiener process on a Hilbert space $V$ is a family of mappings $\check{W}(t,\omega) : V \to L_2(\Omega, \mathcal{F}, P; \mathbb{R})$ such that for every $u \in V$, the real-valued random variable $\check{W}(t,\omega)(v) := \langle \check{W}(t,\omega), v \rangle$ follows the centered normal distribution $N(0,t)$ and $\mathbb{E}[\check{W}(t,\omega)(v_1)\check{W}(t,\omega)(v_2)] = t(v_1,v_2)_V$.*

We would also like to think of our Wiener process as a cylindrical Wiener process defined on a real separable Hilbert space $G$.
We do not even need any special features of the space $G$ except that it is infinite dimensional. Because the eigenvalues of a HilbertSchmidt operator must decay fast enough for the sum of their squares to be finite, assuming that G is infinite dimensional does not disagree with the principal component analysis, used to justify the introduction of models with finitely many factors or HJM models with finite rank volatility. No generality would be lost letting $G = \ell_2$ and the reader is free to substitute $\ell_2$ everywhere $G$ appears in what follows. Of course, choosing $G = \ell_2$ is equivalent to fixing a basis for $G$ and working with the coordinates of vectors expressed in this basis. We prefer, though, to keep our presentation free of coordinates whenever possible. Also, keeping $G$ unspecified allows for the possibility that the Wiener process takes values in a function space. Equivalently, the infinite dimensional Wiener process may be viewed as a two parameter random field with a tensor covariant structure. In any event, we pick our favorite $G$ once and for all, and fix it for

the remainder of the chapter. To simplify the presentation, we will always identify $G$ with its dual $G^*$.

### 7.2.1 Function Spaces for Forward Curves

The first ask is term structure modelling is to choose the state space $F$ for the forward rate dynamics in such a way that the mathematical analysis of Eq. (7.1) is clean. However this space should be general enough to accommodate as large a family of models as possible. We now list the assumptions that we use to carry out this analysis.

**Assumption 7.1**. 1. The space $F$ is a separable Hilbert space and the elements of $F$ are continuous, real-valued functions. The domain $\chi$ of these functions is either a bounded interval $[0, x_{max}]$ or the half-line $\mathbb{R}_+$. We also assume that for every $x \epsilon \chi$, the evaluation functional:

$$\delta_x(f) = f(x)$$

is well-defined, and is in fact a continuous linear function on $F$, i.e. an element of the dual space $F^*$.

2. The semigroup $\{S_t\}_{t\geq 0}$ where $\{S_t\}$ is the left shift operator , is strongly continuous and defined by:

$$(S_t f)(x) = f(t + x) \tag{7.3}$$

The generator of $\{S_t\}_{t\geq 0}$ is the (possibly unbounded) operator $A$.

3. The map $F_{HJM}$ is measurable from some non-empty subset $D \subset \mathcal{L}_{HS}(G, F)$ into $F$ where the HJM map $F_{HJM}$ is defined by

$$F_{HJM}(\sigma)(x) = \langle \sigma^* \delta_x, \sigma^* I_x \rangle_G$$

for each $\sigma \epsilon \mathcal{L}_{HS}(G, F)$, where $G$ is a given real separable Hilbert space and The definite integration functional $I_x$ defined by

$$I_x(f) = \int_0^x f(s)ds$$

is continuous on $F$ for each $x \epsilon \chi$ since

$$\left| \int_0^x f(s)ds \right| \leq x \sup_{s\epsilon[0,x]} |f(x)| \leq x \sup_{s\epsilon[0,x]} ||\delta_s||_{F^*} ||f||_F$$

and $\sup_{s\epsilon[0,x]} ||\delta_s||_{F^*}$ is finite by the Banach-Steinhaus theorem.

Let us remark on these assumptions. The most important property that the space $F$ should have is that elements of $F$ should be locally integrable functions indeed, the formula for the bond price:

$$P(t,T) = \exp\left( - \int_0^{T-t} f_t(s)ds \right),$$

should make sense. For instance, the classical Lebesgue spaces $L_p(\mathbb{R}_+)$ have this property. Recall, however, that space $L_p(\mathbb{R}_+)$ is in fact a space of equivalent classes of functions. As such, its elements are only defined almost everywhere, and they cannot be evaluated on a set of measure zero.
In our analysis, we will find it necessary to be able to evaluate a forward curve (i.e. an element of the space $F$) at a given time to maturity.
Fortunately, almost everyone working with the term structure of interest rates would agree that the forward curves should be smooth functions of the time to maturity. Hence our Assumption 7.1.1 is reasonable. Of course, the elements of $F$ are locally integrable, but more is true.

**Remark** : Let $X, Y$ be normed spaces and let $\mathcal{S} \subset \mathcal{B}(X,Y)$. Assume that

$$D := \{x \epsilon X | \sup_{T\epsilon\mathcal{S}} ||Tx|| < \infty\}$$

is fat in $X$. In particular, $D = X$.

The financial implication of Assumption 7.1.1 is that the short interest rate $r_t$ is well-defined as:

$$r_t = f_t(0)$$

Once the short rate is defined, the money-market account is defined by:

$$B_t = \exp\left( \int_0^t r_s ds \right)$$

It is the solution of the ordinary differential equation $dB_t = r_t B_t dt$ which satisfies the initial condition $B_0 = 1$. It is a traded asset that pays the floating interest rate $r_t$ continuously compounded. We shall use it as a numeraire, i.e. the unit in which the prices of all the other assets are expressed. Prices expressed in units of the numeraire are denoted with a tilde and are called discounted prices:

$$\tilde{P}(x) = B_t^{-1} P_t(x) = \exp\left( - \int_0^t r_s ds - \int_0^x f_t(y) dy \right). \tag{7.4}$$

We should mention that the assumption that $F$ has the structure of a separable Hilbert space is motivated rather by mathematical convenience than financial considerations.

The left shift operator $\{S_t\}_{t \geq 0}$ defined in Assumption 7.1.2 allows us to pass from the time of maturity notation $f(t,T)$ to Musielas time to maturity notation $f_t(x)$ where $f_t(x) = f(t, t+x)$. Note that all of the evaluation functionals $\delta_x = S_x^* \delta_0$ can be recovered by a left shift of the functional $\delta_0$. The connection between the shift operators and the presence of enough smooth functions relies on the fact that the shift operators form a semigroup of operators whose infinitesimal generator $A$ should be the operator of differentiation, in the sense that one should have $Af = f$ whenever $f$ is differentiable.
Assumption 7.1.3 is intimately related to the no-arbitrage principle. In particular, we will need the function $F_{HJM}$ in order to define the drift term of an abstract HJM model.

Note that since the elements of $F$ are continuous, the function $x \mapsto F_{HJM}(\sigma)(x)$ is continuous for all $\sigma \, \epsilon \, \mathcal{L}_{HS}(G,F)$. However, it is not necessarily true that $F_{HJM}(\sigma)$ is an element of $F$. In fact, for the spaces we shall consider, it is generally false that $F_{HJM}(\sigma)$ is an element of $F$ unless the operator $\sigma$ is an element of a proper subset $D \subset \mathcal{L}_{HS}(G,F)$.
Assumption 7.1.]3) is usually hard to check in practice. We give a sufficient condition.

**Assumption 7.2** The space $F$ satisfies Assumption 7.1.1 and 7.1.2. Furthermore, there exists a subspace $F^0 \subset F$ such that the binary operator $\star$ defined by the formula

$$(f \star g)(x) = f(x) \int_0^x g(s)ds$$

maps $F^0 \times F^0$ into $F$, and is such that for all $f, g \,\epsilon\, F^0$ the following bounded holds:

$$||f \star g||_F \leq C||f||_F||g||_F$$

for some constant $C > 0$.

**Proposition 7.1.** *Let the space $F$ satisfy Assumption 7.2. Then the map $F_{HJM}$ satisfies the local Lipschitz bound*

$$||F_{HJM}(\sigma_1) - F_{HJM}(\sigma_2)||_F \leq C||\sigma_1 + \sigma_2||_{\mathcal{L}_{HS}(G,F)}||\sigma_1 - \sigma_2||_{\mathcal{L}_{HS}(G,S)}$$

*for all Hilbert-Schmidt operators $\sigma_1, \sigma_2 \,\epsilon\, \mathcal{L}_{HS}(G, F^0)$ with ranges contained in $F^0$. In particular, the map $F_{HJM}$ is measurable from $D = \mathcal{L}_{HS}(G, F^0)$ into $F$.*

*Proof.* We have the simple estimate:

$$\begin{aligned} ||f \star f - g \star g|| &= \frac{1}{2}||(f-g) \star (f+g) + (f+g) \star (f-g)|| \\ &\leq C||f-g||||f+g||. \end{aligned}$$

Notice that the HJM function $F_{HJM}$ is then recovered by the norm convergent series

$$F_{HJM}(\sigma) = \sum_{i=1}^{\infty} (\sigma g_i) \star (\sigma g_i)$$

for $\sigma \,\epsilon\, \mathcal{L}_{HS}(G, F^0)$, where $\{g_i\}_{i \epsilon \mathbb{N}}$ is a complete orthonormal system for $G$. The proof is now complete since we have

$$||F_{HJM}(\sigma_1) - F_{HJM}(\sigma_2)||_F \leq \sum_{i=1}^{\infty} ||(\sigma_1 g_i) \star (\sigma_2 g_i) - (\sigma_2 g_i) \star (\sigma_1 g_i)||_F$$

$$= C \sum_{i=1}^{\infty} ||(\sigma_1 - \sigma_2 g_i)||_F ||(\sigma_1 + \sigma_2) g_i||_F \leq C ||\sigma_1 - \sigma_2||_{\mathcal{L}_{HS}(G,F)} ||\sigma_1 + \sigma_2||_{\mathcal{L}_{HS}(G,F)}$$

by the triangle and Cauchy-Schwarz inequalities.

□

In the same way that the short rate is defined as the value of the forward rate curve at the left hand point of the time to maturity interval $[0, x_{max}]$, the long interest rate $\ell_t$ is defined as the value of the forward rate curve at the right end point of the domain $\chi$. This is possible when $\chi = [0, x_{max}]$ is bounded, in which case:

$$\ell_t = f_t(x_{max}),$$

but it requires a special property of the space $F$ when the domain $\chi = \mathbb{R}_+$ is the halfline. Indeed, in order to define:

$$\ell_t = f_t(\infty)$$

we need to make sure that, for all $f \in F$, the limit:

$$f(\infty) = \lim_{x \to \infty} f(x)$$

exists.

## 7.3 The Abstract HJM Model

In this section, we formulate a precise definition of an HJM model in a function space $F$. We assume that $F$ satisfies Assumption 7.1. We fix a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual conditions and such that there exists a Wiener process $W$ defined cylindrically on the separable Hilbert space $G$. Let $\mathcal{P}$ be the predictable sigma-field on $\mathbb{R}_+ \times \Omega$. We now state a definition of an HJM model for the forward rate:

**Definition 7.2.** *An HJM model on $F$ is a pair of functions $(\lambda, \sigma)$ where:*
*(i) $\lambda$ is a measurable function from $(\mathbb{R}_+ \times \Omega \times F, \mathcal{P} \otimes \mathcal{B}_F)$ into $(G, \mathcal{B}_G)$,*
*(ii) $\sigma$ is a measurable function from $(\mathbb{R}_+ \times \Omega \times F, \mathcal{P} \otimes \mathcal{B}_F)$ into $(D, \mathcal{B}_{\mathcal{L}_{HS}(G,S)})$,*
*such that there exists a non-empty set of initial conditions $f_0 \in F$ for which there exists a unique, continuous mild $F$-valued solution $\{f_t\}_{t \geq 0}$ of the HJM equation:*

$$df_t = (Af_t + a(t, \cdot, f_t))dt + \sigma(t, \cdot, f_t)dW_t \tag{7.5}$$

*where*

$$a(t, \omega, f) = F_{HJM} \circ \sigma(t, \omega, f) + \sigma(t, \omega, f)\lambda(t, \omega, f).$$

If $(\sigma, \lambda)$ is an abstract HJM model on the space $F$ with initial condition $f_0 \in F$, then the forward rate process $\{f_t\}_{t \geq 0}$ satisfies the integral equation

$$f_t = S_t f_0 + \int_0^t S_{t-s} a(s, \cdot, f_s)ds + \int_0^t S_{t-s}\sigma(s, \cdot, f_s)dW_s. \tag{7.6}$$

We now use the Proposition 7.1 to give a sufficient condition for the existence of an HJM model.

**Proposition 7.2.** *Suppose that the state space $F$ satisfies Assumption 7.2, and let the closed subspace $F^0 \subset F$ be such that $||f \star g||_F \leq C||f||_F||g||_F$ for $f, g \in F^0$. Assume that for every $(t, \omega, f) \in \mathbb{R}_+ \times \Omega \times F$ the range of the operator $\sigma(t, \omega, f)$ is contained in the subspace $F^0$. If $\sigma$ is bounded and if the Lipschitz bounds*

$$||\sigma(t, \omega, f) - \sigma(t, \omega, g)||_{\mathcal{L}_{HS}(G,F)} \leq C||f - g||_F$$

$$||\sigma(t, \omega, f)\lambda(t, \omega, f) - \sigma(t, \omega, g)\lambda(t, \omega, g)||_F \leq C||f - g||_F$$

*are satisfied for some constant $C > 0$ and all $t \geq 0$, $\omega \in \Omega$ and $f, g \in F$, then the pair $(\lambda, \sigma)$ is an HJM model on $F$. Furthermore, for any initial forward curve $f_0 \in F$ there exists a unique, continuous solution to the Eq. (7.5) such that $\mathbb{E}\{\sup_{t \in [0,T]} ||f_t||_F^p\} < \infty$ for all finite $T \geq 0$ and $p \geq 0$.*

### 7.3.1 Drift Condition and Absence of Arbitrage

We now fix an HJM model $(\sigma, \lambda)$ with initial condition $f_0 \epsilon F$, and we denote by $\{f_t\}_{t\geq 0}$ the unique solution to Eq. (7.5). To simplify the notation, let $\lambda_t = \lambda(t, \omega, f_t)$ and $\sigma_t = \sigma(t, \omega, f_t)$.

**Theorem 7.1.** *If we have*

$$\mathbb{E}\{\exp\left(-\frac{1}{2}\int_0^t ||\lambda_s||_G^2 + \int_0^t \lambda_s dW_s\right)\} = 1$$

*and if*

$$\int_0^t \mathbb{E}\{\int_0^t ||\sigma_s^* \delta_{s-u}||_G^2 du\}^{1/2} ds < +\infty$$

*for all $t \geq 0$ then the market given by the HJM model $(\sigma, \lambda)$ admits no arbitrage.*

*Proof.* We compute the dynamics of the discounted bond price $\tilde{P}(t,T) = B_t^{-1} P(t,T)$. We will make use of the relation $S_a^* I_u = I_{u+a} - I_a$, which is revealed in the following calculation:

$$(S_a^* I_u)g = \int_0^u (S_a g)(s) ds = \int_0^u g(s+a) ds$$

$$= \int_a^{u+a} g(s) ds = (I_{u+a} - I_a)g.$$

Let us compute the dynamics of the bond price:

$$-\log P(t,T) = I_T - t f_t = I_{T-t} S_t f_0 + \int_0^t T_{T-t} S_{t-s} a_s ds + \int_0^t \sigma_s^* S_{t-s} I_{T-t} dW_s$$

$$= I_T f_0 - I_t r_0 + \int_0^t I_{T-t} a_s ds - \int_0^t I_{t-s} a_s ds + \int_0^t I_{T-s} \sigma_s dW_s - \int_0^t \sigma_s^* I_{t-s} dW_s$$

Now by the stochastic Fubini theorem and the assumption of the theorem we have

$$\int_0^t \left( \int_0^t \sigma(s)^* \delta_{t-u} dW_s \right) du = \int_0^t \sigma^* I_{t-s} dW_s.$$

Using $\log P(0,t) = -I_t f_0$ and

$$\int_0^t r_s(0) ds = I_t f_0 + \int_0^t I_{t-s} a_s ds + \int_0^t \sigma_s^* I_{t-s} dW_s.$$

we conclude that

$$\log P(t,T) = \log P(0,t) + \int_0^t (f_s(0) - I_{T_s} a_s) ds + \int_0^t \sigma_s^* I_{T_s} dW_s.$$

Now in $\{P(t,T)\}_{t\epsilon[0,T]}$ is in the form of an $It\hat{o}$ process. Applying $It\hat{o}'s$ formula yields

$$P(t,T) = P(0,T) + \int_0^t P(s,T) \left( f_s(0) - I_{T-s} a_s + \frac{1}{2} ||\sigma_s^* I_{T-s}||_G^2 \right) ds$$

$$- \int_0^t P(s,T) \sigma_s^* I_{T-s} dW_s$$

Finally, substituting $a_s = F_{HJM}(\sigma_s) + \sigma_s \lambda_s$, the discounted bond prices are given by

$$\tilde{P}(t,T) = P(0,T) - \int_0^t P(s,T) I_{T-s} \sigma_s d\tilde{W}_s$$

where $\tilde{W}_t = W_t + \int_0^t \lambda_s ds$. But the Cameron $-$ Martin Girsanov theorem says that there exists a measure $\mathbb{Q}$, locally equivalent to $\mathbb{P}$ such that the process $\tilde{W}_t = W_t + \int_0^t \lambda_s ds$ defines a cylindrical Wiener process on $G$ for the measure $\mathbb{Q}$. We will find that for each $T > 0$ the discounted bond prices are local martingales under the measure $\mathbb{Q}$. Hence, by the fundamental theorem, there is no arbitrage.

□

The current framework may be too general for practical needs. At this level of generality, we only know that the discounted bond prices are local martingales. They are bona-fide martingales if

$$\mathbb{E}^{\mathbb{Q}} \left[ \exp \left( -\frac{1}{2} \int_0^T ||\sigma_s^* I_{T-s}||_G^2 + \int_0^t \sigma_s^* I_{T-s} d\tilde{W}_s \right) \right] = 1.$$

We can ensure that the discounted bond prices are martingales if we can check the well-known Novikov condition

$$\mathbb{E}^{\mathbb{Q}}\left[\exp\left(\frac{1}{2}\int_0^T ||\sigma_s^* I_{T-s}||_G^2\right)\right] < +\infty.$$

Alternatively, we can ensure the discounted bond prices are martingales if the forward rates are positive almost surely, since if the rates are positive, the discounted bond prices $\tilde{P}(t,T) = \exp(-\int_0^t f_s(0)ds - \int_0^{T-t} f_t(s)ds)$ are clearly bounded by one.

### 7.3.2 Long Rates Never Fall

There are some differences in modeling the forward rate as a function on a bounded interval $[0, x_{max}]$ versus the half line $\mathbb{R}_+$. In particular, when we work on the half-line and define the long rate by the limit $\ell_t = \lim_{x\to\infty} f_t(x)$, an unexpected phenomenon is found: The long rate never falls. We give an account of this result in the context of the abstract HJM models studied in this chapter.
Let $F$ be the state space. Throughout this subsection, we grant Assumption 7.1, as well as one additional assumption

**Assumption** 7.3 Every $f \epsilon F$ is a function $f : \mathbb{R}_+ \to \mathbb{R}$ such that the limit $f(\infty) = \lim_{x\to\infty} f(x)$ exists, and the functional $\delta_\infty : f \hookrightarrow f(\infty)$ is an element on $F^*$.

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\{f_t\}_{t\geq 0}$ be an $F$-valued forward rate process given by an abstract HJM model, and let $\ell_t = f_t(\infty)$ be the long rate. We prove that the long rate is almost surely increasing.

**Theorem 7.2.** *For $0 \leq s \leq t$, the inequality $\ell_s \leq \ell_t$ holds almost surely.*

*Proof.* We use the following observation: For fixed $(t, \omega)$ we have $\lim_{T\to\infty} \tilde{P}(t,T)^{1/T} = e^{-\ell_t}$, where

$$\tilde{P}(t,T) = \exp\left(-\int_0^t f_s(0)ds - \int_0^{T-t} f_t(x)ds\right)$$

are the discounted bond prices. Since we are interested in an almost sure property of the forward rate process, we may work with any measure which is equivalent to the given measure $\mathbb{P}$. In particular, from the discussion of the previous section, there exists a measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that the discounted bond price processes $\{\tilde{P}(t,T)\}_{t\epsilon[0,T]}$ are local martingales simultaneously for all $T > 0$. All expected values will be calculated under this measure $\mathbb{Q}$.
Let $\xi$ be a positive and bounded random variable. By the conditional versions of Fatous lemma (*Let* $f_n \geq 0$, then $\int_D \liminf f_n d\mu \leq \liminf_n \int_D f_n d\mu$) and Holders inequality we have

$$\mathbb{E}\{e^{-\ell_t}\xi\} = \mathbb{E}\{\lim_{T\to\infty} \tilde{P}(t,T)^{1/T}\xi\} = \mathbb{E}\{\mathbb{E}\{\lim_{T\to\infty} \tilde{P}(t,T)^{1/T}\xi|\mathcal{F}_s\}\}$$

$$\leq \mathbb{E}\{\liminf_{T\to\infty} \mathbb{E}\{\tilde{P}(t,T)^{1/T}\xi|\mathcal{F}_s\}\} \leq \mathbb{E}\{\liminf_{T\to\infty} \mathbb{E}\{\tilde{P}(t,T)|\mathcal{F}_s\}^{1/T}\mathbb{E}\{\xi^{T/(T-1)}|\mathcal{F}_s\}^{(T-1)/T}\}$$

$$\leq \mathbb{E}\{\liminf_{T\to\infty} \tilde{P}(s,T)^{1/T}\mathbb{E}\{\xi^{T/(T-1)}|\mathcal{F}_s\}^{(T-1)/T}\} \leq \mathbb{E}\{e^{-\ell_s}\xi\}.$$

We have used the fact that $\{\tilde{P}(t,T)\}_{t\epsilon[0,T]}$ is a super-martingale for $\mathbb{Q}$. Since $\xi$ is positive but arbitrary, the result follows.

□

Notice that the above proof needs very little of the structure of the abstract HJM models introduced earlier. In fact, it is easy to see that the result holds in discrete time and with models with jumps. All that is assumed is that the long rate exists. We note that the popular short rate models produce

constant long rates. For instance, for the Vasicek model the short interest rate satisfies the SDE

$$dr_t = (a - \beta r_t)dt + \sigma dw_t$$

for a scalar Wiener process $\{w_t\}_{t\geq 0}$. The forward rates are given by

$$f_t(x) = e^{-\beta x} r_t + (1 - e^{-\beta x})\frac{a}{\beta} - \frac{a^2}{2\beta^2}(1 - e^{-\beta x})^2$$

where $f_t(0) = r_t$. Note that not only the long rate is well-defined, but it is explicitly given by the constant

$$\ell_t = \frac{a}{\beta} - \frac{a^2}{2\beta^2}$$

independent of $(t, \omega)$. There do exist models for which the long rate is strictly increasing. Consider an HJM model driven by a scalar Wiener process $\{w_t\}_{t\geq 0}$ with a constant volatility function given by $\sigma(x) = \sigma_0(x+1)^{1/2}$. Since $F_{HJM} \circ \sigma(x) = 2\sigma_0^2(1(x+1)^{1/2})$ and $\int_0^t (x+t-s)^{-1/2} dw_s$ converges to zero a.s., it follows that the long rate for this model is the increasing process

$$\ell_t = \ell_0 + 2\sigma_0^2 t.$$

# Bibliography


[1] A.N.Giannakopoulos, Stochastic Analysis and Appliations in Finance, University of Aegean, 2010

[2] I.Spiliotis,Notes of Measure Analysis, Technical University of Athens, 2010

[3] A. Katavolos, Intoduction to Operators theory, Symmetria, 2003

[4] Anton Bovier, Stochastic Process I, Uni Bonn , 2009

[5] Ivan F.Wilde, Distribution Theory Notes

[6] IV Melnikova, AI Filinkov, UA Anufrieva, Abstract Stochastic Equations, Journal of Mathematical Sciences, Springer, 2002

[7] V. Serov, M. Harju, Inverse Problems , iopscience,2007

[8] SL LeVarge Semigroups of Linear Operators, math.arizona.edu, 2003

[9] RA Carmona, MR Tehranchi, Interest Rate Models: Infinite Dimensional Stochastic Analysis Perspective, 2006

[10] A. Giannopoulos, Notes of Functional Analysis, University of Crete,2003