

# Diffusion processes and Stochastic calculus

February 13, 2019

## Preface

This book originates from a set of lecture notes for graduate classes I delivered since 2005, first at the Paul Sabatier University in Toulouse, and then at Purdue University in West Lafayette. The lecture notes benefited much from the input and criticism from several students and have been modified and expanded numerous times before reaching the final form of this book.

My motivation is to present at the graduate level and in a concise but complete way the most important tools and ideas of the classical theory of continuous time processes and at the same time introduce the readers to more advanced theories : the theory of Dirichlet forms, the Malliavin calculus and the Lyons rough paths theory.

Several exercises of various levels are distributed throughout the text in order to test the understanding of the reader. Results proved in these exercises are sometimes used in later parts of the text so I really encourage the reader to have a dynamic approach in his reading and to try to solve the exercises.

I included at the end of each chapter a short section listing references for the reader wishing to complement his reading or looking for more advanced theories and topics.

Chapters 1,2,5 and 6 are essentially independent from Chapters 3 and 4. I often used the materials in Chapters 1,2,5 and 6 as a graduate course on Stochastic calculus and Chapters 3 and 4 on their own as a course on Markov processes and Markov semigroups assuming some of the basic results of Chapter 1. Chapter 7 is almost entirely independent from the other chapters. It is an introduction to Lyons rough paths theory which offers a deterministic approach to understand differential equations driven by very irregular processes including as a special case Brownian motion.

To conclude, I would like to express my gratitude to the students who pointed out typos and inaccuracies in various versions of the lecture notes leading to this book and to my colleague Cheng Ouyang for a detailed reading of an early draft of the manuscript. Of course, all the remaining typos and mistakes are my own responsibility and a list of corrections will be kept online on my personal blog. Finally, I thank Igor Kortchmeski for letting me use his nice picture of a random stable tree on the cover of the book.

*West Lafayette, January 2014*

*Fabrice Baudoin*

# Contents

Preface	v
<b>Introduction</b>	<b>1</b>
<b>1 Stochastic processes</b>	<b>7</b>
1 Measure theory in function spaces . . . . .	7
2 Stochastic processes . . . . .	9
3 Filtrations . . . . .	10
4 The Daniell-Kolmogorov extension theorem . . . . .	11
5 The Kolmogorov continuity theorem . . . . .	17
6 Stopping times . . . . .	19
7 Martingales . . . . .	21
8 Martingale inequalities . . . . .	33
Notes and comments . . . . .	35
<b>2 Brownian motion</b>	<b>36</b>
1 Definition and basic properties . . . . .	36
2 Basic properties . . . . .	40
3 The law of iterated logarithm . . . . .	44
4 Symmetric random walks . . . . .	47
5 Donsker theorem . . . . .	57
Notes and Comments . . . . .	64
<b>3 Markov processes</b>	<b>65</b>
1 Markov processes . . . . .	65
2 Strong Markov processes . . . . .	75
3 Feller-Dynkin diffusions . . . . .	78
4 Lévy processes . . . . .	89
Notes and Comments . . . . .	98
<b>4 Symmetric diffusion semigroups</b>	<b>100</b>
1 Essential self-adjointness, Spectral theorem . . . . .	100
2 Existence and regularity of the heat kernel . . . . .	109
3 The submarkov property . . . . .	111
4 $L^p$ theory: The interpolation method . . . . .	115
5 $L^p$ theory: The Hille-Yosida method . . . . .	117
6 Diffusion semigroups as solutions of a parabolic Cauchy problem	128
7 The Dirichlet semigroup . . . . .	130
8 The Neumann semigroup . . . . .	135

9	Symmetric diffusion processes . . . . .	136
	Notes and Comments . . . . .	139
<b>5</b>	<b>Itô calculus</b>	<b>141</b>
1	Variation of the Brownian paths . . . . .	141
2	Itô integral . . . . .	143
3	Square integrable martingales and quadratic variations . . . . .	149
4	Local martingales, Semimartingales and Integrators . . . . .	157
5	Döblin-Itô formula . . . . .	164
6	Recurrence and transience of the Brownian motion in higher dimensions . . . . .	168
7	Itô representation theorem . . . . .	171
8	Time changed martingales and planar Brownian motion . . . . .	174
9	Burkholder-Davis-Gundy inequalities . . . . .	178
10	Girsanov theorem . . . . .	181
	Notes and Comments . . . . .	186
<b>6</b>	<b>Stochastic differential equations and Malliavin calculus</b>	<b>189</b>
1	Existence and uniqueness of solutions . . . . .	189
2	Continuity and differentiability of stochastic flows . . . . .	194
3	The Feynman-Kac formula . . . . .	197
4	The strong Markov property of solutions . . . . .	203
5	Stratonovitch stochastic differential equations and the language of vector fields . . . . .	204
6	Malliavin calculus . . . . .	209
7	Existence of a smooth density . . . . .	220
	Notes and Comments . . . . .	223
<b>7</b>	<b>An introduction to Lyons rough paths theory</b>	<b>225</b>
1	Continuous paths with finite $p$ -variation . . . . .	225
2	The signature of a bounded variation path . . . . .	230
3	Estimating iterated integrals . . . . .	232
4	Rough differential equations . . . . .	243
5	The Brownian motion as a rough path . . . . .	249
	Notes and Comments . . . . .	258
	<b>Appendix A. Unbounded operators</b>	<b>260</b>
	<b>Appendix B. Regularity theory</b>	<b>265</b>
	References	270
	Index	270
	Conventions and frequently used notations	270

## Introduction

The first stochastic process that has been extensively studied is the Brownian motion, named in honor of the botanist Robert Brown, who observed and described in 1828 the random movement of particles suspended in a liquid or gas. One of the first mathematical studies of this process goes back to the mathematician Louis Bachelier, in 1900, who, in his thesis [?], presented a stochastic modelling of the stock and option markets. But, mainly due to the lack of rigorous foundations of probability theory at that time, the seminal work of Bachelier has been ignored for a long time by mathematicians. However, in his 1905 paper, Albert Einstein brought this stochastic process to the attention of physicists by presenting it as a way to indirectly confirm the existence of atoms and molecules. The rigorous mathematical study of stochastic processes really began with the mathematician Andrei Kolmogorov. His monograph [?] published in Russian in 1933 built up probability theory in a rigorous way from fundamental axioms in a way comparable to Euclid's treatment of geometry. From this axiomatic, Kolmogorov gives a precise definition of stochastic processes. His point of view stresses the fact that a stochastic process is nothing else but a random variable valued in a space of functions (or a space of curves). For instance, if an economist reads a financial newspaper because he is interested in the prices of barrel of oil for last year, then he will focus on the curve of these prices. According to Kolmogorov's point of view, saying that these prices form a stochastic process is then equivalent to saying that the curve that is seen is the realization of a random variable defined on a suitable probability space. This point of view is mathematically quite deep and provides existence results for stochastic processes as well as pathwise regularity results.

Joseph Doob writes in his introduction to his famous book "Stochastic processes" [?]:

*A stochastic process is any process running along in time and controlled by probability laws...more precisely any family of random variables where a random variable ... is simply a measurable function...*

Doob's point of view, which is consistent with Kolmogorov's and built on the work by Paul Lévy, is nowadays commonly given as a definition of a stochastic process. Relying on this point of view that emphasizes the role of time, Doob's work, developed during the 1940's and the 1950's has quickly become one of the most powerful tools available to study stochastic processes.

Let us now describe the seminal considerations of Bachelier. Let  $X_t$  denote the price at time  $t$  of a given asset on a financial market (Bachelier considered

a given quantity of wheat). We will assume that  $X_0 = 0$  (otherwise, we work with  $X_t - X_0$ ). The first observation is that the price  $X_t$  can not be predicted with absolute certainty. It seems therefore reasonable to assume that  $X_t$  is a random variable defined on some probability space. One of the initial problems of Bachelier was to understand the distribution of prices at given times, that is the distribution of the random variable  $(X_{t_1}, \dots, X_{t_n})$ , where  $t_1, \dots, t_n$  are fixed.

The two following fundamental observations of Bachelier were based on empirical observations:

- If  $\tau$  is very small then, in absolute value, the price variation  $X_{t+\tau} - X_t$  is of order  $\sigma\sqrt{\tau}$ , where  $\sigma$  is a positive parameter (nowadays called the volatility of the asset);
- The expectation of a speculator is always zero<sup>1</sup> (nowadays, a generalization of this principle is called the absence of arbitrage).

Next, Bachelier assumes that for every  $t > 0$ ,  $X_t$  has a density with respect to the Lebesgue measure, let us say  $p(t, x)$ . It means that if  $[x - \varepsilon, x + \varepsilon]$  is a small interval around  $x$ , then

$$\mathbb{P}(X_t \in [x - \varepsilon, x + \varepsilon]) \simeq 2\varepsilon p(t, x).$$

The two above observations imply that for  $\tau$  small,

$$p(t + \tau, x) \simeq \frac{1}{2}p(t, x - \sigma\sqrt{\tau}) + \frac{1}{2}p(t, x + \sigma\sqrt{\tau}).$$

Indeed, due to the first observation, if the price is  $x$  at time  $t + \tau$ , it means that at time  $t$  the price was equal to  $x - \sigma\sqrt{\tau}$  or to  $x + \sigma\sqrt{\tau}$ . According to the second observation, each of this case occurs with probability  $\frac{1}{2}$ .

Now Bachelier assumes that  $p(t, x)$  is regular enough and uses the following approximations coming from a Taylor expansion:

$$\begin{aligned} p(t + \tau, x) &\simeq p(t, x) + \tau \frac{\partial p}{\partial t}(t, x) \\ p(t, x - \sigma\sqrt{\tau}) &\simeq p(t, x) - \sigma\sqrt{\tau} \frac{\partial p}{\partial x}(t, x) + \frac{1}{2}\sigma^2\tau \frac{\partial^2 p}{\partial x^2}(t, x) \\ p(t, x + \sigma\sqrt{\tau}) &\simeq p(t, x) + \sigma\sqrt{\tau} \frac{\partial p}{\partial x}(t, x) + \frac{1}{2}\sigma^2\tau \frac{\partial^2 p}{\partial x^2}(t, x). \end{aligned}$$

This yields the equation

$$\frac{\partial p}{\partial t} = \frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2}(t, x). \tag{0.1}$$

---

<sup>1</sup>Quoted and translated from French: *It seems that the market, the aggregate of speculators, can believe in neither a market rise nor a market fall, since, for each quoted price, there are as many buyers as sellers..*

This is the so-called heat equation, which is the primary example of a diffusion equation. Explicit solutions to this equation are obtained using the Fourier transform, and by using the fact that at time 0,  $p$  is the Dirac distribution at 0, it is computed that:

$$p(t, x) = \frac{e^{-\frac{x^2}{2\sigma^2 t}}}{\sigma\sqrt{2\pi t}}.$$

It means that  $X_t$  has a Gaussian distribution with mean 0 and variance  $\sigma^2$ . Let now  $0 < t_1 < \dots < t_n$  be fixed times and  $I_1, \dots, I_n$  be fixed intervals. In order to compute  $\mathbb{P}(X_{t_1} \in I_1, \dots, X_{t_n} \in I_n)$  the next step is to assume that the above analysis did not depend on the origin of time, or more precisely that the best information available at time  $t$  is given by the price  $X_t$ . That leads first to the following computation

$$\begin{aligned} \mathbb{P}(X_{t_1} \in I_1, X_{t_2} \in I_2) &= \int_{I_1} \mathbb{P}(X_{t_2} \in I_2 \mid X_{t_1} = x_1) p(t_1, x_1) dx_1 \\ &= \int_{I_1} \mathbb{P}(X_{t_2-t_1} + x_1 \in I_2 \mid X_{t_1} = x_1) p(t_1, x_1) dx_1 \\ &= \int_{I_1 \times I_2} p(t_2 - t_1, x_2 - x_1) p(t_1, x_1) dx_1 dx_2, \end{aligned}$$

which is easily generalized to

$$\begin{aligned} \mathbb{P}(X_{t_1} \in I_1, \dots, X_{t_n} \in I_n) & \tag{0.2} \\ &= \int_{I_1 \times \dots \times I_n} p(t_n - t_{n-1}, x_n - x_{n-1}) \cdots p(t_2 - t_1, x_2 - x_1) p(t_1, x_1) dx_1 dx_2 \cdots dx_n. \end{aligned}$$

In many regards, the previous computations were not rigorous but heuristic. One of our first motivations is to provide a rigorous construction of this object  $X_t$  on which worked Bachelier and that is called a Brownian motion.

From a rigorous point of view the question is: Does there exist a sequence of random variables  $\{X_t, t \geq 0\}$  such that  $t \rightarrow X_t$  is continuous and such that the property (0.2) is satisfied? Chapter 1 will give a positive answer to this question. We will see how to define and construct processes. In particular we will prove the existence of Brownian motion and then study several of its properties.

Chapter 1 sets the foundations. It deals with the basic definitions and results that are required to rigorously deal with stochastic processes. We introduce the relevant  $\sigma$ -fields and prove the fundamental Daniell-Kolmogorov theorem which may be seen as an infinite dimensional version of the Carathéodory extension of measure theorem. It is the basic theorem to prove the existence of a stochastic process. However, despite its importance and usefulness, the Daniell-Kolmogorov result relies on the axiom of choice and as such is non-constructive and does not give any information or insight about the stochastic

process that has been proved to exist. The Kolmogorov continuity theorem fills one of those gaps and gives a useful criterion to ensure that we can work with a process whose sample paths are continuous. Chapter 1 also includes a thorough study of continuous martingales. We focus on Doob's theorems: The stopping theorem, the regularization result and the maximal inequalities. Martingale techniques are essential to study stochastic processes. They give the tools to handle stopping times which are naturally associated to processes and provide the inequalities which are the cornerstones of the theory of stochastic integration which is presented in Chapter 5.

Chapter 2 is devoted to the study of the most important stochastic process: The Brownian motion. As a consequence of the Daniell-Kolmogorov and Kolmogorov continuity theorems, we prove the existence of this process and then study some of its most fundamental properties. From many point of views, Brownian motion can be seen as the continuous random walk in continuous time. This is made precise at the end of the chapter where we give an alternative proof of the existence of the Brownian motion as a limit of suitably rescaled random walks.

Chapter 3 is devoted to the study of Markov processes. Our goal is to emphasize the role of the theory of semigroups when studying Markov processes. More precisely, we wish to understand how one can construct a Markov process from a semigroup and then see what are the properties inherited from the semigroup to the sample paths properties of the process. We will particularly focus on the class of Feller-Dynkin Markov processes which are a class of Markov processes enjoying several nice properties, like the existence of regular versions and the strong Markov property. We finish the chapter with a study of the Lévy processes which form an important subset of the class of Feller-Dynkin Markov processes.

Chapter 4 can be thought an introduction to the theory of symmetric Dirichlet forms. As we will see, this theory and the set of tools attached of it belong much more to functional analysis than probability theory. The basic problem is the construction of a Markov semigroup and of a Markov process only from the generator. More precisely, the question is: Given a diffusion operator  $L$ , does  $L$  generate Markov semigroup  $\mathbf{P}_t$  and if yes, is this semigroup the transition semigroup of a continuous Markov process ? We will answer positively this question in quite a general framework under the basic assumption that  $L$  is elliptic and essentially self-adjoint with respect to some Borel measure.

Chapter 5 is about stochastic calculus and its applications to the study of Brownian motion. Stochastic calculus is an integral and differential calculus with respect to Brownian motion or more generally with respect to martingales. It allows to give a meaningful sense to integrals with respect to Brownian



motion paths and to define differential equations driven by such paths. It is not a straightforward extension of the usual calculus because Brownian motion paths are not regular enough, they are only  $\gamma$ -Hölder continuous with  $\gamma < 1/2$ . The range of applications of the stochastic calculus is huge and still growing today. We mention in particular that the applications to mathematical finance have drawn a lot of attention on this calculus. Actually, most of the modern pricing theories for derivatives on financial markets are based on Itô-Döblin's formula which is the chain rule for stochastic calculus.

Chapter 6 deals with the theory of stochastic differential equations. Stochastic differential equations are the differential equations that correspond to Itô's integration theory. They give a very powerful tool to construct a Markov process from its generator. We will prove the basic existence and uniqueness results for such equations and quickly turn to the basic properties of the solution. One of the problems we are the most interested in is the existence of a smooth density for the solution of a stochastic differential equation. This problem gave birth to the so-called Malliavin calculus which is the study of the Sobolev regularity of Brownian functionals and that we study in some details.

Chapter 7 is an introduction to the theory of rough paths that was developed by Lyons in the 1990's. The theory is deterministic. It allows to give a sense to solutions of differential equations driven by very irregular paths. Stochastic differential equations driven by Brownian motions are then seen as a very special case of rough differential equations. The advantage of the theory is that it goes much beyond the scope of Itô calculus when dealing with differential systems driven by random noises and comes with powerful estimates.

# Chapter 1

## Stochastic processes

In this chapter we set the rigorous foundations on which the theory of general stochastic processes is built. Our first task will be to build a measure theory in infinite dimension and then to study sample paths properties of stochastic processes. The second part of the chapter is devoted to the study of martingales. Martingales techniques will be used in almost every parts of this book.

### 1 Measure theory in function spaces

Stochastic processes can be seen as random variables taking their values in a function space. It is therefore important to understand the naturally associated  $\sigma$ -algebras.

Let  $\mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$ ,  $d \geq 1$ , be the set of functions  $\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$ . We denote by  $\mathcal{T}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$  the  $\sigma$ -algebra generated by the so-called cylindrical sets

$$\{f \in \mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}^d), f(t_1) \in I_1, \dots, f(t_n) \in I_n\}$$

where

$$t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$$

and where  $I_1, \dots, I_n$  are products of intervals:  $I_i = \prod_{k=1}^d (a_i^k, b_i^k]$ .

**Remark 1.1.** As a  $\sigma$ -algebra  $\mathcal{T}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$  is also generated by the following families:

- $$\{f \in \mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}^d), f(t_1) \in B_1, \dots, f(t_n) \in B_n\}$$
 where  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$  and where  $B_1, \dots, B_n$  are Borel sets in  $\mathbb{R}^d$ .

- $$\{f \in \mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}^d), (f(t_1), \dots, f(t_n)) \in B\}$$
 where  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$  and where  $B$  is a Borel set in  $(\mathbb{R}^d)^{\otimes n}$ .

**Exercise 1.2.** Show that the following sets are not in  $\mathcal{T}([0, 1], \mathbb{R})$ :

1. 
$$\{f \in \mathcal{A}([0, 1], \mathbb{R}), \sup_{t \in [0, 1]} f(t) < 1\}$$

2. 
$$\{f \in \mathcal{A}([0, 1], \mathbb{R}), \exists t \in [0, 1] f(t) = 0\}$$

The above exercise shows that the  $\sigma$ -algebra  $\mathcal{T}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$  is not rich enough to include *natural* events; this is due to the fact that the space  $\mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$  is by far too big.

In this book, we shall mainly be interested in processes with continuous paths. In that case, we use the space of continuous functions  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$  endowed with the  $\sigma$ -algebra  $\mathcal{B}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$  generated by

$$\{f \in \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}^d), f(t_1) \in I_1, \dots, f(t_n) \in I_n\}$$

where  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$  and where  $I_1, \dots, I_n$  are products of intervals  $\prod_{k=1}^d (a_i^k, b_i^k]$ . This  $\sigma$ -algebra enjoys nice properties. It is generated by the open sets of the (metric) topology of uniform convergence on compact sets.

**Proposition 1.3.** *The  $\sigma$ -algebra  $\mathcal{B}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$  is generated by the open sets of the topology of uniform convergence on compact sets.*

*Proof.* We make the proof when the dimension  $d = 1$  and let the reader adapt it in higher dimension. Let us first recall that, on  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$  the topology of uniform convergence on compact sets is given by the distance

$$d(f, g) = \sum_{n=1}^{+\infty} \frac{1}{2^n} \min\left(\sup_{0 \leq t \leq n} |f(t) - g(t)|, 1\right).$$

This distance endows  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$  with the structure of a complete, separable, metric space. Let us denote by  $\mathcal{O}$  the  $\sigma$ -field generated by the open sets of this metric space. First, it is clear that the cylinders

$$\{f \in \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}), f(t_1) < a_1, \dots, f(t_n) < a_n\}$$

are open sets that generate  $\mathcal{B}(\mathbb{R}_{\geq 0}, \mathbb{R})$ . Thus, we have  $\mathcal{B}(\mathbb{R}_{\geq 0}, \mathbb{R}) \subset \mathcal{O}$ . On the other hand, since for every  $g \in \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$ ,  $n \in \mathbb{N}$ ,  $n \geq 1$  and  $\rho > 0$

$$\begin{aligned} & \{f \in \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}), \sup_{0 \leq t \leq n} |f(t) - g(t)| \leq \rho\} \\ &= \bigcap_{t \in \mathbb{Q}, 0 \leq t \leq n} \{f \in \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}), |f(t) - g(t)| \leq \rho\}, \end{aligned}$$

we deduce that

$$\{f \in \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}), \sup_{0 \leq t \leq n} |f(t) - g(t)| \leq \rho\} \in \mathcal{B}(\mathbb{R}_{\geq 0}, \mathbb{R}).$$

Since  $\mathcal{O}$  is generated by the above sets, this implies  $\mathcal{O} \subset \mathcal{B}(\mathbb{R}_{\geq 0}, \mathbb{R})$ .  $\square$

**Exercise 1.4.** *Show that the following sets are in  $\mathcal{B}([0, 1], \mathbb{R})$ :*

1.

$$\{f \in \mathcal{C}([0, 1], \mathbb{R}), \sup_{t \in [0, 1]} f(t) < 1\}.$$

2.

$$\{f \in \mathcal{C}([0, 1], \mathbb{R}), \exists t \in [0, 1] f(t) = 0\}.$$

## 2 Stochastic processes

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space.

**Definition 1.5.** On  $(\Omega, \mathcal{F}, \mathbb{P})$ , a ( $d$ -dimensional) stochastic process is a sequence  $(X_t)_{t \geq 0}$  of  $\mathbb{R}^d$ -valued random variables that are measurable with respect to  $\mathcal{F}$ .

A process  $(X_t)_{t \geq 0}$  can also be seen as an application

$$X(\omega) \in \mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}^d), t \rightarrow X_t(\omega).$$

The applications  $t \rightarrow X_t(\omega)$  are called the paths of the process. The application  $X : (\Omega, \mathcal{F}) \rightarrow (\mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}^d), \mathcal{T}(\mathbb{R}_{\geq 0}, \mathbb{R}^d))$  is measurable. The probability measure defined by

$$\mu(A) = \mathbb{P}(X^{-1}(A)), A \in \mathcal{T}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$$

is then called the law (or distribution) of  $(X_t)_{t \geq 0}$ .

For  $t \geq 0$ , we denote by  $\pi_t$  the application that transforms  $f \in \mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$  into  $f(t)$ :  $\pi_t : f \rightarrow f(t)$ . The stochastic process  $(\pi_t)_{t \in \mathbb{R}_{\geq 0}}$  which is thus defined on the probability space  $(\mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}), \mathcal{T}(\mathbb{R}_{\geq 0}, \mathbb{R}^d), \mu)$  is called the canonical process associated to  $X$ . It is a process with distribution  $\mu$ .

**Definition 1.6.** A process  $(X_t)_{t \geq 0}$  is said to be measurable if the application

$$(t, \omega) \rightarrow X_t(\omega)$$

is measurable with respect to the  $\sigma$ -algebra  $\mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{F}$  that is, if

$$\forall A \in \mathcal{B}(\mathbb{R}^d), \{(t, \omega), X_t(\omega) \in A\} \in \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{F}.$$

The paths of a measurable process are, of course, measurable functions  $\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$ .

**Definition 1.7.** If a process  $X$  takes its values in  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$ , that is if the paths of  $X$  are continuous functions, then we say that  $X$  is a continuous process.

If  $(X_t)_{t \geq 0}$  is a continuous process then the application

$$X : (\Omega, \mathcal{F}) \rightarrow (\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}^d), \mathcal{B}(\mathbb{R}_{\geq 0}, \mathbb{R}))$$

is measurable and the distribution of  $X$  is a probability measure on the space  $(\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}^d), \mathcal{B}(\mathbb{R}_{\geq 0}, \mathbb{R}^d))$ . Moreover, a continuous process is measurable in the sense of the Definition 1.6:

**Proposition 1.8.** *A continuous stochastic process is measurable.*

*Proof.* Let  $(X_t)_{t \geq 0}$  be a continuous process. Let us first prove that if  $A$  is a Borel set in  $\mathbb{R}$ , then

$$\{(t, \omega) \in [0, 1] \times \Omega, X_t(\omega) \in A\} \in \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{F}.$$

For  $n \in \mathbb{N}$ , let

$$X_t^n = X_{\lfloor \frac{2^n t}{2^n} \rfloor}, t \in [0, 1],$$

where  $\lfloor x \rfloor$  denotes the integer part of  $x$ . Since the paths of  $X^n$  are piecewise constant, we have

$$\{(t, \omega) \in [0, 1] \times \Omega, X_t^n(\omega) \in A\} \in \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{F}.$$

Moreover,  $\forall t \in [0, 1], \omega \in \Omega$ , we have

$$\lim_{n \rightarrow +\infty} X_t^n(\omega) = X_t(\omega),$$

which implies

$$\{(t, \omega) \in [0, 1] \times \Omega, X_t(\omega) \in A\} \in \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{F}.$$

In the same way we obtain that  $\forall k \in \mathbb{N}$ ,

$$\{(t, \omega) \in [k, k+1] \times \Omega, X_t(\omega) \in A\} \in \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{F}.$$

Observing

$$\{(t, \omega) \in \mathbb{R} \times \Omega, X_t(\omega) \in A\} = \cup_{k \in \mathbb{N}} \{(t, \omega) \in [k, k+1] \times \Omega, X_t(\omega) \in A\},$$

yields the sought of conclusion.  $\square$

### 3 Filtrations

A stochastic process  $(X_t)_{t \geq 0}$  may also be seen as a random system evolving in time. This system carries some information. More precisely, if one observes the paths of a stochastic process up to a time  $t > 0$ , one is able to decide if an event

$$A \in \sigma(X_s, s \leq t)$$

has occurred (here and in the sequel  $\sigma(X_s, s \leq t)$  denotes the smallest  $\sigma$ -field that makes all the random variables  $\{(X_{t_1}, \dots, X_{t_n}), 0 \leq t_1 \leq \dots \leq t_n \leq t\}$  measurable). This notion of information carried by a stochastic process is modeled by filtrations.

**Definition 1.9.** Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. A filtration  $(\mathcal{F}_t)_{t \geq 0}$  is a non-decreasing family of sub- $\sigma$ -algebras of  $\mathcal{F}$ .

As a basic example, if  $(X_t)_{t \geq 0}$  is a stochastic process defined on  $(\Omega, \mathcal{F}, \mathbb{P})$ , then

$$\mathcal{F}_t = \sigma(X_s, s \leq t)$$

is a filtration. This filtration is called the natural filtration of the process  $X$  and often denoted by  $(\mathcal{F}_t^X)_{t \geq 0}$ .

**Definition 1.10.** A stochastic process  $(X_t)_{t \geq 0}$  is said to be adapted to a filtration  $(\mathcal{F}_t)_{t \geq 0}$  if for every  $t \geq 0$ , the random variable  $X_t$  is measurable with respect to  $\mathcal{F}_t$ .

Of course, a stochastic process is always adapted with respect to its natural filtration. We may observe that if a stochastic process  $(X_t)_{t \geq 0}$  is adapted to a filtration  $(\mathcal{F}_t)_{t \geq 0}$  and that if  $\mathcal{F}_0$  contains all the subsets of  $\mathcal{F}$  that have a zero probability, then every process  $(\tilde{X}_t)_{t \geq 0}$  that satisfies

$$\mathbb{P}(\tilde{X}_t = X_t) = 1, \quad t \geq 0,$$

is still adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .

We previously defined the notion of measurability for a stochastic process. In order to take into account the dynamic aspect associated to a filtration, the notion of progressive measurability is needed.

**Definition 1.11.** A stochastic process  $(X_t)_{t \geq 0}$  that is adapted to a filtration  $(\mathcal{F}_t)_{t \geq 0}$ , is said to be progressively measurable with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  if for every  $t \geq 0$ ,

$$\forall A \in \mathcal{B}(\mathbb{R}), \{(s, \omega) \in [0, t] \times \Omega, X_s(\omega) \in A\} \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t.$$

By using the diagonal method, it is possible to construct adapted but not progressively measurable processes. However, the next proposition whose proof is let as an exercise to the reader shows that an adapted and continuous stochastic process is automatically progressively measurable.

**Proposition 1.12.** *A continuous stochastic process  $(X_t)_{t \geq 0}$ , that is adapted with respect to a filtration  $(\mathcal{F}_t)_{t \geq 0}$ , is also progressively measurable with respect to it.*

## 4 The Daniell-Kolmogorov extension theorem

The Daniell-Kolmogorov extension theorem is one of the first deep theorems of the theory of stochastic processes. It provides existence results for nice probability measures on path (function) spaces. It is however non-constructive and relies on the axiom of choice. In what follows, in order to avoid heavy notations we restrict the presentation to the one dimensional case  $d = 1$ . The multidimensional extension is straightforward.

**Definition 1.13.** Let  $(X_t)_{t \geq 0}$  be a stochastic process. For  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$  we denote by  $\mu_{t_1, \dots, t_n}$  the probability distribution of the random variable

$$(X_{t_1}, \dots, X_{t_n}).$$

It is therefore a probability measure on  $\mathbb{R}^n$ . This probability measure is called a finite dimensional distribution of the process  $(X_t)_{t \geq 0}$ .

If two processes have the same finite dimensional distributions, then it is clear that the two processes induce the same distribution on the path space  $\mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R})$  because cylinders generate the  $\sigma$ -algebra  $\mathcal{T}(\mathbb{R}_{\geq 0}, \mathbb{R})$ .

The finite dimensional distributions of a given process satisfy the two following properties: If  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$  and if  $\tau$  is a permutation of the set  $\{1, \dots, n\}$ , then:

1.

$$\mu_{t_1, \dots, t_n}(A_1 \times \dots \times A_n) = \mu_{t_{\tau(1)}, \dots, t_{\tau(n)}}(A_{\tau(1)} \times \dots \times A_{\tau(n)}), \quad A_i \in \mathcal{B}(\mathbb{R}).$$

2.

$$\mu_{t_1, \dots, t_n}(A_1 \times \dots \times A_{n-1} \times \mathbb{R}) = \mu_{t_1, \dots, t_{n-1}}(A_1 \times \dots \times A_{n-1}), \quad A_i \in \mathcal{B}(\mathbb{R}).$$

Conversely, we have

**Theorem 1.14** (Daniell-Kolmogorov theorem). *Assume that we are given for every  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$  a probability measure  $\mu_{t_1, \dots, t_n}$  on  $\mathbb{R}^n$ . Let us assume that these probability measures satisfy:*

1.

$$\mu_{t_1, \dots, t_n}(A_1 \times \dots \times A_n) = \mu_{t_{\tau(1)}, \dots, t_{\tau(n)}}(A_{\tau(1)} \times \dots \times A_{\tau(n)}), \quad A_i \in \mathcal{B}(\mathbb{R}).$$

2.

$$\mu_{t_1, \dots, t_n}(A_1 \times \dots \times A_{n-1} \times \mathbb{R}) = \mu_{t_1, \dots, t_{n-1}}(A_1 \times \dots \times A_{n-1}), \quad A_i \in \mathcal{B}(\mathbb{R}).$$

*Then, there is a unique probability measure  $\mu$  on  $(\mathcal{A}(\mathbb{R}_+, \mathbb{R}), \mathcal{T}(\mathbb{R}_+, \mathbb{R}))$  such that for  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$ ,  $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R})$ :*

$$\mu(\pi_{t_1} \in A_1, \dots, \pi_{t_n} \in A_n) = \mu_{t_1, \dots, t_n}(A_1 \times \dots \times A_n).$$

The Daniell-Kolmogorov theorem is often used to construct processes thanks to the following corollary:

**Corollary 1.15.** *Assume given for every  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$  a probability measure  $\mu_{t_1, \dots, t_n}$  on  $\mathbb{R}^n$ . Let us further assume that these measures satisfy the assumptions of the Daniell-Kolmogorov theorem. Then, there exists a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  as well as a process  $(X_t)_{t \geq 0}$  defined on this space such that the finite dimensional distributions of  $(X_t)_{t \geq 0}$  are given by the  $\mu_{t_1, \dots, t_n}$ 's.*

*Proof.* As a probability space we chose

$$(\Omega, \mathcal{F}, \mathbb{P}) = (\mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}), \mathcal{T}(\mathbb{R}_{\geq 0}, \mathbb{R}), \mu)$$

where  $\mu$  is the probability measure given by the Daniell-Kolmogorov theorem. The canonical process  $(\pi_t)_{t \geq 0}$  defined on  $\mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R})$  by  $\pi_t(f) = f(t)$  satisfies the required property.  $\square$

We now turn to the proof of the Daniell-Kolmogorov theorem. This proof proceeds in several steps.

As a first step, let us recall the Carathéodory extension theorem that is often useful for the effective construction of measures (for instance the construction of the Lebesgue measure on  $\mathbb{R}$ ):

**Theorem 1.16** (Carathéodory theorem). *Let  $\Omega$  be a non-empty set and let  $\mathcal{A}$  be a family of subsets that satisfy:*

1.  $\Omega \in \mathcal{A}$ ;
2. If  $A, B \in \mathcal{A}$ ,  $A \cup B \in \mathcal{A}$ ;
3. If  $A \in \mathcal{A}$ ,  $\Omega \setminus A \in \mathcal{A}$ .

*Let  $\sigma(\mathcal{A})$  be the  $\sigma$ -algebra generated by  $\mathcal{A}$ . If  $\mu_0$  is  $\sigma$ -additive measure on  $(\Omega, \mathcal{A})$ , then there exists a unique  $\sigma$ -additive measure  $\mu$  on  $(\Omega, \sigma(\mathcal{A}))$  such that for  $A \in \mathcal{A}$ ,*

$$\mu_0(A) = \mu(A).$$

As a second step, we prove the following fact:

**Lemma 1.17.** *Let  $B_n \subset \mathbb{R}^n$ ,  $n \in \mathbb{N}$  be a sequence of Borel sets that satisfy*

$$B_{n+1} \subset B_n \times \mathbb{R}.$$

*Let us assume that for every  $n \in \mathbb{N}$  a probability measure  $\mu_n$  is given on  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  and that these probability measures are compatible in the sense that*

$$\mu_n(A_1 \times \dots \times A_{n-1} \times \mathbb{R}) = \mu_{n-1}(A_1 \times \dots \times A_{n-1}), \quad A_i \in \mathcal{B}(\mathbb{R}).$$

*and satisfy:*

$$\mu_n(B_n) > \varepsilon,$$

*where  $0 < \varepsilon < 1$ . There exists a sequence of compact sets  $K_n \subset \mathbb{R}^n$ ,  $n \in \mathbb{N}$ , such that:*

- $K_n \subset B_n$



- $K_{n+1} \subset K_n \times \mathbb{R}$ .
- $\mu_n(K_n) \geq \frac{\varepsilon}{2}$ .

*Proof.* For every  $n$ , we can find a compact set  $K_n^* \subset \mathbb{R}^n$  such that

$$K_n^* \subset B_n$$

and

$$\mu_n(B_n \setminus K_n^*) \leq \frac{\varepsilon}{2^{n+1}}.$$

Let us consider

$$K_n = (K_1^* \times \mathbb{R}^{n-1}) \cap \dots \cap (K_{n-1}^* \times \mathbb{R}) \cap K_n^*.$$

It is easily checked that:

- $K_n \subset B_n$
- $K_{n+1} \subset K_n \times \mathbb{R}$ .

Moreover, we have

$$\begin{aligned} & \mu_n(K_n) \\ = & \mu_n(B_n) - \mu_n(B_n \setminus K_n) \\ = & \mu_n(B_n) - \mu_n(B_n \setminus ((K_1^* \times \mathbb{R}^{n-1}) \cap \dots \cap (K_{n-1}^* \times \mathbb{R}) \cap K_n^*)) \\ \geq & \mu_n(B_n) - \mu_n(B_n \setminus ((K_1^* \times \mathbb{R}^{n-1}))) - \dots - \mu_n(B_n \setminus (K_{n-1}^* \times \mathbb{R})) - \mu_n(B_n \setminus K_n^*) \\ \geq & \mu_n(B_n) - \mu_1(B_1 \setminus K_1^*) - \dots - \mu_n(B_n \setminus K_n^*) \\ \geq & \varepsilon - \frac{\varepsilon}{4} - \dots - \frac{\varepsilon}{2^{n+1}} \\ \geq & \frac{\varepsilon}{2}. \end{aligned}$$

□

With this in hands, we can now turn to the proof of the Daniell-Kolmogorov theorem.

*Proof.* For the cylinder

$$\mathcal{C}_{t_1, \dots, t_n}(B) = \{f \in \mathcal{A}(\mathbb{R}_+, \mathbb{R}), (f(t_1), \dots, f(t_n)) \in B\}$$

where

$$t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$$

and where  $B$  is a Borel subset of  $\mathbb{R}^n$ , we define

$$\mu(\mathcal{C}_{t_1, \dots, t_n}(B)) = \mu_{t_1, \dots, t_n}(B).$$

Thanks to the assumptions on the  $\mu_{t_1, \dots, t_n}$ 's, it is seen that such a  $\mu$  is well defined and satisfies:

$$\mu(\mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R})) = 1.$$

The set  $\mathcal{A}$  of all the possible cylinders  $\mathcal{C}_{t_1, \dots, t_n}(B)$  satisfies the assumption of Carathéodory theorem. Therefore, in order to conclude, we have to show that  $\mu$  is  $\sigma$ -additive, that is, if  $(C_n)_{n \in \mathbb{N}}$  is a sequence of pairwise disjoint cylinders and if  $C = \cup_{n \in \mathbb{N}} C_n$  is a cylinder then

$$\mu(C) = \sum_{n=0}^{+\infty} \mu(C_n).$$

This is the difficult part of the theorem. Since for  $N \in \mathbb{N}$ ,

$$\mu(C) = \mu(C \setminus \cup_{n=0}^N C_n) + \mu(\cup_{n=0}^N C_n),$$

we just have to show that

$$\lim_{N \rightarrow +\infty} \mu(D_N) = 0.$$

where  $D_N = C \setminus \cup_{n=0}^N C_n$ .

The sequence  $(\mu(D_N))_{N \in \mathbb{N}}$  is positive decreasing and therefore converges. Let assume that it converges toward  $\varepsilon > 0$ . We shall prove that in that case

$$\cap_{N \in \mathbb{N}} D_N \neq \emptyset,$$

which is clearly absurd.

Since  $D_N$  is a cylinder, the event  $\cup_{N \in \mathbb{N}} D_N$  only involves a countable sequence of times  $t_1 < \dots < t_n < \dots$  and we may assume (otherwise we can add convenient other sets in the sequence of the  $D_N$ 's) that every  $D_N$  can be described as follows

$$D_N = \{f \in \mathcal{A}(\mathbb{R}_{\geq 0}, \mathbb{R}), (f(t_1), \dots, f(t_N)) \in B_N\}$$

where  $B_n \subset \mathbb{R}^n$ ,  $n \in \mathbb{N}$ , is a sequence of Borel sets such that

$$B_{n+1} \subset B_n \times \mathbb{R}.$$

Since we assumed  $\mu(D_N) \geq \varepsilon$ , we can use the previous lemma to construct a sequence of compact sets  $K_n \subset \mathbb{R}^n$ ,  $n \in \mathbb{N}$ , such that:

- $K_n \subset B_n$
- $K_{n+1} \subset K_n \times \mathbb{R}$ .
- $\mu_{t_1, \dots, t_n}(K_n) \geq \frac{\varepsilon}{2}$ .

Since  $K_n$  is non-empty, we pick

$$(x_1^n, \dots, x_n^n) \in K_n.$$

The sequence  $(x_1^n)_{n \in \mathbb{N}}$  has a convergent subsequence  $(x_1^{j_1(n)})_{n \in \mathbb{N}}$  that converges toward  $x_1 \in K_1$ . The sequence  $((x_1^{j_1(n)}, x_2^{j_1(n)})_{n \in \mathbb{N}}$  has a convergent subsequence that converges toward  $(x_1, x_2) \in K_2$ . By pursuing this process<sup>1</sup> we obtain a sequence  $(x_n)_{n \in \mathbb{N}}$  such that for every  $n$ ,

$$(x_1, \dots, x_n) \in K_n.$$

The event

$$\{f \in \mathcal{A}(\mathbb{R}_+, \mathbb{R}), (f(t_1), \dots, f(t_N)) = (x_1, \dots, x_N)\}$$

is in  $D_N$ , this leads to the expected contradiction. Therefore, the sequence  $(\mu(D_N))_{N \in \mathbb{N}}$  converges toward 0, which implies the  $\sigma$ -additivity of  $\mu$ .  $\square$

The Daniell-Kolmogorov theorem is the basic tool to prove the existence of a stochastic process with given finite dimensional distributions. As an example, let us illustrate how it may be used to prove the existence of the so-called Gaussian processes.

**Definition 1.18.** A real-valued stochastic process  $(X_t)_{t \geq 0}$  defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  is said to be a Gaussian process if all the finite dimensional distributions of  $X$  are Gaussian random variables.

If  $(X_t)_{t \geq 0}$  is a Gaussian process, its finite dimensional distributions can be characterized, through Fourier transform, by its mean function

$$m(t) = \mathbb{E}(X_t)$$

and its covariance function

$$R(s, t) = \mathbb{E}((X_t - m(t))(X_s - m(s))).$$

We can observe that the covariance function  $R(s, t)$  is symmetric ( $R(s, t) = R(t, s)$ ) and positive, that is for  $a_1, \dots, a_n \in \mathbb{R}$  and  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$ ,

$$\begin{aligned} \sum_{1 \leq i, j \leq n} a_i a_j R(t_i, t_j) &= \sum_{1 \leq i, j \leq n} a_i a_j \mathbb{E}((X_{t_i} - m(t_i))(X_{t_j} - m(t_j))) \\ &= \mathbb{E} \left( \left( \sum_{i=1}^n (X_{t_i} - m(t_i)) \right)^2 \right) \\ &\geq 0. \end{aligned}$$

Conversely, as an application of the Daniell-Kolmogorov theorem, we let the reader prove as an exercise the following proposition.

---

<sup>1</sup>This is here that the axiom of choice is needed.

**Proposition 1.19.** *Let  $m : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  and let  $R : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  be a symmetric and positive function. There exists a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a Gaussian process  $(X_t)_{t \geq 0}$  defined on it, whose mean function is  $m$  and whose covariance function is  $R$ .*

**Exercise 1.20.** *Let  $(X_t)_{0 \leq t \leq T}$  be a continuous Gaussian process. Show that the random variable  $\int_0^T X_s ds$  is a Gaussian random variable.*

## 5 The Kolmogorov continuity theorem

The Daniell-Kolmogorov theorem is a very useful tool since it provides existence results for stochastic processes. However this theorem does not say anything about the paths of this process. The following theorem, due to Kolmogorov, precises that, under mild conditions, we can work with processes whose paths are quite regular.

**Definition 1.21.** A function  $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$  is said to be Hölder with exponent  $\alpha > 0$  if there exists a constant  $C > 0$  such that for  $s, t \in \mathbb{R}_{\geq 0}$ ,

$$\|f(t) - f(s)\| \leq C |t - s|^\alpha .$$

Hölder functions are in particular continuous.

**Definition 1.22.** A stochastic process  $(\tilde{X}_t)_{t \geq 0}$  is called a modification of the process  $(X_t)_{t \geq 0}$  if for  $t \geq 0$ ,

$$\mathbb{P}(X_t = \tilde{X}_t) = 1.$$

**Remark 1.23.** We can observe that if  $(\tilde{X}_t)_{t \geq 0}$  is a modification of  $(X_t)_{t \geq 0}$  then  $(\tilde{X}_t)_{t \geq 0}$  has the same distribution as  $(X_t)_{t \geq 0}$ .

**Theorem 1.24** (Kolmogorov continuity theorem). *Let  $\alpha, \varepsilon, c > 0$ . If a  $d$ -dimensional process  $(X_t)_{t \in [0,1]}$  defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  satisfies for  $s, t \in [0, 1]$ ,*

$$\mathbb{E}(\|X_t - X_s\|^\alpha) \leq c |t - s|^{1+\varepsilon},$$

*then there exists a modification of the process  $(X_t)_{t \in [0,1]}$  that is a continuous process and whose paths are  $\gamma$ -Hölder for every  $\gamma \in [0, \frac{\varepsilon}{\alpha})$ .*

*Proof.* We make the proof for  $d = 1$  and let the reader extend it as an exercise to the case  $d \geq 2$ . For  $n \in \mathbb{N}$ , we denote

$$\mathcal{D}_n = \left\{ \frac{k}{2^n}, k = 0, \dots, 2^n \right\}$$

and

$$\mathcal{D} = \cup_{n \in \mathbb{N}} \mathcal{D}_n.$$

Let  $\gamma \in [0, \frac{\varepsilon}{\alpha})$ . From Chebychev inequality:

$$\begin{aligned} \mathbb{P}\left(\max_{1 \leq k \leq 2^n} |X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| \geq 2^{-\gamma n}\right) &= \mathbb{P}\left(\cup_{1 \leq k \leq 2^n} |X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| \geq 2^{-\gamma n}\right) \\ &\leq \sum_{k=1}^{2^n} \mathbb{P}\left(|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| \geq 2^{-\gamma n}\right) \\ &\leq \sum_{k=1}^{2^n} \frac{\mathbb{E}\left(|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}|^\alpha\right)}{2^{-\gamma \alpha n}} \\ &\leq c 2^{-n(\varepsilon - \gamma \alpha)} \end{aligned}$$

Therefore, since  $\gamma \alpha < \varepsilon$ , we deduce

$$\sum_{n=1}^{+\infty} \mathbb{P}\left(\max_{1 \leq k \leq 2^n} |X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| \geq 2^{-\gamma n}\right) < +\infty.$$

From the Borel-Cantelli lemma, we can thus find a set  $\Omega^* \in \mathcal{F}$  such that  $\mathbb{P}(\Omega^*) = 1$  and such that for  $\omega \in \Omega^*$ , there exists  $N(\omega)$  such that for  $n \geq N(\omega)$ ,

$$\max_{1 \leq k \leq 2^n} |X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega)| \leq 2^{-\gamma n}.$$

In particular, there exists an almost surely finite random variable  $C$  such that for every  $n \geq 0$ ,

$$\max_{1 \leq k \leq 2^n} |X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega)| \leq C 2^{-\gamma n}$$

We now claim that the paths of the restricted process  $X_{/\Omega^*}$  are consequently  $\gamma$ -Hölder on  $\mathcal{D}$ . Indeed, let  $s, t \in \mathcal{D}$ . We can find  $n$  such that

$$|s - t| \leq \frac{1}{2^n}.$$

We now pick an increasing and stationary sequence  $(s_k)_{k \geq n}$  converging toward  $s$ , such that  $s_k \in \mathcal{D}_k$  and

$$|s_{k+1} - s_k| = 2^{-(k+1)} \quad \text{or} \quad 0.$$

In the same way, we can find an analogue sequence  $(t_k)_{k \geq n}$  that converges toward  $t$  and such that  $s_n$  and  $t_n$  are neighbors in  $\mathcal{D}_n$ . We have then:

$$X_t - X_s = \sum_{i=n}^{+\infty} (X_{s_{i+1}} - X_{s_i}) + (X_{s_n} - X_{t_n}) + \sum_{i=n}^{+\infty} (X_{t_i} - X_{t_{i+1}}),$$

where the above sums are actually finite.

Therefore,

$$\begin{aligned} |X_t - X_s| &\leq C2^{-\gamma n} + 2 \sum_{k=n}^{+\infty} C2^{-\gamma(k+1)} \\ &\leq 2C \sum_{k=n}^{+\infty} 2^{-\gamma k} \\ &\leq \frac{2C}{1 - 2^{-\gamma}} 2^{-\gamma n}. \end{aligned}$$

Hence the paths of  $X_{/\Omega^*}$  are  $\gamma$ -Hölder on the set  $\mathcal{D}$ . For  $\omega \in \Omega^*$ , let  $t \rightarrow \tilde{X}_t(\omega)$  be the unique continuous function that agrees with  $t \rightarrow X_t(\omega)$  on  $\mathcal{D}$ . For  $\omega \notin \Omega^*$ , we set  $\tilde{X}_t(\omega) = 0$ . The process  $(\tilde{X}_t)_{t \in [0,1]}$  is the desired modification of  $(X_t)_{t \in [0,1]}$ .  $\square$

**Exercise 1.25.** Let  $\alpha, \varepsilon, c > 0$ . Let  $(X_t)_{t \in [0,1]}$  be a continuous Gaussian process such that for  $s, t \in [0, 1]$ ,

$$\mathbb{E}(\|X_t - X_s\|^\alpha) \leq c |t - s|^{1+\varepsilon},$$

Show that for every  $\gamma \in [0, \varepsilon/\alpha)$ , there is a positive random variable  $\eta$  such that  $\mathbb{E}(\eta^p) < \infty$ , for every  $p \geq 1$  and such that for every  $s, t \in [0, 1]$ ,

$$\|X_t - X_s\| \leq \eta |t - s|^\gamma, \quad a.s.$$

**Hint:** You may use the Garsia-Rodemich-Rumsey inequality which is stated in Theorem 7.34, Chapter 7, in greater generality: Let  $p \geq 1$  and  $\alpha > p^{-1}$ , then there exists a constant  $C_{\alpha,p} > 0$  such that for any continuous function  $f$  on  $[0, T]$ , and for all  $t, s \in [0, T]$  one has:

$$\|f(t) - f(s)\|^p \leq C_{\alpha,p} |t - s|^{\alpha p - 1} \int_0^T \int_0^T \frac{\|f(x) - f(y)\|^p}{|x - y|^{\alpha p + 1}} dx dy.$$

## 6 Stopping times

In the study of a stochastic process it is often useful to consider some properties of the process that hold up to a random time. A natural question is for instance: For how long is the process less than a given constant ?

**Definition 1.26.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $T$  be a random variable, measurable with respect to  $\mathcal{F}$  and valued in  $\mathbb{R}_{\geq 0} \cup \{+\infty\}$ . We say that  $T$  is a stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$  if for  $t \geq 0$ ,

$$\{T \leq t\} \in \mathcal{F}_t.$$

Often, a stopping time will be the time during which a stochastic process adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  satisfies a given property. The above definition means that for any  $t \geq 0$ , at time  $t$ , one is able to decide if this property is satisfied or not.

Among the most important examples of stopping times, are the (first) hitting times of a closed set by a continuous stochastic process.

**Exercise 1.27** (First hitting time of a closed set by a continuous stochastic process). *Let  $(X_t)_{t \geq 0}$  be a continuous process adapted to a filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Let*

$$T = \inf\{t \geq 0, X_t \in F\},$$

where  $F$  is a closed subset of  $\mathbb{R}$ . Show that  $T$  is a stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .

Given a stopping time  $T$ , we may define the  $\sigma$ -algebra of events that occur before the time  $T$ :

**Proposition 1.28.** *Let  $T$  be a stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Let*

$$\mathcal{F}_T = \{A \in \mathcal{F}, \forall t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_t\}.$$

Then  $\mathcal{F}_T$  is a  $\sigma$ -algebra.

*Proof.* Since for every  $t \geq 0$ ,  $\emptyset \in \mathcal{F}_t$ , we have that  $\emptyset \in \mathcal{F}_T$ . Let us now consider  $A \in \mathcal{F}_T$ . We have

$${}^c A \cap \{T \leq t\} = \{T \leq t\} \setminus (A \cap \{T \leq t\}) \in \mathcal{F}_t,$$

and thus  ${}^c A \in \mathcal{F}_T$ . Finally, if  $(A_n)_{n \in \mathbb{N}}$  is a sequence of subsets of  $\mathcal{F}_T$ ,

$$(\cap_{n \in \mathbb{N}} A_n) \cap \{T \leq t\} = \cap_{n \in \mathbb{N}} (A_n \cap \{T \leq t\}) \in \mathcal{F}_t.$$

□

If  $T$  is a stopping time of a filtration with respect to which a given process is adapted, then it is possible to stop this process in a natural way at the time  $T$ . We let the proof of the corresponding proposition as an exercise to the reader.

**Proposition 1.29.** *Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $T$  be an almost surely finite stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Let  $(X_t)_{t \geq 0}$  be a stochastic process that is adapted and progressively measurable with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . The stopped stochastic process  $(X_{t \wedge T})_{t \geq 0}$  is progressively measurable with respect to the filtration  $(\mathcal{F}_{t \wedge T})_{t \geq 0}$ .*

## 7 Martingales

We introduce and study in this section martingales in continuous time. Such processes were first introduced and extensively studied by Joseph Doob. Together with Markov processes they are among the most important class of stochastic processes and lie at the hearth of the theory of stochastic integration.

**Definition 1.30.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . A process  $(M_t)_{t \geq 0}$  that is adapted to  $(\mathcal{F}_t)_{t \geq 0}$  is called a submartingale with respect to this filtration if:

1. For every  $t \geq 0$ ,  $\mathbb{E}(|M_t|) < +\infty$ ;
2. For every  $t \geq s \geq 0$

$$\mathbb{E}(M_t | \mathcal{F}_s) \geq M_s.$$

A stochastic process  $(M_t)_{t \geq 0}$  that is adapted to  $(\mathcal{F}_t)_{t \geq 0}$  and such that  $(-M_t)_{t \geq 0}$  is a submartingale, is called a supermartingale.

Finally, a stochastic process  $(M_t)_{t \geq 0}$  that is adapted to  $(\mathcal{F}_t)_{t \geq 0}$  and that is at the same time a submartingale and a supermartingale is called a martingale.

The following exercises provide some first properties of these processes.

**Exercise 1.31.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $X$  be an integrable and  $\mathcal{F}$ -measurable random variable. Show that the process  $(\mathbb{E}(X | \mathcal{F}_t))_{t \geq 0}$  is a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .

**Exercise 1.32.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a submartingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Show that the function  $t \rightarrow \mathbb{E}(M_t)$  is non-decreasing.

**Exercise 1.33.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Let now  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  be a convex function such that for  $t \geq 0$ ,  $\mathbb{E}(|\psi(M_t)|) < +\infty$ . Show that the process  $(\psi(M_t))_{t \geq 0}$  is a submartingale.

The following theorem, which is due to Doob, turns out to be extremely useful. It shows that martingales behave in a very nice way with respect to stopping times.

**Proposition 1.34** (Doob stopping theorem). Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a continuous stochastic process that is adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . The following properties are equivalent:

1.  $(M_t)_{t \geq 0}$  is a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ ;



2. For any, almost surely bounded stopping time  $T$  of the filtration  $(\mathcal{F}_t)_{t \geq 0}$  such that  $\mathbb{E}(|M_T|) < +\infty$ , we have

$$\mathbb{E}(M_T) = \mathbb{E}(M_0).$$

*Proof.* Let us assume that  $(M_t)_{t \geq 0}$  is a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Let now  $T$  be a stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$  that is almost surely bounded by  $K > 0$ . Let us first assume that  $T$  takes its values in a finite set:

$$0 \leq t_1 < \dots < t_n \leq K.$$

Thanks to the martingale property, we have

$$\begin{aligned} \mathbb{E}(M_T) &= \mathbb{E}\left(\sum_{i=1}^n M_T 1_{T=t_i}\right) \\ &= \sum_{i=1}^n \mathbb{E}(M_{t_i} 1_{T=t_i}) \\ &= \sum_{i=1}^n \mathbb{E}(M_{t_n} 1_{T=t_i}) \\ &= \mathbb{E}(M_{t_n}) \\ &= \mathbb{E}(M_0). \end{aligned}$$

The theorem is therefore proved whenever  $T$  takes its values in a finite set. If  $T$  takes an infinite number of values, we approximate  $T$  by the following sequence of stopping times:

$$\tau_n = \sum_{k=1}^{2^n} \frac{kK}{2^n} 1_{\left\{\frac{(k-1)K}{2^n} \leq T < \frac{kK}{2^n}\right\}}.$$

The stopping time  $\tau_n$  takes its values in a finite set and when  $n \rightarrow +\infty$ ,  $\tau_n \rightarrow T$ . To conclude the proof of the first part of the proposition, we therefore have to show that

$$\lim_{n \rightarrow +\infty} \mathbb{E}(M_{\tau_n}) = \mathbb{E}(M_T).$$

For this, we are going to show that the family  $(M_{\tau_n})_{n \in \mathbb{N}}$  is uniformly integrable.

Let  $A \geq 0$ . Since  $\tau_n$  takes its values in a finite set, by using the martingale property and Jensen's inequality, it is easily checked that

$$\mathbb{E}(|M_K| 1_{M_{\tau_n} \geq A}) \geq \mathbb{E}(|M_{\tau_n}| 1_{M_{\tau_n} \geq A}).$$

Therefore,

$$\mathbb{E}(M_{\tau_n} 1_{M_{\tau_n} \geq A}) \leq \mathbb{E}(M_K 1_{\sup_{0 \leq s \leq K} M_s \geq A}) \xrightarrow{A \rightarrow +\infty} 0.$$

By uniform integrability, we deduce that

$$\lim_{n \rightarrow +\infty} \mathbb{E}(M_{\tau_n}) = \mathbb{E}(M_T),$$

from which it is concluded

$$\mathbb{E}(M_T) = \mathbb{E}(M_0).$$

Conversely, let us now assume that for any, almost surely bounded stopping time  $T$  of the filtration  $(\mathcal{F}_t)_{t \geq 0}$  such that  $\mathbb{E}(|M_T|) < +\infty$ , we have

$$\mathbb{E}(M_T) = \mathbb{E}(M_0).$$

Let  $0 \leq s \leq t$  and  $A \in \mathcal{F}_s$ . By using the stopping time

$$T = s1_A + t1_{c_A},$$

we are led to

$$\mathbb{E}((M_t - M_s)1_A) = 0,$$

which implies the martingale property for  $(M_t)_{t \geq 0}$ .  $\square$

The hypothesis that the paths of  $(M_t)_{t \geq 0}$  be continuous is actually not strictly necessary, however the hypothesis that the stopping time  $T$  be almost surely bounded is essential, as it is proved in the following exercise.

**Exercise 1.35.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a continuous martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  such that  $M_0 = 0$  almost surely. For  $a > 0$ , we denote  $T_a = \inf\{t > 0, M_t = a\}$ . Show that  $T_a$  is a stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Prove that  $T_a$  is not almost surely bounded.

**Exercise 1.36.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a continuous submartingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . By mimicking the proof of Doob's stopping theorem, show that if  $T_1$  and  $T_2$  are two almost surely bounded stopping times of the filtration  $(\mathcal{F}_t)_{t \geq 0}$  such that  $T_1 \leq T_2$  and  $\mathbb{E}(|M_{T_1}|) < +\infty$ ,  $\mathbb{E}(|M_{T_2}|) < +\infty$ , then,

$$\mathbb{E}(M_{T_1}) \leq \mathbb{E}(M_{T_2}).$$

By using a similar proof, the Doob's stopping theorem is easily extended as follows:

**Proposition 1.37.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a continuous martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . If  $T_1$  and  $T_2$  are two almost surely bounded stopping times of the filtration  $(\mathcal{F}_t)_{t \geq 0}$  such that  $T_1 \leq T_2$  and  $\mathbb{E}(|M_{T_1}|) < +\infty$ ,  $\mathbb{E}(|M_{T_2}|) < +\infty$ , then,

$$\mathbb{E}(M_{T_2} | \mathcal{F}_{T_1}) = M_{T_1}.$$

Finally, as a direct consequence of Doob's stopping theorem, we finally have the following result that shall repeatedly be used in the sequel.

**Proposition 1.38.** *Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a continuous martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . If  $T$  is a bounded stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$  then the stopped process  $(M_{t \wedge T})_{t \geq 0}$  is a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .*

*Proof.* Let  $S$  be a bounded stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . From Doob's stopping theorem, we have  $\mathbb{E}(M_{S \wedge T}) = \mathbb{E}(M_0)$ . Since  $S$  is arbitrary, we conclude also from Doob's theorem that  $(M_{t \wedge T})_{t \geq 0}$  is a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .  $\square$

We now turn to convergence theorems for martingales. These convergence results rely on the notion of uniform integrability that we now remind.

**Definition 1.39.** Let  $(X_i)_{i \in \mathcal{I}}$  be a family of random variables. We say that the family  $(X_i)_{i \in \mathcal{I}}$  is uniformly integrable if for every  $\varepsilon > 0$ , there exists  $K \geq 0$  such that

$$\forall i \in \mathcal{I}, \quad \mathbb{E}(|X_i| \mathbf{1}_{|X_i| > K}) < \varepsilon.$$

We have the following properties:

- A finite family of integrable random variables is uniformly integrable ;
- If the family  $(X_i)_{i \in \mathcal{I}}$  is uniformly integrable then it is bounded in  $L^1$ , that is  $\sup_{\mathcal{I}} \mathbb{E}(|X_i|) < +\infty$ ;
- If the family  $(X_i)_{i \in \mathcal{I}}$  is bounded in  $L^p$  with  $p > 1$ , that is  $\sup_{\mathcal{I}} \mathbb{E}(|X_i|^p) < +\infty$ , then it is uniformly integrable.

Thanks to the following result, the notion of uniform integrability is often used to prove a convergence in  $L^1$ :

**Proposition 1.40.** *Let  $(X_n)_{n \in \mathbb{N}}$  be a sequence of integrable random variables. Let  $X$  be an integrable random variable. The sequence  $(X_n)_{n \in \mathbb{N}}$  converges toward  $X$  in  $L^1$ , that is  $\lim_{n \rightarrow +\infty} \mathbb{E}(|X_n - X|) = 0$ , if and only if:*

1. *In probability,  $X_n \rightarrow_{n \rightarrow +\infty} X$ , that is for every  $\varepsilon > 0$ ,*

$$\lim_{n \rightarrow +\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0;$$

2. *The family  $(X_n)_{n \in \mathbb{N}}$  is uniformly integrable.*

We have seen in Exercise 1.31 that if  $X$  is an integrable random variable defined on a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  then the process  $(\mathbb{E}(X | \mathcal{F}_t))_{t \geq 0}$  is a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . The following theorem characterizes the martingales that are of this form.

**Theorem 1.41** (Doob convergence theorem). *Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  whose paths are left limited and right continuous. The following properties are equivalent:*

1. When  $t \rightarrow +\infty$ ,  $(M_t)_{t \geq 0}$  converges in  $L^1$ ;
2. When  $t \rightarrow +\infty$ ,  $(M_t)_{t \geq 0}$  almost surely converges toward an integrable and  $\mathcal{F}$ -measurable random variable  $X$  that satisfies

$$M_t = \mathbb{E}(X \mid \mathcal{F}_t), t \geq 0;$$

3. The family  $(M_t)_{t \geq 0}$  is uniformly integrable.

*Proof.* As a first step, we show that if the martingale  $(M_t)_{t \geq 0}$  is bounded in  $L^1$ , that is

$$\sup_{t \geq 0} \mathbb{E}(|M_t|) < +\infty$$

then  $(M_t)_{t \geq 0}$  almost surely converges toward an integrable and  $\mathcal{F}$ -measurable random variable  $X$ .

Let us first observe that

$$\{\omega \in \Omega, M_t(\omega) \text{ converges}\} = \left\{ \omega \in \Omega, \limsup_{t \rightarrow +\infty} M_t(\omega) = \liminf_{t \rightarrow +\infty} M_t(\omega) \right\}$$

Therefore, in order to show that  $(M_t)_{t \geq 0}$  almost surely converges when  $t \rightarrow +\infty$ , we may prove that

$$\mathbb{P} \left( \left\{ \omega \in \Omega, \limsup_{t \rightarrow +\infty} M_t(\omega) > \liminf_{t \rightarrow +\infty} M_t(\omega) \right\} \right) = 0.$$

Let us assume that

$$\mathbb{P} \left( \left\{ \omega \in \Omega, \limsup_{t \rightarrow +\infty} M_t(\omega) > \liminf_{t \rightarrow +\infty} M_t(\omega) \right\} \right) > 0.$$

In that case we may find  $a < b$  such that:

$$\mathbb{P} \left( \left\{ \omega \in \Omega, \limsup_{t \rightarrow +\infty} M_t(\omega) > a > b > \liminf_{t \rightarrow +\infty} M_t(\omega) \right\} \right) > 0.$$

The idea now is to study the oscillations of  $(M_t)_{t \geq 0}$  between  $a$  and  $b$ . For  $N \in \mathbb{N}$ ,  $N > 0$  and  $n \in \mathbb{N}$ , we denote

$$\mathcal{D}_{n,N} = \left\{ \frac{kN}{2^n}, 0 \leq k \leq 2^n \right\},$$

and

$$\mathcal{D} = \cup_{n,N} \mathcal{D}_{n,N}.$$

Let  $\mathcal{N}(a, b, n, N)$  be the greatest integer  $k$  for which we may find elements of  $\mathcal{D}_{n,N}$ ,

$$0 \leq q_1 < r_1 < q_2 < r_2 < \dots < q_k < r_k \leq N$$

that satisfy

$$M_{q_i} < a, \quad M_{r_i} > b.$$

Let now

$$Y_{n,N} = \sum_{k=1}^{2^n} C_{\frac{kN}{2^n}} \left( M_{\frac{kN}{2^n}} - M_{\frac{(k-1)N}{2^n}} \right),$$

where  $C_k \in \{0, 1\}$  is recursively defined by:

$$C_1 = 1_{M_0 < a},$$

$$C_k = 1_{C_{k-1}=1} 1_{M_{\frac{(k-1)N}{2^n}} \leq b} + 1_{C_{k-1}=0} 1_{M_{\frac{(k-1)N}{2^n}} < a}.$$

Since  $(M_t)_{t \geq 0}$  is martingale, it is easily checked that

$$\mathbb{E}(Y_{n,N}) = 0.$$

Furthermore, thanks to the very definition of  $\mathcal{N}(a, b, n, N)$ , we have

$$Y_{n,N} \geq (b-a)\mathcal{N}(a, b, n, N) - \max(a - M_N, 0).$$

Therefore, we have

$$\begin{aligned} (b-a)\mathbb{E}(\mathcal{N}(a, b, n, N)) &\leq \mathbb{E}(\max(a - M_N, 0)) \\ &\leq |a| + \mathbb{E}(|M_N|) \\ &\leq |a| + \sup_{t>0} \mathbb{E}(|M_t|), \end{aligned}$$

and thus

$$(b-a)\mathbb{E}\left(\sup_{n,N} \mathcal{N}(a, b, n, N)\right) \leq |a| + \sup_{t>0} \mathbb{E}(|M_t|).$$

This implies that almost surely  $\sup_{n,N} \mathcal{N}(a, b, n, N) < +\infty$ , from which it is deduced

$$\mathbb{P}\left(\left\{\omega \in \Omega, \lim_{t \rightarrow +\infty, t \in \mathcal{D}} \sup M_t(\omega) > a > b > \lim_{t \rightarrow +\infty, t \in \mathcal{D}} \inf M_t(\omega)\right\}\right) = 0.$$

Since the paths of  $(M_t)_{t \geq 0}$  are right continuous, we have

$$\begin{aligned} &\mathbb{P}\left(\left\{\omega \in \Omega, \lim_{t \rightarrow +\infty, t \in \mathcal{D}} \sup M_t(\omega) > a > b > \lim_{t \rightarrow +\infty, t \in \mathcal{D}} \inf M_t(\omega)\right\}\right) \\ &= \mathbb{P}\left(\left\{\omega \in \Omega, \lim_{t \rightarrow +\infty} \sup M_t(\omega) > a > b > \lim_{t \rightarrow +\infty} \inf M_t(\omega)\right\}\right). \end{aligned}$$

This is absurd. Thus, if  $(M_t)_{t \geq 0}$  is bounded in  $L^1$ , it almost surely converges toward an  $\mathcal{F}$ -measurable random variable  $X$ . Fatou's lemma provides the integrability of  $X$ .

With this preliminary result in hands, we can now turn to the proof of the theorem.

Let us assume that  $(M_t)_{t \geq 0}$  converges in  $L^1$ . In that case, it is of course bounded in  $L^1$ , and thus almost surely converges toward an  $\mathcal{F}$ -measurable and integrable random variable  $X$ . Let  $t \geq 0$  and  $A \in \mathcal{F}_t$ , we have for  $s \geq t$ ,

$$\mathbb{E}(M_s 1_A) = \mathbb{E}(M_t 1_A)$$

By letting  $s \rightarrow +\infty$ , the dominated convergence theorem yields

$$\mathbb{E}(X 1_A) = \mathbb{E}(M_t 1_A).$$

Therefore, as expected, we obtain

$$\mathbb{E}(X | \mathcal{F}_t) = M_t.$$

Let us now assume that  $(M_t)_{t \geq 0}$  almost surely converges toward an  $\mathcal{F}$ -measurable and integrable random variable  $X$  that satisfies

$$M_t = \mathbb{E}(X | \mathcal{F}_t), t \geq 0.$$

We almost surely have  $\sup_{t \geq 0} |M_t| < +\infty$  and thus for  $A \geq 0$ ,

$$\begin{aligned} \mathbb{E}(|M_t| 1_{|M_t| \geq A}) &= \mathbb{E}(|\mathbb{E}(X | \mathcal{F}_t)| 1_{|M_t| \geq A}) \\ &\leq \mathbb{E}(|X| 1_{|M_t| \geq A}) \\ &\leq \mathbb{E}(|X| 1_{\sup_{t \geq 0} |M_t| \geq A}). \end{aligned}$$

This implies the uniform integrability for the family  $(M_t)_{t \geq 0}$ .

Finally, if the family  $(M_t)_{t \geq 0}$  is uniformly integrable, then it is bounded in  $L^1$  and therefore almost surely converges. The almost sure convergence, together with the uniform integrability, provides the convergence in  $L^1$ .  $\square$

**Exercise 1.42.** *By using the same reasoning as in the previous proof, show that a right continuous and left limited positive supermartingale needs to converge almost surely when  $t \rightarrow \infty$ .*

When dealing with stochastic processes, it is often important to work with versions of these processes whose paths are as regular as possible. In that direction, the Kolmogorov's continuity theorem (see Theorem 1.24) provided a sufficient condition allowing to work with continuous versions of stochastic processes. For martingales, the possibility of working with regular versions, is related to the regularity properties of the filtration with respect to which the martingale property is satisfied.

**Definition 1.43.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . If the following assumptions are fulfilled :

1. If  $A \in \mathcal{F}$  satisfies  $\mathbb{P}(A) = 0$ , then every subset of  $A$  is in  $\mathcal{F}_0$ ;
2. The filtration  $(\mathcal{F}_t)_{t \geq 0}$  is right continuous , that is for every  $t \geq 0$

$$\mathcal{F}_t = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon},$$

the filtered probability space

$$(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$$

is said to satisfy the usual conditions.

**Remark 1.44.** The above set of assumptions are called the usual conditions because, as we will see it in a next section, these are the conditions under which it is convenient to work in order to properly define the stochastic integral

**Remark 1.45.** Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a (sub, super) martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  whose paths are right continuous and left limited. The filtered probability space

$$(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$$

may canonically be enlarged into a filtered probability space

$$(\Omega, (\mathcal{G}_t)_{t \geq 0}, \mathcal{G}, \mathbb{P})$$

that satisfies the usual conditions. Indeed,  $\mathcal{G}$  can be taken to be the  $\mathbb{P}$ -completion of  $\mathcal{F}$  and

$$\mathcal{G}_t = \bigcap_{u > t} \sigma(\mathcal{F}_u, \mathcal{N})$$

where  $\mathcal{N}$  is the set of events whose probability is zero. Moreover  $(M_t)_{t \geq 0}$  is a (sub, super) martingale with respect to the filtration  $(\mathcal{G}_t)_{t \geq 0}$  (this is not straightforward and let to the reader as an exercise ). The filtered probability space

$$(\Omega, (\mathcal{G}_t)_{t \geq 0}, \mathcal{G}, \mathbb{P})$$

is called the usual completion of

$$(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P}).$$

**Exercise 1.46.** Let  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  be a filtered probability space that satisfies the usual conditions and let  $(X_t)_{t \geq 0}$  be stochastic process adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  whose paths are left limited and right continuous. Let  $K$  be compact subset of  $\mathbb{R}$ . Show that the random time

$$T = \inf\{t \geq 0, X_t \in K\}$$

is a stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .

**Theorem 1.47** (Doob regularization theorem). *Let  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  be a filtered probability space that satisfies the usual conditions and let  $(M_t)_{t \geq 0}$  be a supermartingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Let us assume that the function  $t \rightarrow \mathbb{E}(M_t)$  is right continuous.*

*There exists a modification  $(\tilde{M}_t)_{t \geq 0}$  of  $(M_t)_{t \geq 0}$  such that:*

1.  $(\tilde{M}_t)_{t \geq 0}$  is adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ ;
2. The paths of  $(\tilde{M}_t)_{t \geq 0}$  are locally bounded, right continuous and left limited;
3.  $(\tilde{M}_t)_{t \geq 0}$  is a supermartingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .

*Proof.* As for the proof of Doob's convergence theorem, the idea is to study the oscillations of  $(M_t)_{t \geq 0}$ . In what follows, we will use the notations introduced in the proof of this theorem that we remind below.

For  $N \in \mathbb{N}$ ,  $N > 0$  and  $n \in \mathbb{N}$ , we denote

$$\mathcal{D}_{n,N} = \left\{ \frac{kN}{2^n}, 0 \leq k \leq 2^n \right\},$$

$$\mathcal{D}_N = \cup_n \mathcal{D}_{n,N}$$

and

$$\mathcal{D} = \cup_{n,N} \mathcal{D}_{n,N}.$$

For  $a < b$ , let  $\mathcal{N}(a, b, n, N)$  be the greatest integer  $k$  for which we can find elements of  $\mathcal{D}_{n,N}$ ,

$$0 \leq q_1 < r_1 < q_2 < r_2 < \dots < q_k < r_k \leq N$$

such that

$$M_{q_i} < a, M_{r_i} > b.$$

Let now  $\Omega^*$  be the set of  $\omega \in \Omega$  such that  $\forall t \geq 0, \lim_{s \rightarrow t, s > t, s \in \mathcal{D}} M_s(\omega)$  exists and is finite.

It is easily seen that

$$\Omega^* = \cap_{a,b \in \mathbb{Q}} \cap_{N \in \mathbb{N}^*} \left\{ \omega \in \Omega, \sup_{t \in \mathcal{D}_N} |M_t(\omega)| < +\infty \text{ et } \sup_{n \in \mathbb{N}} \mathcal{N}(a, b, n, N) < +\infty \right\}.$$

Therefore,  $\Omega^* \in \mathcal{F}$ . We may prove, as we proved the Doob's convergence theorem that  $\mathbb{P}(\Omega^*) = 1$ .

For  $t \geq 0$ , we define  $(\tilde{M}_t)_{t \geq 0}$  in the following way:

- If  $\omega \in \Omega^*$ ,

$$\tilde{M}_t(\omega) = \lim_{s \rightarrow t, s > t, s \in \mathcal{D}} M_s(\omega)$$



- If  $\omega \notin \Omega^*$ ,

$$\tilde{M}_t(\omega) = 0.$$

It is clear that the paths of  $(\tilde{M}_t)_{t \geq 0}$  are locally bounded, right continuous and left limited. Let us now show that this process is the expected modification of  $(M_t)_{t \geq 0}$ .

We first observe that for  $t \geq 0$ , the random variable

$$\lim_{s \rightarrow t, s > t, s \in \mathcal{D}} M_s$$

is measurable with respect to  $\cap_{s > t} \mathcal{F}_s = \mathcal{F}_t$ . Furthermore,  $\Omega \setminus \Omega^*$  has a zero probability and is therefore in  $\mathcal{F}_0$ , according to the usual conditions. This shows that the process  $(\tilde{M}_t)_{t \geq 0}$  is adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .

We now show that  $(\tilde{M}_t)_{t \geq 0}$  is a modification of  $(M_t)_{t \geq 0}$ . Let  $t \geq 0$ . We have almost surely

$$\lim_{s \rightarrow t, s > t, s \in \mathcal{D}} M_s = \tilde{M}_t.$$

Let us prove that this convergence also holds in  $L^1$ . To prove this, it is enough to check that for every decreasing family  $(s_n)_{n \in \mathbb{N}}$  such that  $s_n \in \mathcal{D}$  and that converges toward  $t$ , the family  $(M_{s_n})_{n \in \mathbb{N}}$  is uniformly integrable.

Let  $\varepsilon > 0$ . Since  $u \rightarrow \mathbb{E}(M_u)$  is assumed to be right continuous, we can find  $s \in \mathbb{R}$  such that  $t < s$  and such that for every  $s > u > t$ ,

$$0 \leq \mathbb{E}(M_u) - \mathbb{E}(M_s) \leq \frac{\varepsilon}{2}.$$

For  $s > u > t$  and  $\lambda > 0$ , we have:

$$\begin{aligned} \mathbb{E}(|M_u| \mathbf{1}_{|M_u| > \lambda}) &= -\mathbb{E}(M_u \mathbf{1}_{M_u < -\lambda}) + \mathbb{E}(M_u) - \mathbb{E}(M_u \mathbf{1}_{M_u \leq -\lambda}) \\ &\leq -\mathbb{E}(M_s \mathbf{1}_{M_u < -\lambda}) + \mathbb{E}(M_u) - \mathbb{E}(M_s \mathbf{1}_{M_u \leq -\lambda}) \\ &\leq \mathbb{E}(|M_s| \mathbf{1}_{|M_u| > \lambda}) + \frac{\varepsilon}{2}. \end{aligned}$$

Now, since  $M_s \in L^1$ , we can find  $\delta > 0$  such that for every  $F \in \mathcal{F}$  that satisfies  $\mathbb{P}(F) < \delta$ , we have  $\mathbb{E}(|M_s| \mathbf{1}_F) < \frac{\varepsilon}{2}$ . But for  $t < u < s$ ,

$$\mathbb{P}(|M_u| > \lambda) \leq \frac{\mathbb{E}(|M_u|)}{\lambda} = \frac{\mathbb{E}(M_u) + 2\mathbb{E}(\max(-M_u, 0))}{\lambda}.$$

From Jensen inequality, it is seen that the process  $(\max(-M_u, 0))_{t < u < s}$  is a submartingale, therefore

$$\mathbb{E}(\max(-M_u, 0)) \leq \mathbb{E}(\max(-M_s, 0)).$$

We deduce that for  $t < u < s$ ,

$$\mathbb{P}(|M_u| > \lambda) \leq \frac{\mathbb{E}(M_t) + 2\mathbb{E}(\max(-M_s, 0))}{\lambda}.$$

It is thus possible to find  $A > 0$  such that for every  $t < u < s$ ,

$$\mathbb{P}(|M_u| > A) < \delta,$$

For  $t < u < s$ , we have then

$$\mathbb{E}(|M_u| \mathbf{1}_{|M_u| > \lambda}) < \varepsilon.$$

This implies that for every decreasing family  $(s_n)_{n \in \mathbb{N}}$  such that  $s_n \in \mathcal{D}$  and that converges toward  $t$ , the family  $(M_{s_n})_{n \in \mathbb{N}}$  is uniformly integrable. The convergence

$$\lim_{s \rightarrow t, s > t, s \in \mathcal{D}} M_s = \tilde{M}_t$$

thus also holds in  $L^1$ . Now, since  $(M_t)_{t \geq 0}$  is a supermartingale, for  $s \geq t$  we have

$$\mathbb{E}(M_s | \mathcal{F}_t) \leq M_t.$$

This implies,

$$\lim_{s \rightarrow t, s > t, s \in \mathcal{D}} \mathbb{E}(M_s | \mathcal{F}_t) \leq M_t,$$

and

$$\mathbb{E}(\tilde{M}_t | \mathcal{F}_t) \leq M_t.$$

Hence, since  $\tilde{M}_t$  is adapted to the filtration  $\mathcal{F}_t$

$$\tilde{M}_t \leq M_t,$$

Due to the fact that the function  $u \rightarrow \mathbb{E}(M_u)$  is right continuous, we have

$$\lim_{s \rightarrow t, s > t, s \in \mathcal{D}} \mathbb{E}(M_s) = \mathbb{E}(M_t).$$

But from the  $L^1$  convergence, we also have

$$\lim_{s \rightarrow t, s > t, s \in \mathcal{D}} \mathbb{E}(M_s) = \mathbb{E}\left(\lim_{s \rightarrow t, s > t, s \in \mathcal{D}} M_s\right) = \mathbb{E}(\tilde{M}_t),$$

This gives

$$\mathbb{E}(\tilde{M}_t) = \mathbb{E}(M_t).$$

The random variable  $M_t - \tilde{M}_t$  is therefore non-negative and has a zero expectation. This implies that almost surely  $M_t = \tilde{M}_t$ . The stochastic process  $(\tilde{M}_t)_{t \geq 0}$  is therefore a modification of  $(M_t)_{t \geq 0}$ . Finally, since a modification of a supermartingale is still a supermartingale, this concludes the proof of the theorem.  $\square$

The following exercise shows that martingales naturally appear when studying equivalent measures on a filtered probability space. We first recall a basic theorem from measure theory which is known as the Radon-Nikodym theorem. Let  $\Omega$  be a set endowed with a  $\sigma$ -field  $\mathcal{F}$ .

**Definition 1.48.** Let  $\mathbb{P}$  and  $\mathbb{Q}$  be probability measures on  $(\Omega, \mathcal{F})$ . It is said that  $\mathbb{P}$  is absolutely continuous with respect to  $\mathbb{Q}$  if for every  $A \in \mathcal{F}$ ,  $\mathbb{Q}(A) = 0$  implies  $\mathbb{P}(A) = 0$ . We denote then  $\mathbb{P} \ll \mathbb{Q}$ . If  $\mathbb{P} \ll \mathbb{Q}$  and  $\mathbb{Q} \ll \mathbb{P}$ , it is said that  $\mathbb{P}$  and  $\mathbb{Q}$  are equivalent: In that case we denote  $\mathbb{P} \simeq \mathbb{Q}$ .

**Theorem 1.49** (Radon-Nikodym). *Let  $\mathbb{P}$  and  $\mathbb{Q}$  be two probability measures on  $(\Omega, \mathcal{F})$ . We have  $\mathbb{P} \ll \mathbb{Q}$  if and only if there is a random variable  $D$ , which is  $\mathcal{F}$ -measurable and such that for every  $A \in \mathcal{F}$ ,*

$$\mathbb{P}(A) = \int_A D d\mathbb{Q}.$$

$D$  is called the density of  $\mathbb{P}$  with respect to  $\mathbb{Q}$  and we denote

$$D = \frac{d\mathbb{P}}{d\mathbb{Q}}.$$

Moreover, under the same assumptions  $\mathbb{P} \simeq \mathbb{Q}$  if and only if  $D$  is positive  $\mathbb{P}$ -a.s. In that case:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{1}{D}.$$

**Exercise 1.50.** *Let  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  be a filtered probability space that satisfies the usual conditions. We denote*

$$\mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \geq 0)$$

and for  $t \geq 0$ ,  $\mathbb{P}_{/\mathcal{F}_t}$  is the restriction of  $\mathbb{P}$  to  $\mathcal{F}_t$ . Let  $\mathbb{Q}$  be a probability measure on  $\mathcal{F}_\infty$  such that for every  $t \geq 0$ ,

$$\mathbb{Q}_{/\mathcal{F}_t} \ll \mathbb{P}_{/\mathcal{F}_t}.$$

1. Show that there exists a right continuous and left limited martingale  $(D_t)_{t \geq 0}$  such that for every  $t \geq 0$ ,

$$D_t = \frac{d\mathbb{Q}_{/\mathcal{F}_t}}{d\mathbb{P}_{/\mathcal{F}_t}}, \quad \mathbb{P} - a.s.$$

2. Show that the following properties are equivalent:

- (a)  $\mathbb{Q}_{/\mathcal{F}_\infty} \ll \mathbb{P}_{/\mathcal{F}_\infty}$ ;
- (b) The martingale  $(D_t)_{t \geq 0}$  is uniformly integrable;
- (c)  $(D_t)_{t \geq 0}$  converges in  $L^1$ ;
- (d)  $(D_t)_{t \geq 0}$  almost surely converges to an integrable and  $\mathcal{F}_\infty$ -measurable random variable  $D$  such that

$$D_t = \mathbb{E}(D \mid \mathcal{F}_t), \quad t \geq 0.$$

## 8 Martingale inequalities

In this section, we prove some fundamental martingale inequalities that, once again, are due to Doob

**Theorem 1.51** (Doob maximal inequalities). *Let  $(\mathcal{F}_t)_{t \geq 0}$  be a filtration on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and let  $(M_t)_{t \geq 0}$  be a continuous martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .*

1. Let  $p \geq 1$  and  $T > 0$ . If  $\mathbb{E}(|M_T|^p) < +\infty$ , then for every  $\lambda > 0$ ,

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} |M_t| \geq \lambda\right) \leq \frac{\mathbb{E}(|M_T|^p)}{\lambda^p}.$$

2. Let  $p > 1$  and  $T > 0$ . If  $\mathbb{E}(|M_T|^p) < +\infty$ , then

$$\mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^p\right) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}(|M_T|^p).$$

*Proof.*

1. Let  $p \geq 1$  and  $T > 0$ . If  $\mathbb{E}(|M_T|^p) < +\infty$  then, from Jensen inequality the process  $(|M_t|^p)_{0 \leq t \leq T}$  is a submartingale. Let  $\lambda > 0$  and

$$\tau = \inf\{s \geq 0 \text{ such that } |M_s| \geq \lambda\} \wedge T,$$

with the convention that  $\inf \emptyset = +\infty$ . It is seen that  $\tau$  is an almost surely bounded stopping time. Therefore

$$\mathbb{E}(|M_\tau|^p) \leq \mathbb{E}(|M_T|^p).$$

But from the very definition of  $\tau$ ,

$$|M_\tau|^p \geq 1_{\sup_{0 \leq t \leq T} |M_t| \geq \lambda} \lambda^p + 1_{\sup_{0 \leq t \leq T} |M_t| < \lambda} |M_T|^p,$$

which implies,

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} |M_t| \geq \lambda\right) \leq \frac{\mathbb{E}\left(|M_T|^p 1_{\sup_{0 \leq t \leq T} |M_t| \geq \lambda}\right)}{\lambda^p} \leq \frac{\mathbb{E}(|M_T|^p)}{\lambda^p}.$$

2. Let  $p \geq 1$  and  $T > 0$ . Let us first assume that

$$\mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^p\right) < +\infty.$$

The previous proof shows that for  $\lambda > 0$ ,

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} |M_t| \geq \lambda\right) \leq \frac{\mathbb{E}\left(|M_T| \mathbf{1}_{\sup_{0 \leq t \leq T} |M_t| \geq \lambda}\right)}{\lambda}.$$

We deduce,

$$\begin{aligned} & \int_0^{+\infty} \lambda^{p-1} \mathbb{P}\left(\sup_{0 \leq t \leq T} |M_t| \geq \lambda\right) d\lambda \\ & \leq \int_0^{+\infty} \lambda^{p-2} \mathbb{E}\left(|M_T| \mathbf{1}_{\sup_{0 \leq t \leq T} |M_t| \geq \lambda}\right) d\lambda. \end{aligned}$$

From Fubini theorem we now have

$$\begin{aligned} & \int_0^{+\infty} \lambda^{p-1} \mathbb{P}\left(\sup_{0 \leq t \leq T} |M_t| \geq \lambda\right) d\lambda \\ & = \int_{\Omega} \left( \int_0^{\sup_{0 \leq t \leq T} |M_t|(\omega)} \lambda^{p-1} d\lambda \right) d\mathbb{P}(\omega) \\ & = \frac{1}{p} \mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^p\right). \end{aligned}$$

Similarly, we obtain

$$\begin{aligned} & \int_0^{+\infty} \lambda^{p-2} \mathbb{E}\left(|M_T| \mathbf{1}_{\sup_{0 \leq t \leq T} |M_t| \geq \lambda}\right) d\lambda \\ & = \frac{1}{p-1} \mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^{p-1} |M_T|\right). \end{aligned}$$

Hence,

$$\mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^p\right) \leq \frac{p}{p-1} \mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^{p-1} |M_T|\right).$$

By using now Hölder's inequality we obtain,

$$\mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^{p-1} |M_T|\right) \leq \mathbb{E}\left(|M_T|^p\right)^{\frac{1}{p}} \mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^p\right)^{\frac{p-1}{p}},$$

which implies

$$\mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^p\right) \leq \frac{p}{p-1} \mathbb{E}\left(|M_T|^p\right)^{\frac{1}{p}} \mathbb{E}\left(\left(\sup_{0 \leq t \leq T} |M_t|\right)^p\right)^{\frac{p-1}{p}}.$$

As a conclusion if  $\mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^p \right) < +\infty$ , we have:

$$\mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^p \right) \leq \left( \frac{p}{p-1} \right)^p \mathbb{E}(|M_T|^p).$$

Now, if  $\mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^p \right) = +\infty$ , we consider for  $N \in \mathbb{N}$ , the stopping time  $\tau_N = \inf\{t \geq 0, |M_t| \geq N\} \wedge T$ . By using the above result to the martingale  $(M_{t \wedge \tau_N})_{t \geq 0}$ , we obtain

$$\mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_{t \wedge \tau_N}| \right)^p \right) \leq \left( \frac{p}{p-1} \right)^p \mathbb{E}(|M_T|^p),$$

from which we may conclude by using the monotone convergence theorem.

□

## Notes and comments

We mostly restricted our attention to continuous stochastic processes but a general theory of left limited right continuous processes can similarly be developed, see the books by Jacod-Shiryaev [?] and by Protter [?] for a detailed account on jump processes. As pointed out in the text, most of the material concerning filtrations, martingales and stopping times is due to Doob and was exposed in his extremely influential book [?]. Further readings about the general theory of stochastic processes and martingales include the classical references by Dellacherie-Meyer [?], Ikeda-Watanabe [?], Protter [?] and Revuz-Yor [?].

## Chapter 2

### Brownian motion

The chapter is devoted to the study of the Brownian motion. This is without doubt the most important continuous stochastic process and will serve as a canonical example in most of the developments of the next chapters. It is at the same time a Gaussian process, a martingale and a Markov process. In the first part of the chapter we prove the existence of the Brownian motion as a consequence of the Daniell-Kolmogorov and of the Kolmogorov continuity theorem. We apply the martingale techniques developed in Chapter 1 to study basic properties of the Brownian motion paths like the law of iterated logarithm. The second part of the chapter focusses on the point of view that the Brownian motion can be seen as a continuous random walk in continuous time. More precisely, we will prove that the Brownian motion is the limit of suitably rescaled symmetric random walks. The study of random walks will then allow us to obtain several properties of the Brownian motion paths by a limiting procedure.

#### 1 Definition and basic properties

**Definition 2.1.** Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. A continuous real-valued process  $(B_t)_{t \geq 0}$  is called a standard Brownian motion if it is a Gaussian process with mean function

$$\mathbb{E}(B_t) = 0$$

and covariance function

$$\mathbb{E}(B_s B_t) = \min(s, t).$$

**Remark 2.2.** It is seen that  $R(s, t) = \min(s, t)$  is a covariance function, because it is obviously symmetric and for  $a_1, \dots, a_n \in \mathbb{R}$  and  $t_1, \dots, t_n \in \mathbb{R}_{\geq 0}$ ,

$$\begin{aligned} \sum_{1 \leq i, j \leq n} a_i a_j \min(t_i, t_j) &= \sum_{1 \leq i, j \leq n} a_i a_j \int_0^{+\infty} \mathbf{1}_{[0, t_i]}(s) \mathbf{1}_{[0, t_j]}(s) ds \\ &= \int_0^{+\infty} \left( \sum_{i=1}^n a_i \mathbf{1}_{[0, t_i]}(s) \right)^2 ds \geq 0. \end{aligned}$$

**Definition 2.3.** The distribution of a standard Brownian motion, which is thus a probability measure on the space of continuous functions  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$  is called the Wiener measure.

**Remark 2.4.** A  $n$ -dimensional stochastic process  $(B_t)_{t \geq 0}$  is called a standard Brownian motion if

$$(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^n)_{t \geq 0}$$

where the processes  $(B_t^i)_{t \geq 0}$  are independent standard Brownian motions.

Of course, the definition of Brownian motion is worth only because such an object exists.

**Theorem 2.5.** *There exist a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a stochastic process on it which is a standard Brownian motion.*

*Proof.* From the Proposition 1.19, there exists a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a Gaussian process  $(X_t)_{t \geq 0}$  on it, whose mean function is 0 and covariance function is

$$\mathbb{E}(X_s X_t) = \min(s, t).$$

We have for  $n \geq 0$  and  $0 \leq s \leq t$ :

$$\mathbb{E}((X_t - X_s)^{2n}) = \frac{(2n)!}{2^n n!} (t - s)^n.$$

Therefore, by using the Kolmogorov continuity theorem, there exists a modification  $(B_t)_{t \geq 0}$  of  $(X_t)_{t \geq 0}$  whose paths are locally  $\gamma$ -Hölder if  $\gamma \in [0, \frac{n-1}{2n})$ .  $\square$

**Remark 2.6.** From the previous proof, we also deduce that the paths of a standard Brownian motion are locally  $\gamma$ -Hölder for every  $\gamma < \frac{1}{2}$ . It can be shown that they are not  $\frac{1}{2}$ -Hölder (see Exercise 2.12 and Theorem 2.23).

The following exercises give some first basic properties of Brownian motion and study some related processes. In all the exercises,  $(B_t)_{t \geq 0}$  is a standard one-dimensional Brownian motion.

**Exercise 2.7.** *Show the following properties:*

1.  $B_0 = 0$  a.s.;
2. For any  $h \geq 0$ , the process  $(B_{t+h} - B_h)_{t \geq 0}$  is a standard Brownian motion;
3. For any  $t > s \geq 0$ , the random variable  $B_t - B_s$  is independent of the  $\sigma$ -algebra  $\sigma(B_u, u \leq s)$ .

**Exercise 2.8** (Symmetry property of the Brownian motion).

1. Show that the process  $(-B_t)_{t \geq 0}$  is a standard Brownian motion.
2. More generally, show that if

$$(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^d)_{t \geq 0}$$

is a  $d$ -dimensional Brownian motion and if  $M$  is an orthogonal  $d \times d$  matrix, then  $(MB_t)_{t \geq 0}$  is a standard Brownian motion.



**Exercise 2.9** (Scaling property of the Brownian motion). *Show that for every  $c > 0$ , the process  $(B_{ct})_{t \geq 0}$  has the same law as the process  $(\sqrt{c}B_t)_{t \geq 0}$ .*

**Exercise 2.10** (Time inversion property of Brownian motion).

1. *Show that almost surely,  $\lim_{t \rightarrow +\infty} \frac{B_t}{t} = 0$ .*
2. *Deduce that the process  $(tB_{\frac{1}{t}})_{t \geq 0}$  has the same law as the process  $(B_t)_{t \geq 0}$ .*

**Exercise 2.11** (Non-canonical representation of Brownian motion).

1. *Show that for  $t > 0$ , the Riemann integral  $\int_0^t \frac{B_s}{s} ds$  almost surely exists.*
2. *Show that the process  $(B_t - \int_0^t \frac{B_s}{s} ds)_{t \geq 0}$  is a standard Brownian motion.*

**Exercise 2.12.** *Show that*

$$\mathbb{P} \left( \sup_{s, t \in [0, 1]} \frac{|B_t - B_s|}{\sqrt{t - s}} = +\infty \right) = 1.$$

**Hint:** *Divide the interval  $[0, 1]$  in subintervals  $[k/n, (k+1)/n]$  to bound from below  $\sup_{s, t \in [0, 1]} \frac{|B_t - B_s|}{\sqrt{t - s}}$  by the supremum of the absolute value of independent Gaussian random variables.*

**Exercise 2.13** (Non-differentiability of the Brownian paths).

1. *Show that if  $f : (0, 1) \rightarrow \mathbb{R}$  is differentiable at  $t \in (0, 1)$ , then there exist an interval  $(t - \delta, t + \delta)$  and a constant  $C > 0$  such that for  $s \in (t - \delta, t + \delta)$ ,*

$$|f(t) - f(s)| \leq C|t - s|.$$

2. *For  $n \geq 1$ , let*

$$M_n = \min_{1 \leq k \leq n} \{ \max \{ |B_{k/n} - B_{(k-1)/n}|, |B_{(k+1)/n} - B_{k/n}|, |B_{(k+2)/n} - B_{(k+1)/n}| \} \}.$$

*Show that  $\lim_{n \rightarrow +\infty} \mathbb{P}(M_n) = 0$ .*

3. *Deduce that*

$$\mathbb{P}(\exists t \in (0, 1), B \text{ is differentiable at } t) = 0.$$

**Exercise 2.14** (Fractional Brownian motion). *Let  $0 < H < 1$ .*

1. Show that for  $s \in \mathbb{R}$ , the function

$$f_s(t) = (|t-s|^{H-\frac{1}{2}} - \mathbf{1}_{(-\infty,0]}(t)|t|^{H-\frac{1}{2}})\mathbf{1}_{(-\infty,s]}(t)$$

is square integrable on  $\mathbb{R}$ .

2. Deduce that

$$R(s,t) = \frac{1}{2} (s^{2H} + t^{2H} - |t-s|^{2H}), \quad s, t \geq 0$$

is a covariance function.

3. A continuous and centered Gaussian process with covariance function  $R$  is called a fractional Brownian motion with parameter  $H$ . Show that such process exists and study its Hölder sample path regularity.

4. Let  $(B_t)_{t \geq 0}$  be a fractional Brownian motion with parameter  $H$ . Show that for any  $h \geq 0$ , the process  $(B_{t+h} - B_h)_{t \geq 0}$  is a fractional Brownian motion.

5. Show that for every  $c > 0$ , the process  $(B_{ct})_{t \geq 0}$  has the same law as the process  $(c^H B_t)_{t \geq 0}$ .

**Exercise 2.15** (Brownian bridge). Let  $T > 0$  and  $x \in \mathbb{R}$ .

1. Show that the process

$$X_t = \frac{t}{T}x + B_t - \frac{t}{T}B_T, \quad 0 \leq t \leq T,$$

is a Gaussian process. Compute its mean function and its covariance function.

2. Show that  $(X_t)_{0 \leq t \leq T}$  is a Brownian motion conditioned to be  $x$  at time  $T$ , that is for every  $0 \leq t_1 \leq \dots \leq t_n < T$ , and  $A_1, \dots, A_n$  Borel sets of  $\mathbb{R}$ ,

$$\mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n) = \mathbb{P}(B_{t_1} \in A_1, \dots, B_{t_n} \in A_n | B_T = x).$$

3. Let  $(\alpha_n)_{n \geq 0}, (\beta_n)_{n \geq 1}$  be two independent sequences of i.i.d. Gaussian random variables with mean 0 and variance 1. By using the Fourier series decomposition of the process  $B_t - tB_1$ , show that the random series

$$X_t = t\alpha_0 + \sqrt{2} \sum_{n=1}^{+\infty} \left( \frac{\alpha_n}{2\pi n} (\cos(2\pi nt) - 1) + \frac{\beta_n}{2\pi n} \sin(2\pi nt) \right)$$

is a Brownian motion on  $[0, 1]$ .

**Exercise 2.16** (Ornstein-Uhlenbeck process). We denote

$$X_t = e^{\frac{t}{2}} B_{1-e^{-t}}, \quad t \geq 0.$$

1. Show that  $(X_t)_{t \geq 0}$  is a Gaussian process. Compute its mean function and its covariance function.
2. Show that the process

$$X_t - \frac{1}{2} \int_0^t X_u du$$

is a Brownian motion.

## 2 Basic properties

In this section, we study some basic properties of the Brownian motion paths.

**Proposition 2.17.** Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion.

$$\mathbb{P} \left( \inf_{t \geq 0} B_t = -\infty, \sup_{t \geq 0} B_t = +\infty \right) = 1.$$

*Proof.* Since the process  $(-B_t)_{t \geq 0}$  is also a Brownian motion, in order to prove that

$$\mathbb{P} \left( \inf_{t \geq 0} B_t = -\infty, \sup_{t \geq 0} B_t = +\infty \right) = 1,$$

we just have to check that

$$\mathbb{P} \left( \sup_{t \geq 0} B_t = +\infty \right) = 1.$$

Let  $N \in \mathbb{N}$ . From the scaling property of Brownian motion we have

$$\mathbb{P} \left( c \sup_{t \geq 0} B_t \leq N \right) = \mathbb{P} \left( \sup_{t \geq 0} B_t \leq N \right), \quad c > 0.$$

Therefore we have

$$\mathbb{P} \left( \sup_{t \geq 0} B_t \leq N \right) = \mathbb{P} \left( \sup_{t \geq 0} B_t = 0 \right).$$

Now, we may observe that

$$\begin{aligned} & \mathbb{P} \left( \sup_{t \geq 0} B_t = 0 \right) \\ & \leq \mathbb{P} \left( B_1 \leq 0, \sup_{t \geq 1} B_t \leq 0 \right) \\ & = \mathbb{P} \left( B_1 \leq 0, \sup_{t \geq 0} (B_{t+1} - B_1) \leq -B_1 \right). \end{aligned}$$

Since the process  $(B_{t+1} - B_1)_{t \geq 0}$  is a Brownian motion independent of  $B_1$  (see Exercise 2.7), we have, as before, for  $c > 0$ ,

$$\mathbb{P} \left( B_1 \leq 0, \sup_{t \geq 0} (B_{t+1} - B_1) \leq -B_1 \right) = \mathbb{P} \left( B_1 \leq 0, c \sup_{t \geq 0} (B_{t+1} - B_1) \leq -B_1 \right).$$

Therefore we get

$$\begin{aligned} \mathbb{P} \left( B_1 \leq 0, \sup_{t \geq 0} (B_{t+1} - B_1) \leq -B_1 \right) &= \mathbb{P} \left( B_1 \leq 0, \sup_{t \geq 0} (B_{t+1} - B_1) = 0 \right) \\ &= \mathbb{P}(B_1 \leq 0) \mathbb{P} \left( \sup_{t \geq 0} (B_{t+1} - B_1) = 0 \right) \\ &= \frac{1}{2} \mathbb{P} \left( \sup_{t \geq 0} B_t = 0 \right). \end{aligned}$$

Thus,

$$\mathbb{P} \left( \sup_{t \geq 0} B_t = 0 \right) \leq \frac{1}{2} \mathbb{P} \left( \sup_{t \geq 0} B_t = 0 \right),$$

and we can deduce that

$$\mathbb{P} \left( \sup_{t \geq 0} B_t = 0 \right) = 0,$$

and

$$\mathbb{P} \left( \sup_{t \geq 0} B_t \leq N \right) = 0.$$

Since this holds for every  $N$ , it implies that

$$\mathbb{P} \left( \sup_{t \geq 0} B_t = +\infty \right) = 1.$$

□

By using this proposition in combination with Exercise 2.7 we deduce the following proposition whose proof is left as an exercise to the reader.

**Proposition 2.18** (Recurrence property of Brownian motion). *Let  $(B_t)_{t \geq 0}$  be a Brownian motion. For every  $t \geq 0$  and  $x \in \mathbb{R}$ ,*

$$\mathbb{P}(\exists s \geq t, B_s = x) = 1.$$

Martingale theory provides powerful tools to study Brownian motion. We list in the Proposition below some martingales naturally associated with the Brownian motion.

**Proposition 2.19.** *Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. The following processes are martingales (with respect to their natural filtration):*

1.  $(B_t)_{t \geq 0}$ ;
2.  $(B_t^2 - t)_{t \geq 0}$ ;
3.  $(e^{\lambda B_t - \frac{\lambda^2}{2}t})_{t \geq 0}$ ,  $\lambda \in \mathbb{C}$ .

*Proof.*

1. First, we note that for  $t \geq 0$ ,  $\mathbb{E}(|B_t|) < +\infty$  because  $B_t$  is a Gaussian random variable. Now for  $t \geq s$ ,

$$\mathbb{E}(B_t - B_s \mid \mathcal{F}_s) = \mathbb{E}(B_t - B_s) = 0,$$

therefore we get

$$\mathbb{E}(B_t \mid \mathcal{F}_s) = B_s.$$

2. For  $t \geq 0$ ,  $\mathbb{E}(B_t^2) = t < +\infty$  and for  $t \geq s$ ,

$$\mathbb{E}((B_t - B_s)^2 \mid \mathcal{F}_s) = \mathbb{E}((B_t - B_s)^2) = t - s,$$

therefore we obtain

$$\mathbb{E}(B_t^2 - t \mid \mathcal{F}_s) = B_s^2 - s.$$

3. For  $t \geq 0$ ,  $\mathbb{E}\left(\left|e^{\lambda B_t - \frac{\lambda^2}{2}t}\right|\right) < +\infty$ , because  $B_t$  is a Gaussian random variable. Then we have for  $t \geq s$ ,

$$\mathbb{E}(e^{\lambda(B_t - B_s)} \mid \mathcal{F}_s) = \mathbb{E}(e^{\lambda(B_t - B_s)}) = e^{\frac{\lambda^2}{2}(t-s)}$$

, and therefore

$$\mathbb{E}\left(e^{\lambda B_t - \frac{\lambda^2}{2}t} \mid \mathcal{F}_s\right) = e^{\lambda B_s - \frac{\lambda^2}{2}s}.$$

□

**Remark 2.20.** The curious reader may wonder how the previous martingales have been constructed. A first hint is that the functions  $(t, x) \rightarrow x$ ,  $(t, x) \rightarrow x^2 - t$ , and  $(t, x) \rightarrow \exp\left(\lambda x - \frac{\lambda^2}{2}t\right)$  have in common that they satisfy the following partial differential equation

$$\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} = 0.$$

The full explanation will be given in the Theorem 3.9 in Chapter 3.

**Exercise 2.21.** Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. Let  $a < 0 < b$ . We denote

$$T = \inf\{t \geq 0, B_t \notin (a, b)\}.$$

1. Show that  $T$  is a stopping time.

2. Show that

$$\mathbb{P}(B_T = a) = \frac{b}{b-a}, \quad \mathbb{P}(B_T = b) = \frac{-a}{b-a}.$$

3. Deduce another proof of the recurrence property of the standard Brownian motion, that is for every  $t \geq 0$  and  $x \in \mathbb{R}$ ,

$$\mathbb{P}(\exists s \geq t, B_s = x) = 1.$$

The previous martingales may be used to explicitly compute the distribution of some functionals associated to the Brownian motion.

**Proposition 2.22.** Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. We denote for  $a > 0$ ,

$$T_a = \inf\{t \geq 0, B_t = a\}.$$

For every  $\lambda > 0$ , we have

$$\mathbb{E}(e^{-\lambda T_a}) = e^{-a\sqrt{2\lambda}}.$$

Therefore, the distribution of  $T_a$  is given by the density function

$$\mathbb{P}(T_a \in dt) = \frac{a}{(2\pi t)^{3/2}} e^{-\frac{a^2}{2t}} dt, \quad t > 0.$$

*Proof.* Let  $\alpha > 0$ . For  $N \geq 1$ , we denote by  $T_N$  the almost surely bounded stopping time:

$$T_N = T_a \wedge N.$$

Applying the Doob's stopping theorem to the martingale  $(e^{\alpha B_t - \frac{\alpha^2}{2}t})_{t \geq 0}$  yields:

$$\mathbb{E}\left(e^{\alpha B_{T_a \wedge N} - \frac{\alpha^2}{2}(T_a \wedge N)}\right) = 1.$$

But for  $N \geq 1$ , we have

$$e^{\alpha B_{T_a \wedge N} - \frac{\alpha^2}{2}(T_a \wedge N)} \leq e^{\alpha a}.$$

Therefore from Lebesgue dominated convergence theorem,  $n \rightarrow +\infty$ , we obtain

$$\mathbb{E}\left(e^{\alpha B_{T_a} - \frac{\alpha^2}{2}T_a}\right) = 1.$$

Since by continuity of the Brownian paths we have,

$$B_{T_a} = a,$$

we conclude,

$$\mathbb{E}\left(e^{-\frac{\alpha^2}{2}T_a}\right) = e^{-\alpha a}.$$

The formula for the density function of  $T_a$  is then obtained by inverting the previous Laplace transform.  $\square$

### 3 The law of iterated logarithm

We already observed that as a consequence of Kolmogorov's continuity theorem, the Brownian paths are  $\gamma$ -Hölder continuous for every  $\gamma \in (0, \frac{1}{2})$ . The next proposition, which is known as the law of iterated logarithm shows in particular that Brownian paths are not  $\frac{1}{2}$ -Hölder continuous.

**Theorem 2.23** (Law of iterated logarithm). *Let  $(B_t)_{t \geq 0}$  be a Brownian motion. For  $s \geq 0$ ,*

$$\mathbb{P} \left( \liminf_{t \rightarrow 0} \frac{B_{t+s} - B_s}{\sqrt{2t \ln \ln \frac{1}{t}}} = -1, \limsup_{t \rightarrow 0} \frac{B_{t+s} - B_s}{\sqrt{2t \ln \ln \frac{1}{t}}} = 1 \right) = 1.$$

*Proof.* Thanks to the symmetry and invariance by translation of the Brownian paths, it suffices to show that :

$$\mathbb{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \ln \ln \frac{1}{t}}} = 1 \right) = 1.$$

Let us first prove that

$$\mathbb{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \ln \ln \frac{1}{t}}} \leq 1 \right) = 1.$$

Let us denote

$$h(t) = \sqrt{2t \ln \ln \frac{1}{t}}.$$

Let  $\alpha, \beta > 0$ , from Doob's maximal inequality applied to the exponential martingale  $(e^{\alpha B_t - \frac{\alpha^2}{2}t})_{t \geq 0}$ , we have for  $t \geq 0$ :

$$\mathbb{P} \left( \sup_{0 \leq s \leq t} \left( B_s - \frac{\alpha}{2}s \right) > \beta \right) = \mathbb{P} \left( \sup_{0 \leq s \leq t} e^{\alpha B_s - \frac{\alpha^2}{2}s} > e^{\alpha\beta} \right) \leq e^{-\alpha\beta}.$$

Let now  $\theta, \delta \in (0, 1)$ . Using the previous inequality for every  $n \in \mathbb{N}$  with

$$t = \theta^n, \alpha = \frac{(1 + \delta)h(\theta^n)}{\theta^n}, \beta = \frac{1}{2}h(\theta^n),$$

yields when  $n \rightarrow +\infty$ ,

$$\mathbb{P} \left( \sup_{0 \leq s \leq \theta^n} \left( B_s - \frac{(1 + \delta)h(\theta^n)}{2\theta^n}s \right) > \frac{1}{2}h(\theta^n) \right) = O \left( \frac{1}{n^{1+\delta}} \right).$$

Therefore from Borel-Cantelli lemma, for almost every  $\omega \in \Omega$ , we may find  $N(\omega) \in \mathbb{N}$  such that for  $n \geq N(\omega)$ ,

$$\sup_{0 \leq s \leq \theta^n} \left( B_s(\omega) - \frac{(1 + \delta)h(\theta^n)}{2\theta^n} s \right) \leq \frac{1}{2}h(\theta^n).$$

But

$$\sup_{0 \leq s \leq \theta^n} \left( B_s(\omega) - \frac{(1 + \delta)h(\theta^n)}{2\theta^n} s \right) \leq \frac{1}{2}h(\theta^n)$$

implies that for  $\theta^{n+1} < t \leq \theta^n$ ,

$$B_t(\omega) \leq \sup_{0 \leq s \leq \theta^n} B_s(\omega) \leq \frac{1}{2}(2 + \delta)h(\theta^n) \leq \frac{(2 + \delta)h(t)}{2\sqrt{\theta}}.$$

We conclude

$$\mathbb{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \ln \ln \frac{1}{t}}} \leq \frac{2 + \delta}{2\sqrt{\theta}} \right) = 1.$$

Letting now  $\theta \rightarrow 1$  and  $\delta \rightarrow 0$  yields

$$\mathbb{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \ln \ln \frac{1}{t}}} \leq 1 \right) = 1.$$

Let us now prove that

$$\mathbb{P} \left( \limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \ln \ln \frac{1}{t}}} \geq 1 \right) = 1.$$

Let  $\theta \in (0, 1)$ . For  $n \in \mathbb{N}$ , we denote

$$A_n = \left\{ \omega, B_{\theta^n}(\omega) - B_{\theta^{n+1}}(\omega) \geq (1 - \sqrt{\theta})h(\theta^n) \right\}.$$

Let us prove that

$$\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) = +\infty.$$

The basic inequality

$$\int_a^{+\infty} e^{-\frac{u^2}{2}} du \geq \frac{a}{1 + a^2} e^{-\frac{a^2}{2}}, \quad a > 0,$$

which is obtained by integrating by parts the left side of

$$\int_a^{+\infty} \frac{1}{u^2} e^{-\frac{u^2}{2}} du \leq \frac{1}{a^2} \int_a^{+\infty} e^{-\frac{u^2}{2}} du,$$



implies

$$\mathbb{P}(A_n) = \frac{1}{\sqrt{2\pi}} \int_{a_n}^{+\infty} e^{-\frac{u^2}{2}} du \geq \frac{a_n}{1+a_n^2} e^{-\frac{a_n^2}{2}},$$

with

$$a_n = \frac{(1 - \sqrt{\theta})h(\theta^n)}{\theta^{n/2}\sqrt{1-\theta}}.$$

When  $n \rightarrow +\infty$ ,

$$\frac{a_n}{1+a_n^2} e^{-\frac{a_n^2}{2}} = O\left(\frac{1}{n^{\frac{1+\theta-2\sqrt{\theta}}{1-\theta}}}\right),$$

therefore we proved,

$$\sum \mathbb{P}(A_n) = +\infty.$$

As a consequence of the independence of the increments of the Brownian motion and of Borel-Cantelli lemma, the event

$$B_{\theta^n} - B_{\theta^{n+1}} \geq (1 - \sqrt{\theta})h(\theta^n)$$

will occur almost surely for infinitely many  $n$ 's. But, thanks to the first part of the proof, for almost every  $\omega$ , we may find  $N(\omega)$  such that for  $n \geq N(\omega)$

$$B_{\theta^{n+1}} > -2h(\theta^{n+1}) \geq -2\sqrt{\theta}h(\theta^n).$$

Thus, almost surely, the event

$$B_{\theta^n} > h(\theta^n)(1 - 3\sqrt{\theta})$$

will occur for infinitely many  $n$ 's. This implies

$$\mathbb{P}\left(\limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \ln \ln \frac{1}{t}}} \geq 1 - 3\sqrt{\theta}\right) = 1.$$

We finally get

$$\mathbb{P}\left(\limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \ln \ln \frac{1}{t}}} \geq 1\right) = 1.$$

by letting  $\theta \rightarrow 0$ . □

As a straightforward consequence, we may observe that the time inversion invariance property of Brownian motion implies:

**Corollary 2.24.** *Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion.*

$$\mathbb{P}\left(\liminf_{t \rightarrow +\infty} \frac{B_t}{\sqrt{2t \ln \ln t}} = -1, \limsup_{t \rightarrow +\infty} \frac{B_t}{\sqrt{2t \ln \ln t}} = 1\right) = 1.$$

## 4 Symmetric random walks

In the previous section, the existence of the Brownian motion was proven as a consequence of Daniell-Kolmogorov theorem. However, as it has been stressed, the proof of Daniell-Kolmogorov theorem relies on the axiom of choice. As a consequence, it does not provide any insight of how Brownian motion may explicitly be constructed or simulated by computers. In this section, we provide an *explicit construction* of Brownian motion as a limit of a sequence of suitably rescaled symmetric random walks.

**Definition 2.25.** A random walk on  $\mathbb{Z}$  is a sequence of  $\mathbb{Z}$ -valued random variables  $(S_n)_{n \geq 0}$  that are defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  for which the following two properties hold:

1.  $(S_n)_{n \geq 0}$  has stationary increments, that is for  $m, n \in \mathbb{N}$ ,  $S_{m+n} - S_m$  has the same distribution as  $S_n$
2.  $(S_n)_{n \geq 0}$  has independent increments, that is for  $m, n \in \mathbb{N}$ ,  $S_{m+n} - S_m$  is independent from the  $\sigma$ -field  $\sigma(S_0, \dots, S_m)$ .

For instance, let us consider a game in which we toss a fair coin. If this is heads, the player earns 1\$, if this is tails, the player loses 1\$. Then the algebraic wealth of the player after  $n$  tosses is a random walk on  $\mathbb{Z}$ .

More precisely, let us consider a sequence of independent random variables  $(X_i)_{i \geq 1}$  that are defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and that satisfy

$$\mathbb{P}(X_i = -1) = \mathbb{P}(X_i = 1) = \frac{1}{2}$$

Consider now the sequence  $(S_n)_{n \geq 0}$  such that  $S_0 = 0$  and for  $n \geq 1$

$$S_n = \sum_{i=1}^n X_i$$

For  $n \geq 0$ , we will denote by  $\mathcal{F}_n$  the  $\sigma$ -field generated by the random variables  $S_0, \dots, S_n$ .

**Proposition 2.26.** *The sequences  $(S_n)_{n \geq 0}$  and  $(-S_n)_{n \geq 0}$  are random walks on  $\mathbb{Z}$ .*

*Proof.* As a consequence of the independence of the  $X_i$ 's, the following equality holds in distribution

$$(X_1, \dots, X_n) = (X_{m+1}, \dots, X_{m+n}).$$

Therefore, in distribution we have

$$S_{n+m} - S_m = \sum_{k=m+1}^{n+m} X_k = \sum_{k=1}^n X_k = S_n.$$

The independence of the increments is shown in the very same way and the proof that  $(-S_n)_{n \geq 0}$  is also a random walk is identical.  $\square$

The random walk  $(S_n)_{n \geq 0}$  is called the symmetric random walk on  $\mathbb{Z}$  and is the one which shall be the most important for us. Its distribution may be computed by standard combinatorial arguments.

**Proposition 2.27.** *For  $n \in \mathbb{N}$ ,  $k \in \mathbb{Z}$ :*

1. *If the integers  $n$  and  $k$  have the same parity:*

$$\mathbb{P}(S_n = k) = \frac{1}{2^n} \binom{n}{\frac{n+k}{2}},$$

where  $\binom{n}{p} = \frac{n!}{p!(n-p)!}$  is the binomial coefficient.

2. *If the integers  $n$  and  $k$  do not have the same parity*

$$\mathbb{P}(S_n = k) = 0.$$

*Proof.* We observe that the random variable  $\frac{n+S_n}{2}$  is the sum of  $n$  independent Bernoulli random variables, therefore the distribution of  $\frac{n+S_n}{2}$  is binomial with parameters  $(0, n)$ . The result follows directly.  $\square$

**Exercise 2.28.** *For  $x \in \mathbb{R}$ , by using Stirling's formula ( $n! \sim e^{-n} n^n \sqrt{2\pi n}$ ) give an equivalent when  $n \rightarrow \infty$  of  $\mathbb{P}(S_{2n} = 2[x])$ , where  $[x]$  denotes the integer part of  $x$ .*

Several interesting martingales are naturally associated to the symmetric random walk. They play the same role in the study of  $(S_n)_{n \geq 0}$  as the martingales in Theorem 2.19 for the Brownian motion.

**Proposition 2.29.** *The following sequences of random variables are martingales with respect to the filtration  $(\mathcal{F}_n)_{n \geq 0}$ :*

1.  $(S_n)_{n \geq 0}$
2.  $(S_n^2 - n)_{n \geq 0}$
3.  $\exp(-\lambda S_n - n \ln(\cosh(\lambda))), n \geq 1, \lambda > 0$ .

*Proof.*

1. For  $n \geq m$

$$\begin{aligned}\mathbb{E}(S_n | \mathcal{F}_m) &= \mathbb{E}(S_n - S_m | \mathcal{F}_m) + \mathbb{E}(S_m | \mathcal{F}_m) \\ &= \mathbb{E}(S_n - S_m) + S_m \\ &= S_m\end{aligned}$$

2. First, let us observe that

$$\begin{aligned}\mathbb{E}(S_n^2) &= \mathbb{E}\left(\left(\sum_{i=1}^n X_i\right)^2\right) \\ &= \mathbb{E}\left(\sum_{i,j=1}^n X_i X_j\right) \\ &= \sum_{i=1}^n \mathbb{E}(X_i^2) + \sum_{i,j=1, i \neq j}^n \mathbb{E}(X_i) \mathbb{E}(X_j) \\ &= n\end{aligned}$$

Now, for  $n \geq m$ , we have

$$\mathbb{E}\left((S_n - S_m)^2 | \mathcal{F}_m\right) = \mathbb{E}\left((S_n - S_m)^2\right) = \mathbb{E}(S_{n-m}^2) = n - m$$

But, on the other hand we have

$$\begin{aligned}\mathbb{E}\left((S_n - S_m)^2 | \mathcal{F}_m\right) &= \mathbb{E}(S_n^2 | \mathcal{F}_m) - 2\mathbb{E}(S_n S_m | \mathcal{F}_m) + \mathbb{E}(S_m^2 | \mathcal{F}_m) \\ &= \mathbb{E}(S_n^2 | \mathcal{F}_m) - 2S_m^2 + S_m^2 \\ &= \mathbb{E}(S_n^2 | \mathcal{F}_m) - S_m^2.\end{aligned}$$

We therefore conclude

$$\mathbb{E}(S_n^2 - n | \mathcal{F}_m) = S_m^2 - m.$$

3. Similarly, we have for  $n \geq m$

$$\begin{aligned}\mathbb{E}\left(e^{-\lambda(S_n - S_m)} | \mathcal{F}_m\right) &= \mathbb{E}\left(e^{-\lambda(S_n - S_m)}\right) \\ &= \mathbb{E}\left(e^{-\lambda S_{n-m}}\right) \\ &= \mathbb{E}\left(e^{-\lambda X_1}\right)^{n-m} \\ &= (\cosh \lambda)^{n-m},\end{aligned}$$

which yields the expected result. □

The following proposition shows that the symmetric random walk is a Markov process.

**Proposition 2.30** (Markov property of the random walk). *For  $m \leq n$ ,  $k \in \mathbb{Z}$ ,*

$$\mathbb{P}(S_n = k | \mathcal{F}_m) = \mathbb{P}(S_n = k | S_m).$$

*Proof.* We have for  $\lambda > 0$ ,

$$\mathbb{E}(e^{-\lambda S_n} \mid \mathcal{F}_m) = (\cosh \lambda)^{n-m} e^{-\lambda S_m}.$$

Thus, we have

$$\mathbb{E}(e^{-\lambda S_n} \mid \mathcal{F}_m) = \mathbb{E}(e^{-\lambda S_n} \mid S_m).$$

Since the conditional distribution of  $S_n$  given  $\mathcal{F}_m$  is completely characterized by its conditional Laplace transform, we obtain the claimed result.  $\square$

A random variable  $T$ , valued in  $\mathbb{N} \cup \{+\infty\}$ , is called a stopping time for the random walk  $(S_n)_{n \geq 0}$  if for any  $m$ , the event  $\{T \leq m\}$  belongs to the  $\sigma$ -field  $\mathcal{F}_m$ . The set

$$\mathcal{F}_T = \{A \in \mathcal{F}, \forall m \in \mathbb{N}, A \cap \{T \leq m\} \in \mathcal{F}_m\}$$

is then a sub  $\sigma$ -field of  $\mathcal{F}$ .

**Proposition 2.31.** *Let  $T$  be a stopping time for the random walk  $(S_n)_{n \geq 0}$  such that  $\mathbb{P}(T < +\infty) = 1$ . The sequence  $(S_{n+T} - S_T)_{n \geq 0}$  is a symmetric random walk on  $\mathbb{Z}$  that is independent from  $\mathcal{F}_T$ .*

*Proof.* Let us denote

$$\tilde{S}_n = S_{n+T} - S_T$$

and consider the stopping time

$$T_m = T + m.$$

From the Doob's stopping theorem applied to the martingale  $\left( (\cos \lambda)^{-n} e^{i\lambda S_n} \right)_{n \geq 0}$  and the bounded stopping time  $T_m \wedge N$  where  $N \in \mathbb{N}$ , we get,  $\forall n \in \mathbb{N}$

$$\mathbb{E} \left( (\cos \lambda)^{-n} e^{i\lambda(S_{n+T_m \wedge N} - S_{T_m \wedge N})} \mid \mathcal{F}_{T_m \wedge N} \right) = 1.$$

Letting  $N \rightarrow +\infty$  yields  $\forall n \in \mathbb{N}$

$$\mathbb{E} \left( (\cos \lambda)^{-n} e^{i\lambda(\tilde{S}_{n+m} - \tilde{S}_m)} \mid \mathcal{F}_{T+m} \right) = 1. \quad (2.1)$$

This implies that the increments of  $\tilde{S}_n$  are stationary and independent. Hence  $(\tilde{S}_n)_{n \geq 0}$  is a random walk on  $\mathbb{Z}$  which is independent from  $\mathcal{F}_T$ .

Let us finally prove that this random walk is symmetric. The random variable  $\tilde{S}_{n+1} - \tilde{S}_n$  is valued in the set  $\{-1, 1\}$  and satisfies,

$$\mathbb{E} \left( e^{i\lambda(\tilde{S}_{n+1} - \tilde{S}_n)} \right) = \cos \lambda.$$

This implies

$$\mathbb{P} \left( \tilde{S}_{n+1} - \tilde{S}_n = -1 \right) = \mathbb{P} \left( \tilde{S}_{n+1} - \tilde{S}_n = 1 \right) = \frac{1}{2}.$$

$\square$

As a corollary of the previous result, we obtain the so-called strong Markov property that reinforces the Markov property.

**Corollary 2.32** (Strong Markov property). *Let  $T$  be a stopping time for the random walk  $(S_n)_{n \geq 0}$  such that  $\mathbb{P}(T < +\infty) = 1$ . For every  $k \in \mathbb{Z}$*

$$\mathbb{P}(S_{T+1} = k \mid \mathcal{F}_T) = \mathbb{P}(S_{T+1} = k \mid S_T).$$

*Proof.* Since the sequence of random variables  $(S_{n+T} - S_T)_{n \geq 0}$  is a symmetric random walk on  $\mathbb{Z}$  that is independent from  $\mathcal{F}_T$ , for  $\lambda > 0$ ,

$$\mathbb{E}(e^{-\lambda S_{T+1}} \mid \mathcal{F}_T) = (\cosh \lambda) e^{-\lambda S_T}.$$

Therefore

$$\mathbb{E}(e^{-\lambda S_{T+1}} \mid \mathcal{F}_T) = \mathbb{E}(e^{-\lambda S_{T+1}} \mid S_T),$$

and we conclude as in the proof of Proposition 2.30.  $\square$

The next proposition shows that with probability 1, the symmetric random walk visits each integer of  $\mathbb{Z}$ . This property is called the recurrence property of the symmetric random walk. A similar result for the Brownian motion was proved in Proposition 2.18.

**Proposition 2.33** (Recurrence property of the symmetric random walk).

$$\forall k \in \mathbb{Z}, \forall m \in \mathbb{N}, \mathbb{P}(\exists n \geq m, S_n = k) = 1$$

*Proof.* We do not give the simplest or the most direct proof (see for instance Exercise 2.34 for a simpler argument and Exercise 2.35 for still another proof), but the following proof has the advantage to provide several useful further informations on the random walk.

We first show that

$$\mathbb{P}(\forall k \in [1, 2n], S_k \neq 0) = \mathbb{P}(S_{2n} = 0).$$

We have

$$\begin{aligned} & \mathbb{P}(\forall k \in [1, 2n], S_k \neq 0) \\ &= 2\mathbb{P}(\forall k \in [1, 2n], S_k > 0) \\ &= 2 \sum_{j=1}^n \mathbb{P}(\forall k \in [1, 2n], S_k > 0 \mid S_{2n} = 2j) \mathbb{P}(S_{2n} = 2j) \\ &= 2 \sum_{j=1}^n \frac{j}{n} \mathbb{P}(S_{2n} = 2j) \\ &= 2 \sum_{j=1}^n \frac{j}{n} \frac{1}{2^{2n}} \binom{2n}{n+j} \end{aligned}$$

Now, we have the following binomial identity which may, for instance, be proved inductively,

$$\sum_{j=1}^n j \binom{2n}{n+j} = \frac{n}{2} \binom{2n}{n}.$$

This yields

$$\mathbb{P}(\forall k \in [1, 2n], S_k \neq 0) = \mathbb{P}(S_{2n} = 0).$$

Since

$$\mathbb{P}(S_{2n} = 0) = \frac{(2n)!}{(n!)^2} \frac{1}{2^{2n}} \xrightarrow{n \rightarrow +\infty} 0,$$

we deduce that

$$\mathbb{P}(\exists n > 1, S_n = 0) = 1.$$

The random variable

$$Z_1 = \inf\{n > 1, S_n = 0\},$$

is therefore finite. Let us denote

$$T_1 = \inf\{n \geq 0, S_n = 1\}$$

and

$$T_{-1} = \inf\{n \geq 0, S_n = -1\}.$$

It is clear that  $T_1$  and  $T_{-1}$  are stopping times for the random walk  $(S_n)_{n \geq 0}$ . Our goal will be to prove that

$$\mathbb{P}(T_1 < +\infty) = 1$$

Thanks to the Markov property, for  $n \geq 1$ , we obtain

$$\begin{aligned} & \mathbb{P}(Z_1 \geq n) \\ &= \mathbb{P}(Z_1 \geq n \mid S_1 = 1) \mathbb{P}(S_1 = 1) + \mathbb{P}(Z_1 \geq n \mid S_1 = -1) \mathbb{P}(S_1 = -1) \\ &= \frac{1}{2} \mathbb{P}(Z_1 \geq n \mid S_1 = 1) + \frac{1}{2} \mathbb{P}(Z_1 \geq n \mid S_1 = -1) \\ &= \frac{1}{2} \mathbb{P}(T_1 \geq n - 1) + \frac{1}{2} \mathbb{P}(T_{-1} \geq n - 1) \end{aligned}$$

Now, since we have the following equality in distribution

$$(S_n)_{n \geq 0} = (-S_n)_{n \geq 0},$$

it is clear that  $\mathbb{P}(T_1 < +\infty) = 1$  if and only if  $\mathbb{P}(T_{-1} < +\infty) = 1$ , in which case the following equality in distribution holds

$$T_1 = T_{-1}.$$

Thus, from

$$\mathbb{P}(Z_1 < +\infty) = 1$$

we deduce that

$$\mathbb{P}(T_1 < +\infty) = 1.$$

Now, we may apply the previous result to the random walk  $(S_{n+T_1} - 1)_{n \geq 0}$ . This gives

$$\mathbb{P}(\exists n > 1, S_n = 2) = 1.$$

By a straightforward induction, we obtain

$$\forall k \in \mathbb{N}, \mathbb{P}(\exists n > 1, S_n = k) = 1.$$

Then, by symmetry we get

$$\forall k \in \mathbb{N}, \mathbb{P}(\exists n > 1, S_n = -k) = 1.$$

By using this with the random walk  $(S_{n+m} - S_m)_{n \geq 0}$ , we finally get the expected

$$\forall k \in \mathbb{Z}, \forall m \in \mathbb{N}, \mathbb{P}(\exists n \geq m, S_n = k) = 1.$$

□

**Exercise 2.34.** *By using a method similar to the method of Exercise 2.21, give another proof of the recurrence property of the symmetric random walk.*

**Exercise 2.35** (Law of iterated logarithm for the random walk).

1. Let  $(B_t)_{t \geq 0}$  be a one dimensional Brownian motion. Consider the sequence of stopping times inductively defined by  $T_0 = 0$  and

$$T_{n+1} = \inf\{t \geq T_n, |B_t - B_{T_n}| = 1\}.$$

Consider then the process  $(S_n)_{n \geq 0}$  defined by  $S_n = B_{T_n}$ . Show that  $(S_n)_{n \geq 0}$  is a symmetric random walk.

2. Deduce that if  $(S_n)_{n \geq 0}$  is a symmetric random walk,

$$\mathbb{P}\left(\liminf_{n \rightarrow +\infty} \frac{S_n}{\sqrt{2n \ln \ln n}} = -1, \limsup_{n \rightarrow +\infty} \frac{S_n}{\sqrt{2n \ln \ln n}} = 1\right) = 1.$$

Let us consider for  $k \in \mathbb{Z}$ , the stopping hitting time

$$T_k = \inf\{n \geq 1, S_n = k\}.$$

A straightforward consequence of the previous proposition is that

$$\mathbb{P}(T_k < +\infty) = 1.$$

This sequence of hitting times  $(T_k)_{k \geq 0}$  enjoys several nice properties, the first being that it is itself a random walk.

**Proposition 2.36.** *The sequence  $(T_k)_{k \geq 0}$  is a random walk on  $\mathbb{Z}$ .*

*Proof.* For  $0 \leq a < b$ , we have

$$T_b = T_a + \inf\{n \geq 1, S_{n+T_a} = b\}.$$

Now, since  $(S_{n+T_a} - a)_{n \geq 0}$  is a random walk on  $\mathbb{Z}$  which is independent from  $T_a$ , we deduce that:



1.  $\inf\{n \geq 1, S_{n+T_a} = b\}$  is independent from  $\mathcal{F}_{T_a}$ .
2. The following equality holds in distribution

$$\inf\{n \geq 1, S_{n+T_a} = b\} = T_{b-a}.$$

□

The distribution of  $T_k$  may explicitly be computed by using the method of generating series.

**Proposition 2.37.** For  $0 < x < 1$ , and  $k \in \mathbb{Z}$ ,

$$\sum_{n=0}^{\infty} x^n \mathbb{P}(T_k = n) = \left( \frac{1 - \sqrt{1 - x^2}}{x} \right)^k.$$

*Proof.* Let  $\lambda > 0$ . Applying the Doob's stopping theorem to the martingale

$$\left( (\cosh \lambda)^{-n} e^{-\lambda S_n} \right)_{n \geq 0}$$

with the stopping time  $T_k \wedge N$ ,  $N \in \mathbb{N}$ , yields

$$\mathbb{E} \left( e^{-\lambda S_{N \wedge T_k} - (T_k \wedge N) \ln \cosh \lambda} \right) = 1.$$

The dominated convergence theorem implies then:

$$\mathbb{E} \left( e^{-\lambda k - T_k \ln \cosh \lambda} \right) = 1.$$

This yields,

$$\mathbb{E} \left( e^{-(\ln \cosh \lambda) T_k} \right) = e^{\lambda k}.$$

By denoting

$$x = e^{-\ln \cosh \lambda}$$

we get therefore

$$\mathbb{E} (x^{T_k}) = \left( \frac{1 - \sqrt{1 - x^2}}{x} \right)^k,$$

which is the result we want to prove. □

From the previous generating series, we deduce the following explicit distribution:

**Corollary 2.38.** For  $k \in \mathbb{N}$ ,  $n \in \mathbb{N}$ ,

$$\mathbb{P}(T_k = n) = \frac{1}{2^n} \sum_I \frac{1}{2i_1 - 1} \cdots \frac{1}{2i_k - 1} \binom{2i_1 - 1}{i_1} \cdots \binom{2i_k - 1}{i_k},$$

where

$$I = \left\{ (i_1, \dots, i_k) \in \mathbb{N}^k, i_1 + \dots + i_k = \frac{n-k}{2} \right\}.$$

In particular, for  $n \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{P}(T_1 = 2n) &= 0 \\ \mathbb{P}(T_1 = 2n - 1) &= \frac{1}{2n - 1} \binom{2n - 1}{n} \frac{1}{2^{2n-1}}. \end{aligned}$$

*Proof.* The proof reduces to the computation of the coefficients in the Taylor expansion of the function

$$x \rightarrow \left( \frac{1 - \sqrt{1 - x^2}}{x} \right)^k.$$

□

**Exercise 2.39.** Show that

$$\mathbb{P}(T_1 = 2n - 1) \sim_{n \rightarrow +\infty} \frac{C}{n^{3/2}},$$

where  $C > 0$  is a constant to be computed.

**Exercise 2.40.** Let

$$Z_0 = 0, Z_{k+1} = \inf\{n > Z_k, S_n = 0\}$$

Show that:

1. The sequence  $(Z_k)_{k \geq 0}$  is a random walk on  $\mathbb{Z}$ .
2. For  $k \geq 1$ , the following equality in distribution holds

$$Z_k = k + T_k.$$

The following proposition is known as the reflection principle for the random walk.

**Proposition 2.41** (Reflection principle). For  $n \geq 0$ , let

$$M_n = \max_{0 \leq k \leq n} S_k.$$

We have for  $k, n \in \mathbb{N}$ ,

$$\mathbb{P}(M_n \geq k, S_n < k) = \frac{1}{2} (\mathbb{P}(T_k \leq n) - \mathbb{P}(S_n = k)).$$

*Proof.* Let  $k, n \in \mathbb{N}$ . We first observe the following equality between events,

$$\{T_k \leq n\} = \{M_n \geq k\}.$$

Let us now consider the random walk  $(\tilde{S}_m)_{m \geq 0}$  defined by

$$\tilde{S}_m = S_{m+T_k} - k.$$

As we have already seen it, this sequence is a symmetric random walk for which  $n - T_k \wedge n$  is a stopping time. Now, the following equality holds in distribution

$$\tilde{S}_{n-T_k \wedge n} = -\tilde{S}_{n-T_k \wedge n}.$$

Therefore

$$\mathbb{P}(T_k \leq n, \tilde{S}_{n-T_k} < 0) = \mathbb{P}(T_k \leq n, \tilde{S}_{n-T_k} > 0),$$

which may be rewritten

$$\mathbb{P}(T_k \leq n, S_n < k) = \mathbb{P}(T_k \leq n, S_n > k).$$

We finally conclude

$$2\mathbb{P}(T_k \leq n, S_n < k) + \mathbb{P}(S_n = k) = \mathbb{P}(T_k \leq n),$$

which is our claim.  $\square$

As a corollary of the reflection principle, we immediately obtain the following expression for the distribution of  $M_n$ .

**Corollary 2.42.** For  $k, n \in \mathbb{N}$

$$\mathbb{P}(M_n \geq k) = 2\mathbb{P}(S_n \geq k) - \mathbb{P}(S_n = k).$$

**Exercise 2.43** (Local time of the symmetric random walk). For  $n \in \mathbb{N}$  we denote

$$l_n^0 = \mathbf{Card}\{0 \leq i \leq n, S_i = 0\}.$$

1. Show that for  $n \in \mathbb{N}$  and  $k \in \mathbb{N}^*$

$$\mathbb{P}(l_{2n}^0 = k) = \mathbb{P}(l_{2n+1}^0 = k) = \mathbb{P}(S_{2n+1-k} = k-1)$$

2. Deduce that when  $n \rightarrow +\infty$  the following convergence takes place in distribution

$$\frac{l_n^0}{\sqrt{n}} \rightarrow |\mathcal{N}(0, 1)|$$

## 5 Donsker theorem

After this study of the symmetric random walk on  $\mathbb{Z}$ , we now turn to the Donsker invariance principle which says that the Brownian motion may be constructed as a limit of conveniently rescaled random walks. Before stating this theorem, we need some general results about weak convergence of probability measures in infinitely dimensional spaces. Let us first recall the definition of convergence in distribution for a sequence of random variables that are valued in a Polish space<sup>1</sup>:

**Definition 2.44.** A sequence  $(X_n)_{n \in \mathbb{N}}$  of random variables, valued in a Polish space  $E$ , is said to converge in distribution toward a random variable  $X$  if for every continuous and bounded function  $f : E \rightarrow \mathbb{R}$ , we have:

$$\mathbb{E}(f(X_n)) \rightarrow_{n \rightarrow +\infty} \mathbb{E}(f(X)).$$

Equivalently, the sequence of the distributions of the  $X_n$ 's is said to weakly converge to the distribution of  $X$ .

You may observe that the random variables  $X_n$  do not need to be defined on the same probability space. A usual strategy to show that a sequence  $(X_n)_{n \in \mathbb{N}}$  converges in distribution is to prove that :

1. The family  $(X_n)_{n \in \mathbb{N}}$  is relatively compact in the weak convergence topology ;
2. The sequence  $(X_n)_{n \in \mathbb{N}}$  has a unique cluster point.

Since a stochastic process is nothing else but a random variable that is valued in  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$ , we have a notion of convergence in distribution for a sequence of stochastic processes and, due to the fact that the distribution of stochastic process is fully described by its finite dimensional distributions, we can prove the following useful fact:

**Proposition 2.45.** *Let  $(X^n)_{n \in \mathbb{N}}$  be a sequence of continuous processes and let  $X$  be a continuous process. Let us assume that:*

1. *The sequence of the distributions of the  $X^n$ 's is relatively compact in the weak convergence topology;*
2. *For every  $t_1, \dots, t_k \in \mathbb{R}^k$ , the following convergence holds in distribution*

$$(X_{t_1}^n, \dots, X_{t_k}^n) \rightarrow_{n \rightarrow +\infty} (X_{t_1}, \dots, X_{t_k}).$$

*Then, the sequence  $(X^n)_{n \in \mathbb{N}}$  converges in distribution toward  $X$ .*

---

<sup>1</sup>A Polish space is a complete and separable metric space

In order to efficiently use the previous proposition, we therefore need to characterize the relatively compact sequences in the weak convergence topology. A first ingredient is the Prokhorov theorem .

**Theorem 2.46** (Prokhorov theorem). *Let  $\mathcal{P}$  be a family of probability measures on a Polish space  $E$  endowed with its Borel  $\sigma$ -field  $\mathcal{E}$ . Then,  $\mathcal{P}$  is a relatively compact set for the weak convergence of probability measures topology if and only if the family  $\mathcal{P}$  is tight, that is for every  $\varepsilon \in (0, 1)$ , we can find a compact set  $K_\varepsilon \subset E$  such that for every  $\mathbb{P} \in \mathcal{P}$ ,*

$$\mathbb{P}(K_\varepsilon) \geq 1 - \varepsilon.$$

The second ingredient is Ascoli theorem that describes the relatively compact sets in  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$  for the topology of uniform convergence on compact sets.

**Theorem 2.47** (Ascoli theorem). *For  $N \in \mathbb{N}$ ,  $f \in \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$  and  $\delta > 0$ , we denote:*

$$V^N(f, \delta) = \sup\{|f(t) - f(s)|, |t - s| \leq \delta, s, t \leq N\}.$$

*A set  $K \subset \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$  is relatively compact if and only if:*

1. *The set  $\{f(0), f \in K\}$  is bounded;*
2. *For every  $N \in \mathbb{N}$ ,*

$$\limsup_{\delta \rightarrow 0} \sup_{f \in K} V^N(f, \delta) = 0.$$

As usual, we denote  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$  the space of continuous functions and the associated Borel  $\sigma$ -field is denoted  $\mathcal{B}(\mathbb{R}_{\geq 0}, \mathbb{R})$ . Also,  $(\pi_t)_{t \geq 0}$  denotes the coordinate process. Combining Prokhorov and Ascoli's theorems we obtain:

**Proposition 2.48.** *On the space  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$ , a sequence of probability measures  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  is relatively compact in the weak convergence topology if and only if:*

1. *For every  $\varepsilon > 0$ , there exist  $A > 0$  and  $n_0 \in \mathbb{N}$  such that for every  $n \geq n_0$ ,*

$$\mathbb{P}_n(|\pi_0| > A) \leq \varepsilon;$$

2. *For every  $\eta, \varepsilon > 0$  and  $N \in \mathbb{N}$ , we may find  $\delta > 0$  and  $n_0 \in \mathbb{N}$  such that for  $n \geq n_0$ ,*

$$\mathbb{P}_n(V^n(\pi, \delta) > \eta) \leq \varepsilon.$$

*Proof.* Assume that the sequence  $(\mathbb{P}_n)_{n \in \mathbb{N}}$  is relatively compact in the topology of weak convergence. From Prokhorov's theorem, this sequence is tight, that is for  $\varepsilon \in (0, 1)$ , we can find a compact set  $K_\varepsilon \subset \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ , such that for  $n \in \mathbb{N}$ :

$$\mathbb{P}_n(K_\varepsilon) \geq 1 - \varepsilon.$$

By writing  $K_\varepsilon$  in the form given by Ascoli's theorem, it is easily checked that properties (1) and (2) are satisfied with  $n_0 = 0$ .

Let us now assume that (1) and (2) are satisfied. First, as a finite sequence is relatively compact, we may assume that the properties (1) and (2) are fulfilled with  $n_0 = 0$ . Also, thanks to Prokhorov's theorem, we only need to show that the sequence is tight. Let  $\varepsilon > 0$  and  $N \in \mathbb{N}$ . For every  $k \geq 1$ , we can find  $A_N > 0$  and  $\delta_{N,k}$  such that:

$$\sup_{n \in \mathbb{N}} \mathbb{P}_n (|\pi_0| > A_N) \leq \frac{\varepsilon}{2^{N+1}}$$

$$\sup_{n \in \mathbb{N}} \mathbb{P}_n \left( V^N(\pi, \delta_{N,k}) > \frac{1}{k} \right) \leq \frac{\varepsilon}{2^{N+k+1}}$$

We consider then

$$K_\varepsilon = \bigcap_{N \in \mathbb{N}} \left\{ f \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}), |f(0)| \leq A_N, V^N(\pi, \delta_{N,k}) \leq \frac{1}{k}, \forall k \geq 1 \right\}.$$

Ascoli's theorem implies that  $K_\varepsilon$  is relatively compact, and it is easy to see that for  $n \geq 0$ ,

$$\mathbb{P}_n(K_\varepsilon) \geq 1 - \varepsilon.$$

□

**Exercise 2.49.** Let  $(X^n)_{n \in \mathbb{N}}$  be a sequence of continuous stochastic processes such that:

1. The family formed by the distributions of the  $X_0^n$ 's,  $n \in \mathbb{N}$  is tight;
2. There exist  $\alpha, \beta, \gamma > 0$  such that for  $s, t \geq 0$  and  $n \geq 0$ ,

$$\mathbb{E}(|X_t^n - X_s^n|^\alpha) \leq \beta |t - s|^{1+\gamma}.$$

Show that the family of the distributions of the  $X^n$ 's  $n \in \mathbb{N}$  is relatively compact in the topology of the weak convergence.

We are now in position to prove the Donsker's approximation theorem, that constructs Brownian motion as a limit of symmetric random walks on  $\mathbb{Z}$ .

Let  $(S_n)_{n \in \mathbb{N}}$  be a symmetric random walk on  $\mathbb{Z}$ . We define the sequence  $(S_t^n)_{t \in [0,1]}$ ,  $n \in \mathbb{N}$ , as follows:

$$S_t^n = \sqrt{n} \left( \left( t - \frac{k}{n} \right) S_{k+1} + \left( \frac{k+1}{n} - t \right) S_k \right), \quad \frac{k}{n} \leq t \leq \frac{k+1}{n}.$$

The process  $(S_t^n)_{t \geq 0}$  is therefore the piecewise affine continuous interpolation of the rescaled discrete sequence  $(S_n)_{n \geq 0}$ .

**Theorem 2.50** (Donsker theorem). *The sequence of processes  $(S_t^n)_{t \in [0,1]}$ ,  $n \in \mathbb{N}$ , converges in distribution to a standard Brownian motion  $(B_t)_{t \in [0,1]}$ .*

*Proof.* We need to check two things:

1. For every  $t_1, \dots, t_k \in [0, 1]$ , the following convergence in distribution takes place

$$(S_{t_1}^n, \dots, S_{t_k}^n) \rightarrow_{n \rightarrow +\infty} (B_{t_1}, \dots, B_{t_k}).$$

2. The family of the distributions of the  $(S_t^n)_{t \in [0,1]}$ ,  $n \in \mathbb{N}$ , is relatively compact in the weak convergence topology.

The first point (1) is left as an exercise to the reader: It is basically a consequence of the multidimensional central limit theorem. Let us however point out that in order to simplify the computations, it will be easier to prove that for every  $t_1 \leq \dots \leq t_k \in [0, 1]$ , the following convergence in distribution takes place

$$(S_{t_1}^n, S_{t_2}^n - S_{t_1}^n, \dots, S_{t_k}^n - S_{t_{k-1}}^n) \rightarrow_{n \rightarrow +\infty} (B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_k} - B_{t_{k-1}}).$$

So, we turn to the second point (2). Let  $\lambda > 0$ . The process  $(S_n^4)_{n \in \mathbb{N}}$  is a submartingale, therefore from Doob's maximal inequality, for  $n \geq 1$ ,

$$\mathbb{P} \left( \max_{k \leq n} |S_k| > \lambda \sqrt{n} \right) \leq \frac{\mathbb{E}(S_n^4)}{\lambda^4 n^2} = \frac{3n^2 - 2n}{\lambda^4 n^2} \leq \frac{3}{\lambda^4}.$$

Thanks to the stationarity of the increments of  $(S_n)_{n \in \mathbb{N}}$ , we deduce that for  $k, n \geq 1$ ,  $\lambda > 0$ ,

$$\mathbb{P} \left( \max_{i \leq n} |S_{i+k} - S_k| > \lambda \sqrt{n} \right) \leq \frac{3}{\lambda^4}.$$

Let now  $\varepsilon, \eta \in (0, 1)$ . From the previous inequality we can find  $\delta > 0$  such that for every  $k, n \geq 1$

$$\mathbb{P} \left( \max_{i \leq [n\delta]} |S_{i+k} - S_k| \geq \varepsilon \sqrt{n} \right) \leq \eta \delta.$$

Let  $N, \varepsilon, \eta > 0$ . From the definition of  $S^n$  we deduce that we can find  $\delta \in (0, 1)$  such that for every  $n \geq 1$  and  $t \leq N$ ,

$$\mathbb{P} \left( \sup_{t \leq s \leq t+\delta} |S_s^n - S_t^n| \geq \eta \right) \leq \varepsilon \delta.$$

For  $0 \leq i < \frac{N}{\delta}$  et  $n \geq 1$ , let

$$A_i^n = \left\{ \omega \in \Omega, \sup_{i\delta \leq s \leq (i+1)\delta \wedge N} |S_{i\delta}^n - S_s^n| \geq \eta \right\}.$$

We may check that

$$\cap_i {}^c A_i^n \subset \{\sup\{|S_t^n - S_s^n|, |t - s| \leq \delta, s, t \leq N\} < 3\eta\}.$$

Therefore, for every  $n \geq 1$ ,

$$\begin{aligned} & \mathbb{P}(\sup\{|S_t^n - S_s^n|, |t - s| \leq \delta, s, t \leq N\} \geq 3\eta) \\ & \leq \mathbb{P}\left(\bigcup_i A_i^n\right) \\ & \leq (1 + [N\delta^{-1}]) \delta\varepsilon \\ & < (N + 1)\varepsilon. \end{aligned}$$

This implies the expected relative compactness property thanks to Proposition 2.48.  $\square$

Besides its importance to simulate Brownian motion on a computer for instance, it is interesting that Donsker's theorem may also be used to explicitly compute some distributions of Brownian functionals.

**Theorem 2.51** (Arcsine law). *Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. For  $t \geq 0$ , we denote*

$$A_t = \int_0^t 1_{[0, +\infty)}(B_s) ds.$$

We have for  $x \leq t$ :

$$\mathbb{P}(A_t \leq x) = \frac{2}{\pi} \arcsin \sqrt{\frac{x}{t}}.$$

*Proof.* Let  $(S_n)_{n \in \mathbb{N}}$  be a symmetric random walk on  $\mathbb{Z}$ . We denote by  $T_1$  the hitting time of 1 by  $(S_n)_{n \in \mathbb{N}}$ , and

$$Z_1 = \inf\{n \geq 1, S_n = 0\}.$$

Exercise 2.40 proved that the following equality in distribution takes place

$$Z_1 = T_1 + 1.$$

Let us denote

$$A_n = \mathbf{Card}\{1 \leq i \leq n, \max(S_{i-1}, S_i) > 0\}.$$

We will prove by induction that for  $0 \leq k \leq n$ ,

$$\mathbb{P}(A_{2n} = 2k) = \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2n-2k} = 0)$$

First,

$$\mathbb{P}(A_{2n} = 0) = \mathbb{P}(T_1 \geq 2n + 1) = \mathbb{P}(Z_1 \geq 2n + 2) = \mathbb{P}(S_1 \neq 0, \dots, S_{2n} \neq 0).$$



As seen before:

$$\mathbb{P}(S_1 \neq 0, \dots, S_{2n} \neq 0) = \mathbb{P}(S_{2n} = 0).$$

Thus,

$$\mathbb{P}(A_{2n} = 0) = \mathbb{P}(S_{2n} = 0).$$

In the same way

$$\mathbb{P}(A_{2n} = 2n) = \mathbb{P}(S_{2n} = 0).$$

For  $1 \leq k \leq n-1$ ,

$$\{A_{2n} = 2k\} \subset \{Z_1 \leq 2k\},$$

thus:

$$\begin{aligned} \mathbb{P}(A_{2n} = 2k) &= \sum_{i=1}^k \mathbb{P}(A_{2n} = 2k \mid Z_1 = 2i) \mathbb{P}(Z_1 = 2i) \\ &= \frac{1}{2} \sum_{i=1}^k \mathbb{P}(Z_1 = 2i) (\mathbb{P}(A_{2n-2i} = 2k) + \mathbb{P}(A_{2n-2i} = 2k - 2i)). \end{aligned}$$

Thanks to the induction assumption at the step  $n$  we have,

$$\begin{aligned} &\mathbb{P}(A_{2n} = 2k) \\ &= \frac{1}{2} \sum_{i=1}^k \mathbb{P}(Z_1 = 2i) (\mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2n-2k-2i} = 0) \\ &\quad + \mathbb{P}(S_{2k-2i} = 0) \mathbb{P}(S_{2n-2k} = 0)). \end{aligned}$$

The Markov property then implies

$$\begin{aligned} \sum_{i=1}^k \mathbb{P}(Z_1 = 2i) \mathbb{P}(S_{2k-2i} = 0) &= \sum_{i=1}^k \mathbb{P}(Z_1 = 2i) \mathbb{P}(S_{2k} = 0 \mid Z_1 = 2i) \\ &= \mathbb{P}(S_{2k} = 0). \end{aligned}$$

In the same way, we can prove

$$\sum_{i=1}^k \mathbb{P}(Z_1 = 2i) \mathbb{P}(S_{2n-2k-2i} = 0) = \mathbb{P}(S_{2n-2k} = 0).$$

Thus

$$\mathbb{P}(A_{2n} = 2k) = \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2n-2k} = 0)$$

which completes the induction. As a conclusion, we have for  $0 \leq k \leq n$ ,

$$\mathbb{P}(A_{2n} = 2k) = \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2n-2k} = 0) = \frac{1}{2^n} \frac{(2n-2k)! (2k)!}{(n-k)!^2 (k)!^2}.$$

The Stirling's formula

$$n! \sim_{n \rightarrow \infty} n^n e^{-n} \sqrt{2\pi n},$$

implies that for  $x \in [0, 1]$ :

$$\mathbb{P}\left(\frac{A_{2n}}{2n} \leq x\right) \sim_{n \rightarrow +\infty} \frac{1}{\pi} \sum_{k=0}^{\lfloor nx \rfloor} \frac{1}{\sqrt{k(n-k)}}.$$

Since

$$\int_0^x \frac{du}{\pi \sqrt{u(1-u)}} = \frac{2}{\pi} \arcsin \sqrt{x}$$

we deduce that for  $x \in [0, 1]$ :

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{A_{2n}}{2n} \leq x\right) = \frac{2}{\pi} \arcsin \sqrt{x}.$$

It is now time to use Donsker's theorem. We consider the sequence

$$S_t^n = \sqrt{n} \left( \left(t - \frac{k}{n}\right) S_{k+1} + \left(\frac{k+1}{n} - t\right) S_k \right), \quad \frac{k}{n} \leq t \leq \frac{k+1}{n}.$$

It is easy to see that

$$\int_0^1 1_{[0, +\infty)}(S_t^n) dt = \frac{A_{2n}}{2n}.$$

Donsker's theorem implies therefore that

$$\mathbb{P}(A_1 \leq x) = \frac{2}{\pi} \arcsin \sqrt{x}.$$

The distribution of  $A_t$  is then finally deduced from the distribution of  $A_1$  by using the scaling property of Brownian motion.  $\square$

### Exercise 2.52.

1. Let  $(S_n)_{n \in \mathbb{N}}$  be a symmetric random walk on  $\mathbb{Z}$ . We denote

$$A_n = \mathbf{Card}\{1 \leq i \leq n, \max(S_{i-1}, S_i) > 0\}.$$

(a) Show that for  $0 < x, y < 1$

$$\sum_{n=0}^{+\infty} \sum_{k=0}^n \mathbb{P}(A_{2n} = 2k, S_{2n} = 0) x^{2k} y^{2n} = \frac{2}{\sqrt{1-y^2} + \sqrt{1-x^2 y^2}}.$$

(b) Deduce that, conditionally to  $S_{2n} = 0$ , the random variable  $A_{2n}$  is uniformly distributed on the set  $\{0, 2, 4, \dots, 2n\}$ .

2. Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. By using Donsker's theorem show that conditionally to  $B_1 = 0$ , the random variable

$$\int_0^1 1_{[0, +\infty)}(B_s) ds,$$

is uniformly distributed on  $[0, 1]$ .

**Exercise 2.53.** Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. The goal of the exercise is to compute the distribution of the random variable

$$g_1 = \sup\{t \in [0, 1], B_t = 0\}.$$

1. Let  $(S_n)_{n \in \mathbb{N}}$  be a symmetric random walk on  $\mathbb{Z}$ . For  $n \in \mathbb{N}$ , we denote

$$d_n = \max\{1 \leq k \leq n, S_k = 0\}.$$

Show that for  $0 \leq k \leq n$

$$\mathbb{P}(d_{2n} = 2k) = \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2n-2k} = 0).$$

2. Deduce that for  $x \in [0, 1]$

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{d_n}{n} \leq x\right) = \frac{2}{\pi} \arcsin \sqrt{x}.$$

3. Deduce the distribution of  $g_1$ .

## Notes and Comments

There are many rigorous constructions of the Brownian motion. The first one is due to Wiener [?]. A direct construction of the Wiener measure is given by Itô in [?]. Many properties of Brownian motions were known from Paul Lévy [?]. For more details about random walks, we refer to Durrett [?], Lawler-Vlada [?] and the classical reference by Spitzer. For further reading on Donsker theorem and convergence in distribution in path spaces we refer to Section 6, Chapter II in [?] or to the book by Jacod and Shiryaev [?].

## Chapter 3

### Markov processes

In this chapter we study continuous time Markov processes and emphasize the role played by the transition semigroup, to study sample paths properties of Markov processes. The class of processes we shall particularly be interested in are the Feller-Dynkin processes, which are Markov processes admitting regular versions and enjoying the strong Markov property. We will show that Feller-Dynkin processes admit generators and that this generator is a second order differential operator if the process is continuous. We finish the chapter with the study of the Lévy processes which are Feller-Dynkin processes associated to convolution semigroups.

#### 1 Markov processes

Intuitively, a stochastic process defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  is a Markov process if it is memoryless, that is if for every bounded and Borel function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_s^X) = \mathbb{E}(f(X_{t+s}) | X_s), \quad s, t \geq 0.$$

Let us turn to the more precise definition that relies on the notion of transition function. A transition function for Markov process is the analogue in continuous time of the transition matrix associated to a Markov chain.

**Definition 3.1.** A transition function  $\{P_t, t \geq 0\}$  on  $\mathbb{R}^n$  is a family of kernels

$$P_t : \mathbb{R} \times \mathcal{B}(\mathbb{R}^n) \rightarrow [0, 1]$$

such that:

1. For  $t \geq 0$  and  $x \in \mathbb{R}^n$ ,  $P_t(x, \cdot)$  is a probability measure on  $\mathbb{R}^n$ ;
2. For  $t \geq 0$  and  $A$  Borel set in  $\mathbb{R}^n$  the application  $x \rightarrow P_t(x, A)$  is measurable;
3. For  $s, t \geq 0$ ,  $x \in \mathbb{R}^n$  and  $A$  Borel set in  $\mathbb{R}^n$ ,

$$P_{t+s}(x, A) = \int_{\mathbb{R}^n} P_t(y, A) P_s(x, dy). \quad (3.1)$$

The relation (3.1) is called the Chapman-Kolmogorov relation.

Given a transition function, we can define a one parameter family of linear operators  $(\mathbf{P}_t)_{t \geq 0}$  from the space of bounded Borel functions into itself as follows:

$$(\mathbf{P}_t f)(x) = \int_{\mathbb{R}} f(y) P_t(x, dy).$$

If  $\{P_t, t \geq 0\}$  is a transition function, then the following properties are satisfied:

- $\mathbf{P}_t 1 = 1$ ;
- For every  $t \geq 0$ ,  $\mathbf{P}_t$  is a positivity preserving operator, in the sense that if  $f$  is non negative, so is  $\mathbf{P}_t f$ ;
- For every  $t \geq 0$ ,  $\mathbf{P}_t$  is a contraction from the space of bounded Borel functions into itself (that is, it is a continuous operator with a norm smaller than 1);
- The semigroup property holds: For every  $s, t \geq 0$ ,

$$\mathbf{P}_{t+s} = \mathbf{P}_t \mathbf{P}_s, s, t \geq 0.$$

**Definition 3.2.** A stochastic process  $(X_t)_{t \geq 0}$  defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  is called a Markov process if there exists a transition function  $\{P_t, t \geq 0\}$  on  $\mathbb{R}^n$  such that for every bounded and Borel function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_s^X) = \mathbf{P}_t f(X_s), \quad s, t \geq 0,$$

where  $\mathcal{F}^X$  denotes the natural filtration<sup>1</sup> of the process  $(X_t)_{t \geq 0}$ . The family of operators  $(\mathbf{P}_t)_{t \geq 0}$  is called the semigroup of the Markov process.

**Remark 3.3.** We may also speak of the Markov property with respect to a given filtration. A stochastic process  $(X_t)_{t \geq 0}$  defined on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is called a Markov process with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  if there exists a transition function  $\{P_t, t \geq 0\}$  on  $\mathbb{R}^n$  such that for every bounded and Borel function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_s) = \mathbf{P}_t f(X_s), \quad s, t \geq 0.$$

Brownian motion is the primary example of a Markov process.

**Proposition 3.4.** Let  $(B_t)_{t \geq 0}$  be a (one-dimensional) stochastic process defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  such that  $B_0 = 0$  a.s. Then,  $(B_t)_{t \geq 0}$  is a standard Brownian motion if and only if it is a Markov process with semigroup:

$$\mathbf{P}_0 = \text{Id}, \quad (\mathbf{P}_t f)(x) = \int_{\mathbb{R}} f(y) \frac{e^{-\frac{(x-y)^2}{2t}}}{\sqrt{2\pi t}} dy, t > 0, x \in \mathbb{R}.$$

<sup>1</sup>Recall that  $\mathcal{F}_s^X$  is the smallest  $\sigma$ -algebra that makes measurable all the random variables  $(X_{t_1}, \dots, X_{t_m})$ ,  $0 \leq t_1 \leq \dots \leq t_m \leq s$ .

*Proof.* Let  $(B_t)_{t \geq 0}$  be a Markov process defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with semigroup:

$$\mathbf{P}_0 = \mathbf{Id}, \quad (\mathbf{P}_t f)(x) = \int_{\mathbb{R}} f(y) \frac{e^{-\frac{(x-y)^2}{2t}}}{\sqrt{2\pi t}} dy, t > 0, x \in \mathbb{R}.$$

Thanks to the Markov property, by denoting  $\mathcal{F}_t$  the natural filtration of  $(B_t)_{t \geq 0}$ , we have for  $s, t \geq 0$ ,  $\lambda \in \mathbb{R}$ ,

$$\mathbb{E}(e^{i\lambda B_{t+s}} | \mathcal{F}_s) = \int_{\mathbb{R}} e^{i\lambda(B_s+y)} \frac{e^{-\frac{y^2}{2t}}}{\sqrt{2\pi t}} dy.$$

This implies

$$\mathbb{E}(e^{i\lambda(B_{t+s}-B_s)} | \mathcal{F}_s) = \int_{\mathbb{R}} e^{i\lambda y} \frac{e^{-\frac{y^2}{2t}}}{\sqrt{2\pi t}} dy = e^{-\frac{1}{2}\lambda^2 t}.$$

In particular, the increments of  $(B_t)_{t \geq 0}$  are stationary and independent. Now, for  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ ,  $0 < t_1 < \dots < t_n$ :

$$\begin{aligned} \mathbb{E}\left(e^{i\sum_{k=1}^n \lambda_k (B_{t_{k+1}} - B_{t_k})}\right) &= \prod_{k=1}^n \mathbb{E}\left(e^{i\lambda_k (B_{t_{k+1}} - B_{t_k})}\right) \\ &= \prod_{k=1}^n \mathbb{E}\left(e^{i\lambda_k B_{t_{k+1}-t_k}}\right) \\ &= e^{-\frac{1}{2}\sum_{k=1}^n (t_{k+1}-t_k)\lambda_k^2}. \end{aligned}$$

Hence  $(B_t)_{t \geq 0}$  is a standard Brownian motion.

Conversely, let  $(B_t)_{t \geq 0}$  be a standard Brownian motion with natural filtration  $\mathcal{F}_t$ . If  $f$  is a bounded Borel function and  $s, t \geq 0$ , we have:

$$\mathbb{E}(f(B_{t+s}) | \mathcal{F}_s) = \mathbb{E}(f(B_{t+s} - B_s + B_s) | \mathcal{F}_s).$$

Since  $B_{t+s} - B_s$  is independent from  $\mathcal{F}_s$ , we deduce that

$$\mathbb{E}(f(B_{t+s}) | \mathcal{F}_s) = \mathbb{E}(f(B_{t+s}) | B_s).$$

For  $x \in \mathbb{R}$ , we have

$$\mathbb{E}(f(B_{t+s}) | B_s = x) = \mathbb{E}(f(B_{t+s} - B_s + B_s) | B_s = x) = \mathbb{E}(f(X_t + x)),$$

where  $X_t$  is a random variable independent from  $B_s$  and normally distributed with mean 0 and variance  $t$ . Therefore we have

$$\mathbb{E}(f(B_{t+s}) | B_s = x) = \int_{\mathbb{R}} f(x+y) \frac{e^{-\frac{y^2}{2t}}}{\sqrt{2\pi t}} dy$$

and

$$\mathbb{E}(f(B_{t+s}) \mid \mathcal{F}_s) = \int_{\mathbb{R}} f(B_s + y) \frac{e^{-\frac{y^2}{2t}}}{\sqrt{2\pi t}} dy.$$

□

**Exercise 3.5** (Multidimensional Brownian motion). *Let  $(B_t)_{t \geq 0}$  be a  $n$ -dimensional stochastic process defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  such that  $B_0 = 0$  a.s. Show that  $(B_t)_{t \geq 0}$  is a standard Brownian motion on  $\mathbb{R}^n$  if and only if it is a Markov process with semigroup:*

$$\mathbf{P}_0 = \mathbf{Id}, \quad (\mathbf{P}_t f)(x) = \int_{\mathbb{R}^n} f(y) \frac{e^{-\frac{\|x-y\|^2}{2t}}}{(2\pi t)^{n/2}} dy, \quad t > 0, x \in \mathbb{R}^n.$$

**Exercise 3.6** (Ornstein-Uhlenbeck process). *Let  $(B_t)_{t \geq 0}$  be a one-dimensional Brownian motion and let  $\theta \in \mathbb{R} \setminus \{0\}$ . We consider the process*

$$X_t = e^{\theta t} B_{\frac{1-e^{-2\theta t}}{2\theta}}.$$

Show that  $(X_t)_{t \geq 0}$  is a Markov process with semigroup

$$(\mathbf{P}_t f)(x) = \int_{\mathbb{R}} f\left(e^{\theta t} x + \sqrt{\frac{e^{2\theta t} - 1}{2\theta}} y\right) \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy.$$

**Exercise 3.7** (Black-Scholes process). *Let  $(B_t)_{t \geq 0}$  be a one-dimensional Brownian motion and let  $\mu \in \mathbb{R}$ ,  $\sigma > 0$ . We consider the process*

$$X_t = e^{(\mu - \frac{\sigma^2}{2})t + \sigma B_t}.$$

Show that  $(X_t)_{t \geq 0}$  is a Markov process with semigroup

$$(\mathbf{P}_t f)(x) = \int_{\mathbb{R}} f\left(x e^{(\mu - \frac{\sigma^2}{2})t + \sigma y}\right) \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi t}} dy.$$

**Exercise 3.8** (Bessel process). *Let  $(B_t)_{t \geq 0}$  be a  $n$ -dimensional Brownian motion. We consider*

$$X_t = \|B_t\| = \sqrt{\sum_{i=1}^n (B_t^i)^2}.$$

Show that  $(X_t)_{t \geq 0}$  is a Markov process with semigroup given by

$$(\mathbf{P}_t f)(x) = \frac{1}{t} \int_{\mathbb{R}_{\geq 0}} f(y) \left(\frac{y}{x}\right)^{\frac{n}{2}-1} I_{\frac{n}{2}-1}\left(\frac{xy}{t}\right) y e^{-\frac{x^2+y^2}{2t}} dy$$

for  $x > 0$ , and

$$(\mathbf{P}_t f)(0) = \frac{2^{1-\frac{n}{2}}}{\Gamma(n/2)t^{n/2}} \int_{\mathbb{R}_{\geq 0}} f(y) y^{n-1} e^{-\frac{y^2}{2t}} dy,$$

where  $\Gamma$  is the Euler Gamma function and  $I_{\frac{n}{2}-1}$  is the modified Bessel function of the first kind with index  $\frac{n}{2}-1$ , that is the solution of Bessel's equation

$$y'' + \frac{1}{x}y - \left(1 + \left(\frac{n}{2} - 1\right)^2 \frac{1}{x^2}\right)y = 0,$$

such that

$$y(x) \sim_{x \rightarrow 0} \frac{x^{\frac{n}{2}-1}}{2^{n/2}\Gamma(n/2)}.$$

The semigroup of the Brownian motion is intimately related to solutions of the heat equation. The next proposition shows how to construct martingales from solutions of the backward heat equation.

**Theorem 3.9.** *Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. Let  $f : \mathbb{R}_{\geq 0} \times \mathbb{R} \rightarrow \mathbb{C}$  be such that:*

1.  *$f$  is once continuously differentiable with respect to its first variable and twice continuously differentiable with respect to its second variable (we write  $f \in \mathcal{C}^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R}, \mathbb{C})$ ).*
2. *For  $t \geq 0$ , there exist constants  $K > 0$  and  $\alpha > 0$  such that for every  $x \in \mathbb{R}$*

$$\sup_{0 \leq s \leq t} |f(s, x)| \leq K e^{\alpha|x|}.$$

*The process  $(f(t, B_t))_{t \geq 0}$  is a martingale if and only if:*

$$\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} = 0.$$

*Proof.* Let  $t > 0$ . In what follows, we denote by  $\mathcal{F}$  the natural filtration of the Brownian motion. Thanks to the Markov property, we have for  $s < t$ ,

$$\mathbb{E}(f(t, B_t) \mid \mathcal{F}_s) = \int_{\mathbb{R}} f(t, y) \frac{e^{-\frac{(y-B_s)^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dy$$

Therefore the process  $(f(t, B_t))_{t \geq 0}$  is a martingale if and only if for  $0 < s < t$  and  $x \in \mathbb{R}$ ,

$$\int_{\mathbb{R}} f(t, y) \frac{e^{-\frac{(y-x)^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dy = f(s, x).$$



We are thus let to characterize the functions  $f$  that satisfy the above functional equation and the given growth conditions.

So, let  $f$  be a function that satisfies the required regularity and growth conditions and such that for  $0 < s < t$  and  $x \in \mathbb{R}$ ,

$$\int_{\mathbb{R}} f(t, y) \frac{e^{-\frac{(y-x)^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dy = f(s, x).$$

For a fixed  $t > 0$ , the function

$$g : (s, x) \rightarrow \int_{\mathbb{R}} f(t, y) \frac{e^{-\frac{(y-x)^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dy$$

which is defined on  $[0, t) \times \mathbb{R}$  is easily seen to satisfy the equation

$$\frac{\partial g}{\partial s} + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} = 0,$$

so that  $f$ , of course, satisfies the same equation.

Now, assume that  $f$  is a function that satisfies the required growth conditions and the equation:

$$\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} = 0.$$

Let  $t > 0$  be fixed. If we still denote

$$g : (s, x) \rightarrow \int_{\mathbb{R}} f(t, y) \frac{e^{-\frac{(y-x)^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dy,$$

we quickly realize that  $h = f - g$  satisfies

$$\frac{\partial h}{\partial s} + \frac{1}{2} \frac{\partial^2 h}{\partial x^2} = 0$$

on  $[0, t) \times \mathbb{R}$  with moreover the boundary condition:

$$\forall x \in \mathbb{R}, \quad \lim_{s \rightarrow t} h(s, x) = 0.$$

From classical uniqueness of solutions results for the heat equation, we deduce that  $h = 0$ .  $\square$

**Exercise 3.10.** Let  $(B_t)_{t \geq 0}$  be a  $n$ -dimensional standard Brownian motion. Let  $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{C}$  be such that:

1.  $f \in \mathcal{C}^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R}^n, \mathbb{C})$ .

2. For  $t \geq 0$ , there exist constants  $K > 0$  and  $\alpha > 0$  such that

$$\sup_{0 \leq s \leq t} |f(s, x)| \leq K e^{\alpha \|x\|}.$$

Show that the process  $(f(t, B_t))_{t \geq 0}$  is a martingale if and only if:

$$\frac{\partial f}{\partial t} + \frac{1}{2} \Delta f = 0.$$

In particular, if  $f$  is a sub-exponential harmonic function (that is  $\Delta f = 0$ ), then the process  $(f(B_t))_{t \geq 0}$  is a martingale.

**Exercise 3.11.** Let  $(X_t)_{t \geq 0}$  be a Markov process with semigroup  $\mathbf{P}_t$ . Show that if  $T > 0$ , the process  $((\mathbf{P}_{T-t}f)(X_t))_{0 \leq t \leq T}$  is a martingale. By using Doob's stopping theorem, deduce that if  $S$  is a stopping time such that  $S \leq T$  almost surely, then

$$\mathbb{E}(f(X_T) | \mathcal{F}_S) = \mathbf{P}_{T-S}f(X_S).$$

**Exercise 3.12** (Self-similar semigroups). A transition function  $\{P_t, t \geq 0\}$  on  $\mathbb{R}^n$  is said to be self-similar with index  $H > 0$  if for every  $t \geq 0$ ,  $c > 0$ ,  $x \in \mathbb{R}^n$  and every Borel set  $A \subset \mathbb{R}^n$ ,

$$P_t(x, \Delta_c A) = P_{t/c} \left( \frac{x}{c^H}, A \right),$$

where  $\Delta_c$  is the dilation of  $\mathbb{R}^n$  defined by  $\Delta_c(y) = c^H y$ .

1. Show that the transition function of the Brownian motion is self-similar with index  $H = \frac{1}{2}$ .
2. Show that if  $(X_t)_{t \geq 0}$  is a Markov process whose transition function is self-similar with index  $H > 0$  such that  $X_0 = 0$ , then the two processes  $(X_{ct})_{t \geq 0}$  and  $(c^H X_t)_{t \geq 0}$  have the same distribution.
3. Show that a transition function  $\{P_t, t \geq 0\}$  is self-similar with index  $H > 0$ , if and only if for every  $t \geq 0$ ,  $c > 0$  and every bounded and Borel function  $f$ ,

$$(\mathbf{P}_{ct}f) \circ \Delta_c = \mathbf{P}_t(f \circ \Delta_c).$$

**Exercise 3.13** (Rotationally invariant semigroups). A transition function  $\{P_t, t \geq 0\}$  on  $\mathbb{R}^n$  is said to be rotationally invariant if for every  $t \geq 0$ ,  $M \in \mathbf{SO}(\mathbb{R}^n)$  and every Borel set  $A \subset \mathbb{R}^n$ ,

$$P_t(x, M \cdot A) = P_t(M^{-1}x, A),$$

where  $\mathbf{SO}(\mathbb{R}^n)$  is the set of  $n \times n$  matrices such that  ${}^t M M = \mathbf{I}_n$  and  $\det(M) = 1$ , and where  $M \cdot A$  is the set  $\{Mx, x \in A\}$ .

1. Show that the transition function of the Brownian motion is rotationally invariant.
2. Show that if  $(X_t)_{t \geq 0}$  is a Markov process whose transition function is rotationally invariant such that  $X_0 = 0$ , then for every  $M \in \mathbf{SO}(\mathbb{R}^n)$  the two processes  $(MX_t)_{t \geq 0}$  and  $(X_t)_{t \geq 0}$  have the same distribution.
3. Show that a transition function  $\{P_t, t \geq 0\}$  is rotationally invariant, if and only if for every  $t \geq 0$ ,  $M \in \mathbf{SO}(\mathbb{R}^n)$  and every bounded and Borel function  $f$ ,

$$(\mathbf{P}_t f) \circ M = \mathbf{P}_t(f \circ M).$$

It is remarkable that given any transition function, it is always possible to find a corresponding Markov process.

**Theorem 3.14.** Let  $\{P_t, t \geq 0\}$  be a transition function on  $\mathbb{R}^n$ . Let  $\nu$  be a probability measure on  $\mathbb{R}^n$ . There exist a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a stochastic process  $(X_t)_{t \geq 0}$  such that:

1. The distribution of  $X_0$  is  $\nu$  ;
2. If  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a bounded and Borel function

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_s^X) = (\mathbf{P}_t f)(X_s), \quad s, t \geq 0$$

where  $\mathcal{F}^X$  is the natural filtration of  $X$ .

*Proof.* For  $0 = t_0 < t_1 < \dots < t_m$ ,  $A$  a Borel set in  $\mathbb{R}^n$  and  $B$  Borel set in  $(\mathbb{R}^n)^{\otimes m}$ , we define

$$\mu_{t_0, t_1, \dots, t_m}(A \times B) = \int_A \int_B P_{t_1}(z, dx_1) P_{t_2 - t_1}(x_1, dx_2) \dots P_{t_m - t_{m-1}}(x_{m-1}, dx_m) \nu(dz).$$

The measure  $\mu_{t_0, t_1, \dots, t_m}$  is therefore a probability measure on  $\mathbb{R}^n \times (\mathbb{R}^n)^{\otimes m}$ . Since for a Borel set  $C$  in  $\mathbb{R}^n$  and  $x \in \mathbb{R}^n$  we have

$$P_{t+s}(x, C) = \int_{\mathbb{R}^n} P_t(y, C) P_s(x, dy),$$

we deduce that this family of probability satisfies the assumptions of the Daniell-Kolmogorov theorem. Therefore, we can find a process  $(X_t)_{t \geq 0}$  defined on some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  whose finite dimensional distributions are given by the  $\mu_{t_0, t_1, \dots, t_n}$ 's. Let us now prove that this process satisfies the property stated in the theorem. First, the distribution of  $X_0$  is  $\nu$  because

$$\mu_0(A) = \int_A \nu(dz) = \nu(A), \quad A \in \mathcal{B}(\mathbb{R}^n).$$

We now have to prove that if  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a bounded and Borel function and if  $0 < s, t$ , then

$$\mathbb{E}(f(X_{s+t}) \mid \mathcal{F}_s^X) = (\mathbf{P}_t f)(X_s).$$

For this, we have to prove that if  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $F : (\mathbb{R}^n)^{\otimes m} \rightarrow \mathbb{R}$ , are bounded and Borel functions and if  $0 = t_0 < t_1 < \dots < t_m$ , then

$$\mathbb{E}(f(X_{t_m})F(X_{t_0}, \dots, X_{t_{m-1}})) = \mathbb{E}((\mathbf{P}_{t_m - t_{m-1}} f)(X_{t_{m-1}})F(X_{t_0}, \dots, X_{t_{m-1}})).$$

But thanks to Fubini's theorem, we have

$$\begin{aligned} & \mathbb{E}(f(X_{t_m})F(X_{t_0}, \dots, X_{t_{m-1}})) \\ &= \int_{(\mathbb{R}^n)^{\otimes (m+1)}} f(x_m)F(z, x_1, \dots, x_{m-1}) \\ & \quad P_{t_1}(z, dx_1)P_{t_2 - t_1}(x_1, dx_2) \cdots P_{t_m - t_{m-1}}(x_{m-1}, dx_m)\nu(dz) \\ &= \int_{\mathbb{R}^n} \int_{(\mathbb{R}^n)^{\otimes m}} (\mathbf{P}_{t_m - t_{m-1}} f)(x_{m-1})F(z, x_1, \dots, x_{m-1}) \\ & \quad P_{t_1}(z, dx_1) \cdots P_{t_{m-1} - t_{m-2}}(x_{m-2}, dx_{m-1})\nu(dz) \\ &= \mathbb{E}((\mathbf{P}_{t_m - t_{m-1}} f)(X_{t_{m-1}})F(X_{t_0}, \dots, X_{t_{m-1}})). \end{aligned}$$

This concludes the proof of the theorem.  $\square$

**Remark 3.15.** Observe the degree of freedom we have on the distribution of  $X_0$ : This reflects the fact that the transition function characterizes the distribution of a Markov process up to its initial distribution.

To finish the section, we introduce the notion of sub-Markov process, a notion which is very useful in many situations. A family of operators  $(\mathbf{P}_t)_{t \geq 0}$  from the space of bounded Borel functions into itself is called a sub-Markov semigroup if the following conditions are satisfied:

- For every  $t \geq 0$ ,  $\mathbf{P}_t$  is a positivity preserving operator;
- For every  $t \geq 0$ ,  $\mathbf{P}_t$  is a contraction from the space of bounded Borel functions into itself;
- The semigroup property holds: For every  $s, t \geq 0$ ,

$$\mathbf{P}_{t+s} = \mathbf{P}_t \mathbf{P}_s, s, t \geq 0.$$

Observe that unlike Markov semigroups which are associated with Markov processes, we do not necessarily have  $\mathbf{P}_t 1 = 1$  but only  $\mathbf{P}_t 1 \leq 1$ . It turns out that a sub-Markov semigroup  $(\mathbf{P}_t)_{t \geq 0}$  can always be embedded into a Markov semigroup which is defined on a larger space. Indeed let us consider a

cemetery point  $\star$ . If  $\tilde{f}$  is a bounded Borel function on  $\mathbb{R}^n$ , then we extend  $\mathbf{P}_t f$  to  $\mathbb{R}^n \cup \{\star\} \rightarrow \mathbb{R}$  by defining

$$\mathbf{P}_t f(\star) = 0.$$

Also any smooth and compactly supported function  $f$  on  $\mathbb{R}^n$  is automatically extended to  $\mathbb{R}^n \cup \{\star\} \rightarrow \mathbb{R}$  by setting  $f(\star) = 0$ . If  $f$  is a bounded Borel function on  $\mathbb{R}^n \cup \{\star\} \rightarrow \mathbb{R}$ , we then define for  $x \in \mathbb{R}^n$

$$\mathbf{Q}_t f(x) = \mathbf{P}_t \tilde{f}(x) + f(\star)(1 - \mathbf{P}_t 1(x)),$$

where  $\tilde{f}$  is the restriction of  $f$  to  $\mathbb{R}^n$ . It is an easy exercise to check that  $\mathbf{Q}_t$  is now a Markov semigroup on the space of bounded Borel functions  $\mathbb{R}^n \cup \{\star\} \rightarrow \mathbb{R}$ . By definition, a sub-Markov process  $(X_t)_{t \geq 0}$  on  $\mathbb{R}^n$  with transition semigroup  $(\mathbf{P}_t)_{t \geq 0}$  is a Markov process on  $\mathbb{R}^n \cup \{\star\} \rightarrow \mathbb{R}$  with transition semigroup  $(\mathbf{Q}_t)_{t \geq 0}$ . If  $(\bar{X}_t^x)_{t \geq 0}$  is a sub-Markov process with transition semigroup  $(\mathbf{P}_t)_{t \geq 0}$  such that  $X_0^x = x$  and if we define  $\mathbf{e}(x) = \inf\{t \geq 0, X_t^x = \star\}$ , then we have for any bounded and Borel function  $f$  on  $\mathbb{R}^n$ ,

$$\mathbf{P}_t f(x) = \mathbb{E}(f(X_t^x) 1_{t < \mathbf{e}(x)}),$$

in particular, observe that we have

$$\mathbf{P}_t 1(x) = \mathbb{P}(t < \mathbf{e}(x)).$$

The random time  $\mathbf{e}(x)$  is called the extinction time of  $(X_t^x)_{t \geq 0}$ . Killed Markov processes are canonical examples of sub-Markov processes.

**Exercise 3.16** (Killed Markov process). *Let  $(X_t)_{t \geq 0}$  be a Markov process in  $\mathbb{R}^n$  and  $K \subset \mathbb{R}^n$  be a non-empty set. Let  $T_K = \inf\{t \geq 0, X_t \in K\}$ . Consider the following process  $(Y_t)_{t \geq 0}$  such that:*

$$Y_t = \begin{cases} X_t, & t \leq T_K \\ \star, & t > T_K. \end{cases}$$

*Show that  $(Y_t)_{t \geq 0}$  is a sub-Markov process.*

**Exercise 3.17** (Killing at an exponential time). *Let  $(X_t)_{t \geq 0}$  be a Markov process in  $\mathbb{R}^n$ . Let  $T$  be an exponential random variable with parameter  $\lambda$  which is independent from  $(X_t)_{t \geq 0}$ . Consider the following process  $(Y_t)_{t \geq 0}$  such that:*

$$Y_t = \begin{cases} X_t, & t \leq T \\ \star, & t > T. \end{cases}$$

*Show that  $(Y_t)_{t \geq 0}$  is a sub-Markov process and compute its semigroup in terms of the semigroup of  $(X_t)_{t \geq 0}$ .*

## 2 Strong Markov processes

In the study of martingales, we already stressed and illustrated that in the study of stochastic processes, it is often important to know how the process behaves with respect to the stopping times of the underlying filtration. For Markov processes, this naturally leads to the notion of strong Markov process.

**Definition 3.18.** Let  $(X_t)_{t \geq 0}$  be a Markov process with transition function  $\{P_t, t \geq 0\}$ . We say that  $(X_t)_{t \geq 0}$  is a strong Markov process if for any bounded Borel function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , and any finite stopping time  $S$  of the filtration  $(\mathcal{F}_t^X)_{t \geq 0}$ , we have:

$$\mathbb{E}(f(X_{S+t}) \mid \mathcal{F}_S^X) = (\mathbf{P}_t f)(X_S), \quad t > 0.$$

**Remark 3.19.** As for the definition of the Markov property, we may of course define the strong Markov property of a process with respect to a filtration that does not need to be the natural filtration of the process.

In general, it is not straightforward to prove that a given Markov process satisfies the strong Markov property and not all Markov processes enjoy the strong Markov property.

Let us first focus on the Brownian motion case. In that case, as for the symmetric random walks, the strong Markov property is a consequence of the following result.

**Proposition 3.20.** *Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion and let  $T$  be a finite stopping time. The process,*

$$(B_{T+t} - B_T)_{t \geq 0}$$

*is a standard Brownian motion independent from  $\mathcal{F}_T^B$ .*

*Proof.* Let  $T$  be a finite stopping time of the filtration  $(\mathcal{F}_t^B)_{t \geq 0}$ . We first assume  $T$  bounded. Let us consider the process

$$\tilde{B}_t = B_{T+t} - B_T, \quad t \geq 0.$$

Let  $\lambda \in \mathbb{R}$ ,  $0 \leq s \leq t$ . Applying Doob's stopping theorem to the martingale

$$\left( e^{i\lambda B_t + \frac{\lambda^2}{2}t} \right)_{t \geq 0},$$

with the stopping times  $t + T$  and  $s + T$ , yields:

$$\mathbb{E} \left( e^{i\lambda B_{T+t} + \frac{\lambda^2}{2}(T+t)} \mid \mathcal{F}_{T+s} \right) = e^{i\lambda B_{T+s} + \frac{\lambda^2}{2}(T+s)}.$$

Therefore we have

$$\mathbb{E} \left( e^{i\lambda(B_{T+t} - B_{T+s})} \mid \mathcal{F}_{T+s} \right) = e^{-\frac{\lambda^2}{2}(t-s)}.$$

The increments of  $(\tilde{B}_t)_{t \geq 0}$  are therefore independent and stationary. The conclusion thus easily follows. If  $T$  is not bounded almost surely, then we can consider the stopping time  $T \wedge N$  and from the previous result the finite dimensional distributions  $(B_{t_1+T \wedge N} - B_{T \wedge N}, \dots, B_{t_n+T \wedge N} - B_{T \wedge N})$  do not depend on  $N$  and are the same as a Brownian motion. We can then let  $N \rightarrow +\infty$  to conclude.  $\square$

As a corollary, we obtain the strong Markov property of the Brownian motion:

**Corollary 3.21.** *Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. Then  $(B_t)_{t \geq 0}$  is a strong Markov process.*

*Proof.* Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a bounded and Borel function  $t \geq 0$  and let  $S$  be a finite stopping time. From the previous proposition, we have:

$$\mathbb{E}(f(B_{t+S}) \mid \mathcal{F}_S) = \mathbb{E}(f(B_{t+S} - B_S + B_S) \mid \mathcal{F}_S).$$

Since  $B_{t+S} - B_S$  is independent from  $\mathcal{F}_S$ , we first deduce that

$$\mathbb{E}(f(B_{t+S}) \mid \mathcal{F}_S) = \mathbb{E}(f(B_{t+S}) \mid B_S).$$

Now for  $x \in \mathbb{R}$ ,

$$\mathbb{E}(f(B_{t+S}) \mid B_S = x) = \mathbb{E}(f(B_{t+S} - B_S + B_S) \mid B_S = x) = \mathbb{E}(f(X_t + x)),$$

where  $X_t$  is a Gaussian random variable independent from  $B_S$  which has mean 0 and variance  $t$ . Thus,

$$\mathbb{E}(f(B_{t+S}) \mid B_S = x) = \int_{\mathbb{R}} f(x+y) \frac{e^{-\frac{y^2}{2t}}}{\sqrt{2\pi t}} dy$$

and

$$\mathbb{E}(f(B_{t+S}) \mid \mathcal{F}_S) = \int_{\mathbb{R}} f(B_S + y) \frac{e^{-\frac{y^2}{2t}}}{\sqrt{2\pi t}} dy.$$

$\square$

The following exercises show some applications of the strong Markov property.

In the following exercises,  $(B_t)_{t \geq 0}$  is a standard Brownian motion and for  $a \in \mathbb{R}$ , we denote

$$T_a = \inf\{t > 0, B_t = a\},$$

and for  $t \geq 0$ ,

$$S_t = \sup\{B_s, s \leq t\}.$$

**Exercise 3.22** (Reflection principle for the Brownian paths). *Let  $a \in \mathbb{R}$ . Show that the process*

$$\begin{aligned}\tilde{B}_t &= B_t, & t < T_a \\ &= 2a - B_t, & t \geq T_a,\end{aligned}$$

*is a Brownian motion.*

**Exercise 3.23.** *Show that for  $t \geq 0$ ,  $a \geq 0$ ,  $x \leq a$ ,*

$$\mathbb{P}(S_t \in da, B_t \in dx) = \frac{2(2a-x)}{\sqrt{2\pi t^3}} e^{-\frac{(2a-x)^2}{2t}} dadx.$$

**Exercise 3.24.** *Show that the two processes  $(S_t - B_t)_{t \geq 0}$  and  $(|B_t|)_{t \geq 0}$  have the same distribution and that  $(|B_t|)_{t \geq 0}$  is a strong Markov process whose semigroup is given by:*

$$(\mathbf{P}_t f)(x) = \frac{2}{\sqrt{2\pi t}} \int_0^{+\infty} e^{-\frac{x^2+y^2}{2t}} \cosh\left(\frac{xy}{t}\right) f(y) dy, \quad t > 0.$$

**Exercise 3.25** (Local time of the Brownian motion).

1. *Let  $\varepsilon > 0$ . We denote  $U = \inf\{t \geq 0, S_t - B_t > \varepsilon\}$ . Show that the random variable  $S_U$  has an exponential distribution.*
2. *We recursively define the following sequence of stopping times:*

$$T'_1(\varepsilon) = 0, \quad T_n(\varepsilon) = \inf\{t > T'_n(\varepsilon), S_t - B_t > \varepsilon\},$$

$$T'_{n+1}(\varepsilon) = \inf\{t > T_n(\varepsilon), S_t - B_t = 0\}.$$

*We denote*

$$U(t, \varepsilon) = \max\{n, T_n(\varepsilon) \leq t\}.$$

*Show that, almost surely,*

$$\lim_{n \rightarrow +\infty} 2^{-n} U(t, 2^{-n}) = S_t.$$

3. *Deduce that there exists a continuous and non-decreasing process  $(L_t)_{t \geq 0}$  such that  $(|B_t| - L_t)_{t \geq 0}$  is a Brownian motion. Show that the process  $(L_t)_{t \geq 0}$  increases only when  $B_t = 0$ . The process  $(L_t)_{t \geq 0}$  is called the local time of Brownian motion at 0.*
4. *Compute the distribution of  $L_t$ .*



### 3 Feller-Dynkin diffusions

From their very definition, Markov processes are associated to semigroups. More precisely, a multi-dimensional stochastic process  $(X_t)_{t \geq 0}$  is a Markov process if and only if there exists a contraction semigroup of operators  $(\mathbf{P}_t)_{t \geq 0}$  on the Banach space  $\mathbf{L}^\infty(\mathbb{R}^n, \mathbb{R})$  such that:

- For  $0 \leq f \leq 1$ ,  $0 \leq \mathbf{P}_t f \leq 1$ ;
- $\mathbf{P}_t 1 = 1$ ;

and

$$\mathbb{E}(f(X_{t+s}) \mid \mathcal{F}_s^X) = \mathbf{P}_t f(X_s), \quad s, t \geq 0.$$

In general for an arbitrary  $f \in \mathbf{L}^\infty(\mathbb{R}^n, \mathbb{R})$ , the map  $t \rightarrow \mathbf{P}_t f$  fails to be continuous in the strong topology. However, for many interesting examples of Markov processes, this continuity issue is solved by restricting  $(\mathbf{P}_t)_{t \geq 0}$  to a closed subspace  $\mathbb{X}$  of  $\mathbf{L}^\infty(\mathbb{R}^n, \mathbb{R})$  that densely contains the set  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  of smooth and compactly supported functions  $\mathbb{R}^n \rightarrow \mathbb{R}$ . In what follows, we denote by  $\mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$  the Banach space of continuous functions  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $\lim_{\|x\| \rightarrow +\infty} f(x) = 0$ .

**Exercise 3.26.** *Let*

$$(\mathbf{P}_t f)(x) = \int_{\mathbb{R}} f(y) \frac{e^{-\frac{(x-y)^2}{2t}}}{\sqrt{2\pi t}} dy, \quad t > 0, x \in \mathbb{R},$$

*be the semigroup of the Brownian motion.*

1. *Give an example of  $f \in \mathbf{L}^\infty(\mathbb{R}, \mathbb{R})$  such that the map  $t \rightarrow \mathbf{P}_t f$  fails to be continuous in the strong topology.*
2. *Show that  $\forall f \in \mathcal{C}_0(\mathbb{R}, \mathbb{R})$*

$$\lim_{t \rightarrow 0} \|\mathbf{P}_t f - f\|_\infty = 0.$$

The previous exercise leads to the following definitions.

**Definition 3.27** (Feller-Dynkin semigroups).

- Let  $(\mathbf{P}_t)_{t \geq 0}$  be a contraction semigroup of operators on the Banach space  $\mathbf{L}^\infty(\mathbb{R}^n, \mathbb{R})$  such that:
  1. For  $0 \leq f \leq 1$ ,  $0 \leq \mathbf{P}_t f \leq 1$ ;
  2.  $\mathbf{P}_t 1 = 1$ .

We say that  $(\mathbf{P}_t)_{t \geq 0}$  is a Feller-Dynkin semigroup if it satisfies the following additional properties:

1.  $\mathbf{P}_t : \mathcal{C}_0(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$ ;
2.  $\forall f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$ 

$$\lim_{t \rightarrow 0} \|\mathbf{P}_t f - f\|_\infty = 0.$$

- A Markov process  $(X_t)_{t \geq 0}$  is said to be a Feller-Dynkin process if its semigroup is a Feller-Dynkin semigroup.

We know from Theorem 3.14 that it is always possible to associate a Markov process to a transition function. For Feller-Dynkin semigroups, we can moreover work with regular versions that enjoy the strong Markov property.

**Theorem 3.28.** *Let  $\{P_t, t \geq 0\}$  be a Feller-Dynkin transition function. For every probability measure  $\nu$  on  $\mathbb{R}^n$ , there exist a filtered probability space and a stochastic process  $(X_t)_{t \geq 0}$  defined on that space such that:*

1. *The distribution of  $X_0$  is  $\nu$ ;*
2. *The paths of  $(X_t)_{t \geq 0}$  are right continuous and left limited;*
3. *With respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ ,  $(X_t)_{t \geq 0}$  is a strong Markov process with transition function  $\{P_t, t \geq 0\}$ .*

*Proof.* Let  $\{P_t, t \geq 0\}$  be a Feller-Dynkin transition function. We denote by  $\mathbf{P}_t$  the corresponding semigroup. We already know from the Theorem 3.14 that there exist a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  and a stochastic process  $(X_t)_{t \geq 0}$  defined on that space such that:

1. The distribution of  $X_0$  is  $\nu$ ;
2. With respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ ,  $(X_t)_{t \geq 0}$  is a Markov process with transition function  $\{P_t, t \geq 0\}$ .

So, we need to prove that  $(X_t)_{t \geq 0}$  admits a right continuous and left limited modification. The idea is to use the so called resolvent functions of the semigroup. For  $\alpha > 0$  and  $f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$ , consider

$$U_\alpha f(x) = \int_0^{+\infty} e^{-\alpha t} \mathbf{P}_t f(x) dt.$$

If  $f \geq 0$ , we have for  $t \geq s$ ,

$$\begin{aligned} \mathbb{E}(e^{-\alpha t} U_\alpha f(X_t) \mid \mathcal{F}_s) &= e^{-\alpha t} \mathbf{P}_{t-s} U_\alpha f(X_s) \\ &= e^{-\alpha t} \int_0^{+\infty} e^{-\alpha u} \mathbf{P}_{t-s+u} f(X_s) du \\ &= e^{-\alpha s} \int_{t-s}^{+\infty} e^{-\alpha u} \mathbf{P}_u f(X_s) du \\ &\leq e^{-\alpha s} U_\alpha f(X_s). \end{aligned}$$

Therefore, if  $f \geq 0$ , then  $e^{-\alpha t}U_\alpha f(X_t)$  is a supermartingale. From the Doob's regularization theorem (Theorem 1.47), we deduce that for every  $\alpha > 0$ , and every  $f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$ ,  $f \geq 0$ , the process  $U_\alpha f(X_t)$  has almost surely right limits along  $\mathbb{Q}$ . Let now  $\mathcal{I}$  be a countable set of non negative functions in  $\mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$  that separate the points in the sense that for  $x, y \in \mathbb{R}^n$  with  $x \neq y$ , we always can find  $f \in \mathcal{I}$  such that  $f(x) \neq f(y)$ . Since, it is easily proved that for  $f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$ , we always have

$$\lim_{\alpha \rightarrow \infty} \|\alpha U_\alpha f - f\|_\infty = 0,$$

we deduce that the countable set

$$\mathcal{A} = \{U_\alpha f, \alpha \in \mathbb{N}, f \in \mathcal{I}\}$$

also separate points. Since for every  $a \in \mathcal{A}$ , the process  $a(X_t)$  has almost surely right limits along  $\mathbb{Q}$ , we conclude that the process  $X_t$  itself has almost surely right limits along  $\mathbb{Q}$ . This allows to define

$$\tilde{X}_t = \lim_{s \searrow t, s \in \mathbb{Q}} X_s.$$

For every bounded functions  $g, f$  on  $\mathbb{R}^n$  and every  $0 \leq t \leq s$ , we have

$$\mathbb{E}(g(X_t)f(X_s)) = (g(X_t)\mathbf{P}_{t-s}f(X_t)).$$

Therefore, by letting  $s \searrow t, s \in \mathbb{Q}$ , we deduce

$$\mathbb{E}\left(g(X_t)f(\tilde{X}_t)\right) = (g(X_t)f(X_t)).$$

Since the previous equality should hold for every bounded functions  $f$  and  $g$ , we deduce easily by the monotone class theorem that for every  $t \geq 0$ ,  $\tilde{X}_t = X_t$  almost surely. As a conclusion  $(\tilde{X}_t)_{t \geq 0}$  is a right-continuous modification of  $(X_t)_{t \geq 0}$ . Finally, by using once again the Doob's regularization theorem for the supermartingales  $(a(\tilde{X}_t))_{t \geq 0}$ ,  $a \in \mathcal{A}$ , we conclude that  $(\tilde{X}_t)_{t \geq 0}$  almost surely has left limits at any points.

Let us now prove that  $(\tilde{X}_t)_{t \geq 0}$  is a strong Markov process. We need to check that for any Borel function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , and any finite stopping time  $S$  of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ , we have:

$$\mathbb{E}(f(\tilde{X}_{S+t}) | \mathcal{F}_S) = (\mathbf{P}_t f)(\tilde{X}_S), \quad t > 0.$$

Without loss of generality we may assume  $S$  to be bounded (otherwise, just consider  $S_n = S \wedge n$  and let  $n \rightarrow \infty$ ).

As a first step, let us assume that  $S$  takes a finite number of values  $s_1, \dots, s_n$ . Let  $Z$  be a bounded random variable measurable with respect to  $\mathcal{F}_S$ . We observe that the random variable  $1_{\{S=s_k\}}Z$  is measurable with respect to  $\mathcal{F}_{s_k}$  and

deduce that

$$\begin{aligned}\mathbb{E}(f(\tilde{X}_{S+t})Z) &= \sum_{k=1}^n \mathbb{E}\left(f(\tilde{X}_{s_k+t})Z1_{\{S=s_k\}}\right) \\ &= \sum_{k=1}^n \mathbb{E}\left((\mathbf{P}_t f)(\tilde{X}_{s_k})Z1_{\{S=s_k\}}\right) \\ &= \mathbb{E}\left((\mathbf{P}_t f)(\tilde{X}_S)Z\right).\end{aligned}$$

This yields

$$\mathbb{E}(f(\tilde{X}_{S+t}) \mid \mathcal{F}_S) = (\mathbf{P}_t f)(\tilde{X}_S), \quad t > 0.$$

If  $S$  takes an infinite number of values, we approximate  $S$  by the following sequence of stopping times:

$$\tau_n = \sum_{k=1}^{2^n} \frac{kK}{2^n} 1_{\left\{\frac{(k-1)K}{2^n} \leq S < \frac{kK}{2^n}\right\}},$$

where  $K$  is such that  $S \leq K$  almost surely. The stopping time  $\tau_n$  takes its values in a finite set and when  $n \rightarrow +\infty$ ,  $\tau_n \searrow S$ . As before, let  $Z$  be a bounded random variable measurable with respect to  $\mathcal{F}_S$  and let  $f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$ . Since  $\tau_n \geq S$ , we have from the above computation

$$\mathbb{E}(f(\tilde{X}_{\tau_n+t})Z) = \mathbb{E}\left((\mathbf{P}_t f)(\tilde{X}_{\tau_n})Z\right).$$

Since  $\mathbf{P}_t$  is a Feller semigroup,  $\mathbf{P}_t f$  is a bounded continuous function. By letting  $n \rightarrow +\infty$  and using the right continuity of  $(\tilde{X}_t)_{t \geq 0}$  in combination with the dominated convergence we deduce that

$$\mathbb{E}(f(\tilde{X}_{S+t})Z) = \mathbb{E}\left((\mathbf{P}_t f)(\tilde{X}_S)Z\right).$$

By using the monotone class theorem, we see that the previous equality then also holds for every bounded function  $f$ . This concludes the proof of the theorem.  $\square$

The Markov process of the previous theorem may not be continuous. Actually, many interesting examples of Feller-Dynkin process are discontinuous. This is for instance the case of the Lévy processes that will be later studied.

**Exercise 3.29.** (*Quasi-left continuity*) Let  $(X_t)_{t \geq 0}$  be the Feller-Dynkin Markov process constructed in Theorem 3.28.

1. Let  $T$  be a stopping time of the natural filtration of  $(X_t)_{t \geq 0}$ . Show that if  $T_n$  is an increasing sequence of stopping times almost surely converging to  $T$ , then almost surely on the set  $\{T < +\infty\}$ , we have

$$\lim_{n \rightarrow +\infty} X_{T_n} = X_T.$$

2. Deduce that for any  $T \geq 0$ , the set  $\{X_t(\omega), 0 \leq t \leq T\}$  is almost surely bounded.

The following proposition gives a useful criterion for the continuity of the Markov process which is associated to a Feller-Dynkin transition function.

**Proposition 3.30.** *For  $x \in \mathbb{R}^n$  and  $\varepsilon > 0$ , we denote by  $\mathbf{B}(x, \varepsilon)$  the open ball in  $\mathbb{R}^n$  with center  $x$  and radius  $\varepsilon$ . Let  $\{P_t, t \geq 0\}$  be a Feller-Dynkin transition function that satisfies the condition*

$$\limsup_{t \rightarrow 0} \sup_{x \in K} \frac{1}{t} P_t(x, {}^c \mathbf{B}(x, \varepsilon)) = 0,$$

for every  $\varepsilon > 0$  and every compact set  $K$ . The stochastic process  $(X_t)_{t \geq 0}$  given by Theorem 3.28 is then continuous.

*Proof.* If  $K$  is a compact set,  $n \in \mathbb{N}$ ,  $n \geq 1$  and  $\varepsilon > 0$ , we define

$$A(n, K, \varepsilon) = \left\{ \omega, \max_{0 \leq k \leq n-1} \left\| X_{\frac{k+1}{n}}(\omega) - X_{\frac{k}{n}}(\omega) \right\| > \varepsilon, \forall s \in [0, 1], X_s(\omega) \in K \right\}.$$

By using the Markov property, it is easily checked that

$$\mathbb{P}(A(n, K, \varepsilon)) \leq n \sup_{x \in K} P_{1/n}(x, {}^c \mathbf{B}(x, \varepsilon)).$$

As a consequence, if

$$\limsup_{t \rightarrow 0} \sup_{x \in K} \frac{1}{t} P_t(x, {}^c \mathbf{B}(x, \varepsilon)) = 0,$$

we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(A(n, K, \varepsilon)) = 0.$$

Since from Exercise 3.29 the set  $\{X_t(\omega), 0 \leq t \leq 1\}$  is almost surely bounded and thus contained in a compact set, we deduce from the Borel-Cantelli lemma that with probability 1, the set of  $n$ 's such that

$$\max_{0 \leq k \leq n-1} \left\| X_{\frac{k+1}{n}}(\omega) - X_{\frac{k}{n}}(\omega) \right\| > \varepsilon,$$

is finite. This implies that  $(X_t)_{0 \leq t \leq 1}$  is continuous.  $\square$

A fundamental property of Feller-Dynkin semigroups is that they admit generators.

**Proposition 3.31.** *Let  $(\mathbf{P}_t)_{t \geq 0}$  be a Feller-Dynkin semigroup. There exists a densely defined operator*

$$L : \mathcal{D}(L) \subset \mathcal{C}_0(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$$

where

$$\mathcal{D}(L) = \left\{ f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R}), \quad \lim_{t \rightarrow 0} \frac{\mathbf{P}_t f - f}{t} \text{ exists} \right\},$$

such that for  $f \in \mathcal{D}(L)$ ,

$$\lim_{t \rightarrow 0} \left\| \frac{\mathbf{P}_t f - f}{t} - Lf \right\|_{\infty} = 0.$$

The operator  $L$  is called the generator of the semigroup  $(\mathbf{P}_t)_{t \geq 0}$ . We also say that  $L$  generates  $(\mathbf{P}_t)_{t \geq 0}$ .

*Proof.* Let us consider the following bounded operators on  $\mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$  :

$$A_t = \frac{1}{t} \int_0^t \mathbf{P}_s ds.$$

For  $f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$  and  $h > 0$ , we have

$$\begin{aligned} \frac{1}{t} (\mathbf{P}_t A_h f - A_h f) &= \frac{1}{ht} \int_0^h (\mathbf{P}_{s+t} f - \mathbf{P}_s f) ds \\ &= \frac{1}{ht} \int_0^t (\mathbf{P}_{s+h} f - \mathbf{P}_s f) ds. \end{aligned}$$

Therefore, we obtain

$$\lim_{t \rightarrow 0} \frac{1}{t} (\mathbf{P}_t A_h f - A_h f) = \frac{1}{h} (\mathbf{P}_h f - f).$$

This implies,

$$\{A_h f, f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R}), h > 0\} \subset \left\{ f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R}), \quad \lim_{t \rightarrow 0} \frac{\mathbf{P}_t f - f}{t} \text{ exists} \right\}$$

Since  $\lim_{h \rightarrow 0} A_h f = f$ , we deduce that

$$\left\{ f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R}), \quad \lim_{t \rightarrow 0} \frac{\mathbf{P}_t f - f}{t} \text{ exists} \right\}$$

is dense in  $\mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$ . We can then consider

$$Lf = \lim_{t \rightarrow 0} \frac{\mathbf{P}_t f - f}{t},$$

which is of course defined on the domain

$$\mathcal{D}(L) = \left\{ f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R}), \quad \lim_{t \rightarrow 0} \frac{\mathbf{P}_t f - f}{t} \text{ exists} \right\}.$$

□

If  $(X_t)_{t \geq 0}$  is a Feller-Dynkin process, the generator of its semigroup is also called the generator of the Markov process  $(X_t)_{t \geq 0}$ .

**Exercise 3.32.** Let  $(B_t)_{t \geq 0}$  be a  $n$ -dimensional standard Brownian motion. Show that  $(B_t)_{t \geq 0}$  is a Feller-Dynkin process, that the domain of its generator  $L$  contains  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  and that for  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,

$$Lf = \frac{1}{2} \Delta f.$$

An interesting sub-class of Feller-Dynkin processes is the class of Feller-Dynkin diffusion processes. Let us recall that we denote by  $\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}^n)$  the space of continuous functions  $\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^n$ . As usual  $(\pi_t)_{t \geq 0}$  will denote the coordinate process on this path space and :

$$\mathcal{G}_t = \sigma(\pi_s, 0 \leq s \leq t), \quad t \geq 0, \quad \mathcal{G}_\infty = \sigma(\pi_s, s \geq 0).$$

**Definition 3.33** (Diffusion process).

- Let  $(\mathbf{P}_t)_{t \geq 0}$  be a Feller-Dynkin semigroup. We say that  $(\mathbf{P}_t)_{t \geq 0}$  is a (Feller-Dynkin) diffusion semigroup, if the domain of its generator contains  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  and if for every probability measure  $\nu$  on  $\mathbb{R}^n$ , there exists a probability measure  $\mathbb{P}^\nu$  on  $\mathcal{G}_\infty$  such that:
  1. The distribution of  $\pi_0$  under  $\mathbb{P}^\nu$  is  $\nu$ ;
  2. On the filtered probability space  $(\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}^n), (\mathcal{G}_t)_{t \geq 0}, \mathcal{G}_\infty, \mathbb{P}^\nu)$ ,  $(\pi_t)_{t \geq 0}$  is a Markov process with semigroup  $(\mathbf{P}_t)_{t \geq 0}$ .
- A continuous Markov process  $(X_t)_{t \geq 0}$  is said to be a (Feller-Dynkin) diffusion process if its semigroup is a Feller-Dynkin diffusion semigroup.

As a consequence of Theorems 3.28 and 3.30, it is immediate that if  $(\mathbf{P}_t)_{t \geq 0}$  is a Feller-Dynkin semigroup such that

$$\limsup_{t \rightarrow 0} \sup_{x \in K} \frac{1}{t} P_t(x, \mathbf{B}(x, \varepsilon)^c) = 0,$$

for every  $\varepsilon > 0$  and every compact set  $K$ , then  $(\mathbf{P}_t)_{t \geq 0}$  is a diffusion semigroup.

**Exercise 3.34** (Brownian motion with drift). Let  $(B_t)_{t \geq 0}$  be a Brownian motion on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . For  $\mu \in \mathbb{R}$ , show that the process  $(B_t + \mu t)_{t \geq 0}$  is diffusion process with infinitesimal generator  $L$  such that for  $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$ ,

$$Lf = \mu \frac{df}{dx} + \frac{1}{2} \frac{d^2 f}{dx^2}.$$

**Exercise 3.35** (Ornstein-Uhlenbeck process). Let  $(B_t)_{t \geq 0}$  be a Brownian motion on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . Let  $\theta \in \mathbb{R} \setminus \{0\}$  and consider the process

$$X_t = e^{\theta t} B_{\frac{1-e^{-2\theta t}}{2\theta}}.$$

Show that  $(X_t)_{t \geq 0}$  is a diffusion with infinitesimal generator  $L$  such that for  $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$ ,

$$Lf = \theta x \frac{df}{dx} + \frac{1}{2} \frac{d^2 f}{dx^2}.$$

Diffusion processes admits canonical martingales.

**Proposition 3.36.** Let  $(X_t)_{t \geq 0}$  be a diffusion process defined on the probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . Let us denote by  $\{P_t, t \geq 0\}$  its transition function and by  $L$  its generator. For  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , the process

$$\left( f(X_t) - \int_0^t (Lf)(X_s) ds \right)_{t \geq 0}$$

is a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .

*Proof.* For  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  and  $t \geq 0$ , we have

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathbf{P}_{t+\varepsilon} f - \mathbf{P}_t f}{\varepsilon} = \mathbf{P}_t \left( \lim_{\varepsilon \rightarrow 0} \frac{\mathbf{P}_\varepsilon f - f}{\varepsilon} \right) = \mathbf{P}_t Lf.$$

Thus, we get

$$\mathbf{P}_t f = f + \int_0^t \mathbf{P}_u Lf du.$$

This yields:

$$\begin{aligned} \mathbb{E}(f(X_t) | \mathcal{F}_s) &= (\mathbf{P}_{t-s} f)(X_s) \\ &= f(X_s) + \int_0^{t-s} (\mathbf{P}_u Lf)(X_s) du \\ &= f(X_s) + \int_s^t (\mathbf{P}_{u-s} Lf)(X_s) du \\ &= f(X_s) + \int_s^t \mathbb{E}((Lf)(X_u) | \mathcal{F}_s) \\ &= f(X_s) + \mathbb{E} \left( \int_s^t (Lf)(X_u) du | \mathcal{F}_s \right). \end{aligned}$$

□

The following very nice theorem which is due to Dynkin states that infinitesimal generators of diffusion semigroups need to be second order differential operators.



**Theorem 3.37** (Dynkin theorem). *Let  $(X_t)_{t \geq 0}$  be a  $n$ -dimensional diffusion process with generator  $L$ . There exist continuous functions  $b : \mathbb{R}^n \rightarrow \mathbb{R}$  and  $a_{ij} : \mathbb{R}^n \rightarrow \mathbb{R}$ , such that the matrix  $(a_{ij}(x))_{1 \leq i, j \leq n}$  is semidefinite non negative and such that for  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ :*

$$Lf = \sum_{i=1}^n b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i, j=1}^n a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

*Proof.* We make the proof in dimension  $n = 1$  and let the reader extend it as an exercise in higher dimension. Let  $(X_t)_{t \geq 0}$  be a one-dimensional diffusion process with generator  $L$  which is defined on a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . Our strategy will be to prove that  $L$  needs to satisfy the following three properties

1.  $L : \mathcal{C}_c(\mathbb{R}, \mathbb{R}) \rightarrow \mathcal{C}_0(\mathbb{R}, \mathbb{R})$  is a linear operator;
2.  $L$  is a local operator, i.e. if  $f, g \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$  agree on a neighborhood of  $x \in \mathbb{R}$ , then  $(Lf)(x) = (Lg)(x)$ ;
3.  $L$  satisfies the positive maximum principle: If  $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$  attains a maximum at  $x \in \mathbb{R}$  with  $f(x) \geq 0$ , then  $(Lf)(x) \leq 0$ .

Then we will show that the only operators that satisfy the three above properties are second order differential operators.

The linearity of  $L$  is obvious so let us focus on the local property of  $L$ .

Let  $f, g \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$  that agree on a neighborhood of  $x \in \mathbb{R}$ . We have

$$(\mathbf{P}_t f)(x) = \mathbb{E}^x(f(\pi_t)),$$

where  $\mathbb{E}^x$  is the expectation under the probability measure  $\mathbb{P}^x$  such that:

- Under  $\mathbb{P}^x$ , the distribution of  $\pi_0$  is the Dirac mass at  $x$ ;
- On the probability space  $(\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}), (\mathcal{G}_t)_{t \geq 0}, \mathcal{G}_{\infty}, \mathbb{P}^x)$ ,  $(\pi_t)_{t \geq 0}$  is a Markov process with transition function  $\{P_t, t \geq 0\}$ .

We also have,

$$(\mathbf{P}_t g)(x) = \mathbb{E}^x(g(\pi_t)).$$

Since  $(\pi_t)_{t \geq 0}$  is a continuous process, we deduce that there is a positive and finite,  $\mathbb{P}^x$  almost surely, stopping time  $T$ , such that

$$f(\pi_t) = g(\pi_t), \quad t < T.$$

This implies

$$\lim_{t \rightarrow 0} \frac{\mathbb{E}^x(f(\pi_t)) - \mathbb{E}^x(g(\pi_t))}{t} = \lim_{t \rightarrow 0} \frac{\mathbb{E}^x(f(1_{t < T} \pi_t)) - \mathbb{E}^x(g(1_{t < T} \pi_t))}{t} = 0.$$

On the other hand, we have

$$\lim_{t \rightarrow 0} \frac{\mathbb{E}^x(f(\pi_t)) - \mathbb{E}^x(g(\pi_t))}{t} = (Lf)(x) - (Lg)(x).$$

We deduce that

$$(Lf)(x) = (Lg)(x),$$

so that  $L$  is indeed a local operator.

Let us now show that  $L$  satisfies the positive maximum principle. Let  $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$  that attains a maximum at  $x \in \mathbb{R}$  with  $f(x) \geq 0$ . As before, let  $\mathbb{P}^x$  be the probability measure such that:

- Under  $\mathbb{P}^x$ , the distribution of  $\pi_0$  is the Dirac mass at  $x$ ;
- On the probability space  $(\mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R}), (\mathcal{G}_t)_{t \geq 0}, \mathcal{G}_\infty, \mathbb{P}^\nu)$ ,  $(\pi_t)_{t \geq 0}$  is a Markov process with transition function  $\{P_t, t \geq 0\}$ .

From the previous proposition, under  $\mathbb{P}^x$ , the process

$$\left( f(\pi_t) - \int_0^t (Lf)(\pi_s) ds \right)_{t \geq 0}$$

is a martingale with respect to the filtration  $(\mathcal{G}_t)_{t \geq 0}$ . Therefore, for every  $t \geq 0$ ,

$$\mathbb{E}^x(f(\pi_t)) = f(x) + \int_0^t \mathbb{E}^x((Lf)(\pi_u)) du.$$

Since for every  $t \geq 0$ ,

$$\mathbb{E}^x(f(\pi_t)) \leq f(x),$$

we deduce that

$$\frac{1}{t} \int_0^t \mathbb{E}^x((Lf)(\pi_u)) du \leq 0.$$

Letting  $t \rightarrow 0$  yields:

$$(Lf)(x) \leq 0.$$

As a conclusion,  $L$  is linear and local operator that satisfies the positive maximum principle. Let us now show that such operators are second order differential operators. Let  $x \in \mathbb{R}$ . Let  $\psi_0$  be a compactly supported  $C^\infty$  function such that in a neighborhood of  $x$ ,  $\psi_0 = 1$ . Since  $P_t 1 = 1$ , it is easy to deduce from the local property of  $L$  that:

$$(L\psi_0)(x) = 0.$$

Let now  $\psi_1$  be a compactly supported  $C^\infty$  function such that in a neighborhood of  $x$ ,  $\psi_1(y) = y - x$ . Let us denote

$$b(x) = (L\psi_1)(x).$$

The function  $\psi_1^2$  attains a local minimum in  $x$ , therefore

$$(L\psi_1^2)(x) \geq 0,$$

and we can define  $\sigma$  such that

$$\sigma^2(x) = (L\psi_1^2)(x).$$

These functions  $b$  and  $\sigma$  are well defined and seen to be continuous. Let us now prove that for every  $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$ ,

$$(Lf)(x) = b(x)f'(x) + \frac{1}{2}\sigma(x)^2f''(x).$$

Let  $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$ . From the Taylor expansion formula, we can write in a neighborhood of  $x$ ,

$$f(y) = f(x)\psi_0(y) + f'(x)\psi_1(y) + \frac{1}{2}f''(x)\psi_1^2(y) + R(y)\psi_1^3(y)$$

where  $R$  is a continuous function.

We therefore have

$$(Lf)(x) = f(x)(L\psi_0)(x) + f'(x)L\psi_1(x) + \frac{1}{2}f''(x)(L\psi_1^2)(x) + (LR\psi_1^3)(x).$$

Since we already know that  $L\psi_0 = 0$ , it remains to prove that

$$(LR\psi_1^3)(x) = 0.$$

For  $\varepsilon > 0$  which is small enough,

$$y \rightarrow R(y)\psi_1^3(y) - \varepsilon(y - x)^2$$

has a local maximum in  $x$ , thus

$$(LR\psi_1^3)(x) \leq \varepsilon\sigma^2(x).$$

By letting  $\varepsilon \rightarrow 0$ , we get therefore

$$(LR\psi_1^3)(x) \leq 0.$$

In the very same way, considering the function

$$y \rightarrow R(y)\psi_1^3(y) + \varepsilon(y - x)^2,$$

we obtain that

$$(LR\psi_1^3)(x) \geq 0.$$

As a consequence we established that

$$(LR\psi_1^3)(x) = 0,$$

which concludes the proof of the theorem. □

**Exercise 3.38.** Let  $L$  be the generator of a Feller-Dynkin semigroup. We assume that  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R}) \subset \mathcal{D}(L)$ . Show that  $L$  satisfies the positive maximum principle: If  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  attains a maximum at  $x \in \mathbb{R}^n$  with  $f(x) \geq 0$ , then  $(Lf)(x) \leq 0$ .

According to the previous exercise, if  $L$  is the domain of a Feller-Dynkin process and if  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R}) \subset \mathcal{D}(L)$ , then  $L$  satisfies the positive maximum principle (see the previous exercise for the definition). Operators satisfying the positive maximum principle have been classified by Courrège in a theorem generalizing Dynkin's and that we mention without proof. The interested reader will find the proof in the book by Jacob [?].

In the following statement  $\mathcal{B}(\mathbb{R}^n)$  denotes the set of Borel sets on  $\mathbb{R}^n$  and a kernel  $\mu$  on  $\mathbb{R}^n \times \mathcal{B}(\mathbb{R}^n)$  is a family  $\{\mu(x, \cdot), x \in \mathbb{R}^n\}$  of Borel measures.

**Theorem 3.39** (Courrège theorem). Let  $(X_t)_{t \geq 0}$  be a Feller-Dynkin process with generator  $L$ . If  $\mathcal{D}(L)$  contains  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , then there exist a symmetric and non negative matrix  $(\sigma_{ij}(x))_{1 \leq i, j \leq n}$ , functions  $b_i, c : \mathbb{R}^n \rightarrow \mathbb{R}$ , with  $c \geq 0$  and a kernel  $\mu$  on  $\mathbb{R}^n \times \mathcal{B}(\mathbb{R}^n)$  such that for every  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  and  $x \in \mathbb{R}^n$ ,

$$Lf(x) = \sum_{i,j=1}^n \sigma_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial f}{\partial x_i} - c(x)f(x) \\ + \int_{\mathbb{R}^n} \left( f(y) - \chi(y-x)f(x) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) \chi(y-x)(y_j - x_j) \right) \mu(x, dy),$$

where  $\chi \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,  $0 \leq \chi \leq 1$ , takes the constant value 1 on the ball  $\mathbf{B}(0, 1)$ . In addition, the functions  $b_j$  and  $c$  are continuous and for every  $y \in \mathbb{R}^n$ , the function  $x \rightarrow \sum_{i,j} \sigma_{ij}(x) y_i y_j$  is upper semicontinuous.

## 4 Lévy processes

Lévy processes form a fundamental class of Feller-Dynkin processes. They are the Markov processes associated with the so-called convolution semigroups.

In what follows, we consider a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ .

**Definition 3.40** (Lévy process). Let  $(X_t)_{t \geq 0}$  be a stochastic process. It is said that  $(X_t)_{t \geq 0}$  is a Lévy process on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  if the following conditions are fulfilled:

1. Almost surely  $X_0 = 0$ ;
2. The paths of  $(X_t)_{t \geq 0}$  are left limited and right continuous;
3.  $(X_t)_{t \geq 0}$  is adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ ;

4. For every  $T \geq 0$ , the process  $(X_{t+T} - X_T)_{t \geq 0}$  is independent of the  $\sigma$ -algebra  $\mathcal{F}_T$ ;
5. For every  $t, T \geq 0$ ,  $X_{t+T} - X_T$  has the same distribution as  $X_t$ .

**Remark 3.41.** Of course, the notion of Lévy process in  $\mathbb{R}^n$  is similarly defined.

**Exercise 3.42.** Show that if  $(X_t)_{t \geq 0}$  is a Lévy process defined on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ , then it is also a Lévy process on the space  $(\Omega, (\mathcal{F}_t^X)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ , where  $(\mathcal{F}_t^X)_{t \geq 0}$  is the natural filtration of  $(X_t)_{t \geq 0}$

The two following exercises provide fundamental examples of Lévy processes.

**Exercise 3.43.** (Brownian motion) Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. Show that it is a Lévy process with respect to its natural filtration.

**Exercise 3.44** (Compound Poisson process). Let  $(T_n)_{n \in \mathbb{N}}$  be an i.i.d. sequence of exponential random variables with parameter  $\lambda$ :

$$\mathbb{P}(T_n \in dt) = \lambda e^{-\lambda t} dt, t \geq 0, n \in \mathbb{N},$$

defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . We denote

$$S_n = \sum_{i=1}^n T_i, \quad n \geq 1,$$

and  $S_0 = 0$ . For  $t \geq 0$ , let

$$N_t = \max\{n \geq 0, S_n \leq t\}.$$

1. Show that the process  $(N_t)_{t \geq 0}$  is a Lévy process on the space  $(\Omega, (\mathcal{F}_t^N)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ , where  $(\mathcal{F}_t^N)_{t \geq 0}$  is the natural filtration of  $(N_t)_{t \geq 0}$ .
2. Show that for  $t \geq 0$ , the random variable  $N_t$  is distributed as a Poisson random variable with parameter  $\lambda t$ , that is:

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, n \in \mathbb{N}.$$

The process  $(N_t)_{t \geq 0}$  is called a Poisson process with intensity  $\lambda > 0$ .

3. Let now  $(Y_n)_{n \geq 0}$  be an i.i.d. sequence of random variables with distribution  $\mu$  and independent from  $(N_t)_{t \geq 0}$ . We define for  $t \geq 0$ ,

$$X_t = 1_{N_t \geq 1} \left( \sum_{i=1}^{N_t} Y_i \right).$$

Show that the process  $(X_t)_{t \geq 0}$  is a Lévy process on the space  $(\Omega, (\mathcal{F}_t^X)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ , where  $(\mathcal{F}_t^X)_{t \geq 0}$  is the natural filtration of  $(X_t)_{t \geq 0}$ .

4. Show that for  $\theta \in \mathbb{R}$  and  $t \geq 0$ :

$$\mathbb{E}(e^{i\theta X_t}) = \exp\left(t\lambda \int_{\mathbb{R}} (e^{i\theta x} - 1)\mu(dx)\right).$$

The process  $(X_t)_{t \geq 0}$  is called a compound Poisson process.

In the study of Lévy processes, one of the most important tools is the Lévy-Khinchin theorem that we remind below. We first recall the following basic definition.

**Definition 3.45.** A random variable  $X$  is said to be infinitely divisible if for every  $n \geq 1$ , we may find independent and identically distributed random variables  $X_1, \dots, X_n$  such that in distribution

$$X_1 + \dots + X_n \stackrel{\text{law}}{=} X.$$

The Lévy-Khinchin theorem, which is admitted here, completely characterizes infinitely divisible random variables in terms of their characteristic function. A proof using Lévy processes may be found in the book by Applebaum [?].

**Theorem 3.46** (Lévy-Khinchin theorem). *Let  $X$  be an infinitely divisible random variable. There exist  $\mu \in \mathbb{R}$ ,  $\sigma > 0$  and a Borel measure  $\nu$  on  $\mathbb{R} \setminus \{0\}$  such that*

$$\int_{\mathbb{R}} (1 \wedge x^2)\nu(dx) < +\infty,$$

$$\forall \lambda \in \mathbb{R}, \quad \mathbb{E}(e^{i\lambda X}) = \exp\left(i\lambda\mu - \frac{1}{2}\sigma^2\lambda^2 + \int_{\mathbb{R}} (e^{i\lambda x} - 1 - i\lambda x 1_{|x| \leq 1})\nu(dx)\right).$$

Conversely, let  $\mu \in \mathbb{R}$ ,  $\sigma > 0$  and  $\nu$  be a Borel measure on  $\mathbb{R} \setminus \{0\}$  such that

$$\int_{\mathbb{R}} (1 \wedge x^2)\nu(dx) < +\infty.$$

The function:

$$\lambda \rightarrow \exp\left(i\lambda\mu - \frac{1}{2}\sigma^2\lambda^2 + \int_{\mathbb{R}} (e^{i\lambda x} - 1 - i\lambda x 1_{|x| \leq 1})\nu(dx)\right)$$

is the characteristic function of the distribution of an infinitely divisible random variable.

In what follows, we consider a Lévy process  $(X_t)_{t \geq 0}$ . Since for every  $n \geq 1$ ,

$$X_1 = \sum_{k=1}^n (X_{\frac{k}{n}} - X_{\frac{k-1}{n}}),$$

and since the increments  $X_{\frac{k}{n}} - X_{\frac{k-1}{n}}$  are independent and identically distributed, we deduce that the random variable  $X_1$  is infinitely divisible. Therefore, from the Lévy-Khinchin theorem, there exist  $\sigma > 0$ ,  $\mu \in \mathbb{R}$  and a Borel measure  $\nu$  on  $\mathbb{R} \setminus \{0\}$  such that:

$$\int_{\mathbb{R}} (1 \wedge x^2) \nu(dx) < +\infty$$

and

$$\mathbb{E}(e^{i\lambda X_1}) = e^{\psi(\lambda)}, \lambda \in \mathbb{R},$$

where

$$\psi(\lambda) = i\mu\lambda - \frac{1}{2}\sigma^2\lambda^2 + \int_{\mathbb{R}} (e^{i\lambda x} - 1 - i\lambda x 1_{|x| \leq 1}) \nu(dx).$$

The function  $\psi$  is called the characteristic exponent of the Lévy process, and the measure  $\nu$  the Lévy measure. For instance, for a Brownian motion, we have

$$\mu = 0, \quad \sigma = 1, \quad \nu = 0,$$

whereas for a compound Poisson process as defined in Exercise 3.44,

$$\mu = - \int_{-1}^1 x \mu(dx), \quad \sigma = 0, \quad \nu(dx) = \mu(dx).$$

The characteristic exponent of a Lévy process characterizes the law of such a process. Indeed, let  $\lambda \in \mathbb{R}$  and let us consider the following function:

$$f(t) = \mathbb{E}(e^{i\lambda X_t}), \quad t \geq 0.$$

Since  $(X_t)_{t \geq 0}$  is a Lévy process, for  $s, t \geq 0$ ,

$$\begin{aligned} f(t+s) &= \mathbb{E}(e^{i\lambda X_{t+s}}) \\ &= \mathbb{E}(e^{i\lambda(X_{t+s}-X_t)+X_t}) \\ &= \mathbb{E}(e^{i\lambda(X_{t+s}-X_t)}) \mathbb{E}(e^{i\lambda X_t}) \\ &= \mathbb{E}(e^{i\lambda X_s}) \mathbb{E}(e^{i\lambda X_t}) \\ &= f(t)f(s). \end{aligned}$$

Moreover, since  $(X_t)_{t \geq 0}$  has right continuous paths at 0, the function  $f$  is itself right continuous at 0. From  $f(1) = e^{\psi(\lambda)}$ , we may then deduce that for  $t \geq 0$ ,

$$f(t) = e^{t\psi(\lambda)}, \quad t \geq 0.$$

We finally conclude that for  $\lambda \in \mathbb{R}$  and  $t \geq 0$ ,

$$\mathbb{E}(e^{i\lambda X_t}) = e^{t\psi(\lambda)}. \tag{3.2}$$

With this in hands we can now turn to the Feller-Dynkin property of Lévy processes and also characterize the Lévy processes as being the Markov processes associated to the convolution semigroups.

**Proposition 3.47.** *Let  $(X_t)_{t \geq 0}$  be a Lévy process. For  $t \geq 0$ , let  $p_t(dx)$  be the distribution of the random variable  $X_t$ . The family  $(p_t(dx))_{t \geq 0}$  is a convolution semigroup of probability measures, that is*

$$p_{t+s} = p_t * p_s, \quad s, t \geq 0.$$

Moreover the process  $(X_t)_{t \geq 0}$  is a Feller-Dynkin process with semigroup

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}} f(x+y) p_t(dy).$$

Conversely, let  $(p_t(dx))_{t \geq 0}$  be a convolution semigroup of probability measures that is right continuous at 0 in the topology of convergence in distribution with  $p_0(dx) = \delta_0$  (Dirac distribution at 0). Then, there exists a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and a Lévy process  $(X_t)_{t \geq 0}$  on it such that the distribution of  $X_t$  is  $p_t(dx)$ .

*Proof.* The family  $(p_t(dx))_{t \geq 0}$  is seen to be a convolution semigroup of probability measures from (3.47), by taking inverse Fourier transform. From this last property, we deduce that the family of operators  $(\mathbf{P}_t)_{t \geq 0}$  defined by

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}} f(x+y) p_t(dy)$$

enjoys the semigroup property. Finally due to the fact that  $p_t(dx)$  is a probability measure, it is easily checked that  $(\mathbf{P}_t)_{t \geq 0}$  is a Feller-Dynkin semigroup.

Let us now prove that it is the semigroup corresponding to the process  $(X_t)_{t \geq 0}$ . Let  $f$  be a bounded and Borel function, we have for  $s, t \geq 0$ ,

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_s) = \mathbb{E}(f((X_{t+s} - X_s) + X_s) | \mathcal{F}_s).$$

But the random variable  $X_{t+s} - X_s$  is independent from  $\mathcal{F}_s$  and distributed as  $p_t(dx)$ , therefore

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_s) = \int_{\mathbb{R}} f(y + X_s) p_t(dy) = \mathbf{P}_t f(X_s).$$

We now turn to the proof of the second part of the proposition. As above, the family of operators  $(\mathbf{P}_t)_{t \geq 0}$  defined by

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}} f(x+y) p_t(dy).$$

defines a Feller-Dynkin semigroup. From Theorem 3.28, there exists a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a stochastic process  $(X_t)_{t \geq 0}$  such that:

- The distribution of  $X_0$  is  $\delta_0$ ;



- $(X_t)_{t \geq 0}$  is a Markov process with transition function  $\{P_t, t \geq 0\}$ .
- The paths of  $(X_t)_{t \geq 0}$  are right continuous and left-limited.

Let us check that this process  $(X_t)_{t \geq 0}$  is a Lévy process with respect to its natural filtration. Let  $\lambda \in \mathbb{R}$  and  $s, t \geq 0$ , we have

$$\begin{aligned} \mathbb{E} \left( e^{i\lambda(X_{t+s} - X_s)} \mid \mathcal{F}_s^X \right) &= e^{-i\lambda X_s} \mathbb{E} \left( e^{i\lambda X_{t+s}} \mid \mathcal{F}_s^X \right) \\ &= e^{-i\lambda X_s} \int_{\mathbb{R}} e^{i\lambda(X_s + y)} p_t(dy) \\ &= \int_{\mathbb{R}} e^{i\lambda y} p_t(dy). \end{aligned}$$

We deduce therefore that if  $f$  is a bounded and Borel function, then

$$\mathbb{E} \left( f(X_{t+s} - X_s) \mid \mathcal{F}_s^X \right) = \mathbf{P}_t f(0).$$

We conclude that, as expected, the process  $(X_t)_{t \geq 0}$  is a Lévy process.  $\square$

Since Lévy processes are Feller-Dynkin processes, according to Proposition 3.31, they have an infinitesimal generator. The next proposition computes the generator in terms of the Lévy measure.

**Proposition 3.48.** *Let  $(X_t)_{t \geq 0}$  be a Lévy process with characteristic exponent*

$$\psi(\lambda) = i\mu\lambda - \frac{1}{2}\sigma^2\lambda^2 + \int_{\mathbb{R}} (e^{i\lambda x} - 1 - i\lambda x 1_{|x| \leq 1}) \nu(dx).$$

*The domain  $\mathcal{D}(L)$  of the infinitesimal generator  $L$  of  $(X_t)_{t \geq 0}$  contains the space  $\mathcal{S}$  of smooth and rapidly decreasing functions and for  $f \in \mathcal{S}$ ,*

$$Lf(x) = \mu f'(x) + \frac{1}{2}\sigma^2 f''(x) + \int_{\mathbb{R}} (f(x+y) - f(x) - yf'(x) 1_{|y| \leq 1}) \nu(dy).$$

*Proof.* Let  $\lambda \in \mathbb{R}$  and denote

$$e_\lambda(x) = e^{i\lambda x}.$$

We have

$$\mathbf{P}_t e_\lambda(x) = \int_{\mathbb{R}} e^{i\lambda(x+y)} p_t(dy) = e_\lambda(x) e^{t\psi(\lambda)}.$$

Therefore, we obtain

$$\lim_{t \rightarrow 0} \frac{\mathbf{P}_t e_\lambda - e_\lambda}{t} = \psi(\lambda) e_\lambda.$$

This last equality proves the proposition by using the inverse Fourier transform that maps  $\mathcal{S}$  into itself.  $\square$

**Exercise 3.49.** Let  $(X_t)_{t \geq 0}$  be a Lévy process in  $\mathbb{R}^n$ . Show that  $(X_t)_{t \geq 0}$  is a Feller-Dynkin process whose generator can be written on the space of rapidly decreasing functions

$$Lf(x) = \langle \mu, \nabla f(x) \rangle + \frac{1}{2} \sum_{i,j=1}^n a_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \int_{\mathbb{R}^n} (f(x+y) - f(x) - \langle y, \nabla f(x) \rangle) 1_{\|y\| \leq 1} \nu(dy),$$

where  $\mu \in \mathbb{R}^n$ ,  $(a_{ij})_{0 \leq i,j \leq n}$  is a non negative symmetric matrix and  $\nu$  is a Borel measure on  $\mathbb{R}^n - \{0\}$ .

An important class of Lévy processes is the class of subordinators.

**Definition 3.50** (Subordinators). A non-negative Lévy process  $(S_t)_{t \geq 0}$  is called a subordinator.

For instance the Poisson process (see Exercise 3.44) is a subordinator. Since for Lévy processes the following identity in distribution holds:  $S_t - S_s \stackrel{law}{=} S_{t-s}$ ,  $t \geq s$ , we deduce that the paths  $t \rightarrow S_t$  of a subordinator are almost surely non decreasing. If  $(S_t)_{t \geq 0}$  is a subordinator, then there exists a function  $\Psi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  such that for every  $t, \lambda \geq 0$ ,

$$\mathbb{E}(e^{-\lambda S_t}) = e^{-t\Psi(\lambda)}. \quad (3.3)$$

Actually, we have  $\Psi(\lambda) = \psi(i\lambda)$ , where  $\psi$  is the analytic extension of the characteristic exponent of  $(S_t)_{t \geq 0}$  on the upper half plane of the complex plane. From the representation (3.3), we can see that  $\Psi$  needs to be a Bernstein function, that is for every  $\lambda > 0$  and  $k \geq 1$ ,

$$(-1)^k \Psi^{(k)}(\lambda) \leq 0.$$

Bernstein functions are characterized by the following well-know theorem of Bernstein (see [?] for a proof).

**Theorem 3.51** (Bernstein theorem). A smooth function  $f : (0, +\infty) \rightarrow \mathbb{R}_{\geq 0}$  is completely monotone, that is for every  $\lambda > 0$  and  $k \geq 0$ ,

$$(-1)^k f^{(k)}(\lambda) \geq 0,$$

if and only if there exists a non negative Borel measure  $m$  on  $[0, +\infty)$  such that

$$f(\lambda) = \int_0^{+\infty} e^{-\lambda z} m(dz), \quad \lambda > 0.$$

As consequence of the Bernstein theorem, any Bernstein function  $\Psi$  such that  $\Psi(0) = 0$  can be written as

$$\Psi(\lambda) = a\lambda + \int_0^{+\infty} (1 - e^{-\lambda z})\mu(dz),$$

where  $a \geq 0$  and where  $\mu$  is a Borel measure on  $(0, +\infty)$  such that

$$\int_0^{+\infty} (1 \wedge z)\mu(dz) < +\infty.$$

Conversely, given any Bernstein function  $\Psi$ , we can find a corresponding subordinator  $(S_t)_{t \geq 0}$ . Indeed, let  $\Psi$  be a Bernstein function such that  $\Psi(0) = 0$ , then for every  $t > 0$ , the function  $\lambda \rightarrow e^{-t\Psi(\lambda)}$  is completely monotone and, as a consequence of the Bernstein theorem, we can find a probability measure  $m_t$  on such that

$$\int_0^{+\infty} e^{-\lambda z} m_t(dz) = e^{-t\Psi(\lambda)}.$$

It is then seen that  $(m_t)_{t \geq 0}$  is a convolution semigroup of probability measure. As a consequence of Proposition 3.47 we can thus find a Lévy process associated to  $m_t$ . Since  $m_t$  is supported on  $[0, +\infty)$ , this process needs to be a subordinator. As a conclusion, there is a one to one correspondence between subordinators and Bernstein functions.

**Exercise 3.52** ( $\alpha$ -stable subordinators). *Let  $0 < \alpha < 1$ .*

1. *Show that  $\Psi_\alpha(\lambda) = \lambda^\alpha$  is a Bernstein function.*
2. *A subordinator  $(S_t^\alpha)_{t \geq 0}$  associated to the Bernstein function  $\Psi_\alpha$  is called an  $\alpha$ -stable subordinator. Show that for  $c > 0$ , the processes  $(S_{ct}^\alpha)_{t \geq 0}$  and  $(c^\alpha S_t^\alpha)_{t \geq 0}$  have the same distribution.*
3. *Show the subordination identity:*

$$e^{-y|\alpha|} = \frac{y}{2\sqrt{\pi}} \int_0^{+\infty} \frac{e^{-\frac{y^2}{4t} - t\alpha^2}}{t^{3/2}} dt, \quad y > 0, \alpha \in \mathbb{R},$$

*and deduce the distribution of  $S_t^{1/2}$ ,  $t > 0$ . We mention that there is no such simple expression for the distribution of  $S_t^\alpha$ ,  $\alpha \neq 1/2$ .*

**Exercise 3.53** (Subordinated Brownian motion). *Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion in  $\mathbb{R}^n$  and let  $(S_t)_{t \geq 0}$  be an independent subordinator. The process*

$$X_t = B_{S_t}, \quad t \geq 0,$$

*is called a subordinated Brownian motion.*

1. Show that  $(X_t)_{t \geq 0}$  is a Lévy process with characteristic exponent  $\Psi\left(\frac{\|\lambda\|^2}{2}\right)$ , where  $\Psi$  is the Bernstein function of the subordinator.
2. Let  $L$  be the generator of  $(X_t)_{t \geq 0}$ . Compute the Fourier transform of  $Lf$  in terms of the Fourier transform of  $f$  when  $f$  is a smooth and rapidly decreasing function.
3. Let  $m_t$  be the probability distribution of  $S_t$  at time  $t > 0$  show that the distribution of  $X_t$  has a density which is given by

$$\int_0^{+\infty} \frac{e^{-\frac{\|x\|^2}{2s}}}{(2\pi s)^{n/2}} m_t(ds)$$

if the integral is convergent. Compute this density if  $(S_t)_{t \geq 0}$  is a  $1/2$ -stable subordinator (see Exercise 3.52).

The following theorem which finishes the section characterizes the continuous Lévy processes. It is remarkable that the only such Lévy processes are, up to a renormalization factor, the Brownian motions with drift.

**Theorem 3.54** (Lévy theorem). *Let  $(X_t)_{t \geq 0}$  be a continuous Lévy process defined on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . Then, the Lévy measure of  $(X_t)_{t \geq 0}$  is 0, and therefore  $(X_t)_{t \geq 0}$  may be written*

$$X_t = \sigma B_t + \mu t,$$

where  $(B_t)_{t \geq 0}$  is a standard Brownian motion.

*Proof.* Let

$$\psi(\lambda) = i\mu\lambda - \frac{1}{2}\sigma^2\lambda^2 + \int_{\mathbb{R}} (e^{i\lambda x} - 1 - i\lambda x 1_{|x| \leq 1}) \nu(dx)$$

be the characteristic exponent of  $(X_t)_{t \geq 0}$ . For  $\varepsilon \in (0, 1)$ , we have

$$\psi = \psi_\varepsilon + \phi_\varepsilon,$$

where

$$\psi_\varepsilon(\lambda) = i\mu\lambda - \frac{1}{2}\sigma^2\lambda^2 + \int_{|x| \leq \varepsilon} (e^{i\lambda x} - 1 - i\lambda x) \nu(dx),$$

and

$$\phi_\varepsilon(\lambda) = \int_{|x| > \varepsilon} (e^{i\lambda x} - 1 - i\lambda x 1_{|x| \leq 1}) \nu(dx).$$

This decomposition of  $\psi$  will actually correspond to a pathwise decomposition of  $(X_t)_{t \geq 0}$ .

For  $t \geq 0$ , let  $\mu_t$  be the probability measure on  $\mathbb{R}$  with characteristic function:

$$\int_{\mathbb{R}} e^{i\lambda x} \mu_t(dx) = e^{t\psi_\varepsilon(\lambda)}, \lambda \in \mathbb{R}.$$

We have for  $s, t \geq 0$

$$\mu_t * \mu_s = \mu_{t+s}.$$

From Proposition 3.47, we therefore can find a probability space  $(\tilde{\Omega}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  and a process  $(Y_t)_{t \geq 0}$  that is a Lévy process with characteristics exponent  $\psi_\varepsilon(\lambda)$ .

In the very same way, we may, by enlarging the filtered probability space  $(\tilde{\Omega}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ , construct a Lévy process  $(Z_t)_{t \geq 0}$  that is independent of  $(Y_t)_{t \geq 0}$  and whose characteristic exponent is  $\phi_\varepsilon$ . We have

$$\begin{aligned} \phi_\varepsilon(\lambda) &= \int_{|x| > \varepsilon} (e^{i\lambda x} - 1 - i\lambda x 1_{|x| \leq 1}) \nu(dx) \\ &= \int_{|x| > \varepsilon} (e^{i\lambda x} - 1) \nu(dx) - i\lambda \int_{|x| > \varepsilon} x 1_{|x| \leq 1} \nu(dx), \end{aligned}$$

and

$$\nu(\{x, |x| > \varepsilon\}) < +\infty.$$

Therefore, from Exercise 3.44

$$Z_t = L_t - t \int_{|x| > \varepsilon} x 1_{|x| \leq 1} \nu(dx),$$

where  $(L_t)_{t \geq 0}$  is a compound Lévy process. We deduce that the paths of the process  $(Z_t)_{t \geq 0}$  only have a finite number of jumps within any finite time interval. Moreover the size of these possible has to be larger than  $\varepsilon$ .

Since the two processes  $(Y_t)_{t \geq 0}$  and  $(Z_t)_{t \geq 0}$  are independent, the process

$$\tilde{X}_t = Y_t + Z_t$$

is a Lévy process that has the same distribution as  $(X_t)_{t \geq 0}$ . Also, almost surely, the possible jumps of  $(Y_t)_{t \geq 0}$  and  $(Z_t)_{t \geq 0}$  do not intersect. Thus, every jump of  $(Y_t)_{t \geq 0}$  induces a jump of  $(\tilde{X}_t)_{t \geq 0}$ . Since  $(X_t)_{t \geq 0}$  is assumed to be continuous,  $(Y_t)_{t \geq 0}$  has no jumps. This implies

$$\nu(\{x, |x| > \varepsilon\}) = 0.$$

Since it should hold for every  $\varepsilon$ , we deduce  $\nu = 0$ . □

**Exercise 3.55.** *State and prove a multidimensional extension of Theorem 3.54.*

## Notes and Comments

**Sections 1,2,3.** One of the first major actors of the rigorous and general theory of continuous time Markov processes as presented here, is Dynkin, we refer to his book [?] for an overview of the origins of the theory. The theory of Markov processes and diffusion processes is now very rich and numerous references are available to the interested reader. We wish to mention in particular the books by Bass [?], Chung [?], and Ethier and Kurtz [?]. For the connection between pseudo-differential operators satisfying maximum principle and Markov processes, we refer to the books by Jacob [?] and [?].

**Section 4.** Lévy processes have been widely and extensively studied and several results of Chapter 2 can be extended to Lévy processes. We refer to the books by Applebaum [?], Bertoin [?] or Sato [?] for further details.

## Chapter 4

# Symmetric diffusion semigroups

As we have seen in Theorem 3.37, the generator of a Feller-Dynkin diffusion process is of the form

$$L = \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i} + \sum_{i,j=1}^n \sigma_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j},$$

where  $b_i$  and  $\sigma_{ij}$  are continuous functions on  $\mathbb{R}^n$  such that for every  $x \in \mathbb{R}^n$ , the matrix  $(\sigma_{ij}(x))_{1 \leq i,j \leq n}$  is a symmetric and non negative matrix. Such second order differential operators are generically called diffusion operators.

Conversely, it is often important to know if given a diffusion operator, there is a unique transition function that admits this operator as a generator. This problem is difficult to answer in all generality. As we will discuss it in Chapter 6, the theory of stochastic differential equations provides a fantastic probabilistic tool to study this question (see Theorem 6.15 and Proposition 6.16 in Chapter 6) but the theory of Dirichlet forms and associated symmetric diffusion semigroups that we introduce in this chapter applies in much more general situations. We assume here some basic knowledge about the theory of unbounded operators on Hilbert or Banach spaces and at some points some basic knowledge about the local regularity theory for elliptic operators. Elements of these theories are given in the appendices at the end of the book.

## 1 Essential self-adjointness, Spectral theorem

Throughout the section, we consider a second order differential operator that can be written

$$L = \sum_{i,j=1}^n \sigma_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i},$$

where  $b_i$  and  $\sigma_{ij}$  are continuous functions on  $\mathbb{R}^n$  and for every  $x \in \mathbb{R}^n$ , the matrix  $(\sigma_{ij}(x))_{1 \leq i,j \leq n}$  is a symmetric and non negative matrix. Such operator is called a diffusion operator.

We will assume that there is Borel measure  $\mu$  which is equivalent to the Lebesgue measure and that symmetrizes  $L$  in the sense that for every smooth and compactly supported functions  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\int_{\mathbb{R}^n} g L f d\mu = \int_{\mathbb{R}^n} f L g d\mu.$$

In what follows, as usual, we denote by  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  the set of smooth and compactly supported functions  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

**Exercise 4.1.** Show that if  $\mu$  is a symmetric measure for  $L$  as above, then in the sense of distributions

$$L'\mu = 0,$$

where  $L'$  is the adjoint of  $L$  in distribution sense.

**Exercise 4.2.** Show that if  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a smooth function and if  $g \in C_c(\mathbb{R}^n, \mathbb{R})$ , then we still have the formula,

$$\int_{\mathbb{R}^n} fLg d\mu = \int_{\mathbb{R}^n} gLf d\mu.$$

**Exercise 4.3.** On  $C_c(\mathbb{R}^n, \mathbb{R})$ , let us consider the operator

$$L = \Delta + \langle \nabla U, \nabla \cdot \rangle,$$

where  $U : \mathbb{R}^n \rightarrow \mathbb{R}$  is a  $C^1$  function. Show that  $L$  is symmetric with respect to the measure

$$\mu(dx) = e^{U(x)} dx.$$

**Exercise 4.4** (Divergence form operator). On  $C_c(\mathbb{R}^n, \mathbb{R})$ , let us consider the operator

$$Lf = \mathbf{div}(\sigma \nabla f),$$

where  $\mathbf{div}$  is the divergence operator defined on a  $C^1$  function  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  by

$$\mathbf{div} \phi = \sum_{i=1}^n \frac{\partial \phi_i}{\partial x_i}$$

and where  $\sigma$  is a  $C^1$  field of non negative and symmetric matrices. Show that  $L$  is a diffusion operator which is symmetric with respect to the Lebesgue measure of  $\mathbb{R}^n$ .

For every smooth functions  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ , let us define the so-called *carré du champ*<sup>1</sup>, which is the symmetric first-order differential form defined by:

$$\Gamma(f, g) = \frac{1}{2} (L(fg) - fLg - gLf).$$

A straightforward computation shows that

$$\Gamma(f, g) = \sum_{i,j=1}^n \sigma_{ij}(x) \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j},$$

so that for every smooth function  $f$ ,

$$\Gamma(f, f) \geq 0.$$

---

<sup>1</sup>The litteral translation from French is *square of the field*.



**Exercise 4.5.**

1. Show that if  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$  are  $C^1$  functions and  $\phi_1, \phi_2 : \mathbb{R} \rightarrow \mathbb{R}$  are also  $C^1$  then,

$$\Gamma(\phi_1(f), \phi_2(g)) = \phi_1'(f)\phi_2'(g)\Gamma(f, g).$$

2. Show that if  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a  $C^2$  function and  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  is also  $C^2$ ,

$$L\phi(f) = \phi'(f)Lf + \phi''(f)\Gamma(f, f).$$

In the sequel we shall consider the bilinear form given for  $f, g \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  by

$$\mathcal{E}(f, g) = \int_{\mathbb{R}^n} \Gamma(f, g) d\mu.$$

This is the so-called Dirichlet form associated to  $L$ . It is readily checked that  $\mathcal{E}$  is symmetric:

$$\mathcal{E}(f, g) = \mathcal{E}(g, f),$$

and non negative

$$\mathcal{E}(f, f) \geq 0.$$

We may observe that thanks to symmetry of  $L$ ,

$$\mathcal{E}(f, g) = - \int_{\mathbb{R}^n} fLg d\mu = - \int_{\mathbb{R}^n} gLf d\mu.$$

The operator  $L$  on its domain  $\mathcal{D}(L) = \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  is a densely defined non positive symmetric operator on the Hilbert space  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ . However, it is of course not self-adjoint, indeed Exercise 4.2 easily shows that

$$\left\{ f \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}), \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} + \|Lf\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} < \infty \right\} \subset \mathcal{D}(L^*).$$

A famous theorem of Von Neumann asserts that any non negative and symmetric operator may be extended into a self-adjoint operator. The following construction, due to Friedrichs, provides a canonical non negative self-adjoint extension.

**Theorem 4.6** (Friedrichs extension). *On the Hilbert space  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ , there exists a densely defined non positive self-adjoint extension of  $L$ .*

*Proof.* The idea is to work with a Sobolev type norm associated to the energy form  $\mathcal{E}$ . On  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , let us consider the following norm

$$\|f\|_{\mathcal{E}}^2 = \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \mathcal{E}(f, f).$$

By completing  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  with respect to this norm, we get a Hilbert space  $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{E}})$ . Since for  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,  $\|f\|_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})} \leq \|f\|_{\mathcal{E}}$ , the injection map  $\iota : (\mathcal{C}_c(\mathbb{R}^n, \mathbb{R}), \|\cdot\|_{\mathcal{E}}) \rightarrow (\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R}), \|\cdot\|_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})})$  is continuous and it may therefore be extended into a continuous map  $\bar{\iota} : (\mathcal{H}, \|\cdot\|_{\mathcal{E}}) \rightarrow (\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R}), \|\cdot\|_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})})$ . Let us show that  $\bar{\iota}$  is injective so that  $\mathcal{H}$  may be identified with a subspace of  $\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})$ . So, let  $f \in \mathcal{H}$  such that  $\bar{\iota}(f) = 0$ . We can find a sequence  $f_n \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , such that  $\|f_n - f\|_{\mathcal{E}} \rightarrow 0$  and  $\|f_n\|_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})} \rightarrow 0$ . We have

$$\begin{aligned} \|f\|_{\mathcal{E}} &= \lim_{m, n \rightarrow +\infty} \langle f_n, f_m \rangle_{\mathcal{E}} \\ &= \lim_{m \rightarrow +\infty} \lim_{n \rightarrow +\infty} \langle f_n, f_m \rangle_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})} + \mathcal{E}(f_n, f_m) \\ &= \lim_{m \rightarrow +\infty} \lim_{n \rightarrow +\infty} \langle f_n, f_m \rangle_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})} - \langle f_n, Lf_m \rangle_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})} \\ &= 0, \end{aligned}$$

thus  $f = 0$  and  $\bar{\iota}$  is injective. Let us now consider the map

$$B = \bar{\iota} \cdot \bar{\iota}^* : \mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R}).$$

It is well defined due to the fact that since  $\bar{\iota}$  is bounded, it is easily checked that

$$\mathcal{D}(\bar{\iota}^*) = \mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R}).$$

Moreover,  $B$  is easily seen to be symmetric, and thus self-adjoint because its domain is equal to  $\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})$ . Also, it is readily checked that the injectivity of  $\bar{\iota}$  implies the injectivity of  $B$ . Therefore the inverse

$$A = B^{-1} : \mathcal{R}(\bar{\iota} \cdot \bar{\iota}^*) \subset \mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})$$

is a densely defined self-adjoint operator on  $\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})$  (see the appendix on unbounded operators on Banach spaces). Now, we observe that for  $f, g \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,

$$\begin{aligned} &\langle f, g \rangle_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})} - \langle Lf, g \rangle_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})} \\ &= \langle \bar{\iota}^{-1}(f), \bar{\iota}^{-1}(g) \rangle_{\mathcal{E}} \\ &= \langle (\bar{\iota}^{-1})^* \bar{\iota}^{-1} f, g \rangle_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})} \\ &= \langle (\bar{\iota} \bar{\iota}^*)^{-1} f, g \rangle_{\mathbf{L}_{\mu}^2(\mathbb{R}^n, \mathbb{R})} \end{aligned}$$

Thus  $A$  and  $\mathbf{Id} - L$  coincide on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . By defining,

$$-\bar{L} = A - \mathbf{Id},$$

we get the required self-adjoint extension of  $-L$ .  $\square$

**Remark 4.7.** The operator  $\bar{L}$ , as constructed above, is called the Friedrichs extension of  $L$ . Intuitively, it is the minimal self-adjoint extension of  $L$ .

**Definition 4.8.** If  $\bar{L}$  is the unique non positive self-adjoint extension of  $L$ , then the operator  $L$  is said to be essentially self-adjoint on  $C_c(\mathbb{R}^n, \mathbb{R})$ . In that case, there is no ambiguity and we shall denote  $\bar{L} = L$ .

We have the following first criterion for essential self-adjointness.

**Lemma 4.9.** *If for some  $\lambda > 0$ ,*

$$\mathbf{Ker}(-L^* + \lambda \mathbf{Id}) = \{0\},$$

*then the operator  $L$  is essentially self-adjoint on  $C_c(\mathbb{R}^n, \mathbb{R})$ .*

*Proof.* We make the proof for  $\lambda = 1$  and let the reader adapt it for  $\lambda \neq 0$ .

Let  $-\tilde{L}$  be a non negative self-adjoint extension of  $-L$ . We want to prove that actually,  $-\tilde{L} = -\bar{L}$ . The assumption

$$\mathbf{Ker}(-L^* + \mathbf{Id}) = \{0\}$$

implies that  $C_c(\mathbb{R}^n, \mathbb{R})$  is dense in  $\mathcal{D}(-L^*)$  for the norm

$$\|f\|_{\mathcal{E}}^2 = \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 - \langle f, L^* f \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}.$$

Since,  $-\tilde{L}$  is a non negative self-adjoint extension of  $-L$ , we have

$$\mathcal{D}(-\tilde{L}) \subset \mathcal{D}(-L^*).$$

The space  $C_c(\mathbb{R}^n, \mathbb{R})$  is therefore dense in  $\mathcal{D}(-\tilde{L})$  for the norm  $\|\cdot\|_{\mathcal{E}}$ .

At that point, we use some notations introduced in the proof of the Friedrichs extension (Theorem 4.6). Since  $C_c(\mathbb{R}^n, \mathbb{R})$  is dense in  $\mathcal{D}(-\tilde{L})$  for the norm  $\|\cdot\|_{\mathcal{E}}$ , we deduce that the equality

$$\langle f, g \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} - \langle \tilde{L}f, g \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} = \langle \bar{i}^{-1}(f), \bar{i}^{-1}(g) \rangle_{\mathcal{E}},$$

which is obviously satisfied for  $f, g \in C_c(\mathbb{R}^n, \mathbb{R})$  actually also holds for  $f, g \in \mathcal{D}(\tilde{L})$ . From the definition of the Friedrichs extension, we deduce that  $\bar{L}$  and  $\tilde{L}$  coincide on  $\mathcal{D}(\tilde{L})$ . Finally, since these two operators are self adjoint we conclude  $\bar{L} = \tilde{L}$ .  $\square$

**Remark 4.10.** Given the fact that  $-L$  is given here with the domain  $C_c(\mathbb{R}^n, \mathbb{R})$ , the condition

$$\mathbf{Ker}(-L^* + \lambda \mathbf{Id}) = \{0\},$$

is equivalent to the fact that if  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  is a function that satisfies in the sense of distributions

$$-Lf + \lambda f = 0,$$

then  $f = 0$ .

As a corollary of the previous lemma, the following proposition provides a useful sufficient condition for essential self-adjointness that is easy to check for several diffusion operators. We recall that a diffusion operator is said to be elliptic if the matrix  $\sigma$  is invertible.

**Proposition 4.11.** *If the diffusion operator  $L$  is elliptic with smooth coefficients and if there exists an increasing sequence  $h_n \in C_c(\mathbb{R}^n, \mathbb{R})$ ,  $0 \leq h_n \leq 1$ , such that  $h_n \nearrow 1$  on  $\mathbb{R}^n$ , and  $\|\Gamma(h_n, h_n)\|_\infty \rightarrow 0$ , as  $n \rightarrow \infty$ , then the operator  $L$  is essentially self-adjoint on  $C_c(\mathbb{R}^n, \mathbb{R})$ .*

*Proof.* Let  $\lambda > 0$ . According to the previous lemma, it is enough to check that if  $L^*f = \lambda f$  with  $\lambda > 0$ , then  $f = 0$ . As it was observed above,  $L^*f = \lambda f$  is equivalent to the fact that, in sense of distributions,  $Lf = \lambda f$ . From the hypoellipticity of  $L$ , we deduce therefore that  $f$  is a smooth function. Now, for  $h \in C_c(\mathbb{R}^n, \mathbb{R})$ ,

$$\begin{aligned} \int_{\mathbb{R}^n} \Gamma(f, h^2 f) d\mu &= -\langle f, L(h^2 f) \rangle_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})} \\ &= -\langle L^* f, h^2 f \rangle_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})} \\ &= -\lambda \langle f, h^2 f \rangle_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})} \\ &= -\lambda \langle f^2, h^2 \rangle_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})} \\ &\leq 0. \end{aligned}$$

Since

$$\Gamma(f, h^2 f) = h^2 \Gamma(f, f) + 2fh\Gamma(f, h),$$

we deduce that

$$\langle h^2, \Gamma(f, f) \rangle_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})} + 2\langle fh, \Gamma(f, h) \rangle_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})} \leq 0.$$

Therefore, by Cauchy-Schwarz inequality

$$\langle h^2, \Gamma(f, f) \rangle_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})} \leq 4\|f\|_2^2 \|\Gamma(h, h)\|_\infty.$$

If we now use the sequence  $h_n$  and let  $n \rightarrow \infty$ , we obtain  $\Gamma(f, f) = 0$  and therefore  $f = 0$ , as desired.  $\square$

**Exercise 4.12.** *Let*

$$L = \Delta + \langle \nabla U, \nabla \cdot \rangle,$$

where  $U$  is a smooth function on  $\mathbb{R}^n$ . Show that with respect to the measure  $\mu(dx) = e^{U(x)} dx$ , the operator  $L$  is essentially self-adjoint on  $C_c(\mathbb{R}^n, \mathbb{R})$ .

**Exercise 4.13.** *On  $\mathbb{R}^n$ , we consider the divergence form operator*

$$Lf = \operatorname{div}(\sigma \nabla f),$$

where  $\sigma$  is a smooth field of positive and symmetric matrices that satisfies

$$a\|x\|^2 \leq \langle x, \sigma x \rangle \leq b\|x\|^2, \quad x \in \mathbb{R}^n,$$

for some constant  $0 < a \leq b$ . Show that with respect to the Lebesgue measure, the operator  $L$  is essentially self-adjoint on  $C_c(\mathbb{R}^n, \mathbb{R})$

**Exercise 4.14.** On  $\mathbb{R}^n$ , we consider the Schrödinger type operator

$$H = L - V,$$

where  $L$  is a diffusion operator and  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  is a smooth function. We denote

$$\Gamma(f, g) = \frac{1}{2} (L(fg) - fLg - gLf).$$

Show that if there exists an increasing sequence  $h_n \in C_c(\mathbb{R}^n, \mathbb{R})$ ,  $0 \leq h_n \leq 1$ , such that  $h_n \nearrow 1$  on  $\mathbb{R}^n$ , and  $\|\Gamma(h_n, h_n)\|_\infty \rightarrow 0$ , as  $n \rightarrow \infty$  and that if  $V$  is bounded from below then  $H$  is essentially self-adjoint on  $C_c(\mathbb{R}^n, \mathbb{R})$ .

From now on, we assume that the diffusion operator  $L$  is essentially self-adjoint on  $C_c(\mathbb{R}^n, \mathbb{R})$ . Its Friedrichs extension is still denoted by  $L$ . The fact that we are now dealing with a non negative self-adjoint operator allows us to use spectral theory in order to define  $e^{tL}$ . Indeed, we have the following so-called spectral theorem whose proof can be found in [?].

**Theorem 4.15** (Spectral theorem). *Let  $A$  be a non negative self-adjoint operator on a separable Hilbert space  $\mathcal{H}$ . There is a measure space  $(\Omega, \nu)$ , a unitary map  $U : \mathbf{L}_\nu^2(\Omega, \mathbb{R}) \rightarrow \mathcal{H}$  and a non negative real valued measurable function  $\lambda$  on  $\Omega$  such that*

$$U^{-1}AUf(x) = \lambda(x)f(x),$$

for  $x \in \Omega$ ,  $Uf \in \mathcal{D}(A)$ . Moreover, given  $f \in \mathbf{L}_\nu^2(\Omega, \mathbb{R})$ ,  $Uf$  belongs to  $\mathcal{D}(A)$  if only if  $\int_\Omega \lambda^2 f^2 d\nu < +\infty$ .

We may now apply the spectral theorem to  $-L$  in order to define  $e^{tL}$ . More generally, given a Borel function  $g : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  and the spectral decomposition of  $-L$ :

$$U^{-1}LUf(x) = -\lambda(x)f(x),$$

we may define an operator  $g(-L)$  as being the unique operator that satisfies

$$U^{-1}g(-L)Uf(x) = (g \circ \lambda)(x)f(x).$$

We may observe that  $g(-L)$  is a bounded operator if  $g$  is a bounded function.

As a particular case, we define the diffusion semigroup  $(\mathbf{P}_t)_{t \geq 0}$  on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  by the requirement

$$U^{-1}\mathbf{P}_tUf(x) = e^{-t\lambda(x)}f(x).$$

This defines a family of bounded operators  $\mathbf{P}_t : \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  whose following properties are readily checked from the spectral decomposition:

- For  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\|\mathbf{P}_t f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} \leq \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}.$$

- $\mathbf{P}_0 = \mathbf{Id}$  and for  $s, t \geq 0$ ,  $\mathbf{P}_s \mathbf{P}_t = \mathbf{P}_{s+t}$ .
- For  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ , the map  $t \rightarrow \mathbf{P}_t f$  is continuous in  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ .
- For  $f, g \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\int_{\mathbb{R}^n} (\mathbf{P}_t f) g d\mu = \int_{\mathbb{R}^n} f (\mathbf{P}_t g) d\mu$$

We summarize the above properties by saying that  $(\mathbf{P}_t)_{t \geq 0}$  is a self-adjoint strongly continuous contraction semigroup on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ .

From the spectral decomposition, it is also easily checked that the operator  $L$  is furthermore the generator of this semigroup, that is for  $f \in \mathcal{D}(L)$ ,

$$\lim_{t \rightarrow 0} \left\| \frac{\mathbf{P}_t f - f}{t} - Lf \right\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} = 0.$$

From the semigroup property, it implies that for  $t \geq 0$ ,  $\mathbf{P}_t \mathcal{D}(L) \subset \mathcal{D}(L)$ , and that for  $f \in \mathcal{D}(L)$ ,

$$\frac{d}{dt} \mathbf{P}_t f = \mathbf{P}_t Lf = L \mathbf{P}_t f,$$

the derivative on the left hand side of the above equality being taken in  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ .

**Exercise 4.16.** *Let  $L$  be an essentially self-adjoint diffusion operator on  $C_c(\mathbb{R}^n, \mathbb{R})$ . Show that if the constant function  $1 \in \mathcal{D}(L)$  and if  $L1 = 0$ , then*

$$\mathbf{P}_t 1 = 1.$$

**Exercise 4.17.** *Let  $L$  be an essentially self-adjoint diffusion operator on  $C_c(\mathbb{R}^n, \mathbb{R})$ .*

1. *Show that for every  $\lambda > 0$ , the range of the operator  $\lambda \mathbf{Id} - L$  is dense in  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ .*
2. *By using the spectral theorem, show that the following limit holds for the operator norm on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$*

$$\mathbf{P}_t = \lim_{n \rightarrow +\infty} \left( \mathbf{Id} - \frac{t}{n} L \right)^{-n}.$$

**Exercise 4.18.** As usual, we denote by  $\Delta$  the Laplace operator on  $\mathbb{R}^n$ :

$$\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}.$$

The Mac-Donald's function with index  $\nu \in \mathbb{R}$  is defined for  $x \in \mathbb{R} \setminus \{0\}$  by

$$K_\nu(x) = \frac{1}{2} \left(\frac{x}{2}\right)^\nu \int_0^{+\infty} \frac{e^{-\frac{x^2}{4t}-t}}{t^{1+\nu}} dt.$$

1. Show that for  $\lambda \in \mathbb{R}^n$  and  $\alpha > 0$ ,

$$\frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{i\langle \lambda, x \rangle} \left(\frac{\|x\|}{\sqrt{\alpha}}\right)^{1-\frac{n}{2}} K_{\frac{n}{2}-1}(\sqrt{\alpha}\|x\|) dx = \frac{1}{\alpha + \|\lambda\|^2}.$$

2. Show that for  $\nu \in \mathbb{R}$ ,  $K_{-\nu} = K_\nu$ .

3. Show that

$$K_{1/2}(x) = \sqrt{\frac{\pi}{2x}} e^{-x}.$$

4. Prove that for  $f \in \mathbf{L}^2(\mathbb{R}^n, \mathbb{R})$  and  $\alpha > 0$ ,

$$(\alpha \mathbf{Id} - \Delta)^{-1} f(x) = \int_{\mathbb{R}^n} G_\alpha(x-y) f(y) dy,$$

where

$$G_\alpha(x) = \frac{1}{(2\pi)^{n/2}} \left(\frac{\|x\|}{\sqrt{\alpha}}\right)^{1-\frac{n}{2}} K_{\frac{n}{2}-1}(\sqrt{\alpha}\|x\|).$$

(You may use Fourier transform to solve the partial differential equation  $\alpha g - \Delta g = f$ ).

5. Prove that for  $f \in \mathbf{L}^2(\mathbb{R}^n, \mathbb{R})$

$$\lim_{n \rightarrow +\infty} \left(\mathbf{Id} - \frac{t}{n} \Delta\right)^{-n} f = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{\|x-y\|^2}{4t}} f(y) dy.$$

the limit being taken in  $\mathbf{L}^2(\mathbb{R}^n, \mathbb{R})$ . Conclude that almost everywhere,

$$\mathbf{P}_t f(x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{\|x-y\|^2}{4t}} f(y) dy.$$

**Exercise 4.19.**

1. Show the subordination identity:

$$e^{-y|\alpha|} = \frac{y}{2\sqrt{\pi}} \int_0^{+\infty} \frac{e^{-\frac{y^2}{4t}-t\alpha^2}}{t^{3/2}} dt, \quad y > 0, \alpha \in \mathbb{R}.$$

2. The Cauchy's semigroup on  $\mathbb{R}^n$  is defined as  $\mathbf{Q}_t = e^{-t\sqrt{-\Delta}}$ . By using the subordination identity and the heat semigroup on  $\mathbb{R}^n$ , show that for  $f \in \mathbf{L}^2(\mathbb{R}^n, \mathbb{R})$

$$\mathbf{Q}_t f(x) = \int_{\mathbb{R}^n} q(t, x-y) f(y) dy,$$

where

$$q(t, x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}} \frac{t}{(t^2 + \|x\|^2)^{\frac{n+1}{2}}}.$$

## 2 Existence and regularity of the heat kernel

In order to study the regularization properties of diffusion semigroups, we apply the theory of local elliptic regularity which is sketched in the Appendix B.

**Proposition 4.20.** *Let  $L$  be an elliptic diffusion operator with smooth coefficients that is essentially self-adjoint with respect to a measure  $\mu$ . Denote by  $(\mathbf{P}_t)_{t \geq 0}$  the corresponding semigroup on  $\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})$ .*

- If  $K$  is a compact set of  $\mathbb{R}^n$ , there exists a positive constant  $C$  such that for  $f \in \mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})$ :

$$\sup_{x \in K} |\mathbf{P}_t f(x)| \leq C \left(1 + \frac{1}{t^\kappa}\right) \|f\|_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})},$$

where  $\kappa$  is the smallest integer larger than  $\frac{n}{4}$ .

- For  $f \in \mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})$ , the function  $(t, x) \rightarrow \mathbf{P}_t f(x)$  is smooth on  $(0, +\infty) \times \mathbb{R}^n$ ;

*Proof.* Let us first observe that from the spectral theorem that if  $f \in \mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})$  then for every  $k \geq 0$ ,  $L^k \mathbf{P}_t f \in \mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})$  and

$$\|L^k \mathbf{P}_t f\|_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})} \leq \left(\sup_{\lambda \geq 0} \lambda^k e^{-\lambda t}\right) \|f\|_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})}.$$

Now, let  $K$  be a compact set of  $\mathbb{R}^n$ . From Proposition 7.65, there exists therefore a positive constant  $C$  such that

$$\left(\sup_{x \in K} |\mathbf{P}_t f(x)|\right)^2 \leq C \left(\sum_{k=0}^{\kappa} \|L^k \mathbf{P}_t f\|_{\mathbf{L}^2_\mu(\mathbb{R}^n, \mathbb{R})}^2\right).$$

Since it is immediately checked that

$$\sup_{\lambda \geq 0} \lambda^k e^{-\lambda t} = \left(\frac{k}{t}\right)^k e^{-k},$$



the bound

$$\sup_{x \in K} |\mathbf{P}_t f(x)| \leq C \left(1 + \frac{1}{t^\kappa}\right) \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})},$$

easily follows.

We now turn to the second part. Let  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ .

First, we fix  $t > 0$ . As above, from the spectral theorem, for every  $k \geq 0$ ,  $L^k \mathbf{P}_t f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R}) \subset \mathcal{H}_0^{loc}(\Omega)$ , for any bounded open set  $\Omega$ . A recursive application of Lemma 7.62 shows that, therefore  $\mathbf{P}_t f \in \cap_{s>0} \mathcal{H}_s^{loc}(\Omega)$ , which implies from Sobolev lemma (Theorem 7.58) that  $\mathbf{P}_t f$  is a smooth function.

Next, we prove joint continuity in the variables  $(t, x) \in (0, +\infty) \times \mathbb{R}^n$ . It is enough to prove that if  $t_0 > 0$  and if  $K$  is a compact set on  $\mathbb{R}^n$ ,

$$\sup_{x \in K} |\mathbf{P}_t f(x) - \mathbf{P}_{t_0} f(x)| \xrightarrow{t \rightarrow t_0} 0.$$

From Proposition 7.65, there exists a positive constant  $C$  such that

$$\sup_{x \in K} |\mathbf{P}_t f(x) - \mathbf{P}_{t_0} f(x)| \leq C \left( \sum_{k=0}^{\kappa} \|L^k \mathbf{P}_t f - L^k \mathbf{P}_{t_0} f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 \right).$$

Now, again from the spectral theorem, it is checked that

$$\lim_{t \rightarrow t_0} \sum_{k=0}^{\kappa} \|L^k \mathbf{P}_t f - L^k \mathbf{P}_{t_0} f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 = 0.$$

This gives the expected joint continuity in  $(t, x)$ . The joint smoothness in  $(t, x)$  is a consequence of the second part of Proposition 7.65 and the details are let to the reader.  $\square$

**Remark 4.21.** If bound

$$\sup_{x \in K} |\mathbf{P}_t f(x)| \leq C(t) \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}$$

uniformly holds on  $\mathbb{R}^n$ , that is if the operator norm

$$\|\mathbf{P}_t\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathbf{L}_\mu^\infty(\mathbb{R}^n, \mathbb{R})} < \infty,$$

then the semigroup  $(\mathbf{P}_t)_{t \geq 0}$  is said to be ultracontractive. The study of ultracontractive semigroups is intimately related to the beautiful theory of log-Sobolev inequalities. We refer the interested reader to the book [?].

**Exercise 4.22.** Let  $L$  be an elliptic diffusion operator with smooth coefficients that is essentially self-adjoint with respect to a measure  $\mu$ . Let  $\alpha$  be a multi-index. If  $K$  is a compact set of  $\mathbb{R}^n$ , show that there exists a positive constant  $C$  such that for  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ :

$$\sup_{x \in K} |\partial^\alpha \mathbf{P}_t f(x)| \leq C \left(1 + \frac{1}{t^{|\alpha| + \kappa}}\right) \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})},$$

where  $\kappa$  is the smallest integer larger than  $\frac{n}{4}$ .

We now prove the following fundamental theorem:

**Theorem 4.23.** *Let  $L$  be an elliptic and essentially self-adjoint diffusion operator. Denote by  $(\mathbf{P}_t)_{t \geq 0}$  the corresponding semigroup on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ . There is a smooth function  $p(t, x, y)$ ,  $t \in (0, +\infty)$ ,  $x, y \in \mathbb{R}^n$ , such that for every  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  and  $x \in \mathbb{R}^n$ ,*

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}^n} p(t, x, y) f(y) d\mu(y).$$

The function  $p(t, x, y)$  is called the heat kernel associated to  $(\mathbf{P}_t)_{t \geq 0}$ . It satisfies furthermore:

- (Symmetry)  $p(t, x, y) = p(t, y, x)$ ;
- (Chapman-Kolmogorov relation)  $p(t+s, x, y) = \int_{\mathbb{R}^n} p(t, x, z) p(s, z, y) d\mu(z)$ .

*Proof.* Let  $x \in \mathbb{R}^n$  and  $t > 0$ . From the previous proposition, the linear form  $f \rightarrow \mathbf{P}_t f(x)$  is continuous on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ , therefore from the Riesz representation theorem, there is a function  $p(t, x, \cdot) \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ , such that for  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\mathbf{P}_t f(x) = \int_{\mathbb{R}^n} p(t, x, y) f(y) d\mu(y).$$

From the fact that  $\mathbf{P}_t$  is self-adjoint on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ :

$$\int_{\mathbb{R}^n} (\mathbf{P}_t f) g d\mu = \int_{\mathbb{R}^n} f (\mathbf{P}_t g) d\mu,$$

we easily deduce the symmetry property:

$$p(t, x, y) = p(t, y, x).$$

And the Chapman-Kolmogorov relation  $p(t+s, x, y) = \int_{\mathbb{R}^n} p(t, x, z) p(s, z, y) d\mu(z)$  stems from the semigroup property  $\mathbf{P}_{t+s} = \mathbf{P}_t \mathbf{P}_s$ . Finally, from the previous proposition the map  $(t, x) \rightarrow p(t, x, \cdot) \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  is smooth on  $(0, +\infty) \times \mathbb{R}^n$  for the weak topology on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ . This implies that it is also smooth on  $(0, +\infty) \times \mathbb{R}^n$  for the norm topology. Since, from the Chapman-Kolmogorov relation

$$p(t, x, y) = \langle p(t/2, x, \cdot), p(t/2, y, \cdot) \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})},$$

we conclude that  $(t, x, y) \rightarrow p(t, x, y)$  is smooth on  $(0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n$ .  $\square$

### 3 The submarkov property

In the previous section, we have proved that if  $L$  is a diffusion operator that is essentially self-adjoint on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , then by using the spectral theorem we

can define a self-adjoint strongly continuous contraction semigroup  $(\mathbf{P}_t)_{t \geq 0}$  on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  with generator  $L$ .

In this section, we will be interested in the following additional property of the semigroup  $(\mathbf{P}_t)_{t \geq 0}$ : If  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  satisfies almost surely  $0 \leq f \leq 1$ , then almost surely  $0 \leq \mathbf{P}_t f \leq 1$ ,  $t \geq 0$ . This property is called the submarkov property of the semigroup  $(\mathbf{P}_t)_{t \geq 0}$ . The terminology stems from the fact that this is precisely this property that makes the link with probability theory because it is equivalent to the fact that  $(\mathbf{P}_t)_{t \geq 0}$  is associated to the transition function of a sub-Markov process.

As a first result, we prove that the semigroup  $(\mathbf{P}_t)_{t \geq 0}$  is positivity preserving. This property may actually be seen at the level of the diffusion operator  $L$ , and relies on the following distributional inequality that is satisfied by  $L$ .

**Lemma 4.24** (Kato inequality). *Let  $L$  be a diffusion operator on  $\mathbb{R}^n$  with symmetric and invariant measure  $\mu$ . Let  $u \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . Define*

$$\begin{aligned} \mathbf{sgn} u &= 0 && \text{if } u(x) = 0, \\ &= \frac{u(x)}{|u(x)|} && \text{if } u(x) \neq 0. \end{aligned}$$

*In the sense of distributions, we have the following inequality*

$$L|u| \geq (\mathbf{sgn} u)Lu.$$

*Proof.* If  $\phi$  is a smooth and convex function and if  $u$  is assumed to be smooth, it is readily checked that

$$L\phi(u) = \phi'(u)Lu + \phi''(u)\Gamma(u, u) \geq \phi'(u)Lu.$$

By choosing for  $\phi$  the function

$$\phi_\varepsilon(x) = \sqrt{x^2 + \varepsilon^2}, \quad \varepsilon > 0,$$

we deduce that for every smooth function  $u \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,

$$L\phi_\varepsilon(u) \geq \frac{u}{\sqrt{u^2 + \varepsilon^2}}Lu.$$

As a consequence this inequality holds in the sense of distributions, that is for every  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,  $f \geq 0$ ,

$$\int_{\mathbb{R}^n} f L\phi_\varepsilon(u) d\mu \geq \int_{\mathbb{R}^n} f \frac{u}{\sqrt{u^2 + \varepsilon^2}} Lu d\mu$$

Letting  $\varepsilon \rightarrow 0$  gives the expected result.  $\square$

We are now in position to state and prove the positivity preserving theorem.

**Theorem 4.25.** *Let  $L$  be an essentially self adjoint diffusion operator on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . If  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  is almost surely nonnegative ( $f \geq 0$ ), then we have for every  $t \geq 0$ ,  $\mathbf{P}_t f \geq 0$  almost surely.*

*Proof.* The main idea is to prove that for  $\lambda > 0$ , the operator  $(\lambda \mathbf{Id} - L)^{-1}$  which is well defined due to essential self adjointness preserves the positivity of function. Then, we may conclude by the fact that for  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ , in the  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  sense

$$\mathbf{P}_t f = \lim_{n \rightarrow +\infty} \left( \mathbf{Id} - \frac{t}{n} L \right)^{-n} f.$$

Let  $\lambda > 0$ . We consider on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  the norm

$$\begin{aligned} \|f\|_\lambda^2 &= \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \lambda \mathcal{E}(f, f) \\ &= \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \lambda \int_{\mathbb{R}^n} \Gamma(f, f) d\mu. \end{aligned}$$

and denote by  $\mathcal{H}_\lambda$  the completion of  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . Using Kato's inequality, it is seen (multiply the inequality by  $u$  and integrate by parts, details are let to the reader) that if  $u \in \mathcal{H}_\lambda$ , then  $|u| \in \mathcal{H}_\lambda$  and

$$\mathcal{E}(|u|, |u|) \leq \mathcal{E}(u, u). \quad (4.1)$$

Since  $L$  is essentially self-adjoint we can consider the bounded operator

$$\mathbf{R}_\lambda = (\mathbf{Id} - \lambda L)^{-1}$$

that goes from  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  to  $\mathcal{D}(L) \subset \mathcal{H}_\lambda$ . For  $f \in \mathcal{H}_\lambda$  and  $g \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  with  $g \geq 0$ , we have

$$\begin{aligned} \langle |f|, \mathbf{R}_\lambda g \rangle_\lambda &= \langle |f|, \mathbf{R}_\lambda g \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} - \lambda \langle |f|, L \mathbf{R}_\lambda g \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} \\ &= \langle |f|, (\mathbf{Id} - \lambda L) \mathbf{R}_\lambda g \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} \\ &= \langle |f|, g \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} \\ &\geq |\langle f, g \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}| \\ &\geq |\langle f, \mathbf{R}_\lambda g \rangle_\lambda|. \end{aligned}$$

Moreover, from inequality (4.1), for  $f \in \mathcal{H}_\lambda$ ,

$$\begin{aligned} \| |f| \|_\lambda^2 &= \| |f| \|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \lambda \mathcal{E}(|f|, |f|) \\ &\geq \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \lambda \mathcal{E}(f, f) \\ &\geq \|f\|_\lambda^2. \end{aligned}$$

By taking  $f = \mathbf{R}_\lambda g$  in the two above sets of inequalities, we draw the conclusion

$$|\langle \mathbf{R}_\lambda g, \mathbf{R}_\lambda g \rangle_\lambda| \leq \langle \mathbf{R}_\lambda g, \mathbf{R}_\lambda g \rangle_\lambda \leq \| \mathbf{R}_\lambda g \|_\lambda \| \mathbf{R}_\lambda g \|_\lambda \leq |\langle \mathbf{R}_\lambda g, \mathbf{R}_\lambda g \rangle_\lambda|.$$

The above inequalities are therefore equalities which implies

$$\mathbf{R}_\lambda g = |\mathbf{R}_\lambda g|.$$

As a conclusion if  $g \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  is  $\geq 0$ , then for every  $\lambda > 0$ ,  $(\mathbf{Id} - \lambda L)^{-1}g \geq 0$ . Thanks to spectral theorem, in  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\mathbf{P}_t g = \lim_{n \rightarrow +\infty} \left( \mathbf{Id} - \frac{t}{n} L \right)^{-n} g.$$

By passing to a subsequence that converges pointwise almost surely, we deduce that  $\mathbf{P}_t g \geq 0$  almost surely.  $\square$

**Exercise 4.26.** *Let  $L$  be an elliptic diffusion operator with smooth coefficients that is essentially self-adjoint. Denote by  $p(t, x, y)$  the heat kernel of  $\mathbf{P}_t$ . Show that  $p(t, x, y) \geq 0$ . (Remark: It actually possible to prove that  $p(t, x, y) > 0$ ).*

Besides the positivity preserving property, the semigroup is a contraction on  $\mathbf{L}_\mu^\infty(\mathbb{R}^n, \mathbb{R})$ . More precisely we have

**Theorem 4.27.** *Let  $L$  be an essentially self adjoint diffusion operator on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . If  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R}) \cap \mathbf{L}_\mu^\infty(\mathbb{R}^n, \mathbb{R})$ , then  $\mathbf{P}_t f \in \mathbf{L}_\mu^\infty(\mathbb{R}^n, \mathbb{R})$  and*

$$\|\mathbf{P}_t f\|_\infty \leq \|f\|_\infty.$$

*Proof.* The proof is close and relies on the same ideas as the proof of the positivity preserving Theorem 4.25. So, we only list below the main steps and let the reader fills the details.

As before, for  $\lambda > 0$ , we consider on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  the norm

$$\begin{aligned} \|f\|_\lambda^2 &= \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \lambda \mathcal{E}(f, f) \\ &= \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \lambda \int_{\mathbb{R}^n} \Gamma(f, f) d\mu. \end{aligned}$$

and denote by  $\mathcal{H}_\lambda$  the completion of  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ .

- The first step is to show that if  $0 \leq f \in \mathcal{H}_\lambda$ , then  $1 \wedge f$  (minimum between 1 and  $f$ ) also lies in  $\mathcal{H}_\lambda$  and moreover

$$\mathcal{E}(1 \wedge f, 1 \wedge f) \leq \mathcal{E}(f, f).$$

- Let  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  satisfy  $0 \leq f \leq 1$  and put  $g = \mathbf{R}_\lambda f = (\mathbf{Id} - \lambda L)^{-1} f \in \mathcal{H}_\lambda$  and  $h = 1 \wedge g$ . According to the first step,  $h \in \mathcal{H}_\lambda$  and  $\mathcal{E}(h, h) \leq$

$\mathcal{E}(g, g)$ . Now, we observe that:

$$\begin{aligned}
& \|g - h\|_\lambda^2 \\
&= \|g\|_\lambda^2 - 2\langle g, h \rangle_\lambda + \|h\|_\lambda^2 \\
&= \langle \mathbf{R}_\lambda f, f \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} - 2\langle f, h \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} + \|h\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \lambda \mathcal{E}(h, h) \\
&= \langle \mathbf{R}_\lambda f, f \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} - \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \|f - h\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \lambda \mathcal{E}(h, h) \\
&\leq \langle \mathbf{R}_\lambda f, f \rangle_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})} - \|f\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \|f - g\|_{\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})}^2 + \lambda \mathcal{E}(g, g) = 0.
\end{aligned}$$

As a consequence  $g = h$ , that is  $0 \leq g \leq 1$ .

- The previous step shows that if  $f \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  satisfies  $0 \leq f \leq 1$  then for every  $\lambda > 0$ ,  $0 \leq (\mathbf{Id} - \lambda L)^{-1} f \leq 1$ . Thanks to spectral theorem, in  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ ,

$$\mathbf{P}_t f = \lim_{n \rightarrow +\infty} \left( \mathbf{Id} - \frac{t}{n} L \right)^{-n} f.$$

By passing to a subsequence that converges pointwise almost surely, we deduce that  $0 \leq \mathbf{P}_t f \leq 1$  almost surely.

□

## 4 $L^p$ theory: The interpolation method

In the previous section, we have seen that if  $L$  is an essentially self-adjoint diffusion operator with invariant and symmetric measure  $\mu$ , then, by using the spectral theorem we can define a self-adjoint strongly continuous contraction semigroup  $(\mathbf{P}_t)_{t \geq 0}$  on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  with generator  $L$ .

Our goal, in this section, is to define, for  $1 \leq p \leq +\infty$ ,  $\mathbf{P}_t$  on  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ . This may be done in a natural way by using the Riesz-Thorin interpolation theorem that we state below. We refer the reader to the book [?] for a proof.

**Theorem 4.28** (Riesz-Thorin interpolation theorem). *Let  $1 \leq p_0, p_1, q_0, q_1 \leq \infty$ , and  $\theta \in (0, 1)$ . Define  $1 \leq p, q \leq \infty$  by*

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

*If  $T$  is a linear map such that*

$$T : \mathbf{L}_\mu^{p_0} \rightarrow \mathbf{L}_\mu^{q_0}, \quad \|T\|_{\mathbf{L}_\mu^{p_0} \rightarrow \mathbf{L}_\mu^{q_0}} = M_0$$

$$T : \mathbf{L}_\mu^{p_1} \rightarrow \mathbf{L}_\mu^{q_1}, \quad \|T\|_{\mathbf{L}_\mu^{p_1} \rightarrow \mathbf{L}_\mu^{q_1}} = M_1,$$

*then, for every  $f \in \mathbf{L}_\mu^{p_0} \cap \mathbf{L}_\mu^{p_1}$ ,*

$$\|Tf\|_q \leq M_0^{1-\theta} M_1^\theta \|f\|_p.$$

Hence  $T$  extends uniquely as a bounded map from  $\mathbf{L}_\mu^p$  to  $\mathbf{L}_\mu^q$  with

$$\|T\|_{\mathbf{L}_\mu^p \rightarrow \mathbf{L}_\mu^q} \leq M_0^{1-\theta} M_1^\theta.$$

**Remark 4.29.** The statement that  $T$  is a linear map such that

$$T : \mathbf{L}_\mu^{p_0} \rightarrow \mathbf{L}_\mu^{q_0}, \quad \|T\|_{\mathbf{L}_\mu^{p_0} \rightarrow \mathbf{L}_\mu^{q_0}} = M_0$$

$$T : \mathbf{L}_\mu^{p_1} \rightarrow \mathbf{L}_\mu^{q_1}, \quad \|T\|_{\mathbf{L}_\mu^{p_1} \rightarrow \mathbf{L}_\mu^{q_1}} = M_1,$$

means that there exists a map  $T : \mathbf{L}_\mu^{p_0} \cap \mathbf{L}_\mu^{p_1} \rightarrow \mathbf{L}_\mu^{q_0} \cap \mathbf{L}_\mu^{q_1}$  with

$$\sup_{f \in \mathbf{L}_\mu^{p_0} \cap \mathbf{L}_\mu^{p_1}, \|f\|_{p_0} \leq 1} \|Tf\|_{q_0} = M_0$$

and

$$\sup_{f \in \mathbf{L}_\mu^{p_0} \cap \mathbf{L}_\mu^{p_1}, \|f\|_{p_1} \leq 1} \|Tf\|_{q_1} = M_1.$$

In such a case,  $T$  can be uniquely extended to bounded linear maps  $T_0 : \mathbf{L}_\mu^{p_0} \rightarrow \mathbf{L}_\mu^{q_0}$ ,  $T_1 : \mathbf{L}_\mu^{p_1} \rightarrow \mathbf{L}_\mu^{q_1}$ . With a slight abuse of notation, these two maps are both denoted by  $T$  in the theorem.

**Remark 4.30.** If  $f \in \mathbf{L}_\mu^{p_0} \cap \mathbf{L}_\mu^{p_1}$  and  $p$  is defined by  $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ , then by Hölder's inequality,  $f \in \mathbf{L}_\mu^p$  and

$$\|f\|_p \leq \|f\|_{p_0}^{1-\theta} \|f\|_{p_1}^\theta.$$

One of the (numerous) beautiful applications of the Riesz-Thorin theorem is to construct diffusion semigroups on  $L^p$  by interpolation. More precisely, let  $L$  be an essentially self-adjoint diffusion operator. We denote by  $(\mathbf{P}_t)_{t \geq 0}$  the self-adjoint strongly continuous semigroup associated to  $L$  constructed on  $\mathbf{L}_\mu^2$  thanks to the spectral theorem. We recall that  $(\mathbf{P}_t)_{t \geq 0}$  satisfies the submarkov property: That is, if  $0 \leq f \leq 1$  is a function in  $\mathbf{L}_\mu^2$ , then  $0 \leq \mathbf{P}_t f \leq 1$ .

We now are in position to state the following theorem:

**Theorem 4.31.** *The space  $\mathbf{L}_\mu^1 \cap \mathbf{L}_\mu^\infty$  is invariant under  $\mathbf{P}_t$  and  $\mathbf{P}_t$  may be extended from  $\mathbf{L}_\mu^1 \cap \mathbf{L}_\mu^\infty$  to a contraction semigroup  $(\mathbf{P}_t^{(p)})_{t \geq 0}$  on  $\mathbf{L}_\mu^p$  for all  $1 \leq p \leq \infty$ : For  $f \in \mathbf{L}_\mu^p$ ,*

$$\|\mathbf{P}_t f\|_{\mathbf{L}_\mu^p} \leq \|f\|_{\mathbf{L}_\mu^p}.$$

*These semigroups are consistent in the sense that for  $f \in \mathbf{L}_\mu^p \cap \mathbf{L}_\mu^q$ ,*

$$\mathbf{P}_t^{(p)} f = \mathbf{P}_t^{(q)} f.$$

*Proof.* If  $f, g \in \mathbf{L}_\mu^1 \cap \mathbf{L}_\mu^\infty$  which is a subset of  $\mathbf{L}_\mu^1 \cap \mathbf{L}_\mu^\infty$ , then from Theorem 4.27,

$$\begin{aligned} \left| \int_{\mathbb{R}^n} (\mathbf{P}_t f) g d\mu \right| &= \left| \int_{\mathbb{R}^n} f (\mathbf{P}_t g) d\mu \right| \\ &\leq \|f\|_{\mathbf{L}_\mu^1} \|\mathbf{P}_t g\|_{\mathbf{L}_\mu^\infty} \\ &\leq \|f\|_{\mathbf{L}_\mu^1} \|g\|_{\mathbf{L}_\mu^\infty}. \end{aligned}$$

This implies

$$\|\mathbf{P}_t f\|_{\mathbf{L}_\mu^1} \leq \|f\|_{\mathbf{L}_\mu^\infty}.$$

The conclusion follows then from the Riesz-Thorin interpolation theorem.  $\square$

**Exercise 4.32.** Show that if  $f \in \mathbf{L}_\mu^p$  and  $g \in \mathbf{L}_\mu^q$  with  $\frac{1}{p} + \frac{1}{q} = 1$  then,

$$\int_{\mathbb{R}^n} f \mathbf{P}_t^{(q)} g d\mu = \int_{\mathbb{R}^n} g \mathbf{P}_t^{(p)} f d\mu.$$

**Exercise 4.33.**

1. Show that for each  $f \in \mathbf{L}_\mu^1$ , the  $\mathbf{L}_\mu^1$ -valued map  $t \rightarrow \mathbf{P}_t^{(1)} f$  is continuous.
2. Show that for each  $f \in \mathbf{L}_\mu^p$ ,  $1 < p < 2$ , the  $\mathbf{L}_\mu^p$ -valued map  $t \rightarrow \mathbf{P}_t^{(p)} f$  is continuous.
3. Finally, by using the reflexivity of  $\mathbf{L}_\mu^p$ , show that for each  $f \in \mathbf{L}_\mu^p$  and every  $p \geq 1$ , the  $\mathbf{L}_\mu^p$ -valued map  $t \rightarrow \mathbf{P}_t^{(p)} f$  is continuous.

We mention, that in general, the  $\mathbf{L}_\mu^\infty$  valued map  $t \rightarrow \mathbf{P}_t^{(\infty)} f$  is not continuous.

## 5 $L^p$ theory: The Hille-Yosida method

In the previous subsections we learnt how to construct a semigroup from a diffusion operator by using the theory of self-adjoint unbounded operators on Hilbert spaces. In particular, we proved the following result.

**Theorem 4.34.** Let  $L$  be a elliptic diffusion operator with smooth coefficients on  $\mathbb{R}^n$ . If:

- There is a Borel measure  $\mu$ , symmetric for  $L$  on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ;
- There exists an increasing sequence  $h_n \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,  $0 \leq h_n \leq 1$ , such that  $h_n \nearrow 1$  on  $\mathbb{R}^n$ , and  $\|\Gamma(h_n, h_n)\|_\infty \rightarrow 0$ , as  $n \rightarrow \infty$ .

Then, by using the spectral theorem we can define a self-adjoint strongly continuous contraction semigroup  $(\mathbf{P}_t)_{t \geq 0}$  on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  with generator  $L$ . Moreover,  $\mathbf{P}_t$  is a submarkov semigroup.



In this section we will work under the assumptions of the above theorem. Our goal will be to define, for  $1 < p < +\infty$ ,  $\mathbf{P}_t$  on  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ . This can be done by using the interpolation method described in the previous subsection but also by using so-called Hille-Yosida theory of contraction semigroups on Banach spaces.

## 5.1 Semigroups on Banach spaces

Let  $(B, \|\cdot\|)$  be a Banach space (which for us will later be  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ ,  $1 < p < +\infty$ ). We first have the following basic definition.

**Definition 4.35.** A family of bounded operators  $(T_t)_{t \geq 0}$  on  $B$  is called a strongly continuous contraction semigroup if:

- $T_0 = \mathbf{Id}$  and for  $s, t \geq 0$ ,  $T_{s+t} = T_s T_t$ ;
- For each  $x \in B$ , the map  $t \rightarrow T_t x$  is continuous;
- For each  $x \in B$  and  $t \geq 0$ ,  $\|T_t x\| \leq \|x\|$ .

Now, let us recall that a densely defined linear operator

$$A : \mathcal{D}(A) \subset B \rightarrow B$$

is said to be closed if for every sequence  $x_n \in \mathcal{D}(A)$  that converges to  $x \in B$  and such that  $Ax_n \rightarrow y \in B$ , we have  $x \in \mathcal{D}(A)$  and  $y = Ax$ . We observe that any densely defined operator  $A$  may be extended into a closed operator. Indeed, let us consider on  $\mathcal{D}(A)$  the norm

$$\|x\|_A = \|x\| + \|Ax\|,$$

and let us complete  $\mathcal{D}(A)$ , to obtain a Banach space  $\mathcal{B}$ . Since  $\|x\| \leq \|x\|_A$ , we may identify  $\mathcal{B}$  as a subset of  $B$ . Moreover  $\|Ax\| \leq \|x\|_A$ , therefore,  $A$  can be extended into an operator  $\bar{A} : \mathcal{B} \rightarrow B$ , and it is readily checked that  $\bar{A}$  with domain  $\mathcal{B}$  is a closed extension of  $A$ . The operator  $\bar{A}$  is called the closure of  $A$ .

In this situation, we have the following proposition that extends the Theorem 3.31, and whose proof is pretty much the same and therefore let to the reader.

**Proposition 4.36.** *Let  $(T_t)_{t \geq 0}$  be a strongly continuous contraction semigroup on  $B$ . There exists a closed and densely defined operator*

$$A : \mathcal{D}(A) \subset B \rightarrow B$$

where

$$\mathcal{D}(A) = \left\{ x \in B, \quad \lim_{t \rightarrow 0} \frac{T_t x - x}{t} \text{ exists} \right\},$$

such that for  $x \in \mathcal{D}(A)$ ,

$$\lim_{t \rightarrow 0} \left\| \frac{T_t x - x}{t} - Ax \right\| = 0.$$

The operator  $A$  is called the generator of the semigroup  $(T_t)_{t \geq 0}$ . We also say that  $A$  generates  $(T_t)_{t \geq 0}$ .

**Remark 4.37.** We may observe that the proof of the above result does not involve the contraction property of  $(T_t)_{t \geq 0}$ , so that it may be extended to strongly continuous semigroups.

The following important theorem is due to Hille and Yosida and provides, through spectral properties, a characterization of closed operators that are generators of contraction semigroups.

Let  $A : \mathcal{D}(A) \subset B \rightarrow B$  be a densely defined closed operator. A constant  $\lambda \in \mathbb{R}$  is said to be in the spectrum of  $A$  if the operator  $\lambda \mathbf{Id} - A$  is not bijective. In that case, it is a consequence of the closed graph theorem<sup>2</sup> that if  $\lambda$  is not in the spectrum of  $A$ , then the operator  $\lambda \mathbf{Id} - A$  has a bounded inverse. The spectrum of an operator  $A$  shall be denoted  $\rho(A)$ .

**Theorem 4.38** (Hille-Yosida theorem). *A necessary and sufficient condition that a densely defined closed operator  $A$  generates a strongly continuous contraction semigroup is that:*

- $\rho(A) \subset (-\infty, 0]$ ;
- $\|(\lambda \mathbf{Id} - A)^{-1}\| \leq \frac{1}{\lambda}$  for all  $\lambda > 0$ .

*Proof.* Let us first assume that  $A$  generates a strongly continuous contraction semigroup  $(T_t)_{t \geq 0}$ . Let  $\lambda > 0$ . We want to prove that  $\lambda \mathbf{Id} - A$  is a bijective operator  $\mathcal{D}(A) \rightarrow B$ .

The formal Laplace transform formula

$$\int_0^{+\infty} e^{-\lambda t} e^{tA} dt = (\lambda \mathbf{Id} - A)^{-1},$$

suggests that the operator

$$\mathbf{R}_\lambda = \int_0^{+\infty} e^{-\lambda t} T_t dt$$

is the inverse of  $\lambda \mathbf{Id} - A$ . First, let us observe that  $\mathbf{R}_\lambda$  is well-defined as a Riemann integral since  $t \rightarrow T_t$  is continuous and  $\|T_t\| \leq 1$ . We now show that

<sup>2</sup>An everywhere defined operator between two Banach spaces  $A : B_1 \rightarrow B_2$  is bounded if and only if it is closed.

for  $x \in B$ ,  $\mathbf{R}_\lambda x \in \mathcal{D}(A)$ . For  $h > 0$ ,

$$\begin{aligned}
\frac{T_h - \mathbf{Id}}{h} \mathbf{R}_\lambda x &= \int_0^{+\infty} e^{-\lambda t} \frac{T_h - \mathbf{Id}}{h} T_t x dt \\
&= \int_0^{+\infty} e^{-\lambda t} \frac{T_{h+t} - T_t}{h} x dt \\
&= e^{\lambda h} \int_h^{+\infty} e^{-\lambda s} \frac{T_s - T_{s-h}}{h} x ds \\
&= \frac{e^{\lambda h}}{h} \left( \mathbf{R}_\lambda x - \int_0^h e^{-\lambda s} T_s x ds - \int_h^{+\infty} e^{-\lambda s} T_{s-h} x ds \right) \\
&= \frac{e^{\lambda h} - 1}{h} \mathbf{R}_\lambda x - \frac{e^{\lambda h}}{h} \int_0^h e^{-\lambda s} T_s x ds
\end{aligned}$$

By letting  $h \rightarrow 0$ , we deduce that  $\mathbf{R}_\lambda x \in \mathcal{D}(A)$  and moreover

$$A \mathbf{R}_\lambda x = \lambda \mathbf{R}_\lambda x - x.$$

Therefore we proved

$$(\lambda \mathbf{Id} - A) \mathbf{R}_\lambda = \mathbf{Id}.$$

Furthermore, it is readily checked that, since  $A$  is closed, for  $x \in \mathcal{D}(A)$ ,

$$A \mathbf{R}_\lambda x = A \int_0^{+\infty} e^{-\lambda t} T_t x dt = \int_0^{+\infty} e^{-\lambda t} A T_t x dt = \int_0^{+\infty} e^{-\lambda t} T_t A x dt = \mathbf{R}_\lambda A x.$$

We therefore conclude

$$(\lambda \mathbf{Id} - A) \mathbf{R}_\lambda = \mathbf{R}_\lambda (\lambda \mathbf{Id} - A) = \mathbf{Id}.$$

Thus,

$$\mathbf{R}_\lambda = (\lambda \mathbf{Id} - A)^{-1},$$

and it is clear that

$$\|\mathbf{R}_\lambda\| \leq \frac{1}{\lambda}.$$

Let us now assume that  $A$  is a densely defined closed operator such that

- $\rho(A) \subset (-\infty, 0]$ ;
- $\|(\lambda \mathbf{Id} - A)^{-1}\| \leq \frac{1}{\lambda}$  for all  $\lambda > 0$ .

The idea is to consider the following sequence of bounded operators

$$A_n = -n \mathbf{Id} + n^2 (n \mathbf{Id} - A)^{-1},$$

from which it is easy to define a contraction semigroup and then to show that  $A_n \rightarrow A$ . We will then define a contraction semigroup associated to  $A$  as the limit of the contraction semigroups associated to  $A_n$ .

First, for  $x \in \mathcal{D}(A)$ , we have

$$A_n x = n(n\mathbf{Id} - A)^{-1} A x \rightarrow_{n \rightarrow +\infty} 0.$$

Now, since  $A_n$  is a bounded operator, we may define a semigroup  $(T_t^n)_{t \geq 0}$  through the formula

$$T_t^n = \sum_{k=0}^{+\infty} \frac{t^k A_n^k}{k!}.$$

At that point, let us observe that we also have

$$T_t^n = e^{-nt} \sum_{k=0}^{+\infty} \frac{n^{2k} t^k (n\mathbf{Id} - A)^{-k}}{k!}.$$

As a consequence, we have

$$\begin{aligned} \|T_t^n\| &\leq e^{-nt} \sum_{k=0}^{+\infty} \frac{n^{2k} \|(n\mathbf{Id} - A)^{-1}\|^k}{k!} \\ &\leq e^{-nt} \sum_{k=0}^{+\infty} \frac{n^k t^k}{k!} \\ &\leq 1 \end{aligned}$$

and  $(T_t^n)_{t \geq 0}$  is therefore a contraction semigroup. The strong continuity is also easily checked:

$$\begin{aligned} \|T_{t+h}^n - T_t^n\| &= \|T_t^n (T_h^n - \mathbf{Id})\| \\ &\leq \|T_h^n - \mathbf{Id}\| \\ &\leq \sum_{k=1}^{+\infty} \frac{h^k \|A_n\|^k}{k!} \\ &\rightarrow_{h \rightarrow 0} 0. \end{aligned}$$

We now prove that for fixed  $t \geq 0$ ,  $x \in \mathcal{D}(A)$ ,  $(T_t^n x)_{n \geq 1}$  is a Cauchy sequence. We have

$$\begin{aligned} \|T_t^n x - T_t^m x\| &= \left\| \int_0^t \frac{d}{ds} (T_s^n T_{t-s}^m x) ds \right\| \\ &= \left\| \int_0^t T_s^n T_{t-s}^m (A_n x - A_m x) ds \right\| \\ &\leq \int_0^t \|A_n x - A_m x\| ds \\ &\leq t \|A_n x - A_m x\|. \end{aligned}$$

Therefore for  $x \in \mathcal{D}(A)$ ,  $(T_t^n x)_{n \geq 1}$  is a Cauchy sequence and we can define

$$T_t x = \lim_{n \rightarrow +\infty} T_t^n x.$$

Since  $\mathcal{D}(A)$  is dense and the family  $(T_t^n)_{n \geq 1}$  uniformly bounded, the above limit actually exists for every  $x \in B$ , so that  $(T_t)_{t \geq 0}$  is well-defined on  $B$ . It is clear that  $(T_t)_{t \geq 0}$  is a strongly continuous semigroup, inheriting these properties from  $(T_t^n)_{t \geq 0}$  (the details are let to the reader here).

It remains to show that the generator of  $(T_t)_{t \geq 0}$ , call it  $\tilde{A}$  is equal to  $A$ . For every  $t \geq 0$ ,  $x \in \mathcal{D}(A)$  and  $n \geq 1$ ,

$$T_t^n x = x + \int_0^t T_s^n A x ds,$$

therefore

$$T_t^n x = x + \int_0^t T_s^n A x ds.$$

Hence  $\mathcal{D}(A) \subset \mathcal{D}(\tilde{A})$  and for  $x \in \mathcal{D}(A)$ ,  $\tilde{A}x = Ax$ . Finally, since for  $\lambda > 0$ ,  $(\lambda \mathbf{Id} - A)\mathcal{D}(A) = B = (\lambda \mathbf{Id} - \tilde{A})\mathcal{D}(\tilde{A})$ , we conclude  $\mathcal{D}(A) = \mathcal{D}(\tilde{A})$ .  $\square$

**Exercise 4.39.** *By using the proof of Theorem 4.38, show the following fact: If  $A_1$  and  $A_2$  are the generators of contraction semigroups  $(T_t^1)_{t \geq 0}$  and  $(T_t^2)_{t \geq 0}$ , then for  $x \in B$ , the two following statements are equivalent:*

- $\forall \lambda > 0, \quad (\lambda \mathbf{Id} - A_1)^{-1}x = (\lambda \mathbf{Id} - A_2)^{-1}x;$
- $\forall t \geq 0, \quad T_t^1 x = T_t^2 x.$

As powerful as it is, the Hille-Yosida theorem is difficult to directly apply to the theory of diffusion semigroups. We shall need a corollary of it that is more suited to this case.

**Definition 4.40.** A densely defined operator on a Banach space  $B$  is called dissipative if for each  $x \in \mathcal{D}(A)$ , we can find an element  $\phi$  of the dual space  $B^*$ , such that:

- $\|\phi\| = \|x\|;$
- $\phi(x) = \|x\|^2;$
- $\phi(Ax) \leq 0.$

With this new definition in hands, we have the following corollary of the Hille-Yosida theorem:

**Corollary 4.41.** *A closed operator  $A$  on a Banach space  $B$  is the generator of a strongly continuous contraction semigroup, if and only if:*

- $A$  is dissipative;
- For  $\lambda > 0$ , the range of the operator  $\lambda \mathbf{Id} - A$  is  $B$ .

*Proof.* Let us first assume that  $A$  is the generator of a contraction semigroup  $(T_t)_{t \geq 0}$ . From the Hahn-Banach theorem, there exists  $\phi \in B^*$  such that  $\|\phi\| = \|x\|$  and  $\phi(x) = \|x\|^2$ . We have, at  $t = 0$ ,

$$\frac{d}{dt}\phi(T_t x) = \phi(Ax),$$

but

$$|\phi(T_t x)| \leq \|\phi\| \|T_t x\| \leq \|\phi\| \|x\| \leq \|x\|^2 \leq \phi(x),$$

thus, at  $t = 0$ ,

$$\frac{d}{dt}\phi(T_t x) \leq 0,$$

and we conclude

$$\phi(Ax) \leq 0.$$

The fact that for  $\lambda > 0$ , the range of the operator  $\lambda \mathbf{Id} - A$  is  $B$  is a straightforward consequence of Theorem 4.38.

Let us now assume that  $A$  is a densely defined closed operator such that:

- $A$  is dissipative;
- For  $\lambda > 0$ , the range of the operator  $\lambda \mathbf{Id} - A$  is  $B$ .

Let  $x \in \mathcal{D}(A)$  and let  $\phi \in B^*$ , such that:

- $\|\phi\| = \|x\|$ ;
- $\phi(x) = \|x\|^2$ ;
- $\phi(Ax) \leq 0$ .

For  $\lambda > 0$ ,

$$\begin{aligned} \lambda \|x\|^2 &\leq \lambda \phi(x) - \phi(Ax) \\ &\leq \phi((\lambda \mathbf{Id} - A)x) \\ &\leq \|x\| \|(\lambda \mathbf{Id} - A)x\|. \end{aligned}$$

Thus,

$$\|(\lambda \mathbf{Id} - A)x\| \geq \lambda \|x\|.$$

This implies that the range  $\mathcal{R}_\lambda$  of the operator  $\lambda \mathbf{Id} - A$  is closed and that this operator has a bounded inverse from  $\mathcal{R}_\lambda$  to  $\mathcal{D}(A)$  with norm lower than  $\frac{1}{\lambda}$ . Since  $\mathcal{R}_\lambda = B$ , the proof is complete.  $\square$

## 5.2 Applications to diffusion semigroups

After this digression on the theory of contraction semigroups on Banach spaces, we now come back to the case of diffusion semigroups. Let us recall that we consider an elliptic diffusion operator  $L$  on  $\mathbb{R}^n$  with smooth coefficients such that:

- There is a Borel measure  $\mu$ , symmetric and invariant for  $L$  on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ;
- There exists an increasing sequence  $h_n \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,  $0 \leq h_n \leq 1$ , such that  $h_n \nearrow 1$  on  $\mathbb{R}^n$ , and  $\|\Gamma(h_n, h_n)\|_\infty \rightarrow 0$ , as  $n \rightarrow \infty$ .

Let  $1 < p < +\infty$ . We want to apply the previous theorems for the operator  $L$  on the Banach space  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ . We denote by  $L^{(p)}$  the closure in  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$  of the operator  $L$ , densely defined on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ .

First, we have the following lemma.

**Lemma 4.42.** *Let  $1 < p \leq q < +\infty$  and  $\lambda > 0$ . If  $f \in \mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R}) + \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$  and satisfies in the sense of distributions*

$$Lf = \lambda f,$$

then  $f = 0$ .

*Proof.* We use an idea already present in the proof of Proposition 4.11. Let us first observe that if  $f \in \mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R}) + \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$ , satisfies in the sense of distributions

$$Lf = \lambda f,$$

then it is actually a smooth function, due to the ellipticity of  $L$ . Now, let  $h \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,  $h \geq 0$ , and let  $\phi$  be a smooth nonnegative function to be precised later. We have

$$\begin{aligned} \int_{\mathbb{R}^n} \Gamma(f, h^2\phi(f)f) d\mu &= - \int_{\mathbb{R}^n} fL(h^2\phi(f)f) d\mu \\ &= -\lambda \int_{\mathbb{R}^n} h^2\phi(f)f^2 d\mu \\ &\leq 0. \end{aligned}$$

But

$$\begin{aligned} \Gamma(f, h^2\phi(f)f) &= h^2\Gamma(f, \phi(f)f) + 2\phi(f)fh\Gamma(f, h) \\ &= h^2(\phi'(f)f + \phi(f))\Gamma(f, f) + 2\phi(f)fh\Gamma(f, h), \end{aligned}$$

therefore

$$\int_{\mathbb{R}^n} h^2(\phi'(f)f + \phi(f))\Gamma(f, f) + 2\phi(f)fh\Gamma(f, h) d\mu \leq 0,$$

which implies

$$\int_{\mathbb{R}^n} h^2(\phi'(f)f + \phi(f))\Gamma(f, f)d\mu \leq -2 \int_{\mathbb{R}^n} \phi(f)fh\Gamma(f, h)d\mu. \quad (4.2)$$

From the Cauchy Schwarz inequality, we have

$$\left| \int_{\mathbb{R}^n} \phi(f)fh\Gamma(f, h)d\mu \right| \leq 2\sqrt{\|\Gamma(h, h)\|_\infty} \sqrt{\int_{\mathbb{R}^n} \phi(f)h^2\Gamma(f, f)d\mu} \sqrt{\int_K \phi(f)f^2d\mu}, \quad (4.3)$$

where  $K$  is any compact set containing the support of  $h$ . Let us now assume that we may chose the function  $\phi$  in such a way that for every  $x \in \mathbb{R}$ ,

$$\phi'(x)x + \phi(x) \geq C\phi(x),$$

with  $C > 0$ . In that case, from (4.2) and (4.3), we have

$$\int_{\mathbb{R}^n} \phi(f)h^2\Gamma(f, f)d\mu \leq \frac{4}{C^2} \|\Gamma(h, h)\|_\infty \int_K \phi(f)f^2d\mu. \quad (4.4)$$

We now proceed to the choice for the function  $\phi$ . For  $0 < \varepsilon < 1$ , let

$$\begin{aligned} \phi_\varepsilon(x) &= |x|^{p-2}, & |x| &\geq 1 \\ &= (x^2 + \varepsilon^2)^{\frac{q}{2}-1}, & |x| &\leq 1 - \varepsilon. \end{aligned}$$

We then extend  $\phi_\varepsilon$  in a smooth way to  $\mathbb{R}$  in such a way that there is a constant  $C > 0$ , independent of  $\varepsilon$  such that for every  $x \in \mathbb{R}$ ,

$$\phi'_\varepsilon(x)x + \phi_\varepsilon(x) \geq C\phi_\varepsilon(x).$$

The proof of the possibility of such an extension is let to the reader.

By using the inequality (4.4) for  $\phi_\varepsilon$ , we obtain

$$\int_{\mathbb{R}^n} \phi_\varepsilon(f)h^2\Gamma(f, f)d\mu \leq \frac{4}{C^2} \|\Gamma(h, h)\|_\infty \int_K \phi_\varepsilon(f)f^2d\mu.$$

By letting  $\varepsilon \rightarrow 0$ , we draw the conclusion

$$\int_{\mathbb{R}^n} \phi_0(f)h^2\Gamma(f, f)d\mu \leq \frac{4}{C^2} \|\Gamma(h, h)\|_\infty \int_K \phi_0(f)f^2d\mu,$$

where

$$\begin{aligned} \phi_0(x) &= |x|^{p-2}, & |x| &\geq 1 \\ &= |x|^{q-2}, & |x| &\leq 1. \end{aligned}$$



Due to the assumption,  $f \in \mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R}) + \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$ , the function  $\phi_0$  is integrable on  $\mathbb{R}^n$ , thus

$$\int_{\mathbb{R}^n} \phi_0(f) h^2 \Gamma(f, f) d\mu \leq \frac{4}{C^2} \|\Gamma(h, h)\|_\infty \int_{\mathbb{R}^n} \phi_0(f) f^2 d\mu.$$

By using the previous inequality with an increasing sequence  $h_n \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,  $0 \leq h_n \leq 1$ , such that  $h_n \nearrow 1$  on  $\mathbb{R}^n$ , and  $\|\Gamma(h_n, h_n)\|_\infty \rightarrow 0$ , as  $n \rightarrow \infty$ , and letting  $n \rightarrow +\infty$ ,

$$\int_{\mathbb{R}^n} \phi_0(f) \Gamma(f, f) d\mu = 0.$$

As a consequence, since  $\mu$  charges the open sets, if  $f \neq 0$ ,  $\Gamma(f, f) = 0$ . Since  $f = \frac{1}{\lambda} Lf$ , we conclude that  $f = 0$ .  $\square$

We are now in position to prove the main result of this section. For  $1 < p < +\infty$ , we denote by  $L^{(p)}$  the closure in  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$  of the operator  $L$ , originally only densely defined on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ .

**Proposition 4.43.** *The operator  $L^{(p)}$  is the generator of a strongly continuous contraction semigroup  $(\mathbf{P}_t^{(p)})_{t \geq 0}$  on  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ . Moreover, for  $1 < p \leq q < +\infty$  and  $f \in \mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R}) \cap \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$ ,*

$$\mathbf{P}_t^{(p)} f = \mathbf{P}_t^{(q)} f.$$

*Proof.* We apply Corollary 4.41. We have to show that:

- $L^{(p)}$  is dissipative;
- For  $\lambda > 0$ , the range of the operator  $\lambda \mathbf{Id} - L^{(p)}$  is  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ .

We first check that  $L^{(p)}$  is dissipative. Since it is easily seen that the closure of a dissipative operator is dissipative, we have to check that  $L$  is dissipative on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ .

Let  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  and

$$g = \frac{1}{\|f\|_{\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})}^{p/q}} |f|^{p-2} f$$

where  $\frac{1}{p} + \frac{1}{q} = 1$ . We have:

- $\|g\|_{\mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})} = \|f\|_{\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})}$ ;
- $\int_{\mathbb{R}^n} g f d\mu = \|f\|_{\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})}^2$ .

Since  $\mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$  is the dual space of  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ , it remains to prove that  $\int_{\mathbb{R}^n} gLf d\mu \leq 0$ . For  $\varepsilon > 0$ , let

$$\Phi_\varepsilon(x) = (x^2 + \varepsilon^2)^{\frac{p}{2}-1}x.$$

We have

$$\int_{\mathbb{R}^n} \Phi_\varepsilon(f)Lf d\mu = - \int_{\mathbb{R}^n} \Gamma(\Phi_\varepsilon(f), f) d\mu = - \int_{\mathbb{R}^n} \Phi'_\varepsilon(f)\Gamma(f, f) d\mu \leq 0.$$

From the Lebesgue dominated convergence theorem applied when  $\varepsilon \rightarrow 0$ , this implies

$$\int_{\mathbb{R}^n} |f|^{p-2}fLf d\mu \leq 0.$$

As a conclusion,  $L$  is dissipative on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  and  $L^{(p)}$  is dissipative on its domain.

We now show that for  $\lambda > 0$ , the range of the operator  $\lambda\mathbf{Id} - L^{(p)}$  is  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ . If not, we could find a non zero  $g \in \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$ ,  $\frac{1}{p} + \frac{1}{q} = 1$  such that, in the sense of distributions,

$$Lg = \lambda g.$$

But, according to Lemma 4.42, the above equation implies  $g = 0$ . The range of the operator  $\lambda\mathbf{Id} - L^{(p)}$  is therefore  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ .

We conclude from Corollary 4.41, that  $L^{(p)}$  is the generator of a contraction semigroup  $(\mathbf{P}_t^{(p)})_{t \geq 0}$  on  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ .

We finally show that for  $1 < p \leq q < +\infty$ ,  $(\mathbf{P}_t^{(p)})_{t \geq 0}$  and  $(\mathbf{P}_t^{(q)})_{t \geq 0}$  coincide on  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R}) \cap \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$ . Let  $f \in \mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R}) \cap \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$ . For  $\lambda > 0$ , the function

$$g = (\lambda\mathbf{Id} - L^{(p)})^{-1}f - (\lambda\mathbf{Id} - L^{(q)})^{-1}f$$

is a function that belongs to  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R}) + \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$  and that satisfies in the sense of distributions

$$Lg = \lambda g.$$

From Lemma 4.42, we get  $g = 0$ . As a consequence for every  $\lambda > 0$ ,

$$(\lambda\mathbf{Id} - L^{(p)})^{-1}f = (\lambda\mathbf{Id} - L^{(q)})^{-1}f.$$

This implies that for  $t \geq 0$ ,  $\mathbf{P}_t^{(p)}f = \mathbf{P}_t^{(q)}f$  (See Exercise 4.39).  $\square$

**Exercise 4.44.** Show that on  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ , the semigroup  $(\mathbf{P}_t^{(2)})_{t \geq 0}$  coincides with the semigroup  $(\mathbf{P}_t)_{t \geq 0}$  that was constructed with the spectral theorem (You may use Exercise 4.39).

**Remark 4.45.** Let us observe that the Hille-Yosida methods fail in the limiting cases  $p = 1$ ,  $p = +\infty$ .

Since,  $(\mathbf{P}_t^{(p)})_{t \geq 0}$  and  $(\mathbf{P}_t^{(q)})_{t \geq 0}$  coincide on  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R}) \cap \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$ , there is no ambiguity and we can safely remove the superscript  $(p)$  from  $(\mathbf{P}_t^{(p)})_{t \geq 0}$ . So that we can use only the notation  $(\mathbf{P}_t)_{t \geq 0}$ , but will precise on which space it is defined if there is ambiguity. The following duality lemma is then easy to prove.

**Lemma 4.46.** *Let  $f \in \mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$  and let  $g \in \mathbf{L}_\mu^q(\mathbb{R}^n, \mathbb{R})$ , with  $\frac{1}{p} + \frac{1}{q} = 1$ . For  $t \geq 0$ ,*

$$\int_{\mathbb{R}^n} (\mathbf{P}_t f) g d\mu = \int_{\mathbb{R}^n} (\mathbf{P}_t g) f d\mu.$$

*Proof.* Let us first assume that  $g \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . The linear form  $\Phi : \mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R}) \rightarrow \mathbb{R}$ ,  $f \rightarrow \int_{\mathbb{R}^n} (\mathbf{P}_t f) g d\mu - \int_{\mathbb{R}^n} (\mathbf{P}_t g) f d\mu$  is continuous and vanishes on the dense subspace  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , it is therefore zero on  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ . This proves that for  $f \in \mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$  and  $g \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,

$$\int_{\mathbb{R}^n} (\mathbf{P}_t f) g d\mu = \int_{\mathbb{R}^n} (\mathbf{P}_t g) f d\mu.$$

The same density argument applied on the linear form  $g \rightarrow \int_{\mathbb{R}^n} (\mathbf{P}_t f) g d\mu - \int_{\mathbb{R}^n} (\mathbf{P}_t g) f d\mu$  provides the expected result.  $\square$

## 6 Diffusion semigroups as solutions of a parabolic Cauchy problem

In this section, we connect the semigroup associated to a diffusion operator  $L$  to the parabolic following Cauchy problem:

$$\frac{\partial u}{\partial t} = Lu, \quad u(0, x) = f(x).$$

As before, let  $L$  be a elliptic diffusion operator with smooth coefficients on  $\mathbb{R}^n$ , such that

- There is a Borel measure  $\mu$ , symmetric for  $L$  on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ;
- There exists an increasing sequence  $h_n \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,  $0 \leq h_n \leq 1$ , such that  $h_n \nearrow 1$  on  $\mathbb{R}^n$ , and  $\|\Gamma(h_n, h_n)\|_\infty \rightarrow 0$ , as  $n \rightarrow \infty$ .

**Proposition 4.47.** *Let  $f \in \mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ ,  $1 \leq p \leq \infty$ , and let*

$$u(t, x) = P_t f(x), \quad t \geq 0, x \in \mathbb{R}^n.$$

*Then  $u$  is smooth on  $(0, +\infty) \times \mathbb{R}^n$  and is a solution of the Cauchy problem*

$$\frac{\partial u}{\partial t} = Lu, \quad u(0, x) = f(x).$$

*Proof.* For  $\phi \in \mathcal{C}_c((0, +\infty) \times \mathbb{R}^n, \mathbb{R})$ , we have

$$\begin{aligned}
& \int_{\mathbb{R}^n \times \mathbb{R}} \left( \left( -\frac{\partial}{\partial t} - L \right) \phi(t, x) \right) u(t, x) d\mu(x) dt \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}^n} \left( \left( -\frac{\partial}{\partial t} - L \right) \phi(t, x) \right) P_t f(x) dx dt \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}^n} P_t \left( \left( -\frac{\partial}{\partial t} - L \right) \phi(t, x) \right) f(x) dx dt \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}^n} -\frac{\partial}{\partial t} (P_t \phi(t, x) f(x)) dx dt \\
&= 0.
\end{aligned}$$

Therefore  $u$  is a weak solution of the equation  $\frac{\partial u}{\partial t} = Lu$ . Since we already know that  $u$  is smooth it is also a strong solution.  $\square$

We now address the uniqueness question.

**Proposition 4.48.** *Let  $v(x, t)$  be a non negative function such that*

$$\frac{\partial v}{\partial t} \leq Lv, \quad v(x, 0) = 0,$$

and such that for every  $t > 0$ ,

$$\|v(\cdot, t)\|_{\mathbf{L}^p_\mu(\mathbb{R}^n, \mathbb{R})} < +\infty,$$

where  $1 < p < +\infty$ . Then  $v(x, t) = 0$ .

*Proof.* Let  $x_0 \in \mathbb{R}^n$  and  $h \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . Since  $v$  is a subsolution with the zero initial data, for any  $\tau \in (0, T)$ ,

$$\begin{aligned}
& \int_0^\tau \int_{\mathbb{R}^n} h^2(x) v^{p-1}(x, t) Lv(x, t) d\mu(x) dt \\
&\geq \int_0^\tau \int_{\mathbb{R}^n} h^2(x) v^{p-1} \frac{\partial v}{\partial t} d\mu(x) dt \\
&= \frac{1}{p} \int_0^\tau \frac{\partial}{\partial t} \left( \int_{\mathbb{R}^n} h^2(x) v^p d\mu(x) \right) dt \\
&= \frac{1}{p} \int_{\mathbb{R}^n} h^2(x) v^p(x, \tau) d\mu(x).
\end{aligned}$$

On the other hand, integrating by parts yields

$$\begin{aligned}
& \int_0^\tau \int_{\mathbb{R}^n} h^2(x) v^{p-1}(x, t) Lv(x, t) d\mu(x) dt \\
&= - \int_0^\tau \int_{\mathbb{R}^n} 2hv^{p-1} \Gamma(h, v) d\mu dt - \int_0^\tau \int_{\mathbb{R}^n} h^2(p-1)v^{p-2} \Gamma(v) d\mu dt.
\end{aligned}$$

Observing that

$$\begin{aligned} 0 &\leq \left( \sqrt{\frac{2}{p-1}}\Gamma(h)v - \sqrt{\frac{p-1}{2}}\Gamma(v)h \right)^2 \\ &\leq \frac{2}{p-1}\Gamma(h)v^2 + 2\Gamma(h,v)hv + \frac{p-1}{2}\Gamma(v)h^2, \end{aligned}$$

we obtain the following estimate.

$$\begin{aligned} &\int_0^\tau \int_{\mathbb{R}^n} h^2(x)v^{p-1}(x,t)Lv(x,t)d\mu(x)dt \\ &\leq \int_0^\tau \int_{\mathbb{R}^n} \frac{2}{p-1}\Gamma(h)v^p d\mu dt - \int_0^\tau \int_{\mathbb{R}^n} \frac{p-1}{2}h^2v^{p-2}\Gamma(v)d\mu dt \\ &= \int_0^\tau \int_{\mathbb{R}^n} \frac{2}{p-1}\Gamma(h)v^p d\mu dt - \frac{2(p-1)}{p^2} \int_0^\tau \int_{\mathbb{R}^n} h^2\Gamma(v^{p/2})d\mu dt. \end{aligned}$$

Combining with the previous conclusion we obtain ,

$$\begin{aligned} &\int_{\mathbb{R}^n} h^2(x)v^p(x,\tau)d\mu(x) + \frac{2(p-1)}{p} \int_0^\tau \int_{\mathbb{R}^n} h^2\Gamma(v^{p/2})d\mu dt \\ &\leq \frac{2p}{(p-1)}\|\Gamma(h)\|_\infty^2 \int_0^\tau \int_{\mathbb{R}^n} v^p d\mu dt. \end{aligned}$$

By using the previous inequality with an increasing sequence  $h_n \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,  $0 \leq h_n \leq 1$ , such that  $h_n \nearrow 1$  on  $\mathbb{R}^n$ , and  $\|\Gamma(h_n, h_n)\|_\infty \rightarrow 0$ , as  $n \rightarrow \infty$ , and letting  $n \rightarrow +\infty$ , we obtain  $\int_{\mathbb{R}^n} v^p(x, \tau)d\mu(x) = 0$  thus  $v = 0$ .  $\square$

As a consequence of this result, any solution in  $\mathbf{L}_\mu^p(\mathbb{R}^n, \mathbb{R})$ ,  $1 < p < +\infty$  of the heat equation  $\frac{\partial u}{\partial t} = Lu$  is uniquely determined by its initial condition, and is therefore of the form  $u(t, x) = P_t f(x)$  (apply the above lemma with the subsolution  $|u|$ ). We stress that without further conditions, this result fails when  $p = 1$  or  $p = +\infty$ .

## 7 The Dirichlet semigroup

In this section we will construct the Dirichlet semigroup associated to a symmetric diffusion operator which is given on some a relatively compact domain of  $\mathbb{R}^n$ . A later application of this construction will be the proof of the existence of a continuous Markov process associated with the given symmetric diffusion operator.

Let  $L$  be an elliptic diffusion operator on  $\mathbb{R}^n$  with smooth coefficients. We assume that  $L$ , defined on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , is symmetric with respect to a measure with smooth and positive density  $\mu$ . Let  $\Omega \subset \mathbb{R}^n$  be a non empty open set whose closure  $\bar{\Omega}$  is compact and whose boundary  $\partial\Omega$  is smooth. We recall that the following Green's identity holds for  $L$  (see for instance [?]):

**Proposition 4.49** (Green's identity). *There exists a never vanishing first order differential operator  $\nu$  defined on  $\partial\Omega$  such that for every  $C^2$  functions  $f, g : \bar{\Omega} \rightarrow \mathbb{R}$ ,*

$$\int_{\Omega} fLgd\mu = - \int_{\Omega} \Gamma(f, g)d\mu + \int_{\partial\Omega} f(\nu g)d\mu_S,$$

where, as usual,  $\Gamma(f, g) = \frac{1}{2}(L(fg) - fLg - gLf)$  and  $\mu_S$  is the surface measure induced by  $\mu$  on  $\partial\Omega$

The basic result to construct the Dirichlet semigroup associated to  $L$  on  $\Omega$  is the following fact.

**Proposition 4.50.** *The operator  $L$  is essentially self-adjoint on the set*

$$\mathcal{C}_0^\infty(\bar{\Omega}) = \{u : \bar{\Omega} \rightarrow \mathbb{R}, u \text{ smooth on } \bar{\Omega}, u = 0 \text{ on } \partial\Omega\}.$$

*Proof.* According to Lemma 4.9, it is enough to prove that

$$\mathbf{Ker}(-L^* + \mathbf{Id}) = \{0\}.$$

Let  $f \in \mathbf{Ker}(-L^* + \mathbf{Id}) = \{0\}$ . Since  $\mathcal{C}_0^\infty(\bar{\Omega})$  contains the space of smooth and compactly supported functions inside  $\Omega$ , we have in the sense of distributions  $Lf = f$ . By ellipticity of  $L$ , it implies that  $f$  is smooth in  $\Omega$ . From Green's identity, we now have for  $g \in \mathcal{C}_0^\infty(\bar{\Omega})$

$$\int_{\Omega} fLgd\mu - \int_{\Omega} gLf d\mu + \int_{\partial\Omega} f(\nu g)d\mu_S.$$

Since  $f \in \mathcal{D}(L^*)$ , the square of the right hand side must be controlled, for every  $g$ , by  $\|g\|_{\mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})}^2$ . This is only possible when  $f = 0$  on  $\partial\Omega$ . Thus,  $f \in \mathcal{C}_0^\infty(\bar{\Omega})$ . Since  $Lf = f$ , from Green's identity, we have

$$\int_{\Omega} f^2 d\mu = \int_{\Omega} fLf d\mu = - \int_{\Omega} \Gamma(f, f)d\mu \leq 0.$$

This implies  $f = 0$  and, as a conclusion,  $L$  is essentially self-adjoint on  $\mathcal{C}_0^\infty(\bar{\Omega})$ .  $\square$

The self-adjoint extension of the previous proposition is called the Dirichlet extension of  $L$  on  $\Omega$  and shall be denoted by  $L_\Omega$ . The semigroup generated by  $L_\Omega$  is called the Dirichlet semigroup and denoted by  $(\mathbf{P}_t^\Omega)_{t \geq 0}$ . By using the same Sobolev embedding techniques as in the previous sections (see Theorems 4.20 and 4.23), it is seen that the semigroup  $\mathbf{P}_t^\Omega$  has a smooth heat kernel  $p^\Omega(t, x, y)$ . It is then easy to see that for  $f \in \mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$ ,  $\mathbf{P}_t^\Omega f \in \mathcal{C}_0^\infty(\bar{\Omega})$ ,  $t > 0$ . In particular if  $x$  or  $y$  are on  $\partial\Omega$ , then  $p^\Omega(t, x, y) = 0$ ,  $t > 0$ . This also implies that  $\mathbf{P}_t^\Omega 1 \neq 1$ , that is  $\mathbf{P}_t^\Omega$  is a strictly sub-Markov semigroup.

It is important to observe that  $L_\Omega$  is not the unique self-adjoint extension of  $L$  on the space of smooth and compactly supported functions inside  $\Omega$ .

The Neumann extension presented in the following section is also particularly important.

It turns out that the compactness of  $\bar{\Omega}$  implies the compactness of the semigroup  $\mathbf{P}_t^\Omega$  (see Appendix A for further details on compact operators).

**Proposition 4.51.** *For  $t > 0$  the operator  $\mathbf{P}_t^\Omega$  is a compact operator on the Hilbert space  $\mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$ . It is moreover trace class and*

$$\mathbf{Tr}(\mathbf{P}_t^\Omega) = \int_{\Omega} p^\Omega(t, x, x) d\mu(x).$$

*Proof.* From the existence of the heat kernel, we have

$$\mathbf{P}_t^\Omega f(x) = \int_{\Omega} p^\Omega(t, x, y) f(y) d\mu(y).$$

But from the compactness of  $\bar{\Omega}$ , and the continuity of  $p(t, \cdot, \cdot)$  on  $\bar{\Omega} \times \bar{\Omega}$ , we have

$$\int_{\Omega} \int_{\Omega} p^\Omega(t, x, y)^2 d\mu(x) d\mu(y) < +\infty.$$

Therefore, from the Theorem 7.54 in the Appendix A

$$\mathbf{P}_t^\Omega : \mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R}) \rightarrow \mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$$

is a Hilbert-Schmidt operator. It is thus in particular a compact operator.

Since  $\mathbf{P}_t^\Omega = \mathbf{P}_{t/2}^\Omega \mathbf{P}_{t/2}^\Omega$ ,  $\mathbf{P}_t^\Omega$  is a product of two Hilbert-Schmidt operators. It is therefore a class trace operator and,

$$\mathbf{Tr}(\mathbf{P}_t^\Omega) = \int_{\Omega} \int_{\Omega} p^\Omega(t/2, x, y) p^\Omega(t/2, y, x) d\mu(x) d\mu(y).$$

We conclude then by applying the Chapman-Kolmogorov relation that

$$\mathbf{Tr}(\mathbf{P}_t^\Omega) = \int_{\Omega} p^\Omega(t, x, x) d\mu(x).$$

□

We have the following expansion of the Dirichlet heat kernel:

**Theorem 4.52.** *There exists a complete orthonormal basis  $(\phi_n)_{n \geq 1}$  of  $\mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$ , consisting of eigenfunctions of  $-L$ , with  $\phi_n \in \mathcal{C}_0^\infty(\bar{\Omega})$  having an eigenvalue  $\lambda_n$  with finite multiplicity and*

$$0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \leq \dots \nearrow +\infty.$$

Moreover, for  $t > 0$ ,  $x, y \in \bar{\Omega}$ ,

$$p^\Omega(t, x, y) = \sum_{n=1}^{+\infty} e^{-\lambda_n t} \phi_n(x) \phi_n(y),$$

where the convergence is absolute and uniform for each  $t > 0$ .

*Proof.* Let  $t > 0$ . From the Hilbert-Schmidt theorem (see Theorem 7.50 in Appendix A), for the non negative self adjoint compact operator  $\mathbf{P}_t^\Omega$ , there exists a complete orthonormal basis  $(\phi_n(t))_{n \geq 1}$  of  $\mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$  and a non increasing sequence  $\alpha_n(t) \geq 0$ ,  $\alpha_n(t) \searrow 0$  such that

$$\mathbf{P}_t^\Omega \phi_n(t) = \alpha_n(t) \phi_n(t).$$

The semigroup property  $\mathbf{P}_{t+s}^\Omega = \mathbf{P}_t^\Omega \mathbf{P}_s^\Omega$  implies first that for  $k \in \mathbb{N}$ ,  $k \geq 1$ ,

$$\phi_n(k) = \phi_n(1), \alpha_n(k) = \alpha_n(1)^k.$$

The same result is then seen to hold for  $k \in \mathbb{Q}$ ,  $k > 0$  and finally for  $k \in \mathbb{R}$ , due to the strong continuity of the semigroup. Since the map  $t \rightarrow \|\mathbf{P}_t^\Omega\|_{\mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})}$  is decreasing, we deduce that  $\alpha_n(1) \leq 1$ . Thus, there is a  $\lambda_n \geq 0$  such that

$$\alpha_n(1) = e^{-\lambda_n}.$$

As a conclusion, there exists a complete orthonormal basis  $(\phi_n)_{n \geq 1}$  of  $\mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$ , and a sequence  $\lambda_n$  satisfying

$$0 \leq \lambda_1 \leq \lambda_2 \leq \dots \nearrow +\infty,$$

such that

$$\mathbf{P}_t \phi_n = e^{-\lambda_n t} \phi_n.$$

If  $f \in \mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$  is such that  $\mathbf{P}_t^\Omega f = f$ , it is straightforward that  $f \in \mathcal{D}(L_\Omega)$  and that  $L_\Omega f = 0$ , so that  $f = 0$ . Therefore we have  $\lambda_1 > 0$ .

Since  $\mathbf{P}_t^\Omega \phi_n = e^{-\lambda_n t} \phi_n$ , by differentiating as  $t \rightarrow 0$  in  $\mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$ , we obtain furthermore that  $\phi_n \in \mathcal{D}(L_\Omega)$  and that  $L_\Omega \phi_n = -\lambda_n \phi_n$ . By ellipticity of  $L$ , we deduce that  $\phi_n \in \mathcal{C}_0^\infty(\bar{\Omega})$ .

The family  $(x, y) \rightarrow \phi_n(x) \phi_m(y)$  forms an orthonormal basis of  $\mathbf{L}_{\mu \otimes \mu}^2(\bar{\Omega} \times \bar{\Omega}, \mathbb{R})$ . We therefore have a decomposition in  $\mathbf{L}_{\mu \otimes \mu}^2(\bar{\Omega} \times \bar{\Omega}, \mathbb{R})$ ,

$$p^\Omega(t, x, y) = \sum_{m, n \in \mathbb{M}} c_{mn} \phi_m(x) \phi_n(y).$$

Since  $p^\Omega(t, \cdot, \cdot)$  is the kernel of the symmetric semigroup  $\mathbf{P}_t^\Omega$ , it is then straightforward that for  $m \neq n$ ,  $c_{mn} = 0$  and that  $c_{nn} = e^{-\lambda_n t}$ . Therefore in  $\mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$ ,

$$p^\Omega(t, x, y) = \sum_{n=1}^{+\infty} e^{-\lambda_n t} \phi_n(x) \phi_n(y).$$

The continuity of  $p^\Omega$ , together with the positivity of  $\mathbf{P}_t^\Omega$  imply, via Mercer's theorem that actually, the above series is absolutely and uniformly convergent for  $t > 0$ .  $\square$



As we stressed it in the statement of the theorem, in the decomposition

$$p^\Omega(t, x, y) = \sum_{n=1}^{+\infty} e^{-\lambda_n t} \phi_n(x) \phi_n(y),$$

the eigenvalue  $\lambda_n$  is repeated according to its multiplicity. It is often useful to rewrite this decomposition in the form

$$p^\Omega(t, x, y) = \sum_{n=1}^{+\infty} e^{-\alpha_n t} \sum_{k=1}^{d_n} \phi_k^n(x) \phi_k^n(y),$$

where the eigenvalue  $\alpha_n$  is not repeated, that is

$$0 < \alpha_1 < \alpha_2 < \dots$$

In this decomposition,  $d_n$  is the dimension of the eigenspace  $\mathcal{V}_n$  corresponding to the eigenvalue  $\alpha_n$  and  $(\phi_k^n)_{1 \leq k \leq d_n}$  is an orthonormal basis of  $\mathcal{V}_n$ . If we denote,

$$\mathcal{K}_n(x, y) = \sum_{k=1}^{d_n} \phi_k^n(x) \phi_k^n(y),$$

then  $\mathcal{K}_n$  is called the reproducing kernel of the eigenspace  $\mathcal{V}_n$ . It satisfies the following properties whose proofs are let to the reader:

**Proposition 4.53.**

- $\mathcal{K}_n$  does not depend on the choice of the basis  $(\phi_k^n)_{1 \leq k \leq d_n}$ ;
- If  $f \in \mathcal{V}_n$ , then  $\int_{\mathbb{M}} \mathcal{K}_n(x, y) f(y) d\mu(y) = f(x)$ .

We finally observe that from the very definition of the reproducing kernels, we have

$$p^\Omega(t, x, y) = \sum_{n=1}^{+\infty} e^{-\alpha_n t} \mathcal{K}_n(x, y).$$

**Exercise 4.54.**

1. Let  $f \in \mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$ . Show that uniformly on  $\bar{\Omega}$ , when  $t \rightarrow +\infty$ ,  $\mathbf{P}_t^\Omega f \rightarrow 0$ .
2. Let  $f \in \mathcal{C}_0^\infty(\bar{\Omega})$ . Show that uniformly on  $\bar{\Omega}$ , when  $t \rightarrow 0$ ,  $\mathbf{P}_t^\Omega f \rightarrow f$ .

**Exercise 4.55.** Show that if  $K$  is a compact subset of  $\Omega$ ,

$$\limsup_{t \rightarrow 0} \sup_{x \in K} \frac{1}{t} (\mathbf{P}_t^\Omega 1_{\mathbf{B}(x, \varepsilon)^c \cap \Omega})(x) = 0.$$

## 8 The Neumann semigroup

In this section we consider another semigroup which is canonically associated to diffusion operators on sub-domains of  $\mathbb{R}^n$ : The Neumann semigroup. The difference with respect to the Dirichlet semigroup is the boundary condition.

The framework is identical to the framework of the previous section. Let  $L$  be an elliptic diffusion operator on  $\mathbb{R}^n$  with smooth coefficients. We assume that  $L$ , defined on  $C_c(\mathbb{R}^n, \mathbb{R})$ , is symmetric with respect to a measure with smooth and positive density  $\mu$ . Let  $\Omega \subset \mathbb{R}^n$  be a non empty open set whose closure  $\bar{\Omega}$  is compact and whose boundary  $\partial\Omega$  is smooth. We list the main results concerning the Neumann semigroup but leave the arguments as an exercise to the reader since the proofs are almost identical to the ones of the previous section.

**Proposition 4.56.** *The operator  $L$  is essentially self-adjoint on the set*

$$\mathcal{C}_\nu^\infty(\bar{\Omega}) = \{u : \bar{\Omega} \rightarrow \mathbb{R}, u \text{ smooth on } \bar{\Omega}, \nu u = 0 \text{ on } \partial\Omega\},$$

where, as before,  $\nu$  denotes the vector field in the Green's formula Proposition 4.49. The self-adjoint extension of  $L$  on that set is called the Neumann extension of  $L$  and will be denoted by  $L_{\Omega, \nu}$ .

The semigroup generated by  $L_{\Omega, \nu}$  is called the Neumann semigroup and we will denote it by  $(\mathbf{P}_t^{\Omega, \nu})_{t \geq 0}$ . The semigroup  $\mathbf{P}_t^{\Omega, \nu}$  has a smooth heat kernel  $p^{\Omega, \nu}(t, x, y)$  and for  $f \in \bar{\mathbf{L}}_\mu^2(\bar{\Omega}, \mathbb{R})$ ,  $\mathbf{P}_t^\Omega f \in \mathcal{C}_\nu^\infty(\bar{\Omega})$ ,  $t > 0$ . The main difference with the Dirichlet semigroup is that the Neumann semigroup is a Markov semigroup, that is  $\mathbf{P}_t^{\Omega, \nu} 1 = 1$ . It comes from the fact that  $1 \in \mathcal{D}(L_{\Omega, \nu})$ .

**Proposition 4.57.** *For  $t > 0$  the operator  $\mathbf{P}_t^{\Omega, \nu}$  is a compact operator on the Hilbert space  $\bar{\mathbf{L}}_\mu^2(\bar{\Omega}, \mathbb{R})$ . It is moreover trace class and*

$$\text{Tr}(\mathbf{P}_t^{\Omega, \nu}) = \int_{\Omega} p^{\Omega, \nu}(t, x, x) d\mu(x).$$

**Theorem 4.58.** *There exists a complete orthonormal basis  $(\psi_n)_{n \geq 1}$  of  $\bar{\mathbf{L}}_\mu^2(\bar{\Omega}, \mathbb{R})$ , consisting of eigenfunctions of  $-L$ , with  $\psi_n \in \mathcal{C}_\nu^\infty(\bar{\Omega})$  having an eigenvalue  $\lambda_n$  with finite multiplicity and  $\psi_1 = \frac{1}{\sqrt{\mu(\Omega)}}$ ,*

$$0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_n \leq \dots \nearrow +\infty.$$

Moreover, for  $t > 0$ ,  $x, y \in \bar{\Omega}$ ,

$$p^{\Omega, \nu}(t, x, y) = \sum_{n=1}^{+\infty} e^{-\lambda_n t} \psi_n(x) \psi_n(y),$$

where the convergence is absolute and uniform for each  $t > 0$ .

**Exercise 4.59.**

1. Let  $f \in \mathbf{L}_\mu^2(\bar{\Omega}, \mathbb{R})$ . Show that uniformly on  $\bar{\Omega}$ , when  $t \rightarrow +\infty$ ,  $\mathbf{P}_t^{\Omega, \nu} f \rightarrow \frac{1}{\mu(\bar{\Omega})} \int_{\Omega} f d\mu$ .
2. Let  $f \in \mathcal{C}_\nu^\infty(\bar{\Omega})$ . Show that uniformly on  $\bar{\Omega}$ , when  $t \rightarrow 0$ ,  $\mathbf{P}_t^{\Omega, \nu} f \rightarrow f$ .

**Exercise 4.60.** Show that if  $K$  is a compact subset of  $\Omega$ ,

$$\limsup_{t \rightarrow 0} \sup_{x \in K} \frac{1}{t} (\mathbf{P}_t^{\Omega, \nu} \mathbf{1}_{\mathbf{B}(x, \varepsilon)^c \cap \Omega})(x) = 0.$$

## 9 Symmetric diffusion processes

Our goal in this section will be to associate to a symmetric and elliptic diffusion operator  $L$  a somehow canonical continuous Markov process. In general, as we have seen,  $L$  does not generate a Markov semigroup but a sub-Markov semigroup. The corresponding process will thus be a sub-Markov process (see the end of Section 1 in Chapter 3 for a definition).

Let  $L$  be an elliptic diffusion operator on  $\mathbb{R}^n$  with smooth coefficients. We assume that  $L$ , defined on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , is symmetric with respect to a measure  $\mu$  which has a smooth and positive density with respect to the Lebesgue measure.

**Proposition 4.61** (Neumann process). *Let  $\Omega \subset \mathbb{R}^n$  be a non empty open set whose closure  $\bar{\Omega}$  is compact and whose boundary  $\partial\Omega$  is smooth. For every  $x \in \Omega$ , there exists a continuous Markov process  $(X_t^x)_{t \geq 0}$  with semigroup  $(P_t^{\Omega, \nu})_{t \geq 0}$  (Neumann semigroup) such that  $X_0^x = x$ . Moreover, for every  $f \in \mathcal{C}_c(\Omega, \mathbb{R})$ , the process*

$$f(X_t^x) - \int_0^t Lf(X_s^x) ds$$

*is a martingale.*

*Proof.* We can use the same arguments as in the proofs of Theorem 3.28, Proposition 3.30 and Proposition 3.36.  $\square$

**Proposition 4.62** (Dirichlet process). *Let  $\Omega \subset \mathbb{R}^n$  be a non empty open set whose closure  $\bar{\Omega}$  is compact and whose boundary  $\partial\Omega$  is smooth. For every  $x \in \Omega$ , there exists a continuous sub-Markov process  $(X_t^x)_{t \geq 0}$  with semigroup  $(P_t^\Omega)_{t \geq 0}$  (Dirichlet semigroup) and extinction time  $\mathbf{e}_\Omega(x) = \inf\{t \geq 0, X_t^x \in \partial\Omega\}$ . Moreover, for every  $f \in \mathcal{C}_c(\Omega, \mathbb{R})$ , the process*

$$f(X_t^x) - \int_0^t Lf(X_s^x) ds, \quad t < \mathbf{e}_\Omega(x)$$

*is a martingale.*

*Proof.* Let  $\mathcal{O}$  be a non empty open set whose closure  $\bar{\mathcal{O}}$  is compact and whose boundary  $\partial\mathcal{O}$  is smooth. We also assume that  $\Omega \subset \mathcal{O}$ . Let  $x \in \Omega$  and let  $(Y_t^x)_{t \geq 0}$  be the Neumann process in  $\mathcal{O}$  started at  $x$ . Let  $T_x = \inf\{t \geq 0, Y_t^x \in \partial\Omega\}$  and consider the killed process

$$X_t^x = \begin{cases} Y_t^x, & t \leq T_x \\ \star, & t > T_x. \end{cases}$$

It is easily seen that  $(X_t^x)_{t \geq 0}$  is a sub-Markov process with semigroup

$$\mathbf{Q}_t f(x) = \mathbb{E}(f(X_t^x) 1_{t < T_x}).$$

Moreover, for every  $f \in \mathcal{C}_c(\mathcal{O}, \mathbb{R})$ , the process

$$f(Y_t^x) - \int_0^t Lf(Y_s^x) ds$$

is a martingale. Let now  $f$  be a smooth function on  $\Omega$  such  $f$  and  $Lf$  vanish on  $\partial\Omega$ . From the Doob's stopping theorem, we have

$$\mathbb{E} \left( f(Y_{t \wedge T_x}^x) - \int_0^{t \wedge T_x} Lf(Y_s^x) ds \right) = f(x)$$

But,

$$\mathbb{E} (f(Y_{t \wedge T_x}^x)) = \mathbb{E} (f(X_t^x) 1_{t < T_x})$$

and

$$\mathbb{E} \left( \int_0^{t \wedge T_x} Lf(Y_s^x) ds \right) = \mathbb{E} \left( \int_0^t Lf(X_s^x) 1_{s < T_x} ds \right),$$

thus we obtain

$$\mathbf{Q}_t f(x) = f(x) + \int_0^t \mathbf{Q}_s Lf(x) ds.$$

Applying this with an eigenfunction  $\Psi_n$  of the Dirichlet Laplacian  $L_\Omega$ , we get that

$$\mathbf{Q}_t \Psi_n = \Psi_n - \lambda_n \int_0^t \mathbf{Q}_s \Psi_n ds.$$

This implies  $\mathbf{Q}_t \Psi_n = e^{-\lambda_n t} \Psi_n$ . Using then the spectral decomposition of the Dirichlet semigroup  $\mathbf{P}_t^\Omega$  we easily conclude that, actually,  $\mathbf{Q}_t = \mathbf{P}_t^\Omega$ .  $\square$

With these two preliminaries in hands, we are now in position to prove the main result of this section.

**Theorem 4.63.** *For every  $x \in \mathbb{R}^n$ , there exists a continuous sub-Markov process  $(X_t^x)_{t \geq 0}$ , such that  $X_0^x = x$ , a.s. and for every  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , the process*

$$f(X_t^x) - \int_0^t Lf(X_s^x) ds, \quad t < \mathbf{e}(x),$$

*is a martingale. If  $L$  is moreover essentially self-adjoint on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , then  $(X_t^x)_{t \geq 0}$  admits the semigroup generated by  $L$  as a transition semigroup.*

*Proof.* Let  $x \in \mathbb{R}^n$ . Let us denote by  $\mathbf{B}_n = \mathbf{B}(x, n)$  the open ball with center  $x$  and radius  $n$ . Consider now independent continuous sub-Markov processes  $(X_t^{y,n})_{t \geq 0, y \in \mathbf{B}_n}$ ,  $n \geq 1$ , such that the transition semigroup of  $(X_t^{y,n})_{t \geq 0}$  is the Dirichlet semigroup  $\mathbf{P}_t^{\mathbf{B}_n}$  and  $X_0^{y,n} = y$ . Let

$$\mathbf{T}_1 = \inf\{t \geq 0, X_t^{x,1} \in \partial\mathbf{B}_1\}$$

and

$$X_{\partial\mathbf{B}_1}^1 = \lim_{t \rightarrow \mathbf{T}_1} X_t^{x,1}.$$

We define then by induction

$$\mathbf{T}_n = \inf\left\{t \geq 0, X_t^{X_{\partial\mathbf{B}_{n-1},n}^{n-1}} \in \partial\mathbf{B}_n\right\}$$

and

$$X_{\partial\mathbf{B}_n}^n = \lim_{t \rightarrow \mathbf{T}_n} X_t^{X_{\partial\mathbf{B}_{n-1},n}^{n-1}}.$$

We finally define the sub-Markov process  $(X_t^x)_{t \geq 0}$  as follows: For  $0 \leq t < \mathbf{T}_1$ ,  $X_t^x = X_t^{1,x}$ , and for  $\mathbf{T}_n \leq t < \mathbf{T}_{n+1}$ ,  $X_t^x = X_t^{X_{\partial\mathbf{B}_n}^n}$ . By the very construction of  $(X_t^x)_{t \geq 0}$ , it is clear that it is a continuous sub-Markov process such that for every  $n \geq 1$ ,

$$\mathbf{T}_n = \inf\{t \geq 0, X_t^x \in \partial\mathbf{B}_n\},$$

and such that the killed process defined by

$$\tilde{X}_t = \begin{cases} X_t^x, & t < \mathbf{T}_n \\ \star, & t \geq \mathbf{T}_n \end{cases}$$

is a sub-Markov process with semigroup  $\mathbf{P}_t^{\mathbf{B}_n}$ . In particular, for every  $f \in \mathcal{C}_c(\mathbf{B}_n, \mathbb{R})$ , the process

$$f(X_t^x) - \int_0^t Lf(X_s^x) ds, \quad t < \mathbf{T}_n,$$

is a martingale. This easily implies that for every  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ , the process

$$f(X_t^x) - \int_0^t Lf(X_s^x) ds, \quad t < \mathbf{e}(x), \quad (4.5)$$

is a martingale. This proves the first part of the theorem.

Let us now assume that  $L$  is moreover essentially self-adjoint on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . Let us denote by  $\mathbf{Q}_t$  the transition semigroup of  $(X_t^x)_{t \geq 0}$ . The construction of  $(X_t^x)_{t \geq 0}$  shows that  $\mathbf{Q}_t$  is limit in  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  of the self-adjoint operators  $\mathbf{P}_t^{\mathbf{B}^n}$ . This implies that  $\mathbf{Q}_t$  is a contraction semigroup in  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$ . Let  $\bar{L}$  be the generator of this semigroup. By using the martingale (4.5), we easily see that the domain of  $\bar{L}$  contains  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  and that for  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ ,

$$\bar{L}f = Lf.$$

The operator  $\bar{L}$  is therefore the unique self-adjoint extension of  $L$ .  $\square$

## Notes and Comments

**Section 1.** For a detailed account on the theory of self-adjointness on Hilbert spaces we refer the reader to the books by Reed and Simon Vol. I, [?] and Vol. II [?]. The criterion for essential self-adjointness given in Proposition 4.11 appears in a note by Bakry [?] and relies on the proof by Strichartz [?] that the Laplace-Beltrami operator on a complete Riemannian manifold is essentially self-adjoint.

**Section 2.** The existence of the heat kernel can also be proved by using the parametrix method, see the book by Friedman [?]. Our method that relies on a Sobolev embedding theorem has the advantage to extend to more general operators like subelliptic operators.

**Section 3.** The fact that the Kato inequality for the generator is equivalent to the positivity preserving property of the corresponding contraction semigroup is pointed out by Simon in [?] and relies on the theory of Dirichlet forms developed by Beurling and Deny in [?]. Dirichlet forms provide an abstract framework which allows to generalize some of the results of this section to much more general situations, see for instance Sturm [?] for the analysis of Dirichlet forms on metric spaces. We refer the reader to the book by Fukushima-Oshima-Takeda [?] for an extensive account on the general theory of Dirichlet forms and to the Chapter 1 of the book by E.B. Davies [?] for an introduction.

**Section 4.** The Riesz-Thorin theorem was first proved by Riesz in [?] but the idea to use complex methods to prove the theorem goes back to his student Thorin. An extension of this theorem to analytic families is due to Stein [?].

**Section 5.** The Hille-Yosida theorem (Theorem 4.38) was independently proved by Hille in [?], [?] and Yosida in [?], [?]. The construction of the heat

semigroup by using Hille-Yosida theory follows the approach of Strichartz [?] but is extended here in the context of general elliptic operators.

**Section 6.** The connection between partial differential equations and semigroup theory goes back to Hadamard [?], [?], [?] who observed the semigroup property for solutions of the Cauchy problem. However semigroup theory was not applied systematically to partial differential equations until Hille and Yosida developed the analytical tools in the late 1940's. One of the seminal fundamental papers applying semigroup techniques to diffusion equations is the paper by Feller [?].

**Section 9.** By using the theory of Dirichlet forms, it is possible to greatly generalize Theorem 4.63. We refer the interested reader to the book by Fukushima-Oshima-Takeda [?] for the construction of continuous Markov processes associated with local Dirichlet forms. We also mention the reference [?].

## Chapter 5

### Itô calculus

The main goal of the chapter is to construct an integral with respect to the Brownian motion. This integral may not be defined as a usual Riemann-Stieltjes integral because Brownian paths are almost surely of infinite variation. However, by using the full strength of probability methods, it is possible to develop a natural and fruitful integration theory for stochastic processes. Once the stochastic integral is constructed, we will prove the change of variable for this integral: The Döblin-Itô formula. Several applications of the Döblin-Itô formula are then investigated.

#### 1 Variation of the Brownian paths

We first study some properties of the Brownian paths that will be useful to develop an integral with respect to Brownian motion. As a first step, we prove that Brownian paths almost surely have an infinite variation,

If

$$\Delta_n[0, t] = \{0 = t_0^n \leq t_1^n \leq \dots \leq t_n^n = t\}$$

is a subdivision of the time interval  $[0, t]$ , we denote by

$$|\Delta_n[0, t]| = \max\{|t_{k+1}^n - t_k^n|, k = 0, \dots, n-1\},$$

the mesh of this subdivision.

**Proposition 5.1.** *Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. Let  $t \geq 0$ . For every sequence  $\Delta_n[0, t]$  of subdivisions such that*

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

*the following convergence takes place in  $L^2$  (and thus in probability),*

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n (B_{t_k^n} - B_{t_{k-1}^n})^2 = t.$$

*As a consequence, almost surely, Brownian paths have an infinite variation on the time interval  $[0, t]$ .*

*Proof.* Let us denote

$$V_n = \sum_{k=1}^n (B_{t_k^n} - B_{t_{k-1}^n})^2.$$



Thanks to the stationarity and the independence of Brownian increments, we have:

$$\begin{aligned}
& \mathbb{E}((V_n - t)^2) \\
&= \mathbb{E}(V_n^2) - 2t\mathbb{E}(V_n) + t^2 \\
&= \sum_{j,k=1}^n \mathbb{E}\left(\left(B_{t_j^n} - B_{t_{j-1}^n}\right)^2 \left(B_{t_k^n} - B_{t_{k-1}^n}\right)^2\right) - t^2 \\
&= \sum_{k=1}^n \mathbb{E}\left(\left(B_{t_j^n} - B_{t_{j-1}^n}\right)^4\right) + 2 \sum_{1 \leq j < k \leq n} \mathbb{E}\left(\left(B_{t_j^n} - B_{t_{j-1}^n}\right)^2 \left(B_{t_k^n} - B_{t_{k-1}^n}\right)^2\right) - t^2 \\
&= \sum_{k=1}^n (t_k^n - t_{k-1}^n)^2 \mathbb{E}(B_1^4) + 2 \sum_{1 \leq j < k \leq n} (t_j^n - t_{j-1}^n)(t_k^n - t_{k-1}^n) - t^2 \\
&= 3 \sum_{k=1}^n (t_j^n - t_{j-1}^n)^2 + 2 \sum_{1 \leq j < k \leq n} (t_j^n - t_{j-1}^n)(t_k^n - t_{k-1}^n) - t^2 \\
&= 2 \sum_{k=1}^n (t_k^n - t_{k-1}^n)^2 \\
&\leq 2t \mid \Delta_n[0, t] \mid \xrightarrow{n \rightarrow +\infty} 0.
\end{aligned}$$

Let us now prove that, as a consequence of this convergence, the paths of the process  $(B_t)_{t \geq 0}$  almost surely have an infinite variation on the time interval  $[0, t]$ . It suffices to prove that there exists a sequence of subdivisions  $\Delta_n[0, t]$  such that almost surely

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n |B_{t_k^n} - B_{t_{k-1}^n}| = +\infty.$$

Reasoning by absurd, let us assume that the supremum on all the subdivisions of the time interval  $[0, t]$  of the sums

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n |B_{t_k^n} - B_{t_{k-1}^n}|$$

may be bounded from above by some positive  $M$ . From the above result, since the convergence in probability implies the existence of an almost surely convergent subsequence, we can find a sequence of subdivisions  $\Delta_n[0, t]$  whose mesh tends to 0 and such that almost surely,

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n \left(B_{t_k^n} - B_{t_{k-1}^n}\right)^2 = t.$$

We get then

$$\sum_{k=1}^n \left( B_{t_k^n} - B_{t_{k-1}^n} \right)^2 \leq M \sup_{1 \leq k \leq n} |B_{t_k^n} - B_{t_{k-1}^n}| \xrightarrow{n \rightarrow +\infty} 0,$$

which is clearly absurd.  $\square$

**Exercise 5.2.** Let  $(B_t)_{t \geq 0}$  be a Brownian motion. Show that for  $t \geq 0$ , almost surely

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^{2^n} \left( B_{\frac{kt}{2^n}} - B_{\frac{(k-1)t}{2^n}} \right)^2 = t.$$

## 2 Itô integral

Since a Brownian motion  $(B_t)_{t \geq 0}$  does not have absolutely continuous paths, we can not directly use the theory of Riemann-Stieltjes integrals to give a sense to integrals like  $\int_0^t \Theta_s dB_s$  for every continuous stochastic process  $(\Theta_s)_{s \geq 0}$ . However, if  $(\Theta_s)_{s \geq 0}$  is regular enough in the Hölder sense, then  $\int_0^t \Theta_s dB_s$  can still be constructed as a limit of Riemann sums by using the so-called Young integral. In the sequel, we shall denote by  $\mathcal{C}^\alpha(I)$  the space of  $\alpha$ -Hölder continuous functions that are defined on an interval  $I$ .

**Theorem 5.3** (Young integral). *Let  $f \in \mathcal{C}^\beta([0, T])$  and  $g \in \mathcal{C}^\gamma([0, T])$ . If  $\beta + \gamma > 1$ , then for every subdivision  $\Delta_n[0, T]$ , whose mesh tends to 0, the Riemann sums*

$$\sum_{i=0}^{n-1} f(t_i^n) (g(t_{i+1}^n) - g(t_i^n))$$

converge, when  $n \rightarrow \infty$  to a limit which is independent of the subdivision  $t_i^n$ . This limit is denoted  $\int_0^T f dg$  and called the Young's integral of  $f$  with respect to  $g$ .

As a consequence of the previous result, we can therefore use Young's integral to give a sense to the integral  $\int_0^t \Theta_s dB_s$  as soon as the stochastic process  $(\Theta_s)_{s \geq 0}$  has  $\gamma$ -Hölder paths with  $\gamma > 1/2$ . This is not satisfying enough, since for instance the integral  $\int_0^t B_s dB_s$  is not even well defined. The alternative to using Riemann sums and study its almost sure convergence is to take advantage of the quadratic variation of the Brownian motion paths and use the full power of probabilistic methods.

In what follows, we consider a Brownian motion  $(B_t)_{t \geq 0}$  which is defined on a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ .  $(B_t)_{t \geq 0}$  is assumed to be adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . We also assume that  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  satisfies the usual conditions as they were defined in Definition 1.43. These assumptions imply in particular the following facts that we record here for later use :

1. A limit (in  $L^2$  or in probability) of adapted processes is still adapted;
2. A modification of a progressively measurable process is still a progressively measurable process.

**Exercise 5.4.** Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion. We denote by  $(\mathcal{F}_t^B)_{t \geq 0}$  its natural filtration:  $\mathcal{F}_\infty^B = \sigma(B_u, u \geq 0)$  and by  $\mathcal{N}$  the null sets of  $\mathcal{F}_\infty^B$ . Show that the filtration  $(\sigma(\mathcal{F}_t^B, \mathcal{N}))_{t \geq 0}$  satisfies the usual conditions.

We denote by  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  the set of processes  $(u_t)_{t \geq 0}$  that are progressively measurable with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  and such that

$$\mathbb{E} \left( \int_0^{+\infty} u_s^2 ds \right) < +\infty.$$

**Exercise 5.5.** Show that the space  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  endowed with the norm

$$\|u\|^2 = \mathbb{E} \left( \int_0^{+\infty} u_s^2 ds \right)$$

is a Hilbert space.

We now denote by  $\mathcal{E}$  the set of processes  $(u_t)_{t \geq 0}$  that may be written as:

$$u_t = \sum_{i=0}^{n-1} F_i 1_{(t_i, t_{i+1}]}(t),$$

where  $0 \leq t_0 \leq \dots \leq t_n$  and where  $F_i$  is a random variable that is measurable with respect to  $\mathcal{F}_{t_i}$  and such that  $\mathbb{E}(F_i^2) < +\infty$ . The set  $\mathcal{E}$  is often called the set of simple previsible processes. We first observe that it is straightforward to check that

$$\mathcal{E} \subset L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}).$$

The following theorem provides the basic definition of the so-called Itô integral.

**Theorem 5.6** (Itô integral). *There is a unique linear map*

$$\mathcal{I} : L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}) \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P})$$

such that:

1. For  $u = \sum_{i=0}^{n-1} F_i 1_{(t_i, t_{i+1}]} \in \mathcal{E}$ ,

$$\mathcal{I}(u) = \sum_{i=0}^{n-1} F_i (B_{t_{i+1}} - B_{t_i});$$

2. For  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ ,

$$\mathbb{E} (\mathcal{I}(u)^2) = \mathbb{E} \left( \int_0^{+\infty} u_s^2 ds \right).$$

The map  $\mathcal{I}$  is called the Itô integral and for  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , we will use the notation

$$\mathcal{I}(u) = \int_0^{+\infty} u_s dB_s.$$

*Proof.* Since  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  endowed with the norm

$$\|u\|^2 = \mathbb{E} \left( \int_0^{+\infty} u_s^2 ds \right)$$

is a Hilbert space, from the isometries extension theorem we just have to prove that

1. For  $u = \sum_{i=0}^{n-1} F_i 1_{(t_i, t_{i+1}]} \in \mathcal{E}$ ,

$$\mathbb{E} \left( \left( \sum_{i=0}^{n-1} F_i (B_{t_{i+1}} - B_{t_i}) \right)^2 \right) = \mathbb{E} \left( \int_0^{+\infty} u_s^2 ds \right);$$

2. The set  $\mathcal{E}$  is dense in  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .

Let  $u = \sum_{i=0}^{n-1} F_i 1_{(t_i, t_{i+1}]} \in \mathcal{E}$ . Due to the independence of the Brownian motion increments, we have:

$$\begin{aligned} & \mathbb{E} \left( \left( \sum_{i=0}^{n-1} F_i (B_{t_{i+1}} - B_{t_i}) \right)^2 \right) \\ &= \mathbb{E} \left( \sum_{i,j=0}^{n-1} F_i F_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j}) \right) \\ &= \mathbb{E} \left( \sum_{i=0}^{n-1} F_i^2 (B_{t_{i+1}} - B_{t_i})^2 \right) + 2 \mathbb{E} \left( \sum_{0 \leq i < j \leq n-1} F_i F_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j}) \right) \\ &= \mathbb{E} \left( \sum_{i=0}^{n-1} F_i^2 (t_{i+1} - t_i) \right) \\ &= \mathbb{E} \left( \int_0^{+\infty} u_s^2 ds \right). \end{aligned}$$

Let us now prove that  $\mathcal{E}$  is dense in  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . We proceed in several steps. As a first step, let us observe that the set of progressively

measurable bounded processes is dense in  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . Indeed, for  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , the sequence  $(u_t \mathbf{1}_{[0, n]}(|u_t|))_{t \geq 0}$  converges to  $u$ .

As a second step, we remark that if  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is a bounded process, then  $u$  is a limit of bounded processes that are in  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and such that almost every paths are supported in a fixed compact set (consider the sequence  $(u_t \mathbf{1}_{[0, n]}(t))_{t \geq 0}$ ).

As a third step, if  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is a bounded process such that almost every paths are supported in a fixed compact set, then the sequence  $\left(\frac{1}{n} \int_{t-\frac{1}{n}}^t u_s ds \mathbf{1}_{(\frac{1}{n}, +\infty)}(t)\right)_{t \geq 0}$  is seen to converge toward  $u$ . Therefore,  $u$  is a limit of left continuous and bounded processes that are in  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and such that almost every paths are supported in a fixed compact set.

Finally, it suffices to prove that if  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  is a left continuous and bounded process such that almost every paths are supported in a fixed compact set, then  $u$  is a limit of processes that belong to  $\mathcal{E}$ . This may be proved by considering the sequence:

$$u_t^n = \sum_{i=0}^{+\infty} u_{\frac{i}{n}} \mathbf{1}_{(\frac{i}{n}, \frac{i+1}{n}]}(t).$$

□

**Exercise 5.7.** Let  $u, v \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , show that

$$\mathbb{E} \left( \int_0^{+\infty} u_s dB_s \right) = 0$$

and

$$\mathbb{E} \left( \int_0^{+\infty} u_s dB_s \int_0^{+\infty} v_s dB_s \right) = \mathbb{E} \left( \int_0^{+\infty} u_s v_s ds \right).$$

Associated with Itô's integral, we can construct an integral process, its fundamental property is that it is a continuous martingale.

**Proposition 5.8.** Let  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . The process

$$\left( \int_0^t u_s dB_s \right)_{t \geq 0} = \left( \int_0^{+\infty} u_s \mathbf{1}_{[0, t]}(s) dB_s \right)_{t \geq 0}$$

is a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  that admits a continuous modification.

*Proof.* We first prove the martingale property. If

$$u_t = \sum_{i=0}^{n-1} F_i \mathbf{1}_{(t_i, t_{i+1}]}(t)$$

is in  $\mathcal{E}$ , then for every  $t \geq s$ ,

$$\begin{aligned} \mathbb{E} \left( \int_0^t u_v dB_v \mid \mathcal{F}_s \right) &= \mathbb{E} \left( \sum_{i=0}^{n-1} F_i(B_{t_{i+1} \wedge t} - B_{t_i \wedge t}) \mid \mathcal{F}_s \right) \\ &= \sum_{i=0}^{n-1} F_i(B_{t_{i+1} \wedge s} - B_{t_i \wedge s}) \\ &= \int_0^s u_v dB_v. \end{aligned}$$

Thus if  $u \in \mathcal{E}$ , the process

$$\left( \int_0^t u_s dB_s \right)_{t \geq 0} = \left( \int_0^{+\infty} u_s 1_{[0,t]}(s) dB_s \right)_{t \geq 0}$$

is a martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Since  $\mathcal{E}$  is dense in  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , and since it is easily checked that a limit in  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  of martingales is still a martingale, we deduce the expected result.

We now prove the existence of a continuous version.

If  $u \in \mathcal{E}$ , the continuity of the integral process easily stems from the continuity of the Brownian paths. Let  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and let  $u^n$  be a sequence in  $\mathcal{E}$  that converges to  $u$ . From Doob's inequality, we have for  $m, n \geq 0$  and  $\varepsilon > 0$ ,

$$\begin{aligned} \mathbb{P} \left( \sup_{t \geq 0} \left| \int_0^t (u_s^n - u_s^m) dB_s \right| \geq \varepsilon \right) &\leq \frac{\mathbb{E} \left( \left| \int_0^{+\infty} (u_s^n - u_s^m) dB_s \right|^2 \right)}{\varepsilon^2} \\ &\leq \frac{\mathbb{E} \left( \int_0^{+\infty} (u_s^n - u_s^m)^2 ds \right)}{\varepsilon^2}. \end{aligned}$$

There exists thus a sequence  $(n_k)_{k \geq 0}$  such that

$$\mathbb{P} \left( \sup_{t \geq 0} \left| \int_0^t (u_s^{n_{k+1}} - u_s^{n_k}) dB_s \right| \geq \frac{1}{2^k} \right) \leq \frac{1}{2^k}.$$

From Borel-Cantelli lemma, the sequence of processes  $\left( \int_0^t u_s^{n_k} dB_s \right)_{t \geq 0}$  converges then almost surely uniformly to the process  $\left( \int_0^t u_s dB_s \right)_{t \geq 0}$  which is therefore continuous.  $\square$

As a straightforward consequence of the previous proposition and Doob's inequalities, we obtain

**Proposition 5.9.** *Let  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .*

1. For every  $\lambda > 0$ ,

$$\mathbb{P}\left(\sup_{t \geq 0} \left| \int_0^t u_s dB_s \right| \geq \lambda\right) \leq \frac{\mathbb{E}\left(\int_0^{+\infty} u_s^2 ds\right)}{\lambda^2};$$

2.

$$\mathbb{E}\left(\left(\sup_{t \geq 0} \left| \int_0^t u_s dB_s \right|\right)^2\right) \leq 4\mathbb{E}\left(\int_0^{+\infty} u_s^2 ds\right).$$

Once again, we insist on the fact that Itô's integral is not pathwise in the sense that for  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , the Riemann sums

$$\sum_{k=0}^{n-1} u_{\frac{kt}{n}} \left( (B_{\frac{(k+1)t}{n}} - B_{\frac{kt}{n}}) \right),$$

need not to almost surely converge to  $\int_0^t u_s dB_s$ . However the following proposition shows that under continuity assumptions we have a convergence in probability.

**Proposition 5.10.** *Let  $u \in L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a continuous process. Let  $t \geq 0$ . For every sequence of subdivisions  $\Delta_n[0, t]$  such that*

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

*the following convergence holds in probability:*

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} u_{t_k^n} (B_{t_{k+1}^n} - B_{t_k^n}) = \int_0^t u_s dB_s.$$

*Proof.* Let us first assume that  $u$  is bounded almost surely. We have

$$\sum_{k=0}^{n-1} u_{t_k^n} (B_{t_{k+1}^n} - B_{t_k^n}) = \int_0^t u_s^n dB_s,$$

where  $u_s^n = \sum_{k=0}^{n-1} u_{t_k^n} 1_{(t_k^n, t_{k+1}^n]}(s)$ . The Itô isometry and the Lebesgue dominated convergence theorem shows then that  $\int_0^t u_s^n dB_s$  converges to  $\int_0^t u_s dB_s$  in  $L^2$  and therefore in probability. For general  $u$ 's we can use a localization procedure. For  $N \geq 0$ , consider the random time

$$T_N = \inf\{t \geq 0, |u_t| \geq N\}.$$

We have for every  $\varepsilon > 0$ ,

$$\begin{aligned} & \mathbb{P} \left( \left| \sum_{k=0}^{n-1} u_{t_k^n} (B_{t_{k+1}^n} - B_{t_k^n}) - \int_0^t u_s dB_s \right| \geq \varepsilon \right) \\ & \leq \mathbb{P}(T_N \leq t) + \mathbb{P} \left( \left| \sum_{k=0}^{n-1} u_{t_k^n} (B_{t_{k+1}^n} - B_{t_k^n}) - \int_0^t u_s dB_s \right| \geq \varepsilon, T_N \geq t \right) \\ & \leq \mathbb{P}(T_N \leq t) + \mathbb{P} \left( \left| \sum_{k=0}^{n-1} u_{t_k^n} 1_{|u_{t_k^n}| \leq M} (B_{t_{k+1}^n} - B_{t_k^n}) - \int_0^t u_s 1_{|u_s| \leq M} dB_s \right| \geq \varepsilon \right). \end{aligned}$$

This easily implies the convergence in probability. □

**Exercise 5.11.**

1. Show that for  $t \geq 0$ ,

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

What is surprising in this formula ?

2. Show that when  $n \rightarrow +\infty$ , the sequence

$$\sum_{k=0}^{n-1} B_{\frac{(k+1)t}{n}} \left( B_{\frac{(k+1)t}{n}} - B_{\frac{kt}{n}} \right),$$

converges in probability to a random variable that shall be computed.

**Exercise 5.12.** Show that if  $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  is locally square integrable, that is  $\int_0^t f^2(x) dx < +\infty$ ,  $t \geq 0$ , then the process  $\left( \int_0^t f(s) dB_s \right)_{t \geq 0}$  is a Gaussian process. Compute its mean and its covariance.

### 3 Square integrable martingales and quadratic variations

It turns out that stochastic integrals may be defined for other stochastic processes than Brownian motions. The key properties that were used in the above approach were the martingale property and the square integrability of the Brownian motion.

As above, we consider a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  that satisfies the usual conditions. A martingale  $(M_t)_{t \geq 0}$  defined on this space is said to be square integrable if for every  $t \geq 0$ ,  $\mathbb{E}(M_t^2) < +\infty$ .

For instance, if  $(B_t)_{t \geq 0}$  is a Brownian motion on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  and if  $(u_t)_{t \geq 0}$  is a process which is progressively measurable with respect to the



filtration  $(\mathcal{F}_t)_{t \geq 0}$  such that for every  $t \geq 0$ ,  $\mathbb{E}\left(\int_0^t u_s^2 ds\right) < +\infty$  then, the process

$$M_t = \int_0^t u_s dB_s, \quad t \geq 0,$$

is a square integrable martingale.

The most important theorem concerning continuous square integrable martingales is that they admit a quadratic variation. Before proving this theorem, we state a preliminary lemma.

**Lemma 5.13.** *Let  $(M_t)_{0 \leq t \leq T}$  be a continuous martingale such that*

$$\sup_{\Delta_n[0, T]} \sum_{k=0}^{n-1} |M_{t_{k+1}^n} - M_{t_k^n}| < +\infty.$$

*Then  $(M_t)_{0 \leq t \leq T}$  is constant.*

*Proof.* We may assume  $M_0 = 0$ . For  $N \geq 0$ , let us consider the stopping time

$$T_N = \inf \left\{ s \in [0, T], |M_s| \geq N, \sup_{\Delta_n[0, s]} \sum_{k=0}^{n-1} |M_{t_{k+1}^n} - M_{t_k^n}| \geq N \right\} \wedge T.$$

The stopped process  $(M_{t \wedge T_N})_{0 \leq t \leq T}$  is a martingale and therefore for  $s \leq t$ ,

$$\mathbb{E}((M_{t \wedge T_N} - M_{s \wedge T_N})^2) = \mathbb{E}(M_{t \wedge T_N}^2) - \mathbb{E}(M_{s \wedge T_N}^2).$$

Consider now a sequence of subdivisions  $\Delta_n[0, T]$  whose mesh tends to 0. By summing up the above inequality on the subdivision, we obtain

$$\begin{aligned} \mathbb{E}(M_{T_N}^2) &= \sum_{k=0}^{n-1} \left( M_{t_k^n \wedge T_N} - M_{t_{k-1}^n \wedge T_N} \right)^2 \\ &\leq \sup |M_{t_k^n \wedge T_N} - M_{t_{k-1}^n \wedge T_N}| \mathbb{E} \left( \sum_{k=0}^{n-1} \left| M_{t_k^n \wedge T_N} - M_{t_{k-1}^n \wedge T_N} \right| \right) \\ &\leq N \sup |M_{t_k^n \wedge T_N} - M_{t_{k-1}^n \wedge T_N}|. \end{aligned}$$

By letting  $n \rightarrow +\infty$ , we get  $\mathbb{E}(M_{T_N}^2) = 0$ . This implies  $M_{T_N} = 0$ . Letting now  $N \rightarrow \infty$ , we conclude  $M_T = 0$ .  $\square$

**Theorem 5.14** (Quadratic variation of a martingale). *Let  $(M_t)_{t \geq 0}$  be a martingale on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  which is continuous and square integrable and such that  $M_0 = 0$ . There is a unique continuous and increasing process denoted  $(\langle M \rangle_t)_{t \geq 0}$  that satisfies the following properties:*

1.  $\langle M \rangle_0 = 0$ ;

2. The process  $(M_t^2 - \langle M \rangle_t)_{t \geq 0}$  is a martingale.

Actually for every  $t \geq 0$  and for every sequence of subdivisions  $\Delta_n[0, t]$  such that

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

the following convergence takes place in probability:

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n (M_{t_k^n} - M_{t_{k-1}^n})^2 = \langle M \rangle_t.$$

The process  $(\langle M \rangle_t)_{t \geq 0}$  is called the quadratic variation process of  $(M_t)_{t \geq 0}$ .

*Proof.* We first assume that the martingale  $(M_t)_{t \geq 0}$  is bounded and prove that if  $\Delta_n[0, t]$  is a sequence of subdivisions of the interval  $[0, t]$  such that

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

then the limit

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n (M_{t_k^n} - M_{t_{k-1}^n})^2$$

exists in  $L^2$  and thus in probability.

Toward this goal, we introduce some notations. If  $\Delta[0, T]$  is a subdivision of the time interval  $[0, T]$  and if  $(X_t)_{t \geq 0}$  is a stochastic process, then we denote

$$S_t^{\Delta[0, T]}(X) = \sum_{i=0}^{k-1} (X_{t_{i+1}} - X_{t_i})^2 + (X_t - X_{t_k})^2,$$

where  $k$  is such that  $t_k \leq t < t_{k+1}$ .

An easy computation on conditional expectations shows that if  $(X_t)_{t \geq 0}$  is a martingale, then the process

$$X_t^2 - S_t^{\Delta[0, T]}(X), \quad t \leq T$$

is also a martingale. Also, if  $\Delta[0, T]$  and  $\Delta'[0, T]$  are two subdivisions of the time interval  $[0, T]$ , we will denote by  $\Delta \vee \Delta'[0, T]$  the subdivision obtained by putting together the points of  $\Delta[0, T]$  and the points of  $\Delta'[0, T]$ . Let now  $\Delta_n[0, T]$  be a sequence of subdivisions of  $[0, T]$  such that

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, T]| = 0.$$

Let us show that the sequence  $S_T^{\Delta_n[0, T]}(M)$  is a Cauchy sequence in  $L^2$ . Since the process  $S^{\Delta_n[0, T]}(M) - S^{\Delta_p[0, T]}(M)$  is a martingale (as a difference of two

martingales), we deduce that

$$\begin{aligned} & \mathbb{E} \left( \left( S_T^{\Delta_n[0,T]}(M) - S_T^{\Delta_p[0,T]}(M) \right)^2 \right) \\ &= \mathbb{E} \left( S_T^{\Delta_n \vee \Delta_p[0,T]}(S^{\Delta_n[0,T]}(M) - S^{\Delta_p[0,T]}(M)) \right) \\ &\leq 2 \left( \mathbb{E} \left( S_T^{\Delta_n \vee \Delta_p[0,T]}(S^{\Delta_n[0,T]}(M)) \right) + \mathbb{E} \left( S_T^{\Delta_n \vee \Delta_p[0,T]}(S^{\Delta_p[0,T]}(M)) \right) \right). \end{aligned}$$

Let us denote by  $s_k$ 's the points of the subdivision  $\Delta_n \vee \Delta_p[0, T]$  and for fixed  $s_k$ , we denote by  $t_l$  the point of  $\Delta_n[0, T]$  which is the closest to  $s_k$  and such that  $t_l \leq s_k \leq t_{l+1}$ . We have

$$\begin{aligned} S_{s_{k+1}}^{\Delta_n[0,T]}(M) - S_{s_k}^{\Delta_n[0,T]}(M) &= (M_{s_{k+1}} - M_{t_l})^2 - (M_{s_k} - M_{t_l})^2 \\ &= (M_{s_{k+1}} - M_{s_k})(M_{s_{k+1}} + M_{s_k} - 2M_{t_l}). \end{aligned}$$

Therefore, from Cauchy-Schwarz inequality, we have

$$\begin{aligned} & \mathbb{E} \left( S_T^{\Delta_n \vee \Delta_p[0,T]}(S^{\Delta_n[0,T]}(M)) \right) \\ &\leq \mathbb{E} \left( \sup_k (M_{s_{k+1}} + M_{s_k} - 2M_{t_l})^4 \right)^{1/2} \mathbb{E} \left( \left( S_T^{\Delta_n \vee \Delta_p[0,T]}(M) \right)^2 \right)^{1/2}. \end{aligned}$$

Since the martingale  $M$  is assumed to be continuous, when  $n, p \rightarrow +\infty$ ,

$$\mathbb{E} \left( \sup_k (M_{s_{k+1}} + M_{s_k} - 2M_{t_l})^4 \right) \rightarrow 0.$$

Thus, in order to conclude, it suffices to prove that  $\mathbb{E} \left( \left( S_T^{\Delta_n \vee \Delta_p[0,T]}(M) \right)^2 \right)$  is bounded. This fact is a consequence of the fact that  $M$  is assumed to be bounded and we let the reader work out the details of the argument. Therefore, in the  $L^2$  sense the following convergence holds

$$\langle M \rangle_t = \lim_{n \rightarrow +\infty} \sum_{k=1}^n (M_{t_k^n} - M_{t_{k-1}^n})^2.$$

The process  $(M_t^2 - \langle M \rangle_t)_{t \geq 0}$  is seen to be a martingale because for every  $n$  and  $T \geq 0$ , the process

$$M_t^2 - S_t^{\Delta_n[0,T]}(M), \quad t \leq T$$

is a martingale. Let us now show that the obtained process  $\langle M \rangle$  is a continuous process. From Doob's inequality, for  $n, p \geq 0$  and  $\varepsilon > 0$ ,

$$\begin{aligned} & \mathbb{P} \left( \sup_{0 \leq t \leq T} \left( S_t^{\Delta_n[0,T]}(M) - S_t^{\Delta_p[0,T]}(M) \right) > \varepsilon \right) \\ &\leq \frac{\mathbb{E} \left( \left( S_T^{\Delta_n[0,T]}(M) - S_T^{\Delta_p[0,T]}(M) \right)^2 \right)}{\varepsilon^2}. \end{aligned}$$

From Borel-Cantelli lemma, there exists therefore a sequence  $n_k$  such that the sequence of continuous stochastic processes  $\left(S_t^{\Delta_{n_k}[0,T]}(M)\right)_{0 \leq t \leq T}$  almost surely uniformly converges to the process  $(\langle M \rangle_t)_{0 \leq t \leq T}$ . This proves the existence of a continuous version for  $\langle M \rangle$ . Finally, to prove that  $\langle M \rangle$  is increasing, it is enough to consider an increasing sequence of subdivisions whose mesh tends to 0. Let us now prove that  $\langle M \rangle$  is the unique process such that  $M^2 - \langle M \rangle$  is a martingale. Let  $A$  and  $A'$  be two continuous and increasing stochastic processes such that  $A_0 = A'_0 = 0$  and such that  $(M_t^2 - A_t)_{t \geq 0}$  and  $(M_t^2 - A'_t)_{t \geq 0}$  are martingales. The process  $(N_t)_{t \geq 0} = (A_t - A'_t)_{t \geq 0}$  is then seen to be a martingale that has a bounded variation. From the previous lemma, this implies that  $(N_t)_{t \geq 0}$  is constant and therefore equal to 0 due to its initial condition.

We now turn to the case where  $(M_t)_{t \geq 0}$  is not necessarily bounded. Let us introduce the sequence of stopping times:

$$T_N = \inf\{t \geq 0, |M_t| \geq N\}.$$

According to the previous arguments, for every  $N \geq 0$ , there is an increasing process  $A^N$  such that  $(M_{t \wedge T_N}^2 - A_t^N)_{t \geq 0}$  is a martingale. By uniqueness of this process, it is clear that  $A_{t \wedge T_N}^{N+1} = A_t^N$ , therefore we can define a process  $A_t$  by requiring that  $A_t(\omega) = A_t^N(\omega)$  provided that  $T_N(\omega) \geq t$ . By using convergence theorems, it is then checked that  $(M_t^2 - A_t)_{t \geq 0}$  is a martingale.

Finally, let  $\Delta_n[0, t]$  be a sequence of subdivisions whose mesh tends to 0. We have for every  $\varepsilon > 0$ ,

$$\begin{aligned} & \mathbb{P} \left( \left| A_t - \sum_{k=1}^n (M_{t_k^n} - M_{t_{k-1}^n})^2 \right| \geq \varepsilon \right) \\ & \leq \mathbb{P}(T_N \leq t) + \mathbb{P} \left( \left| A_t^N - \sum_{k=1}^n (M_{t_k^n \wedge T_N} - M_{t_{k-1}^n \wedge T_N})^2 \right| \geq \varepsilon \right). \end{aligned}$$

This easily implies the announced convergence in probability of the quadratic variations to  $A_t$ .  $\square$

**Exercise 5.15.** Let  $(M_t)_{t \geq 0}$  be a square integrable martingale on a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . Assume that  $M_0 = 0$ . If  $\Delta[0, T]$  is a subdivision of the time interval  $[0, T]$  and if  $(X_t)_{t \geq 0}$  is a stochastic process, we denote

$$S_t^{\Delta[0,T]}(X) = \sum_{i=0}^{k-1} (X_{t_{i+1}} - X_{t_i})^2 + (X_t - X_{t_k})^2,$$

where  $k$  is such that  $t_k \leq t < t_{k+1}$ . Let  $\Delta_n[0, T]$  be a sequence of subdivisions of  $[0, T]$  such that

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, T]| = 0.$$

Show that the following convergence holds in probability,

$$\lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} \left| S_t^{\Delta[0, T]}(M) - \langle M \rangle_t \right| = 0.$$

Thus, in the previous theorem, the convergence is actually uniform on compact intervals.

We have already pointed out that stochastic integrals with respect to Brownian motion provide an example of square integrable martingale, they therefore have a quadratic variation. The next proposition explicitly computes this variation.

**Proposition 5.16.** *Let  $(B_t)_{t \geq 0}$  be a Brownian motion on a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  that satisfies the usual conditions. Let  $(u_t)_{t \geq 0}$  be a progressively measurable process such that for every  $t \geq 0$ ,  $\mathbb{E} \left( \int_0^t u_s^2 ds \right) < +\infty$ . For  $t \geq 0$ :*

$$\left\langle \int_0^\cdot u_s dB_s \right\rangle_t = \int_0^t u_s^2 ds.$$

*Proof.* Since the process  $\left( \int_0^t u_s^2 ds \right)_{t \geq 0}$  is continuous, increasing and equals 0 when  $t = 0$ , we just need to prove that

$$\left( \int_0^t u_s dB_s \right)^2 - \int_0^t u_s^2 ds$$

is a martingale.

If  $u \in \mathcal{E}$ , is a simple process, it is easily seen that for  $t \geq s$ :

$$\begin{aligned} \mathbb{E} \left( \left( \int_0^t u_v dB_v \right)^2 \mid \mathcal{F}_s \right) &= \mathbb{E} \left( \left( \int_0^s u_v dB_v + \int_s^t u_v dB_v \right)^2 \mid \mathcal{F}_s \right) \\ &= \mathbb{E} \left( \left( \int_0^s u_v dB_v \right)^2 \mid \mathcal{F}_s \right) + \mathbb{E} \left( \left( \int_s^t u_v dB_v \right)^2 \mid \mathcal{F}_s \right) \\ &= \left( \int_0^s u_v dB_v \right)^2 + \mathbb{E} \left( \int_s^t u_v^2 dv \mid \mathcal{F}_s \right). \end{aligned}$$

We may then conclude by using the density of  $\mathcal{E}$  in  $L^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .  $\square$

As a straightforward corollary of the proposition 5.14, we immediately obtain:

**Corollary 5.17.** *Let  $(M_t)_{t \geq 0}$  and  $(N_t)_{t \geq 0}$  be two continuous square integrable martingales on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  such that  $M_0 = N_0 = 0$ . There is a unique continuous process  $((M, N)_t)_{t \geq 0}$  with bounded variation that satisfies:*

1.  $\langle M, N \rangle_0 = 0$ ;
2. The process  $(M_t N_t - \langle M, N \rangle_t)_{t \geq 0}$  is a martingale.

Moreover, for  $t \geq 0$  and for every sequence  $\Delta_n[0, t]$  such that

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

the following convergence holds in probability:

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n (M_{t_k^n} - M_{t_{k-1}^n}) (N_{t_k^n} - N_{t_{k-1}^n}) = \langle M, N \rangle_t.$$

The process  $(\langle M, N \rangle_t)_{t \geq 0}$  is called the quadratic covariation process of  $(M_t)_{t \geq 0}$  and  $(N_t)_{t \geq 0}$ .

*Proof.* We may actually just use the formula

$$\langle M, N \rangle = \frac{1}{4} (\langle M + N \rangle - \langle M - N \rangle),$$

as a definition of the covariation and then check that the above properties are indeed satisfied due to Proposition 5.14.  $\square$

**Exercise 5.18.** Let  $(B_t^1)_{t \geq 0}$  and  $(B_t^2)_{t \geq 0}$  be two independent Brownian motions. Show that

$$\langle B^1, B^2 \rangle_t = 0.$$

In the same way that a stochastic integral with respect to Brownian motion was constructed, a stochastic integral with respect to square integrable martingales may be defined. We shall not repeat this construction, since it was done in the Brownian motion case, but we point out the main results without proofs and let the proofs as an exercise to the reader.

Let  $(M_t)_{t \geq 0}$  be a continuous square integrable martingale on a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  that satisfies the usual conditions. We assume that  $M_0 = 0$ . Let us denote by  $\mathcal{L}_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  the set of processes  $(u_t)_{t \geq 0}$  that are progressively measurable with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  and such that

$$\mathbb{E} \left( \int_0^{+\infty} u_s^2 d\langle M \rangle_s \right) < +\infty.$$

We still denote by  $\mathcal{E}$  the set of simple and predictable processes, that is the set of processes  $(u_t)_{t \geq 0}$  that may be written as:

$$u_t = \sum_{i=0}^{n-1} F_i 1_{(t_i, t_{i+1}]}(t),$$

where  $0 \leq t_0 \leq \dots \leq t_n$  and where  $F_i$  is a random variable that is measurable with respect to  $\mathcal{F}_{t_i}$  and such that  $\mathbb{E}(F_i^2) < +\infty$ . We define an equivalence relation  $\mathcal{R}$  on the set  $\mathcal{L}_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  as follows:

$$u\mathcal{R}v \Leftrightarrow \mathbb{E} \left( \int_0^{+\infty} (u_s - v_s)^2 d\langle M \rangle_s \right) = 0.$$

and denote by

$$L_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}) = \mathcal{L}_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}) / \mathcal{R},$$

the set of equivalence classes. It is easy to check that  $L_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  endowed with the norm

$$\|u\|^2 = \mathbb{E} \left( \int_0^{+\infty} u_s^2 d\langle M \rangle_s \right),$$

is a Hilbert space.

The following theorems are then proved similarly as for the Brownian motion.

**Theorem 5.19.** *There exists a unique linear map*

$$\mathcal{I}_M : L_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}) \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P})$$

such that:

- For  $u = \sum_{i=0}^{n-1} F_i 1_{(t_i, t_{i+1}]} \in \mathcal{E}$ ,

$$\mathcal{I}(u) = \sum_{i=0}^{n-1} F_i (M_{t_{i+1}} - M_{t_i});$$

- For  $u \in L_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ ,

$$\mathbb{E} (\mathcal{I}_M(u)^2) = \mathbb{E} \left( \int_0^{+\infty} u_s^2 d\langle M \rangle_s \right).$$

The map  $\mathcal{I}_M$  is called the Itô integral with respect to the continuous and square integrable martingale  $(M_t)_{t \geq 0}$ . We denote for  $u \in L_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ ,

$$\mathcal{I}_M(u) = \int_0^{+\infty} u_s dM_s.$$

In the case where  $(M_t)_{t \geq 0}$  is itself a stochastic integral with respect to Brownian motion, we can express the stochastic integral with respect to  $(M_t)_{t \geq 0}$  as an integral with respect to the underlying Brownian motion.

**Proposition 5.20.** *Let us assume that  $M_t = \int_0^t \Theta_s dB_s$  where  $(B_t)_{t \geq 0}$  is a Brownian motion on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  and where  $(\Theta_t)_{t \geq 0}$  is a progressively measurable process such that for every  $t \geq 0$ ,  $\mathbb{E} \left( \int_0^t \Theta_s^2 ds \right) < +\infty$ . Then  $u \in L_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  if and only if  $u\Theta \in L_B^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and in that case for every  $t \geq 0$ ,*

$$\int_0^t u_s dM_s = \int_0^t u_s \Theta_s dB_s.$$

It should come as no surprise, that stochastic integrals with respect to martingales are still martingales. This is confirmed in the next proposition.

**Proposition 5.21.** *Let  $(u_t)_{t \geq 0}$  be a stochastic process which is progressively measurable with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  and such that for every  $t \geq 0$ ,  $\mathbb{E} \left( \int_0^t u_s^2 d\langle M \rangle_s \right) < +\infty$ . The process*

$$\left( \int_0^t u_s dM_s \right)_{t \geq 0} = \left( \int_0^{+\infty} u_s 1_{[0,t]}(s) dM_s \right)_{t \geq 0}$$

*is a square integrable martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  that admits a continuous modification and its quadratic variation process is*

$$\left\langle \int_0^\cdot u_s dM_s \right\rangle_t = \int_0^t u_s^2 d\langle M \rangle_s.$$

As in the Brownian motion case, the stochastic integral with respect to a martingale is not an almost sure limit of Riemann sums, it is however a limit in probability.

**Proposition 5.22.** *Let  $u \in L_M^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a continuous stochastic process. Let  $t \geq 0$ . For every sequence of subdivisions  $\Delta_n[0, t]$  such that*

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

*the following convergence holds in probability:*

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} u_{t_k^n} (M_{t_{k+1}^n} - M_{t_k^n}) = \int_0^t u_s dM_s.$$

## 4 Local martingales, Semimartingales and Integrators

The goal of this paragraph is to extend the domain of definition of the Itô's integral with respect to Brownian motion. The idea is to use the fruitful concept of localization. We will then finally be interested in the widest class of processes



for which it is possible to define a stochastic integral satisfying natural probabilistic properties. This will lead to the natural notion of semimartingales.

As before, we consider here a Brownian motion  $(B_t)_{t \geq 0}$  that is defined on a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  that satisfies the usual conditions.

**Definition 5.23.** We define the space  $L_{loc}^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , as the set of the processes  $(u_t)_{t \geq 0}$  that are progressively measurable with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  and such that for every  $t \geq 0$

$$\mathbb{P} \left( \int_0^t u_s^2 ds < +\infty \right) = 1.$$

We first have the following fact whose proof is left as an exercise to the reader:

**Lemma 5.24.** *Let  $u \in L_{loc}^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . There exists an increasing family of stopping times  $(T_n)_{n \geq 0}$  for the filtration  $(\mathcal{F}_t)_{t \geq 0}$  such that:*

1. Almost surely,

$$\lim_{n \rightarrow +\infty} T_n = +\infty;$$

2.

$$\mathbb{E} \left( \int_0^{T_n} u_s^2 ds \right) < +\infty.$$

Thanks to this lemma, it is now easy to naturally define  $\int_0^t u_s dB_s$  for  $u \in L_{loc}^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . Indeed, let  $u \in L_{loc}^2(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and let  $t \geq 0$ . According to the previous lemma, let us now consider an increasing sequence of stopping times  $(T_n)_{n \geq 0}$  such that:

1. Almost surely,

$$\lim_{n \rightarrow +\infty} T_n = +\infty;$$

2.

$$\mathbb{E} \left( \int_0^{T_n} u_s^2 ds \right) < +\infty.$$

Since

$$\mathbb{E} \left( \int_0^{T_n} u_s^2 ds \right) < +\infty,$$

the stochastic integral

$$\int_0^{T_n} u_s dB_s = \int_0^{+\infty} u_s 1_{[0, T_n]}(s) dB_s$$

exists. We may therefore define in a unique way a stochastic process

$$\left( \int_0^t u_s dB_s \right)_{t \geq 0}$$

such that:

1.

$$\left( \int_0^t u_s dB_s \right)_{t \geq 0}$$

is a continuous stochastic process adapted to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ ;

2. The stochastic process

$$\left( \int_0^{t \wedge T_n} u_s dB_s \right)_{t \geq 0}$$

is a uniformly integrable martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  (because it is bounded in  $L^2$ ).

This leads to the following definition:

**Definition 5.25** (Local martingale). A stochastic process  $(M_t)_{t \geq 0}$  is called a local martingale (with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ ) if there is a sequence of stopping times  $(T_n)_{n \geq 0}$  such that:

1. The sequence  $(T_n)_{n \geq 0}$  is increasing and almost surely satisfies  $\lim_{n \rightarrow +\infty} T_n = +\infty$ ;
2. For  $n \geq 1$ , the process  $(M_{t \wedge T_n})_{t \geq 0}$  is a uniformly integrable martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .

Thus, as an example, if  $u \in L^2_{loc}(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  then the process  $\left( \int_0^t u_s dB_s \right)_{t \geq 0}$  is a local martingale. Of course, any martingale turns out to be a local martingale. But, as we will see it later, in general the converse is not true:

The following Exercise gives a useful criterion to prove that a given local martingale is actually martingale.

**Exercise 5.26.** Let  $(M_t)_{t \geq 0}$  be a continuous local martingale such that  $t \geq 0$ ,

$$\mathbb{E} \left( \sup_{s \leq t} |M_s| \right) < +\infty.$$

Show that  $(M_t)_{t \geq 0}$  is a martingale. As a consequence, bounded local martingales necessarily are martingales.

**Exercise 5.27.** Show that a positive local martingale is a supermartingale.

It is interesting to observe that if  $(M_t)_{t \geq 0}$  is a local martingale, then the sequence of stopping times may explicitly be chosen so that the resulting stopped martingales enjoy nice properties.

**Lemma 5.28.** *Let  $(M_t)_{t \geq 0}$  be a continuous local martingale on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  such that  $M_0 = 0$ . Let*

$$T_n = \inf\{t \geq 0, |M_t| \geq n\}.$$

*Then, for  $n \in \mathbb{N}$ , the process  $(M_{t \wedge T_n})_{t \geq 0}$  is a bounded martingale.*

*Proof.* Let  $(S_n)_{n \geq 0}$  be a sequence of stopping times such that:

1. The sequence  $(S_n)_{n \geq 0}$  is increasing and almost surely  $\lim_{n \rightarrow +\infty} S_n = +\infty$ ;
2. For every  $n \geq 1$ , the process  $(M_{t \wedge S_n})_{t \geq 0}$  is a uniformly integrable martingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .

For  $t \geq s$  and  $k, n \geq 0$ , we have:

$$\mathbb{E}(M_{t \wedge S_k \wedge T_n} \mid \mathcal{F}_s) = M_{s \wedge S_k \wedge T_n}.$$

Letting  $k \rightarrow +\infty$  leads then to the expected result.  $\square$

Since bounded martingales are of course square integrable, we easily deduce from the previous Lemma that the following result holds:

**Theorem 5.29.** *Let  $(M_t)_{t \geq 0}$  be a continuous local martingale on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  such that  $M_0 = 0$ . Then, there is a unique continuous increasing process  $(\langle M \rangle_t)_{t \geq 0}$  such that:*

1.  $\langle M \rangle_0 = 0$ ;
2. The process  $(M_t^2 - \langle M \rangle_t)_{t \geq 0}$  is a local martingale.

*Furthermore, for every  $t \geq 0$  and every sequence of subdivisions  $\Delta_n[0, t]$  such that*

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

*the following limit holds in probability:*

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n (M_{t_k^n} - M_{t_{k-1}^n})^2 = \langle M \rangle_t.$$

*The process  $(\langle M \rangle_t)_{t \geq 0}$  is called the quadratic variation of the local martingale. Moreover, if  $u$  is a progressively measurable process such that for every  $t \geq 0$ ,*

$$\mathbb{P}\left(\int_0^t u_s^2 d\langle M \rangle_s < +\infty\right) = 1,$$

then we may define a stochastic integral  $\left(\int_0^t u_s dM_s\right)_{t \geq 0}$  such that the stochastic process  $\left(\int_0^t u_s dM_s\right)_{t \geq 0}$  is a continuous local martingale.

At that point, we already almost found the widest class of stochastic processes with respect to which it was possible to naturally construct a stochastic integral. To go further in that direction, let us first observe that if we add a bounded variation process to a local martingale, then we obtain a process with respect to which a stochastic integral is naturally defined.

More precisely, if  $(X_t)_{t \geq 0}$  may be written under the form:

$$X_t = X_0 + A_t + M_t$$

where  $(A_t)_{t \geq 0}$  is a bounded variation process and where  $(M_t)_{t \geq 0}$  is a continuous local martingale on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  such that  $M_0 = 0$ , then if  $u$  is a progressively measurable process such that for  $t \geq 0$ ,

$$\mathbb{P} \left( \int_0^t u_s^2 d\langle M \rangle_s < +\infty \right) = 1,$$

we may define a stochastic integral as

$$\left( \int_0^t u_s dX_s \right)_{t \geq 0} = \left( \int_0^t u_s dA_s + \int_0^t u_s dM_s \right)_{t \geq 0}$$

where  $\int_0^t u_s dA_s$  is simply understood as the Riemann-Stieltjes integral with respect to the process  $(A_t)_{t \geq 0}$ .

The class of stochastic processes that we obtained is called the class of semimartingales and, as we will see it later, is the most relevant one:

**Definition 5.30** (Semimartingale). Let  $(X_t)_{t \geq 0}$  be an adapted continuous stochastic process on the filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . We say that  $(X_t)_{t \geq 0}$  is a semimartingale with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  if  $(X_t)_{t \geq 0}$  may be written as:

$$X_t = X_0 + A_t + M_t$$

where  $(A_t)_{t \geq 0}$  is a bounded variation process and  $(M_t)_{t \geq 0}$  is a continuous local martingale such that  $M_0 = 0$ . If it exists, the previous decomposition is unique.

**Remark 5.31.** It is possible to prove that if  $(X_t)_{t \geq 0}$  is a semimartingale with respect to a filtration  $(\mathcal{F}_t)_{t \geq 0}$  then it is also semimartingale in its own natural filtration (see [?]).

**Exercise 5.32.** Let  $(M_t)_{t \geq 0}$  be a continuous local martingale on the filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . Show that  $(M_t^2)_{t \geq 0}$  is a semimartingale.

Since a bounded variation process has a zero quadratic variation, it is easy to prove the following theorem:

**Proposition 5.33.** *Let*

$$X_t = X_0 + A_t + M_t, \quad t \geq 0,$$

*be a continuous adapted semimartingale. For every  $t \geq 0$  and every sequence of subdivisions  $\Delta_n[0, t]$  such that*

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

*the following limit holds in probability:*

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n \left( X_{t_k^n} - X_{t_{k-1}^n} \right)^2 = \langle M \rangle_t.$$

*We therefore call  $\langle M \rangle$  the quadratic variation of  $X$  and denote  $\langle X \rangle = \langle M \rangle$ .*

**Exercise 5.34.** *Let  $(X_t)_{t \geq 0}$  be a continuous semimartingale on the filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . If  $\Delta[0, T]$  is a subdivision of the time interval  $[0, T]$ , we denote*

$$S_t^{\Delta[0, T]}(X) = \sum_{i=0}^{k-1} (X_{t_{i+1}} - X_{t_i})^2 + (X_t - X_{t_k})^2,$$

*where  $k$  is such that  $t_k \leq t < t_{k+1}$ . Let  $\Delta_n[0, T]$  be a sequence of subdivisions of  $[0, T]$  such that*

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, T]| = 0.$$

*Show that the following limit holds in probability,*

$$\lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} \left| S_t^{\Delta_n[0, T]}(X) - \langle X \rangle_t \right| = 0.$$

**Exercise 5.35.** *Let  $(X_t)_{t \geq 0}$  be a continuous semimartingale on  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$ . Let  $u^n$  be a sequence of locally bounded and adapted processes almost surely converging toward 0 such that  $u^n \leq u$ , where  $u$  is a locally bounded process. Show that for  $T \geq 0$ , the following limit holds in probability*

$$\lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} \left| \int_0^t u_s^n dX_s \right| = 0.$$

It already has been observed that in the Brownian case, though the stochastic integral is not an almost sure limit of Riemann sums, it is however a limit in probability of such sums. This may be extended to semimartingales in the following way.

**Proposition 5.36.** *Let  $u$  be a continuous and adapted process, let  $(X_t)_{t \geq 0}$  be a continuous and adapted semimartingale and let  $t \geq 0$ . For every sequence of subdivisions  $\Delta_n[0, t]$  such that*

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

*the following limit holds in probability:*

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} u_{t_k^n} (X_{t_{k+1}^n} - X_{t_k^n}) = \int_0^t u_s dX_s.$$

**Exercise 5.37** (Backward and Stratonovitch integrals). *Let  $(X_t)_{t \geq 0}, (Y_t)_{t \geq 0}$  be continuous semimartingales.*

1. *Show that for every sequence of subdivisions  $\Delta_n[0, t]$  such that*

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

*the following limit exists in probability:*

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} \frac{X_{t_k^n} + X_{t_{k+1}^n}}{2} (Y_{t_{k+1}^n} - Y_{t_k^n}).$$

*This limit is called the Stratonovitch integral and denoted  $\int_0^t X_s \circ dY_s$ .*

2. *Show the formula*

$$\int_0^t X_s \circ dY_s = \int_0^t X_s dY_s + \frac{1}{2} \langle X, Y \rangle_t.$$

3. *With the same assumptions as above, show that the following limit exists in probability:*

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} X_{t_{k+1}^n} (Y_{t_{k+1}^n} - Y_{t_k^n}).$$

*The limit is called the backward stochastic integral. Find a formula relating the backward integral to Stratonovitch's.*

As we already suggested it, the class of semimartingales is actually the widest class of stochastic processes with respect to which we may define a stochastic integral that enjoys natural probabilistic properties. Let us more precisely explain what the previous statement means.

We denote by  $\mathcal{E}_b$  the set of processes  $(u_t)_{t \geq 0}$  such that:

$$u_t = \sum_{i=1}^N F_i \mathbf{1}_{(S_i, T_i]}(t),$$

where  $0 \leq S_1 \leq T_1 \leq \dots \leq S_N \leq T_N$  are bounded stopping times and where the  $F_i$ 's are random variable that are bounded and measurable with respect to  $\mathcal{F}_{S_i}$ . If  $(X_t)_{t \geq 0}$  is a continuous and adapted process and if  $u \in \mathcal{E}_b$ , then we naturally define

$$\int_0^t u_s dX_s = \sum_{i=1}^N F_i(X_{T_i \wedge t} - X_{S_i \wedge t}).$$

We have the following theorem whose proof goes back to Meyer and that can be found in the book [?] by Protter:

**Theorem 5.38.** *Let  $(X_t)_{t \geq 0}$  be a continuous and adapted process. The process  $(X_t)_{t \geq 0}$  is a semimartingale if and only if for every sequence  $u^n$  in  $\mathcal{E}_b$  that almost surely converges to 0, we have for every  $t \geq 0$  and  $\varepsilon > 0$ ,*

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( \left| \int_0^t u_s^n dX_s \right| > \varepsilon \right) = 0.$$

If the previous continuity in probability is not asked, then integrals in the rough paths sense of Lyons (see Chapter 7) may be useful in applications. This integral is a natural extension of Young's integral which coincides with the Stratonovitch integral (see Exercise 5.37) for semimartingales but which does not enjoy any nice probabilistic property.

## 5 Döblin-Itô formula

The Döblin-Itô formula is certainly the most important and useful formula of stochastic calculus. It is the change of variable formula for stochastic integrals. It is a very simple formula whose specificity is the appearance of a quadratic variation term. This reflects the fact that semimartingales have a finite quadratic variation.

Due to its importance, we first provide a heuristic argument on how to derive Itô formula. Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a smooth function and  $x : \mathbb{R} \rightarrow \mathbb{R}$  be a  $C^1$  path  $\mathbb{R} \rightarrow \mathbb{R}$ . We have the following heuristic computation:

$$\begin{aligned} f(x_{t+dt}) &= f(x_t + (x_{t+dt} - x_t)) \\ &= f(x_t) + f'(x_t)(x_{t+dt} - x_t) \\ &= f(x_t) + f'(x_t)dx_t. \end{aligned}$$

This suggests, by summation, the following correct formula:

$$f(x_t) = f(x_0) + \int_0^t f'(x_s) dx_s.$$

Let us now try to consider a Brownian motion  $(B_t)_{t \geq 0}$  instead of the smooth path  $x$  and let us try to adapt the previous computation to this case. Since

Brownian motion has quadratic variation which is not zero,  $\langle B \rangle_t = t$ , we need to go at the order 2 in the Taylor expansion of  $f$ . This leads to the following heuristic computation:

$$\begin{aligned} f(B_{t+dt}) &= f(B_t + (B_{t+dt} - B_t)) \\ &= f(B_t) + f'(B_t)(B_{t+dt} - B_t) + \frac{1}{2}f''(B_t)((B_{t+dt} - B_t))^2 \\ &= f(B_t) + f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt. \end{aligned}$$

By summation, we are therefore led to the formula

$$f(B_t) = f(0) + \int_0^t f'(B_s)dB_s + \frac{1}{2} \int_0^t f''(B_s)ds,$$

which is, as we will see it later, perfectly correct.

In what follows, we consider a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  that satisfies the usual conditions. Our starting point to prove Döblin-Itô formula is the following formula which is known as the integration by parts formula for semimartingales:

**Theorem 5.39** (Integration by parts formula). *Let  $(X_t)_{t \geq 0}$  and  $(Y_t)_{t \geq 0}$  be two continuous semimartingales, then the process  $(X_t Y_t)_{t \geq 0}$  is a continuous semimartingale and we have:*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t, \quad t \geq 0.$$

*Proof.* By bilinearity of the multiplication, we may assume  $X = Y$ . Also by considering, if needed,  $X - X_0$  instead of  $X$ , we may assume that  $X_0 = 0$ .

Let  $t \geq 0$ . For every sequence  $\Delta_n[0, t]$  such that

$$\lim_{n \rightarrow +\infty} |\Delta_n[0, t]| = 0,$$

we have

$$\sum_{k=1}^n \left( X_{t_k^n} - X_{t_{k-1}^n} \right)^2 = X_t^2 - 2 \sum_{k=1}^n X_{t_{k-1}^n} \left( X_{t_k^n} - X_{t_{k-1}^n} \right).$$

By letting  $n \rightarrow \infty$ , and using Propositions 5.33 and 5.36 we therefore obtain the following identity which yields the expected result.

$$X_t^2 = 2 \int_0^t X_s dX_s + \langle X \rangle_t.$$

□



**Exercise 5.40.** *Let*

$$L = \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j},$$

be a diffusion operator. Let us assume that there exists a diffusion process  $(X_t)_{t \geq 0}$  with generator  $L$ . Show that for  $f \in C_c(\mathbb{R}^n, \mathbb{R})$ , the quadratic variation of the martingale

$$M_t^f = f(X_t) - \int_0^t Lf(X_s) ds$$

is given by

$$\langle M^f \rangle_t = \int_0^t \Gamma(f)(X_s)^2 ds,$$

where  $\Gamma(f) = \frac{1}{2}(Lf^2 - 2fLf)$  is the carré du champ.

We are now in position to prove the Döblin-Itô formula in its simpler form.

**Theorem 5.41** (Döblin-Itô formula I). *Let  $(X_t)_{t \geq 0}$  be a continuous and adapted semimartingale and let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a function which is twice continuously differentiable. The process  $(f(X_t))_{t \geq 0}$  is a semimartingale and the following change of variable formula holds:*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d\langle X \rangle_s. \quad (5.1)$$

*Proof.* We assume that the semimartingale  $(X_t)_{t \geq 0}$  is bounded. If this is not, we may apply the following arguments to the semimartingale  $(X_{t \wedge T_n})_{t \geq 0}$ , where  $T_n = \inf\{t \geq 0, X_t \geq n\}$  and then let  $n \rightarrow \infty$ .

Let  $\mathcal{A}$  be the set of two times continuously differentiable functions  $f$  for which the formula (5.1) holds. It is straightforward that  $\mathcal{A}$  is a vector space.

Let us show that  $\mathcal{A}$  is also an algebra, that is also let stable by multiplication. Let  $f, g \in \mathcal{A}$ . By using the integration by parts formula with the semimartingales  $(f(X_t))_{t \geq 0}$  and  $(g(X_t))_{t \geq 0}$ , we obtain

$$f(X_t)g(X_t) = f(X_0)g(X_0) + \int_0^t f(X_s) dg(X_s) + \int_0^t g(X_s) df(X_s) + \langle f(X), g(X) \rangle_t.$$

The terms of the previous sum may be separately treated in the following way. Since  $f, g \in \mathcal{A}$ , we get:

$$\int_0^t f(X_s) dg(X_s) = \int_0^t f(X_s) g'(X_s) dX_s + \frac{1}{2} \int_0^t f(X_s) g''(X_s) d\langle X \rangle_s$$

$$\int_0^t g(X_s) df(X_s) = \int_0^t g(X_s) f'(X_s) dX_s + \frac{1}{2} \int_0^t g(X_s) f''(X_s) d\langle X \rangle_s$$

$$\langle f(X), g(X) \rangle_t = \int_0^t f'(X_s)g'(X_s)d\langle X \rangle_s.$$

Therefore,

$$\begin{aligned} f(X_t)g(X_t) &= f(X_0)g(X_0) + \int_0^t f(X_s)g'(X_s)dX_s + \int_0^t g(X_s)f'(X_s)dX_s \\ &\quad + \frac{1}{2} \int_0^t f(X_s)g''(X_s)d\langle X \rangle_s + \int_0^t f'(X_s)g'(X_s)d\langle X \rangle_s \\ &\quad + \frac{1}{2} \int_0^t g(X_s)f''(X_s)d\langle X \rangle_s \\ &= f(X_0)g(X_0) + \int_0^t (fg)'(X_s)dX_s + \frac{1}{2} \int_0^t (fg)''(X_s)d\langle X \rangle_s. \end{aligned}$$

We deduce that  $fg \in \mathcal{A}$ .

As a conclusion,  $\mathcal{A}$  is an algebra of functions. Since  $\mathcal{A}$  contains the function  $x \rightarrow x$ , we deduce that  $\mathcal{A}$  actually contains every polynomial function. Now in order to show that every function  $f$  which is twice continuously differentiable is actually in  $\mathcal{A}$ , we first observe that since  $X$  is assumed to be bounded, it takes its values in a compact set.

It is then possible to find a sequence of polynomials  $P_n$  such that, on this compact set,  $P_n$  uniformly converges toward  $f$ ,  $P'_n$  uniformly converges toward  $f'$  and  $P''_n$  uniformly converges toward  $f''$ . We may then conclude by using the result of Exercise 5.35.  $\square$

As a particular case of the previous formula, if we apply this formula with  $X$  as a Brownian motion, we get the formula that was already pointed out at the beginning of the section: If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is twice continuously differentiable function, then

$$f(B_t) = f(0) + \int_0^t f'(B_s)dB_s + \frac{1}{2} \int_0^t f''(B_s)ds.$$

It is easy to derive the following variations of the Döblin-Itô formula:

**Theorem 5.42** (Döblin-Itô formula II). *Let  $(X_t)_{t \geq 0}$  be a continuous and adapted semimartingale, and let  $(A_t)_{t \geq 0}$  be an adapted bounded variation process. If  $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  is a function that is once continuously differentiable with respect to its first variable and that is twice continuously differentiable with respect to its second variable, then for  $t \geq 0$ :*

$$\begin{aligned} f(A_t, X_t) &= \\ f(A_0, X_0) &+ \int_0^t \frac{\partial f}{\partial t}(A_s, X_s)dA_s + \int_0^t \frac{\partial f}{\partial x}(A_s, X_s)dX_s + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(A_s, X_s)d\langle X \rangle_s. \end{aligned}$$

**Theorem 5.43** (Döblin-Itô formula III). *Let  $(X_t^1)_{t \geq 0}, \dots, (X_t^n)_{t \geq 0}$  be  $n$  adapted and continuous semimartingales and let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a twice continuously differentiable function. We have:*

$$f(X_t^1, \dots, X_t^n) = f(X_0^1, \dots, X_0^n) + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(X_s^1, \dots, X_s^n) dX_s^i + \frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s^1, \dots, X_s^n) d\langle X^i, X^j \rangle_s.$$

**Exercise 5.44.** *Let  $f : \mathbb{R}_{\geq 0} \times \mathbb{R} \rightarrow \mathbb{C}$  be a function that is once continuously differentiable with respect to its first variable and twice continuously differentiable with respect to its second variable that satisfies*

$$\frac{1}{2} \frac{\partial^2 f}{\partial x^2} + \frac{\partial f}{\partial t} = 0.$$

*Show that if  $(M_t)_{t \geq 0}$  is a continuous local martingale, then  $(f(\langle M \rangle_t, M_t))_{t \geq 0}$  is a continuous local martingale. Deduce that for  $\lambda \in \mathbb{C}$ , the process*

$$\left( \exp(\lambda M_t - \frac{1}{2} \lambda^2 \langle M \rangle_t) \right)_{t \geq 0}$$

*is a local martingale.*

**Exercise 5.45.** *The Hermite polynomial of order  $n$  is defined as*

$$H_n(x) = (-1)^n \frac{1}{n!} e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}.$$

1. *Compute  $H_0, H_1, H_2, H_3$ .*
2. *Show that if  $(B_t)_{t \geq 0}$  is a Brownian motion, then the process  $\left( t^{n/2} H_n\left(\frac{B_t}{\sqrt{t}}\right) \right)_{t \geq 0}$  is a martingale.*
3. *Show that*

$$t^{n/2} H_n\left(\frac{B_t}{\sqrt{t}}\right) = \int_0^t \int_0^{t_n} \cdots \int_0^{t_2} dB_{t_1} \cdots dB_{t_n}.$$

## 6 Recurrence and transience of the Brownian motion in higher dimensions

The purpose of the next sections is to illustrate through several applications the power of the stochastic integration theory and particularly of the Döblin-Itô formula. We start with a study of the multidimensional Brownian motion. As already pointed out, a multidimensional stochastic process  $(B_t)_{t \geq 0} =$

$(B_t^1, \dots, B_t^n)_{t \geq 0}$ , is called a Brownian motion if the processes  $(B_t^1)_{t \geq 0}$ ,  $\dots$ ,  $(B_t^n)_{t \geq 0}$  are independent Brownian motions. In the sequel we denote by  $\Delta$  the Laplace operator on  $\mathbb{R}^n$ , that is

$$\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}.$$

The following result is an easy consequence of the Döblin-Itô formula.

**Proposition 5.46.** *Let  $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}$  be a function that is once continuously differentiable with respect to its first variable and twice continuously differentiable with respect to its second variable and let  $(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^n)_{t \geq 0}$  be a  $n$ -dimensional Brownian motion. The process*

$$X_t = f(t, B_t) - \left( \int_0^t \frac{1}{2} \Delta f(s, B_s) + \frac{\partial f}{\partial t}(s, B_s) ds \right)$$

is a local martingale. If moreover  $f$  is such that

$$\sum_{i=1}^n \left( \frac{\partial f}{\partial x_i}(t, x) \right)^2 \leq \phi(t) e^{K\|x\|},$$

for some continuous function  $\phi$  and some constant  $K \in \mathbb{R}$ , then  $(X_t)_{t \geq 0}$  is a martingale.

In particular, if  $f$  is a harmonic function, i.e.  $\Delta f = 0$ , and if  $(B_t)_{t \geq 0}$  is a multidimensional Brownian motion, then the process  $(f(B_t))_{t \geq 0}$  is a local martingale. As we will see it later, this nice fact has many consequences. A first nice application is the study of recurrence or transience of the multidimensional Brownian motion paths. As we have seen before, the one-dimensional Brownian motion is recurrent: It reaches any value with probability 1. In higher dimensions, the situation is more subtle.

Let  $(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^n)_{t \geq 0}$  be a  $n$ -dimensional Brownian motion with  $n \geq 2$ . For  $a > 0$  and  $x \in \mathbb{R}^n$ , we consider the stopping time

$$T_a^x = \inf\{t \geq 0, \|B_t + x\| = a\}.$$

**Proposition 5.47.** *For  $a < \|x\| < b$ ,*

$$\mathbb{P}(T_a^x < T_b^x) = \begin{cases} \frac{\ln b - \ln \|x\|}{\ln b - \ln a}, & n = 2 \\ \frac{\|x\|^{2-n} - b^{2-n}}{a^{2-n} - b^{2-n}}, & n \geq 3. \end{cases}$$

*Proof.* For  $a < \|x\| < b$ , we consider the function

$$f(x) = \Psi(\|x\|) = \begin{cases} \ln \|x\|, & n = 2 \\ \|x\|^{2-n}, & n \geq 3. \end{cases}$$

A straightforward computation shows that  $\Delta f = 0$ . The process  $(f(B_{t \wedge T_a^x \wedge T_b^x}))_{t \geq 0}$  is therefore a martingale, which implies  $\mathbb{E}(f(B_{T_a^x \wedge T_b^x})) = f(x)$ . This yields

$$\Psi(a)\mathbb{P}(T_a^x < T_b^x) + \Psi(b)\mathbb{P}(T_b^x < T_a^x) = f(x).$$

Since

$$\mathbb{P}(T_a^x < T_b^x) + \mathbb{P}(T_b^x < T_a^x) = 1,$$

we deduce that

$$\mathbb{P}(T_a^x < T_b^x) = \begin{cases} \frac{\ln b - \ln \|x\|}{\ln b - \ln a}, & n = 2 \\ \frac{\|x\|^{2-n} - b^{2-n}}{a^{2-n} - b^{2-n}}, & n \geq 3. \end{cases}$$

□

By letting  $b \rightarrow \infty$ , we get

**Corollary 5.48.** For  $0 < a < \|x\|$ ,

$$\mathbb{P}(T_a^x < +\infty) = \begin{cases} 1, & n = 2 \\ \frac{\|x\|^{2-n}}{a^{2-n}}, & n \geq 3. \end{cases}$$

As a consequence, for  $n = 2$  the Brownian motion is recurrent, that is, for every non empty set  $\mathcal{O} \subset \mathbb{R}^2$ ,

$$\mathbb{P}(\exists t \geq 0, B_t \in \mathcal{O}) = 1.$$

Though the two-dimensional Brownian motion is recurrent, points are always polar.

**Proposition 5.49.** For every  $x \in \mathbb{R}^n$ ,  $\mathbb{P}(\exists t \geq 0, B_t = x) = 0$ .

*Proof.* It suffices to prove that for every  $x \in \mathbb{R}^n$ ,  $x \neq 0$ ,  $\mathbb{P}(T_0^x < +\infty) = 0$ . We have

$$\{T_0^x < +\infty\} = \cup_{n \geq 0} \cap_{m \geq \frac{1}{\|x\|}} \{T_{1/m}^x \leq T_n^x\}.$$

Since  $\mathbb{P}\left(\cap_{m \geq \frac{1}{\|x\|}} \{T_{1/m}^x \leq T_n^x\}\right) = \lim_{m \rightarrow \infty} \mathbb{P}\left(T_{1/m}^x \leq T_n^x\right) = 0$ , we get

$$\mathbb{P}(T_0^x < +\infty) = 0.$$

□

As we have just seen, the two-dimensional Brownian motion will hit every non empty open set with probability one. The situation is different in dimension higher than 3: Brownian motion paths will eventually leave any bounded set with probability one.

**Proposition 5.50.** *Let  $(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^n)_{t \geq 0}$  be a  $n$ -dimensional Brownian motion. If  $n \geq 3$  then almost surely*

$$\lim_{t \rightarrow \infty} \|B_t\| = +\infty.$$

*Proof.* Let us assume  $n \geq 3$ . Let  $\Phi(x) = \frac{1}{\|x+a\|^{n-2}}$  where  $a \in \mathbb{R}^n$ ,  $a \neq 0$ . Since  $(B_t)_{t \geq 0}$  will never hit the point  $-a$ , we can consider the process  $(\Phi(B_t))_{t \geq 0}$  which is seen to be a positive local martingale from Itô's formula. A positive local martingale is always a supermartingale. Therefore from the Doob's convergence theorem (see Exercise 1.42), the process  $(\Phi(B_t))_{t \geq 0}$  converges almost surely when  $t \rightarrow \infty$  to an integrable and non negative random variable  $Z$ . From Fatou's lemma, we have  $\mathbb{E}(Z) \leq \liminf_{t \rightarrow +\infty} \mathbb{E}(\Phi(B_t))$ . By the scaling property of the Brownian motion, it is clear that  $\liminf_{t \rightarrow +\infty} \mathbb{E}(\Phi(B_t)) = 0$ . We conclude  $Z = 0$ .  $\square$

**Exercise 5.51.** *(Probabilistic proof of Liouville theorem) By using martingale methods, prove that if  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a bounded harmonic function, then  $f$  is constant.*

**Exercise 5.52.** *Let  $(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^n)_{t \geq 0}$  be a  $n$ -dimensional Brownian motion. Show that for  $n \geq 3$ , the process  $\left(\frac{1}{\|B_t+a\|^{n-2}}\right)_{t \geq 0}$ , is a local martingale which is not a martingale.*

## 7 Itô representation theorem

In this section we show that, remarkably, any square integrable random variable which is measurable with respect to a Brownian motion, can be expressed as a stochastic integral with respect to this Brownian motion. A striking consequence of this result, which is known as Itô's representation theorem, is that any square martingale of the filtration has a continuous version.

Let  $(B_t)_{t \geq 0}$  be a Brownian motion. In the sequel, we consider the filtration  $(\mathcal{F}_t)_{t \geq 0}$  which is the usual completion of the natural filtration of  $(B_t)_{t \geq 0}$  (such a filtration is called a Brownian filtration).

The following lemma is a straightforward consequence of the Döblin-Itô formula.

**Lemma 5.53.** *Let  $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  be a locally square integrable function. The process*

$$\left( \exp \left( \int_0^t f(s) dB_s - \frac{1}{2} \int_0^t f(s)^2 ds \right) \right)_{t \geq 0}$$

*is a square integrable martingale.*

*Proof.* From the Döblin-Itô formula we have

$$\begin{aligned} & \exp\left(\int_0^t f(s)dB_s - \frac{1}{2}\int_0^t f(s)^2 ds\right) \\ &= 1 + \int_0^t f(s) \exp\left(\int_0^s f(u)dB_u - \frac{1}{2}\int_0^s f(u)^2 du\right) dB_s. \end{aligned}$$

The random variable  $\int_0^s f(u)dB_u$  is a Gaussian random variable with mean 0 and variance  $\int_0^s f(u)^2 du$ . As a consequence

$$\mathbb{E}\left(\int_0^t f(s)^2 \exp\left(2\int_0^s f(u)dB_u\right) ds\right) < +\infty$$

and the process

$$\int_0^t f(s) \exp\left(\int_0^s f(u)dB_u - \frac{1}{2}\int_0^s f(u)^2 du\right) dB_s$$

is a martingale.  $\square$

**Lemma 5.54.** *Let  $\mathcal{D}$  be the set of compactly supported and piecewise constant functions  $\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ , i.e. the set of functions  $f$  that can be written as*

$$f = \sum_{i=1}^n a_i \mathbf{1}_{(t_{i-1}, t_i]},$$

for some  $0 \leq t_1 \leq \dots \leq t_n$  and  $a_1, \dots, a_n \in \mathbb{R}$ . The linear span of the family

$$\left\{ \exp\left(\int_0^{+\infty} f(s)dB_s - \frac{1}{2}\int_0^{+\infty} f(s)^2 ds\right), f \in \mathcal{D} \right\}$$

is dense in  $\mathbf{L}^2(\mathcal{F}_\infty, \mathbb{P})$ .

*Proof.* Let  $F \in \mathbf{L}^2(\mathcal{F}_\infty, \mathbb{P})$  such that for every  $f \in \mathcal{D}$ ,

$$\mathbb{E}\left(F \exp\left(\int_0^{+\infty} f(s)dB_s - \frac{1}{2}\int_0^{+\infty} f(s)^2 ds\right)\right) = 0.$$

Let  $t_1, \dots, t_n \geq 0$ . We have for every  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ ,

$$\mathbb{E}\left(F \exp\left(\sum_{i=1}^n \lambda_i (B_{t_i} - B_{t_{i-1}})\right)\right) = 0.$$

By analytic continuation, we see that

$$\mathbb{E}\left(F \exp\left(\sum_{i=1}^n \lambda_i (B_{t_i} - B_{t_{i-1}})\right)\right) = 0.$$

actually also holds for every  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ . By using the Fourier transform, it implies that

$$\mathbb{E}(F \mid B_{t_1}, \dots, B_{t_n}) = 0.$$

Since  $t_1, \dots, t_n$  were arbitrary, we conclude that  $\mathbb{E}(F \mid \mathcal{F}_\infty) = 0$ . As a conclusion  $F = 0$ .  $\square$

We are now in position to state the representation theorem.

**Theorem 5.55** (Itô representation theorem). *For every  $F \in \mathbf{L}^2(\mathcal{F}_\infty, \mathbb{P})$ , there is a unique progressively measurable process  $(u_t)_{t \geq 0}$  such that  $\mathbb{E}(\int_0^\infty u_s^2 ds) < +\infty$  and*

$$F = \mathbb{E}(F) + \int_0^{+\infty} u_s dB_s.$$

*Proof.* The uniqueness is immediate as a consequence of the Itô's isometry for stochastic integrals. Let  $\mathcal{A}$  be the set of random variables  $F \in \mathbf{L}^2(\mathcal{F}_\infty, \mathbb{P})$  such that there exists a progressively measurable process  $(u_t)_{t \geq 0}$  such that  $\mathbb{E}(\int_0^\infty u_s^2 ds) < +\infty$  and

$$F = \mathbb{E}(F) + \int_0^{+\infty} u_s dB_s.$$

From the above lemma, it is clear that  $\mathcal{A}$  contains the set of set of random variables

$$\left\{ \exp \left( \int_0^{+\infty} f(s) dB_s - \frac{1}{2} \int_0^{+\infty} f(s)^2 ds \right), f \in \mathcal{D} \right\}.$$

Since this set is total in  $\mathbf{L}^2(\mathcal{F}_\infty, \mathbb{P})$ , we just need to prove that  $\mathcal{A}$  is closed in  $\mathbf{L}^2(\mathcal{F}_\infty, \mathbb{P})$ . So, let  $(F_n)_{n \in \mathbb{N}}$  be a sequence of random variables such that  $F_n \in \mathcal{A}$  and  $F_n \rightarrow_{n \rightarrow \infty} F$  in  $\mathbf{L}^2(\mathcal{F}_\infty, \mathbb{P})$ . There is a progressively measurable process  $(u_t^n)_{t \geq 0}$  such that  $\mathbb{E}(\int_0^\infty (u_s^n)^2 ds) < +\infty$  and

$$F_n = \mathbb{E}(F_n) + \int_0^{+\infty} u_s^n dB_s.$$

By using Itô's isometry, it is seen that the sequence  $u^n$  is a Cauchy sequence and therefore converges to a process  $u$  which is seen to satisfy

$$F = \mathbb{E}(F_n) + \int_0^{+\infty} u_s^n dB_s.$$

$\square$

As a consequence of the representation theorem, we obtain the following description of the square integrable martingales of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ .



**Corollary 5.56.** *Let  $(M_t)_{t \geq 0}$  be a square integrable martingale of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . There is a unique progressively measurable process  $(u_t)_{t \geq 0}$  such that for every  $t \geq 0$ ,  $\mathbb{E} \left( \int_0^t u_s^2 ds \right) < +\infty$  and*

$$M_t = \mathbb{E}(M_0) + \int_0^t u_s dB_s.$$

*In particular,  $(M_t)_{t \geq 0}$  admits a continuous version.*

**Exercise 5.57.** *Show that if  $(M_t)_{t \geq 0}$  is a local martingale of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ , then there is a unique progressively measurable process  $(u_t)_{t \geq 0}$  such that for every  $t \geq 0$ ,  $\mathbb{P} \left( \int_0^t u_s^2 ds < +\infty \right) = 1$  and*

$$M_t = M_0 + \int_0^t u_s dB_s.$$

## 8 Time changed martingales and planar Brownian motion

In the previous section, we proved that any martingale which is adapted to a Brownian filtration can be written as a stochastic integral. In this section, we prove that any martingale can also be represented as a time changed Brownian motion. To prove this fact, we give first first a characterization of the Brownian motion which is interesting in itself. In this section, we denote by  $(\mathcal{F}_t)_{t \geq 0}$  a filtration that satisfies the usual conditions.

**Proposition 5.58** (Lévy characterization theorem). *Let  $(M_t)_{t \geq 0}$  be a continuous local martingale such that  $M_0 = 0$  and such that for every  $t \geq 0$ ,  $\langle M \rangle_t = t$ . The process  $(M_t)_{t \geq 0}$  is a standard Brownian motion.*

*Proof.* Let  $N_t = e^{i\lambda M_t + \frac{1}{2}\lambda^2 t}$ . By using the Döblin-Itô formula, we obtain that for  $s \leq t$ ,

$$N_t = N_s + \int_s^t N_u dM_u.$$

As a consequence, the process  $(N_t)_{t \geq 0}$  is a martingale and, from the above equality we get

$$\mathbb{E} \left( e^{i\lambda(M_t - M_s)} \mid \mathcal{F}_s \right) = e^{-\frac{1}{2}\lambda^2(t-s)}.$$

The process  $(M_t)_{t \geq 0}$  is therefore a continuous process with stationary and independent increments such that  $M_t$  is normally distributed with mean 0 and variance  $t$ . It is thus a Brownian motion.  $\square$

**Exercise 5.59** (Bessel process). *Let  $(B_t)_{t \geq 0}$  be a  $d$ -dimensional Brownian motion  $d \geq 2$ . We denote*

$$\rho_t = \|x + B_t\|,$$

where  $x \neq 0$ . Show that

$$\rho_t - \int_0^t \frac{d-1}{2\rho_s} ds$$

is a standard Brownian motion.

The next proposition shows that continuous martingales behave in a nice way with respect to time changes.

**Proposition 5.60.** *Let  $(C_t)_{t \geq 0}$  be a continuous and increasing process such that for every  $t \geq 0$ ,  $C_t$  is a finite stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Let  $(M_t)_{t \geq 0}$  be a continuous martingale with respect to  $(\mathcal{F}_t)_{t \geq 0}$ . The process  $(M_{C_t})_{t \geq 0}$  is a local martingale with respect to the filtration  $(\mathcal{F}_{C_t})_{t \geq 0}$ . Moreover  $\langle M_C \rangle = \langle M \rangle_C$ .*

*Proof.* By using localization, if necessary, we can assume that  $C$  is bounded. According to the Doob's stopping theorem (see Theorem 1.34), we need to prove that for every bounded stopping time  $T$  of the filtration  $(\mathcal{F}_{C_t})_{t \geq 0}$ , we have  $\mathbb{E}(M_{C_T}) = 0$ . But  $C_T$  is obviously a bounded stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$  and thus from Doob's stopping theorem we have  $\mathbb{E}(M_{C_T}) = 0$ . The same argument shows that  $M_C^2 - \langle M \rangle_C$ .  $\square$

**Exercise 5.61.** *Let  $(C_t)_{t \geq 0}$  be an increasing and right continuous process such that for every  $t \geq 0$ ,  $C_t$  is a finite stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Let  $(M_t)_{t \geq 0}$  be a continuous martingale with respect to  $(\mathcal{F}_t)_{t \geq 0}$  such that  $M$  is constant on each interval  $[C_{t-}, C_t]$ . Show that the process  $(M_{C_t})_{t \geq 0}$  is a continuous local martingale with respect to the filtration  $(\mathcal{F}_{C_t})_{t \geq 0}$  and that  $\langle M_C \rangle = \langle M \rangle_C$ .*

We can now prove the following nice representation result for martingales.

**Theorem 5.62** (Dambis, Dubins-Schwarz theorem). *Let  $(M_t)_{t \geq 0}$  be a continuous martingale such that  $M_0 = 0$  and  $\langle M \rangle_\infty = +\infty$ . There exists a Brownian motion  $(B_t)_{t \geq 0}$ , such that for every  $t \geq 0$ ,*

$$M_t = B_{\langle M \rangle_t}.$$

*Proof.* Let  $C_t = \inf\{s \geq 0, \langle M \rangle_s > t\}$ .  $(C_t)_{t \geq 0}$  is an increasing and right continuous process such that for every  $t \geq 0$ ,  $C_t$  is a finite stopping time of the filtration  $(\mathcal{F}_t)_{t \geq 0}$  and  $M$  is obviously constant on each interval  $[C_{t-}, C_t]$ . From the previous exercise the process  $(M_{C_t})_{t \geq 0}$  is a continuous local martingale whose quadratic variation is equal to  $t$ . From Lévy's characterization theorem, it is thus a Brownian motion.  $\square$

**Exercise 5.63.** *Show that if  $(M_t)_{t \geq 0}$  is a continuous local martingale such that  $M_0 = 0$  and  $\langle M \rangle_\infty = +\infty$ , there exists a Brownian motion  $(B_t)_{t \geq 0}$ , such that for every  $t \geq 0$ ,*

$$M_t = B_{\langle M \rangle_t}.$$

**Exercise 5.64.** Let  $(u_t)_{t \geq 0}$  be a continuous adapted process and let  $(B_t)_{t \geq 0}$  be a Brownian motion. Show that for every  $T \geq 0$ , the process  $\left(\int_0^t u_s dB_s\right)_{0 \leq t \leq T}$  has  $\frac{1}{2} - \varepsilon$  Hölder paths, where  $0 < \varepsilon \leq \frac{1}{2}$ .

**Exercise 5.65** (Construction of one-dimensional diffusions). Let  $b : \mathbb{R} \rightarrow \mathbb{R}$  and let  $\sigma : \mathbb{R} \rightarrow (0, +\infty)$  be two continuous functions. Let

$$s(x) = \int_0^x \exp\left(-2 \int_0^y \frac{|b(z)|}{\sigma(z)^2} dz\right) dy, \quad x \in \mathbb{R}.$$

and let  $(B_t)_{t \geq 0}$  be a Brownian motion. For  $u \geq 0$ , we denote

$$A_u = \int_0^u \frac{ds}{s'(B_s)\sigma^2(B_s)} = \int_0^u \frac{\exp\left(2 \int_0^{B_s} \frac{|b(z)|}{\sigma(z)^2} dz\right)}{\sigma^2(B_s)} ds.$$

Show that the process

$$\left(s^{-1}\left(B_{\inf\{u, A_u > t\}}\right)\right)_{t \geq 0}$$

is a Markov process whose semigroup is generated by the operator

$$Lf = b(x)f'(x) + \frac{1}{2}\sigma(x)^2 f''(x), \quad f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R}).$$

Therefore, all one-dimensional diffusions may be constructed from the Brownian motion.

The study of the planar Brownian is deeply connected to the theory of analytic functions. The fundamental property of the Brownian curve is that it is a conformal invariant. By definition, a complex Brownian motion is a process  $(B_t)_{t \geq 0}$  in the complex plane that can be decomposed as  $B_t = B_t^1 + iB_t^2$  where  $B^1$  and  $B^2$  are independent Brownian motions. The following proposition is easily proved as a consequence of Döblin-Itô formula and of the Dambins, Dubins-Schwarz theorem.

**Proposition 5.66** (Conformal invariance). Let  $(B_t)_{t \geq 0}$  be a complex Brownian motion and  $f : \mathbb{C} \rightarrow \mathbb{C}$  be an analytic function. Then we have for every  $t \geq 0$ ,

$$f(B_t) = f(0) + \int_0^t f'(B_s) dB_s.$$

As a consequence, there exists a complex Brownian motion  $(\beta_t)_{t \geq 0}$  such that for every  $t \geq 0$ ,

$$f(B_t) = f(0) + \beta \int_0^t |f'(B_s)|^2 ds.$$

To study the complex Brownian motion, it is useful to look at it in polar coordinates. It leads to the so-called skew-product decomposition of the complex Brownian motion that we study in the following exercise.

**Exercise 5.67** (Skew-product decomposition). Let  $(B_t)_{t \geq 0}$  be a complex Brownian motion started at  $z \neq 0$ .

1. Show that for  $t \geq 0$ ,

$$B_t = z \exp \left( \int_0^t \frac{dB_s}{B_s} \right).$$

2. Show that there exists a complex Brownian motion  $(\beta_t)_{t \geq 0}$  such that

$$B_t = z \exp \left( \beta \int_0^t \frac{ds}{\rho_s^2} \right),$$

where  $\rho_t = |B_t|$ .

3. Show that the process  $(\rho_t)_{t \geq 0}$  is independent from the Brownian motion  $(\gamma_t)_{t \geq 0} = (\mathbf{Im}(\beta_t))_{t \geq 0}$ .

4. We denote  $\theta_t = \mathbf{Im} \left( \int_0^t \frac{dB_s}{B_s} \right)$  which can be interpreted as a winding number around 0 of the complex Brownian motion paths. For  $r > |z|$ , we consider the stopping time

$$T_r = \inf\{t \geq 0, |B_t| = r\}.$$

Compute for every  $r > |z|$ , the distribution of the random variable

$$\frac{1}{\ln(r/|z|)} \theta_{T_r}.$$

5. Prove Spitzer theorem: In distribution, we have the following convergence

$$\frac{2\theta_t}{\ln t} \rightarrow_{+\infty} C,$$

where  $C$  is a Cauchy random variable with parameter 1 that is a random variable with density  $\frac{1}{\pi(1+x^2)}$ .

**Exercise 5.68.** Let  $(B_t)_{t \geq 0}$  be a complex Brownian motion started at  $z$  such that  $\mathbf{Im}(z) > 0$ . We consider the stopping time

$$T = \inf\{t \geq 0, \mathbf{Im}(B_t) = 0\}.$$

Compute the distribution of the random variable  $\mathbf{Re}(B_T)$ .

## 9 Burkholder-Davis-Gundy inequalities

In this section, we study some of the most important martingale inequalities: The Burkholder-Davis-Gundy inequalities. Interestingly, the range of application of these inequalities is very large and they play an important role in harmonic analysis and the study of singular integrals. These inequalities admit several proofs. We present here a proof using the Döblin-Itô formula and an interesting domination inequality which is due to Lenglart.

**Proposition 5.69** (Lenglart inequality). *Let  $(N_t)_{t \geq 0}$  be a positive adapted continuous process and  $(A_t)_{t \geq 0}$  be an increasing process such that  $A_0 = 0$ . Assume that for every bounded stopping time  $\tau$ ,*

$$\mathbb{E}(N_\tau) \leq \mathbb{E}(A_\tau).$$

*Then, for every  $k \in (0, 1)$ , and  $T \geq 0$ ,*

$$\mathbb{E} \left( \left( \sup_{0 \leq t \leq T} N_t \right)^k \right) \leq \frac{2-k}{1-k} \mathbb{E}(A_T^k).$$

*Proof.* Let  $x, y > 0$ . We denote  $R = \inf\{0 \leq t \leq T, A_t \geq y\}$  and  $S = \inf\{0 \leq t \leq T, X_t \geq x\}$  with the usual convention that the infimum of the empty set is  $\infty$ . We first observe that

$$\begin{aligned} \mathbb{P} \left( \sup_{0 \leq t \leq T} X_t \geq x, A_T \leq y \right) &\leq \mathbb{P}(X_{T \wedge S \wedge R} \geq x) \\ &\leq \frac{1}{x} \mathbb{E}(A_{T \wedge S \wedge R}) \\ &\leq \frac{1}{x} \mathbb{E}(A_T \wedge y). \end{aligned}$$

Let now  $F$  be a continuous increasing function from  $\mathbb{R}_{>0}$  into  $\mathbb{R}_{>0}$  with  $F(0) = 0$ . By using the previous inequality and Fubini theorem, we have

$$\begin{aligned} \mathbb{E}(F(\sup_{0 \leq t \leq T} X_t)) &= \mathbb{E} \left( \int_0^{+\infty} \mathbf{1}_{\sup_{0 \leq t \leq T} X_t > x} dF(x) \right) \\ &\leq \int_0^{+\infty} \left[ \mathbb{P} \left( \sup_{0 \leq t \leq T} X_t \geq x, A_T \leq x \right) + \mathbb{P}(A_T \leq x) \right] dF(x) \\ &\leq \int_0^{+\infty} \left[ \frac{1}{x} \mathbb{E}(A_T \wedge x) + \mathbb{P}(A_T \leq x) \right] dF(x) \\ &\leq \int_0^{+\infty} \left[ \frac{1}{x} \mathbb{E}(A_T \mathbf{1}_{A_T \leq x}) + 2\mathbb{P}(A_T \leq x) \right] dF(x) \\ &\leq 2\mathbb{E}(F(A_T)) + \mathbb{E} \left( A_T \int_{A_T}^{+\infty} \frac{dF(x)}{x} \right) \end{aligned}$$

Taking  $f(x) = x^k$ , we obtain the desired result.  $\square$

We now turn to the Burkholder-Davis-Gundy inequalities.

**Theorem 5.70** (Burkholder-Davis-Gundy inequalities). *Let  $T > 0$  and  $(M_t)_{0 \leq t \leq T}$  be a continuous local martingale such that  $M_0 = 0$ . For every  $0 < p < \infty$ , there exist universal constants  $c_p$  and  $C_p$ , independent of  $T$  and  $(M_t)_{0 \leq t \leq T}$  such that*

$$c_p \mathbb{E} \left( \langle M \rangle_T^{\frac{p}{2}} \right) \leq \mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^p \right) \leq C_p \mathbb{E} \left( \langle M \rangle_T^{\frac{p}{2}} \right).$$

*Proof.* By stopping it is enough to prove the result for bounded  $M$ . Let  $q \geq 2$ . From the Döblin-Itô formula we have

$$\begin{aligned} d|M_t|^q &= q|M_t|^{q-1} \mathbf{sgn}(M_t) dM_t + \frac{1}{2} q(q-1) |M_t|^{q-2} d\langle M \rangle_t \\ &= q \mathbf{sgn}(M_t) |M_t|^{q-1} dM_t + \frac{1}{2} q(q-1) |M_t|^{q-2} d\langle M \rangle_t. \end{aligned}$$

As a consequence of the Doob's stopping theorem, we get that for every bounded stopping time  $\tau$ ,

$$\mathbb{E} (|M_\tau|^q) \leq \frac{1}{2} q(q-1) \mathbb{E} \left( \int_0^\tau |M_t|^{q-2} d\langle M \rangle_t \right).$$

From the Lenglart's domination inequality, we deduce then that for every  $k \in (0, 1)$ ,

$$\mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t|^q \right)^k \right) \leq \frac{2-k}{1-k} \left( \frac{1}{2} q(q-1) \right)^k \mathbb{E} \left( \left( \int_0^T |M_t|^{q-2} d\langle M \rangle_t \right)^k \right).$$

We now bound

$$\begin{aligned} \mathbb{E} \left( \left( \int_0^T |M_t|^{q-2} d\langle M \rangle_t \right)^k \right) &\leq \mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^{k(q-2)} \left( \int_0^T d\langle M \rangle_t \right)^k \right) \\ &\leq \mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^{kq} \right)^{1-\frac{2}{q}} \mathbb{E} \left( \langle M \rangle_T^{\frac{kq}{2}} \right)^{\frac{2}{q}}. \end{aligned}$$

As a consequence, we obtain

$$\begin{aligned} &\mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t|^q \right)^k \right) \\ &\leq \frac{2-k}{1-k} \left( \frac{1}{2} q(q-1) \right)^k \mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^{kq} \right)^{1-\frac{2}{q}} \mathbb{E} \left( \langle M \rangle_T^{\frac{kq}{2}} \right)^{\frac{2}{q}}. \end{aligned}$$

Letting  $p = qk$  yields the claimed result, that is

$$\mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^p \right) \leq C_p \mathbb{E} \left( \langle M \rangle_T^{\frac{p}{2}} \right).$$

We proceed now to the proof of the left hand side inequality. We have,

$$M_t^2 = \langle M \rangle_t + 2 \int_0^t M_s dM_s.$$

Therefore, we get

$$\mathbb{E} \left( \langle M \rangle_T^{\frac{p}{2}} \right) \leq A_p \left( \mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^p \right) + \mathbb{E} \left( \sup_{0 \leq t \leq T} \left| \int_0^t M_s dM_s \right|^{p/2} \right) \right).$$

By using the previous argument, we now have

$$\begin{aligned} \mathbb{E} \left( \sup_{0 \leq t \leq T} \left| \int_0^t M_s dM_s \right|^{p/2} \right) &\leq B_p \mathbb{E} \left( \left( \int_0^T M_s^2 d\langle M \rangle_s \right)^{p/4} \right) \\ &\leq B_p \mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^{p/2} \langle M \rangle_T^{p/4} \right) \\ &\leq B_p \mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^p \right)^{1/2} \mathbb{E} \left( \langle M \rangle_T^{p/2} \right)^{1/2}. \end{aligned}$$

As a conclusion, we obtained

$$\begin{aligned} &\mathbb{E} \left( \langle M \rangle_T^{\frac{p}{2}} \right) \\ &\leq A_p \left( \mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^p \right) + B_p \mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |M_t| \right)^p \right)^{1/2} \mathbb{E} \left( \langle M \rangle_T^{p/2} \right)^{1/2} \right). \end{aligned}$$

This is an inequality of the form

$$x^2 \leq A_p (y^2 + B_p xy),$$

which easily implies

$$c_p x^2 \leq y^2,$$

thanks to the inequality  $2xy \leq \frac{1}{\delta} x^2 + \delta y^2$ , with a conveniently chosen  $\delta$ .  $\square$

**Exercise 5.71.** Let  $T > 0$  and  $(M_t)_{0 \leq t \leq T}$  be a continuous local martingale. Consider the process

$$Z_t = e^{\int_0^t V_s ds} \int_0^t e^{-\int_0^s V_u du} dM_s,$$

where  $(V_t)_{0 \leq t \leq T}$  is a non positive adapted and continuous process. Show that for every  $0 < p < \infty$ , there is a universal constant  $C_p$ , independent of  $T$ ,  $(M_t)_{0 \leq t \leq T}$  and  $(V_t)_{0 \leq t \leq T}$  such that

$$\mathbb{E} \left( \left( \sup_{0 \leq t \leq T} |Z_t| \right)^p \right) \leq C_p \mathbb{E} \left( \langle M \rangle_T^{\frac{p}{2}} \right).$$

## 10 Girsanov theorem

In this section, we describe a theorem which has far reaching consequences in mathematical finance: The Girsanov theorem. It describes the impact of a probability change on stochastic calculus.

Let  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a filtered probability space. We assume that  $(\mathcal{F}_t)_{t \geq 0}$  is the usual completion of the filtration of a Brownian motion  $(B_t)_{t \geq 0}$ . Let  $\mathbb{Q}$  be a probability measure on  $\mathcal{F}_\infty$  which is equivalent to  $\mathbb{P}$ . We denote by  $D$  the density of  $\mathbb{Q}$  with respect to  $\mathbb{P}$ .

**Theorem 5.72** (Girsanov theorem). *There exists a progressively measurable process  $(\Theta_t)_{t \geq 0}$  such that for every  $t \geq 0$ ,  $\mathbb{P} \left( \int_0^t \Theta_s^2 ds < +\infty \right) = 1$  and*

$$\mathbb{E}(D \mid \mathcal{F}_t) = \exp \left( \int_0^t \Theta_s dB_s - \frac{1}{2} \int_0^t \Theta_s^2 ds \right).$$

Moreover, the process

$$B_t - \int_0^t \Theta_s ds$$

is a Brownian motion on the filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{Q})$ . As a consequence, a continuous and adapted process  $(X_t)_{t \geq 0}$  is a  $\mathbb{P}$ -semimartingale if and only if it is a  $\mathbb{Q}$ -semimartingale.

*Proof.* Since  $\mathbb{P}$  and  $\mathbb{Q}$  are equivalent on  $\mathcal{F}_\infty$ , there are of course also equivalent on  $\mathcal{F}_t$  for every  $t \geq 0$ . The density of  $\mathbb{Q}_{/\mathcal{F}_t}$  with respect to  $\mathbb{P}_{/\mathcal{F}_t}$  is given by  $D_t = \mathbb{E}^{\mathbb{P}}(D \mid \mathcal{F}_t)$ . As a consequence, the process  $D_t$  is a positive martingale. From Itô's representation theorem, we therefore deduce that there exists a progressively measurable process  $(u_t)_{t \geq 0}$  such that

$$D_t = 1 + \int_0^t u_s dB_s.$$

Let now  $\Theta_t = \frac{u_t}{D_t}$ . We have then,

$$D_t = 1 + \int_0^t \Theta_s D_s dB_s.$$



By using Döblin-Itô's formula to the process  $D_t \exp\left(-\int_0^t \Theta_s dB_s + \frac{1}{2} \int_0^t \Theta_s^2 ds\right)$ , we see that it implies

$$D_t = \exp\left(\int_0^t \Theta_s dB_s - \frac{1}{2} \int_0^t \Theta_s^2 ds\right).$$

We now want to prove that the process  $B_t - \int_0^t \Theta_s ds$  is a  $\mathbb{Q}$ -Brownian motion. It is clear the  $\mathbb{Q}$ -quadratic variation of this process is  $t$ . From the Lévy's characterization result, we therefore just need to prove that it is a  $\mathbb{Q}$  local martingale. For this, we are going to prove that the process

$$N_t = \left(B_t - \int_0^t \Theta_s ds\right) \exp\left(\int_0^t \Theta_s dB_s - \frac{1}{2} \int_0^t \Theta_s^2 ds\right)$$

is a  $\mathbb{P}$ -local martingale. Indeed, from the integration by parts formula, it is immediate that

$$dN_t = D_t dB_t + \left(B_t - \int_0^t \Theta_s ds\right) dD_t.$$

Since  $D_t$  is the density of  $\mathbb{Q}_{\mathcal{F}_t}$  with respect to  $\mathbb{P}_{\mathcal{F}_t}$ , it is then easy to deduce that  $N_t$  is a  $\mathbb{P}$ -local martingale and thus a  $\mathbb{P}$  Brownian motion.  $\square$

**Exercise 5.73.** Let  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a filtered probability space that satisfies the usual conditions. As before, let  $\mathbb{Q}$  be a probability measure on  $\mathcal{F}_\infty$  which is equivalent to  $\mathbb{P}$ . We denote by  $D$  the density of  $\mathbb{Q}$  with respect to  $\mathbb{P}$  and  $D_t = \mathbb{E}^{\mathbb{P}}(D | \mathcal{F}_t)$ . Let  $(M_t)_{t \geq 0}$  be a  $\mathbb{P}$  local martingale. Show that the process

$$N_t = M_t - \int_0^t \frac{d\langle M, D \rangle_s}{D_s}$$

is a  $\mathbb{Q}$  local martingale. As a consequence, a continuous and adapted process  $(X_t)_{t \geq 0}$  is a  $\mathbb{P}$ -semimartingale if and only if it is a  $\mathbb{Q}$ -semimartingale.

**Exercise 5.74.** Let  $(B_t)_{t \geq 0}$  be a Brownian motion. We denote by  $\mathbb{P}$  the Wiener measure, by  $(\pi_t)_{t \geq 0}$  the coordinate process and by  $(\mathcal{F}_t)_{t \geq 0}$  its natural filtration.

1. Let  $\mu \in \mathbb{R}$  and  $\mathbb{P}^\mu$  be the distribution of the process  $(B_t + \mu t)_{t \geq 0}$ . Show that for every  $t \geq 0$ ,

$$\mathbb{P}^\mu_{/\mathcal{F}_t} \ll \mathbb{P}_{/\mathcal{F}_t},$$

and that

$$\frac{d\mathbb{P}^\mu_{/\mathcal{F}_t}}{d\mathbb{P}_{/\mathcal{F}_t}} = e^{\mu\pi_t - \frac{\mu^2}{2}t}.$$

2. Is it true that

$$\mathbb{P}^\mu_{/\mathcal{F}_\infty} \ll \mathbb{P}_{/\mathcal{F}_\infty} \quad ?$$

3. For  $a \in \mathbb{R}_{\geq 0}$ , we denote

$$T_a = \inf\{t \geq 0, B_t + \mu t = a\}.$$

Compute the density function of  $T_a$  (You may use the previous question).

4. More generally, let  $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  be a measurable function such that for every  $t \geq 0$ ,  $\int_0^t f^2(s) ds < +\infty$ . We denote by  $\mathbb{P}^f$  the distribution of the process  $\left(B_t + \int_0^t f(s) ds\right)_{t \geq 0}$ . Show that for every  $t \geq 0$ ,

$$\mathbb{P}_{/\mathcal{F}_t}^f \ll \mathbb{P}_{/\mathcal{F}_t},$$

and that

$$\frac{d\mathbb{P}_{/\mathcal{F}_t}^f}{d\mathbb{P}_{/\mathcal{F}_t}} = e^{\int_0^t f(s) d\pi_s - \frac{1}{2} \int_0^t f^2(s) ds}.$$

Let  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  be a filtered probability space that satisfies the usual conditions and let  $(B_t)_{t \geq 0}$  be a Brownian motion on it. Let now  $(\Theta_t)_{t \geq 0}$  be a progressively measurable process such that for every  $t \geq 0$ ,  $\mathbb{P}\left(\int_0^t \Theta_s^2 ds < +\infty\right) = 1$ . We denote

$$Z_t = \exp\left(\int_0^t \Theta_s dB_s - \frac{1}{2} \int_0^t \Theta_s^2 ds\right), \quad t \geq 0.$$

As a consequence of the Döblin-Itô's formula, it is clear that  $(Z_t)_{t \geq 0}$  is a local martingale. If we assume that  $(Z_t)_{t \geq 0}$  is a uniformly integrable martingale, then it is easy to see that on the  $\sigma$ -field  $\mathcal{F}_\infty$ , there is a unique probability measure  $\mathbb{Q}$  equivalent to  $\mathbb{P}$  such that for every  $t \geq 0$ ,

$$\frac{d\mathbb{Q}_{/\mathcal{F}_t}}{d\mathbb{P}_{/\mathcal{F}_t}} = Z_t, \quad \mathbb{P} - p.s.$$

The same argument as before shows then that with respect to  $\mathbb{Q}$ , the process

$$B_t - \int_0^t \Theta_s ds$$

is a Brownian motion.

It is thus important to decide whether or not  $(Z_t)_{t \geq 0}$  is a uniformly martingale. The following two lemmas provide sufficient conditions that it is.

**Lemma 5.75.** *If for every  $t \geq 0$ ,*

$$\mathbb{E}(Z_t) = 1,$$

*then  $(Z_t)_{t \geq 0}$  is a uniformly integrable martingale.*

*Proof.* The process  $Z$  is a non negative local martingale and thus a supermartingale.  $\square$

The second condition is known as the Novikov condition, it is often easier to check in practice than the previous one.

**Lemma 5.76** (Novikov condition). *If*

$$\mathbb{E} \left( \exp \left( \frac{1}{2} \int_0^{+\infty} \Theta_s^2 ds \right) \right) < +\infty,$$

*then  $(Z_t)_{t \geq 0}$  is a uniformly integrable martingale.*

*Proof.* We denote  $M_t = \int_0^t \Theta_s dB_s$ . As a consequence of the integrability condition

$$\mathbb{E} \left( \exp \left( \frac{1}{2} \langle M \rangle_\infty \right) \right) < +\infty,$$

the random variable  $\langle M \rangle_\infty$  has moments of all order. So from Burkholder-Davis-Gundy inequalities,  $\sup_{t \geq 0} |M_t|$  has moments of all orders, which implies that  $M$  is a uniformly integrable martingale. We have then

$$\exp \left( \frac{1}{2} M_\infty \right) = \exp \left( \frac{1}{2} M_\infty - \frac{1}{4} \langle M \rangle_\infty \right) \exp \left( \frac{1}{4} \langle M \rangle_\infty \right).$$

The Cauchy-Schwarz inequality implies then that  $\mathbb{E} \left( \exp \left( \frac{1}{2} M_\infty \right) \right) < +\infty$ . We deduce from the Doob's convergence theorem that the process  $\exp \left( \frac{1}{2} M \right)$  is a uniformly integrable submartingale. Let now  $\eta < 1$ . We have

$$\exp \left( \eta M_t - \frac{\eta^2}{2} \langle M \rangle_t \right) = \left( \exp \left( M_t - \frac{1}{2} \langle M \rangle_t \right) \right)^{\eta^2} \exp \left( \frac{\eta M_t}{1 + \eta} \right)^{1 - \eta^2}.$$

Hölder's inequality shows then that

$$\begin{aligned} & \mathbb{E} \left( \exp \left( \eta M_t - \frac{\eta^2}{2} \langle M \rangle_t \right) \right) \\ & \leq \mathbb{E} \left( \exp \left( M_t - \frac{1}{2} \langle M \rangle_t \right) \right)^{\eta^2} \mathbb{E} \left( \exp \left( \frac{\eta M_t}{1 + \eta} \right) \right)^{1 - \eta^2} \\ & \leq \mathbb{E} \left( \exp \left( M_t - \frac{1}{2} \langle M \rangle_t \right) \right)^{\eta^2} \mathbb{E} \left( \exp \left( \frac{M_t}{2} \right) \right)^{2\eta(1 - \eta)} \\ & \leq \mathbb{E} \left( \exp \left( M_t - \frac{1}{2} \langle M \rangle_t \right) \right)^{\eta^2} \mathbb{E} \left( \exp \left( \frac{M_\infty}{2} \right) \right)^{2\eta(1 - \eta)} \end{aligned}$$

If we can prove that  $\mathbb{E} \left( \exp \left( \eta M_t - \frac{\eta^2}{2} \langle M \rangle_t \right) \right) = 1$ , then by letting  $\eta \rightarrow 1$  in the above inequality, we would get

$$\mathbb{E} \left( \exp \left( M_t - \frac{1}{2} \langle M \rangle_t \right) \right) \geq 1$$

and thus  $\mathbb{E} \left( \exp \left( M_t - \frac{1}{2} \langle M \rangle_t \right) \right) = 1$ .

Let  $p > 1$  such that  $\frac{\eta\sqrt{p}}{\sqrt{p-1}} \leq 1$ . Consider  $r = \frac{\sqrt{p+1}}{\sqrt{p-1}}$  and  $s = \frac{\sqrt{p+1}}{2}$  so that  $1/r + 1/s = 1$ . Using

$$\exp \left( \eta M_t - \frac{\eta^2}{2} \langle M \rangle_t \right)^p = \exp \left( \sqrt{\frac{p}{r}} \eta M_t - \frac{p}{2} \eta^2 \langle M \rangle_t \right) \exp \left( \left( p\eta - \sqrt{\frac{p}{r}} \right) M_t \right)$$

and then Hölder's inequality, shows that there is a constant  $C$  (depending only on  $M$ ) such that for any stopping time  $T$

$$\mathbb{E} \left( \exp \left( \eta M_T - \frac{\eta^2}{2} \langle M \rangle_T \right)^p \right) \leq C.$$

By the Doob's maximal inequality, it implies that the local martingale

$$\exp \left( \eta M_t - \frac{\eta^2}{2} \langle M \rangle_t \right)$$

is actually a true martingale. This implies

$$\mathbb{E} \left( \exp \left( \eta M_t - \frac{\eta^2}{2} \langle M \rangle_t \right) \right) = 1$$

and the desired conclusion follows.  $\square$

**Exercise 5.77** (Lévy area formula). Let  $(B_t^1, B_t^2)_{t \geq 0}$  be a two dimensional Brownian motion. The random variable

$$S_t = \int_0^t B_s^1 dB_s^2 - B_s^2 dB_s^1$$

is two times the algebraic area swept out by the path  $s \rightarrow (B_s^1, B_s^2)$  in the time interval  $[0, t]$ .

1. Show that

$$(B_t^1, B_t^2, S_t)_{t \geq 0} = \left( B_t^1, B_t^2, \beta \int_0^t \rho_s^2 ds \right)_{t \geq 0},$$

where  $(\beta_t)_{t \geq 0}$  is a Brownian motion independent from  $(\rho_t)_{t \geq 0}$  with  $\rho_t^2 = (B_t^1)^2 + (B_t^2)^2$ .

**Hint:** You may write  $S_t = \int_0^t \rho_s d\gamma_s$  where  $(\gamma_t)_{t \geq 0}$  is a Brownian motion independent from  $\int_0^t \frac{B_s^1 dB_s^1 + B_s^2 dB_s^2}{\rho_s}$ .

2. Show that for  $\lambda \in \mathbb{R}$ ,

$$\mathbb{E}(e^{i\lambda S_t} \mid B_t^1 = x, B_t^2 = y) = \mathbb{E}\left(e^{-\frac{\lambda^2}{2} \int_0^t \rho_s^2 ds} \mid \rho_t = \sqrt{x^2 + y^2}\right).$$

3. Show that the process

$$D_t = \exp\left(\frac{\lambda}{2}(\rho_t^2 - 2t) - \frac{\lambda^2}{2} \int_0^t \rho_s^2 ds\right)$$

is a martingale.

4. Consider the probability measure  $\mathbb{Q}_{\mathcal{F}_t} = D_t \mathbb{P}_{\mathcal{F}_t}$ . Show that under  $\mathbb{Q}$ ,  $(B_t^1, B_t^2)_{t \geq 0}$  is a Gaussian process, whose covariance is to be computed.

5. Conclude that

$$\mathbb{E}(e^{i\lambda S_t} \mid B_t^1 = x, B_t^2 = y) = \frac{\lambda}{\sinh \lambda} \exp\left(-\frac{1}{2}(x^2 + y^2)(\lambda \coth \lambda - 1)\right).$$

**Exercise 5.78.** Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion and  $Y$  be an independent random variable such that  $\mathbb{P}(Y = -1) = \frac{1}{2}$  and  $\mathbb{P}(Y = 1) = \frac{1}{2}$ . We consider the process  $X_t = B_t + \mu t Y$ , where  $\mu \in \mathbb{R}$ . By using Girsanov theorem provide the semimartingale decomposition of  $(X_t)_{t \geq 0}$  in its own filtration.

## Notes and Comments

**Section 1.** It can be shown (see instance [?]) that almost surely

$$\sup_{\Pi \in \Delta[0, T]} \sum_{k=0}^{n-1} |B_{t_{k+1}} - B_{t_k}|^2 = +\infty,$$

that is, the 2-variation of Brownian motion paths is almost surely  $+\infty$ .

**Sections 2,3,4.** The study of integrals with respect to Brownian motion has a long history going back at least to Wiener, but the breakthrough of understanding the role of adaptedness of integrands is due to Itô in his famous 1944 paper [?]. The general theory of stochastic integration with respect to martingales, local martingales and semimartingales was later developed and popularized in particular by Watanabe (see [?]) and Meyer (see [?]). The hypothesis of the continuity of the integrator process is not strictly necessary and an integral with respect to a possibly discontinuous local martingale can be defined (see the book by Bichteler [?] or Protter [?]).

**Section 5.** The work by Döbblin has long been unknown to mathematicians. Before being killed during the second world war, in a letter sent to the French academy of sciences in 1940 and only opened in 2000, Döbblin presented a version of the Döbblin-Itô formula. His formula for a Brownian motion  $(B_t)_{t \geq 0}$  writes as

$$f(B_t) = f(0) + \beta \int_0^t f'(B_s)^2 ds + \frac{1}{2} \int_0^t f''(B_s) ds,$$

where  $\beta$  is another Brownian motion. Observe that this formula is of course consistent with the formula we stated in the text since, according to the Dambis, Dubins-Schwarz theorem

$$\int_0^t f'(B_s) dB_s = \beta \int_0^t f'(B_s)^2 ds,$$

for some Brownian motion  $\beta$ . What is remarkable is that Döbblin obtained his formula even before stochastic integration was conceived ! We refer to [?] and [?] for more details about the fascinating story of Döbblin.

**Section 6.** Results of this Section may be extended to more general diffusion processes and hint the deep and rich connection between Markov processes and potential theory. There are several books devoted to the interactions between potential analysis and Markov processes, we mention in particular the classical books by Chung [?] and Doob [?] or to the more recent book by Durrett [?].

**Section 7.** The representation theorem plays a fundamental role in mathematical finance, where the integrand appears as a hedging strategy, see the book by Karatzas and Shreve [?]. A formula for the integrand can be obtained through Malliavin calculus (see Exercise 6.36 in Chapter 6)

**Section 8.** There is a deep interaction between planar Brownian motion and complex analysis, due to the conformal invariance of the planar Brownian motion paths. Conformally invariant processes in the plane play an important role in statistical mechanics, see the book by Lawler [?] for a detailed account.

**Section 9.** Martingales techniques in general and the Burkholder-Davis-Gundy inequalities in particular have been found to play an important role in the analysis of singular integrals where sharp, dimension free estimates can be obtained, see the survey [?] by Bañuelos.

**Section 10.** Girsanov theorem is another pillar of mathematical finance. As illustrated in the Lévy area formula exercise, it may be used to compute distributions of Brownian functionals. This formula for the Fourier transform of the area swept out by the planar Brownian motion is due to Lévy [?]. An

independent and analytic proof may be found in Gaveau [?]. For many other computations related to the stochastic area formula we refer to the book by Yor [?].

## Chapter 6

# Stochastic differential equations and Malliavin calculus

In this chapter we study the stochastic differential equations. These are the equations associated to Itô integral. Stochastic differential equations are a great device to study the existence of a Markov process with a given generator, a problem studied in Chapter 4. In the first part of the chapter, we study existence and uniqueness questions and then prove that solution of stochastic differential equations are Markov processes. The second part of the chapter is an introduction to Malliavin calculus which is the Sobolev regularity theory of functionals defined on the Wiener space. Malliavin calculus is applied to show the existence of a smooth density for solutions of stochastic differential equations under an ellipticity assumption.

### 1 Existence and uniqueness of solutions

As usual, we consider a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  which satisfies the usual conditions and on which is defined a  $n$ -dimensional Brownian motion  $(B_t)_{t \geq 0}$ . Let  $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , and  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  be functions.

**Theorem 6.1.** *Let us assume that there exists  $C > 0$  such that*

$$\|b(x) - b(y)\| + \|\sigma(x) - \sigma(y)\| \leq C\|x - y\|, x, y \in \mathbb{R}^n.$$

*Then, for every  $x_0 \in \mathbb{R}^n$ , there exists a unique continuous and adapted process  $(X_t^{x_0})_{t \geq 0}$  such that for  $t \geq 0$*

$$X_t^{x_0} = x_0 + \int_0^t b(X_s^{x_0}) ds + \int_0^t \sigma(X_s^{x_0}) dB_s. \quad (6.1)$$

*Moreover, for every  $T \geq 0$ , we have*

$$\mathbb{E} \left( \sup_{0 \leq s \leq T} \|X_s\|^2 \right) < +\infty.$$

*Proof.* Let us first observe that from our assumptions, there exists  $K > 0$  such that

1.  $\|b(x) - b(y)\| + \|\sigma(x) - \sigma(y)\| \leq K\|x - y\|, x, y \in \mathbb{R}^n;$
2.  $\|b(x)\| + \|\sigma(x)\| \leq K(1 + \|x\|), x \in \mathbb{R}^n.$



The idea is to apply a fixed point theorem in a convenient Banach space. For  $T > 0$ , let us consider the space  $\mathcal{E}_T$  of continuous and adapted processes such that

$$\mathbb{E} \left( \sup_{0 \leq s \leq T} \|X_s\|^2 \right) < +\infty$$

endowed with the norm

$$\|X\|^2 = \mathbb{E} \left( \sup_{0 \leq s \leq T} \|X_s\|^2 \right).$$

It is easily seen that  $(\mathcal{E}_T, \|\cdot\|)$  is a Banach space.

**Step one:** We first prove that if a continuous and adapted process  $(X_t^{x_0})_{t \geq 0}$  is a solution of the equation (6.1) then, for every  $T > 0$ ,  $(X_t^{x_0})_{0 \leq t \leq T} \in \mathcal{E}_T$ .

Let us fix  $T > 0$  and consider for  $n \in \mathbb{N}$  the stopping times

$$T_n = \inf\{t \geq 0, \|X_t^{x_0}\| > n\},$$

For  $t \leq T$ ,

$$X_{t \wedge T_n}^{x_0} = x_0 + \int_0^{t \wedge T_n} b(X_s^{x_0}) ds + \int_0^{t \wedge T_n} \sigma(X_s^{x_0}) dB_s.$$

Therefore, by using the inequality

$$\|a + b + c\|^2 \leq 3(\|a\|^2 + \|b\|^2 + \|c\|^2),$$

we get

$$\|X_{t \wedge T_n}^{x_0}\|^2 \leq 3 \left( \|x_0\|^2 + \left\| \int_0^{t \wedge T_n} b(X_s^{x_0}) ds \right\|^2 + \left\| \int_0^{t \wedge T_n} \sigma(X_s^{x_0}) dB_s \right\|^2 \right).$$

Thus, we have

$$\begin{aligned} & \mathbb{E} \left( \sup_{0 \leq u \leq t \wedge T_n} \|X_u^{x_0}\|^2 \right) \\ & \leq 3 \left( \|x_0\|^2 + \mathbb{E} \left( \sup_{0 \leq u \leq t \wedge T_n} \left\| \int_0^{u \wedge T_n} b(X_s^{x_0}) ds \right\|^2 \right) \right. \\ & \quad \left. + \mathbb{E} \left( \sup_{0 \leq u \leq t \wedge T_n} \left\| \int_0^{u \wedge T_n} \sigma(X_s^{x_0}) dB_s \right\|^2 \right) \right) \end{aligned}$$

By using our assumptions, we first estimate

$$\mathbb{E} \left( \sup_{0 \leq u \leq t \wedge T_n} \left\| \int_0^{u \wedge T_n} b(X_s^{x_0}) ds \right\|^2 \right) \leq K^2 \mathbb{E} \left( \left( \int_0^{t \wedge T_n} (1 + \|X_s^{x_0}\|) ds \right)^2 \right).$$

By using our assumptions and Doob's inequality, we now estimate

$$\mathbb{E} \left( \sup_{0 \leq u \leq t \wedge T_n} \left\| \int_0^{u \wedge T_n} \sigma(X_s^{x_0}) dB_s \right\|^2 \right) \leq 4K^2 \mathbb{E} \left( \int_0^{t \wedge T_n} (1 + \|X_s\|)^2 ds \right).$$

Therefore, from the inequality  $\|a + b\|^2 \leq 2(\|a\|^2 + \|b\|^2)$ , we get

$$\begin{aligned} & \mathbb{E} \left( \sup_{0 \leq u \leq t \wedge T_n} \|X_u^{x_0}\|^2 \right) \\ & \leq 3 \left( \|x_0\|^2 + 2(K^2T + 4K^2) \int_0^t \left( 1 + \mathbb{E} \left( \sup_{0 \leq u \leq s \wedge T_n} \|X_u^{x_0}\|^2 \right) ds \right) \right) \end{aligned}$$

We may now apply Gronwall's lemma to the function

$$t \rightarrow \mathbb{E} \left( \sup_{0 \leq u \leq t \wedge T_n} \|X_u^{x_0}\|^2 \right)$$

and deduce

$$\mathbb{E} \left( \sup_{0 \leq u \leq T \wedge T_n} \|X_u^{x_0}\|^2 \right) \leq C$$

where  $C$  is a constant that does not depend on  $n$ . Fatou's lemma implies by passing to the limit when  $n \rightarrow +\infty$  that

$$\mathbb{E} \left( \sup_{0 \leq u \leq T} \|X_u^{x_0}\|^2 \right) \leq C.$$

We conclude, as expected, that

$$(X_t^{x_0})_{0 \leq t \leq T} \in \mathcal{E}_T.$$

More generally, by using the same arguments we can observe that if a continuous and adapted process satisfies

$$X_t = X_0 + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dB_s,$$

with  $\mathbb{E}(X_0^2) < +\infty$ , then  $(X_t)_{0 \leq t \leq T} \in \mathcal{E}_T$ .

**Step 2:** We now show existence and uniqueness of solutions for the equation (6.1) on a time interval  $[0, T]$  where  $T$  is small enough.

Let us consider the application  $\Phi$  that sends a continuous and adapted process  $(X_t)_{0 \leq t \leq T}$  to the process

$$\Phi(X)_t = x_0 + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dB_s.$$

By using successively the inequalities  $(a + b)^2 \leq 2(a^2 + b^2)$ , Cauchy-Schwarz inequality and Doob's inequality, we get

$$\|\Phi(X) - \Phi(Y)\|^2 \leq 2(K^2T^2 + 4K^2T) \|X - Y\|^2.$$

Moreover, arguing the same way as above, we can prove

$$\|\Phi(0)\|^2 \leq 3(x_0^2 + K^2T^2 + 4K^2T).$$

Therefore, if  $T$  is small enough  $\Phi$  is a Lipschitz map  $\mathcal{E}_T \rightarrow \mathcal{E}_T$  whose Lipschitz constant is strictly less than 1. Consequently, it has a unique fixed point. This fixed point is, of course the unique solution of (6.1) on the time interval  $[0, T]$ . Here again, we can observe that the same reasoning applies if  $x_0$  is replaced by a random variable  $X_0$  that satisfies  $\mathbb{E}(X_0^2) < +\infty$ .

**Step 3:** In order to get a solution of (6.1) on  $[0, +\infty)$ , we may apply the previous step to get a solution on intervals  $[T_n, T_{n+1}]$ , where  $T_{n+1} - T_n$  is small enough and  $T_n \rightarrow +\infty$ . This will provide a solution of (6.1) on  $[0, +\infty)$ . This solution is unique, from the uniqueness on each interval  $[T_n, T_{n+1}]$ .  $\square$

**Definition 6.2.** An equation like (6.1) is called a stochastic differential equation.

**Remark 6.3.** As it appears from the very definition, solutions of stochastic differential equations are constructed by using stochastic integrals and thus probabilistic techniques. A theory developed in the recent years, the rough paths theory, allows to give a pathwise interpretation of solutions of differential equations driven by rough signals, including the Brownian motion. We refer the reader to Chapter 7 for an overview of this theory.

**Exercise 6.4.** Let  $(X_t^{x_0})_{t \geq 0}$  denote the solution of (6.1). Show that under the same assumptions as the previous theorem, we actually have for every  $T \geq 0$ , and  $p \geq 1$ ,

$$\mathbb{E} \left( \sup_{0 \leq s \leq T} \|X_s\|^p \right) < +\infty.$$

**Exercise 6.5** (Ornstein-Uhlenbeck process). Let  $\theta \in \mathbb{R}$ . We consider the following stochastic differential equation,

$$dX_t = \theta X_t dt + dB_t, \quad X_0 = x.$$

1. Show that it admits a unique solution that is given by

$$X_t = e^{\theta t} x + \int_0^t e^{\theta(t-s)} dB_s.$$

2. Show that  $(X_t)_{t \geq 0}$  is Gaussian process. Compute its mean and covariance function.
3. Show that if  $\theta < 0$  then, when  $t \rightarrow +\infty$ ,  $X_t$  converges in distribution toward a Gaussian distribution.

**Exercise 6.6** (Brownian bridge). We consider for  $0 \leq t < 1$  the following stochastic differential equation

$$dX_t = -\frac{X_t}{1-t}dt + dB_t, \quad X_0 = 0.$$

1. Show that

$$X_t = (1-t) \int_0^t \frac{dB_s}{1-s}.$$

2. Deduce that  $(X_t)_{t \geq 0}$  is Gaussian process. Compute its mean and covariance function.
3. Show that in  $L^2$ , when  $t \rightarrow 1$ ,  $X_t \rightarrow 0$ .

**Exercise 6.7** (Black-Scholes process). Let  $\mu \in \mathbb{R}$  and  $\sigma > 0$ . We consider the following stochastic differential equation,

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad X_0 = x > 0.$$

Show that for every  $t \geq 0$

$$X_t = x e^{\sigma B_t + (\mu - \frac{\sigma^2}{2})t}.$$

**Exercise 6.8** (Explosion time). Let  $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , and  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  be locally Lipschitz functions, that is for every  $N \geq 0$ , there exists  $C_N > 0$  such that

$$\|b(x) - b(y)\| + \|\sigma(x) - \sigma(y)\| \leq C_N \|x - y\|, \quad x, y \in [-N, N]^n.$$

Show that for every  $x \in \mathbb{R}^n$  we can find stopping time  $\mathbf{e}(x)$ , almost surely positive, and a stochastic process  $(X_t^x)_{t < \mathbf{e}(x)}$  such that:

•

$$X_t^x = x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dB_s, \quad t < \mathbf{e}(x). \quad (6.2)$$

- The process  $(X_t^x)_{t < \mathbf{e}(x)}$  is unique in the sense that if  $\mathbf{f}(x)$  is an almost surely positive stopping time and if  $(Y_t^x)_{t < \mathbf{f}(x)}$  is a stochastic process such that

$$Y_t^x = x + \int_0^t b(Y_s^x) ds + \int_0^t \sigma(Y_s^x) dB_s, \quad t < \mathbf{f}(x),$$

then  $\mathbf{f}(x) \leq \mathbf{e}(x)$  and for every  $t \geq 0$ ,  $Y_t^x \mathbf{1}_{t < \mathbf{f}(x)} = X_t^x \mathbf{1}_{t < \mathbf{e}(x)}$  almost surely.

The process  $(X_t^x)_{t < \mathbf{e}(x)}$  is called the solution of the stochastic differential equation (6.2) up to the explosion time  $\mathbf{e}(x)$ .

The next proposition shows that solutions of stochastic differential equations are intrinsically related to a second order differential operator, which will turn out to be the generator of the Markov process  $(X_t^x)_{t \geq 0}$ .

**Proposition 6.9.** *Let  $(X_t^x)_{t \geq 0}$  be the solution of a stochastic differential equation*

$$X_t^{x_0} = x_0 + \int_0^t b(X_s^{x_0}) ds + \int_0^t \sigma(X_s^{x_0}) dB_s,$$

where  $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  are Borel functions. Let now  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a  $C^2$  function. The process

$$M_t^f = f(X_t^x) - \int_0^t Lf(X_s^x) ds,$$

is a local martingale, where  $L$  is the second order differential operator

$$L = \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j},$$

and  $a_{ij}(x) = (\sigma(x)\sigma^*(x))_{ij}$ .

*Proof.* The proof is as an easy application of the Döblin-Itô formula applied to  $f(X_t^x)$ .  $\square$

## 2 Continuity and differentiability of stochastic flows

In this section, we study the regularity of the solution of a stochastic differential equation with respect to its initial condition. The key tool is a multidimensional parameter extension of the Kolmogorov continuity theorem whose proof is almost identical to the one-dimensional case and let to the reader as an exercise.

**Theorem 6.10.** *Let  $(\Theta_x)_{x \in [0,1]^d}$  be a  $n$ -dimensional stochastic process such that there exist positive constants  $\gamma, c, \varepsilon$  such that for every  $x, y \in [0, 1]^d$*

$$\mathbb{E}(\|\Theta_x - \Theta_y\|^\gamma) \leq C\|x - y\|^{d+\varepsilon}.$$

*There exists a modification  $(\tilde{\Theta}_x)_{x \in [0,1]^d}$  of the process  $(\Theta_x)_{x \in [0,1]^d}$  such that for every  $\alpha \in [0, \varepsilon/\gamma)$  there exists a finite random variable  $K_\alpha$  such that for every  $x, y \in [0, 1]^d$*

$$\|\tilde{\Theta}_x - \tilde{\Theta}_y\| \leq K_\alpha \|x - y\|^\alpha.$$

As above, we consider two functions  $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\sigma : \mathbb{R}^{n \times n}$  and we assume that there exists  $C > 0$  such that

$$\|b(x) - b(y)\| + \|\sigma(x) - \sigma(y)\| \leq C\|x - y\|, x, y \in \mathbb{R}^n.$$

As we already know, for every  $x \in \mathbb{R}^n$ , there exists a continuous and adapted process  $(X_t^x)_{t \geq 0}$  such that for  $t \geq 0$ ,

$$X_t^x = x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dB_s. \quad (6.3)$$

**Proposition 6.11.** *Let  $T > 0$ . For every  $p \geq 2$ , there exists a constant  $C_{p,T} > 0$  such that for every  $0 \leq s \leq t \leq T$  and  $x, y \in \mathbb{R}^n$ ,*

$$\mathbb{E}(\|X_t^x - X_s^y\|^p) \leq C_{p,T} \left( \|x - y\|^p + |t - s|^{p/2} \right)$$

As a consequence, there exists a modification  $(\tilde{X}_t^x)_{t \geq 0, x \in \mathbb{R}^n}$  of the process  $(X_t^x)_{t \geq 0, x \in \mathbb{R}^n}$  such that for  $t \geq 0, x \in \mathbb{R}^n$ ,

$$\tilde{X}_t^x = x + \int_0^t b(\tilde{X}_s^x) ds + \int_0^t \sigma(\tilde{X}_s^x) dB_s.$$

and such that  $(t, x) \rightarrow X_t^x(\omega)$  is continuous for almost every  $\omega$ .

*Proof.* As before, we can find  $K > 0$  such that

1.  $\|b(x) - b(y)\| + \|\sigma(x) - \sigma(y)\| \leq K\|x - y\|, x, y \in \mathbb{R}^n;$
2.  $\|b(x)\| + \|\sigma(x)\| \leq K(1 + \|x\|), x \in \mathbb{R}^n.$

We fix  $x, y \in \mathbb{R}^n$  and  $p \geq 2$ . Let

$$h(t) = \mathbb{E}(\|X_t^x - X_t^y\|^p).$$

By using the inequality  $\|a + b + c\|^p \leq 3^{p-1}(\|a\|^p + \|b\|^p + \|c\|^p)$ , we obtain

$$\begin{aligned} & \|X_t^x - X_t^y\|^p \\ & \leq 3^{p-1} \left( \|x - y\|^p + \left( \int_0^t \|b(X_s^x) - b(X_s^y)\| ds \right)^p + \left\| \int_0^t (\sigma(X_s^x) - \sigma(X_s^y)) dB_s \right\|^p \right). \end{aligned}$$

We now have

$$\left( \int_0^t \|b(X_s^x) - b(X_s^y)\| ds \right)^p \leq t^{p-1} \int_0^t \|b(X_s^x) - b(X_s^y)\|^p ds \leq K^p t^{p-1} \int_0^t \|X_s^x - X_s^y\|^p ds,$$

and from Burkholder-Davis-Gundy inequality

$$\begin{aligned} \mathbb{E} \left( \left\| \int_0^t (\sigma(X_s^x) - \sigma(X_s^y)) dB_s \right\|^p \right) &\leq C_p \mathbb{E} \left( \left( \int_0^t \|\sigma(X_s^x) - \sigma(X_s^y)\|^2 ds \right)^{p/2} \right) \\ &\leq C_p K^2 \mathbb{E} \left( \left( \int_0^t \|X_s^x - X_s^y\|^2 ds \right)^{p/2} \right) \\ &\leq C_p K^2 t^{p/2-1} \mathbb{E} \left( \int_0^t \|X_s^x - X_s^y\|^p ds \right). \end{aligned}$$

As a conclusion we obtain

$$h(t) \leq 3^{p-1} \left( \|x - y\|^p + (K^p t^{p-1} + C_p K^2 t^{p/2-1}) \int_0^t h(s) ds \right).$$

Gronwall's inequality yields then

$$h(t) \leq \phi(t) \|x - y\|^p,$$

where  $\phi$  is a continuous function.

On the other side, we have for  $0 \leq s \leq t \leq T$ ,

$$\|X_t^x - X_s^x\|^p \leq 2^{p-1} \left( \left\| \int_s^t b(X_u^x) ds \right\|^p + \left\| \int_s^t \sigma(X_u^x) dB_u \right\|^p \right),$$

and

$$\left\| \int_s^t b(X_u^x) ds \right\|^p \leq K^p (t-s)^p \left( 1 + \sup_{0 \leq s \leq T} \|X_s\| \right)^p,$$

$$\begin{aligned} \mathbb{E} \left( \left\| \int_s^t \sigma(X_u^x) dB_u \right\|^p \right) &\leq C_p \mathbb{E} \left( \left( \int_s^t \|\sigma(X_u^x)\|^2 du \right)^{p/2} \right) \\ &\leq C_p K^p (t-s)^{p/2} \mathbb{E} \left( \left( 1 + \sup_{0 \leq s \leq T} \|X_s\| \right)^p \right). \end{aligned}$$

The conclusion then easily follows by combining the two previous estimates.  $\square$

In the sequel, of course, we shall always work with this bicontinuous version of the solution.

**Definition 6.12.** The continuous process of continuous maps  $\Psi_t : x \rightarrow X_t^x$  is called the stochastic flow associated to the equation (6.3).

If the maps  $b$  and  $\sigma$  are moreover  $C^1$ , then the stochastic flow is itself differentiable and the equation for the derivative can be obtained by formally differentiating the equation with respect to its initial condition. We will admit this result without proof (see [?] for a proof):

**Theorem 6.13.** *Let us assume that  $b$  and  $\sigma$  are  $C^1$  Lipschitz functions, then for every  $t \geq 0$ , the flow  $\Psi_t$  associated to the equation (6.3) is a flow of differentiable maps. Moreover, the first variation process  $\mathbf{J}_t(x)$  which is defined as the Jacobian matrix  $\frac{\partial \Psi_t}{\partial x}(x)$  is the unique solution of the matrix stochastic differential equation:*

$$\mathbf{J}_t(x) = \mathbf{Id} + \int_0^t \frac{\partial b}{\partial x}(X_s^x) \mathbf{J}_s(x) ds + \sum_{i=1}^n \int_0^t \frac{\partial \sigma_i}{\partial x}(X_s^x) \mathbf{J}_s(x) dB_s^i.$$

**Exercise 6.14.** *For  $x \in \mathbb{R}^n$ , let  $(\mathbf{K}_t(x))_{t \geq 0}$  be the unique solution of the matrix linear stochastic differential equation:*

$$\begin{aligned} \mathbf{K}_t(x) = & \mathbf{Id} + \sum_{i=1}^n \int_0^t \mathbf{K}_s(x) \left( \frac{\partial \sigma_i}{\partial x} \right)^2 (X_s^x) ds - \int_0^t \mathbf{K}_s(x) \frac{\partial b}{\partial x}(X_s^x) ds \\ & - \sum_{i=1}^n \int_0^t \mathbf{K}_s(x) \frac{\partial \sigma_i}{\partial x}(X_s^x) dB_s^i. \end{aligned}$$

*Show that the process  $(\mathbf{K}_t(x) \mathbf{J}_t(x))_{t \geq 0}$  is constant and equal to  $\mathbf{Id}$ . As a consequence, the Jacobian matrix  $\mathbf{J}_t(x)$  is always almost surely invertible.*

### 3 The Feynman-Kac formula

It is now time to give some applications of the theory of stochastic differential equations to parabolic second order partial differential equations. In particular we are going to prove that solutions of such equations can be represented by using solutions of stochastic differential equations. This representation formula is called the Feynman-Kac formula. As usual, we consider a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  which satisfies the usual conditions and on which is defined a  $n$ -dimensional Brownian motion  $(B_t)_{t \geq 0}$ . Again, we consider two functions  $b : \mathbb{R}^n \rightarrow \mathbb{R}$  and  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  and we assume that there exists  $C > 0$  such that

$$\|b(x) - b(y)\| + \|\sigma(x) - \sigma(y)\| \leq C \|x - y\|, x, y \in \mathbb{R}^n.$$

Let  $L$  be the diffusion operator

$$L = \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j},$$

where  $a_{ij}(x) = (\sigma(x) \sigma^*(x))_{ij}$ .

As we know, there exists a bicontinuous process  $(X_t^x)_{t \geq 0, x \in \mathbb{R}^d}$  such that for  $t \geq 0$ ,

$$X_t^x = x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dB_s. \quad (6.4)$$



Moreover, as it has been stressed before, for every  $p \geq 1$ , and  $T \geq 0$

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} \|X_t^x\|^p \right) < +\infty.$$

As a consequence, if  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a Borel function with polynomial growth, we can consider the function

$$\mathbf{P}_t f(x) = \mathbb{E}(f(X_t^x)).$$

**Theorem 6.15.** *For every  $x \in \mathbb{R}^n$ ,  $(X_t^x)_{t \geq 0, x \in \mathbb{R}^d}$  is a Markov process with semigroup  $(\mathbf{P}_t)_{t \geq 0}$ . More precisely, for every Borel function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  with polynomial growth and every  $t \geq s$ ,*

$$\mathbb{E}(f(X_t^x) \mid \mathcal{F}_s) = (\mathbf{P}_{t-s} f)(X_s^x).$$

*Proof.* The key point, here, is to observe that solutions are actually adapted to the natural filtration of the Brownian motion  $(B_t)_{t \geq 0}$ . More precisely, there exists on the space of continuous functions  $[0, +\infty) \rightarrow \mathbb{R}^n$  a predictable functional such that for  $t \geq 0$ :

$$X_t^{x_0} = F(x_0, (B_u)_{0 \leq u \leq t}).$$

Indeed, let us first work on  $[0, T]$  where  $T$  is small enough. In that case, as seen previously, the process  $(X_t^{x_0})_{0 \leq t \leq T}$  is the unique fixed point of the application  $\Phi$  defined by

$$\Phi(X)_t = x_0 + \int_0^t b(X_s^{x_0}) ds + \int_0^t \sigma(X_s^{x_0}) dB_s$$

Alternatively, one can interpret this by observing that  $(X_t^{x_0})_{0 \leq t \leq T}$  is the limit of the sequence of processes  $(X_t^n)_{0 \leq t \leq T}$  inductively defined by

$$X^{n+1} = \Phi(X^n), \quad X^0 = x_0.$$

It is easily checked that for each  $X^n$  there is a predictable functional  $F_n$  such that

$$X_t^n = F_n(x_0, (B_u)_{0 \leq u \leq t}),$$

which proves the above claim when  $T$  is small enough. To get the existence of  $F$  for any  $T$ , we can proceed analogously on the intervals  $[T, 2T]$ ,  $[2T, 3T]$  and so on.

With this hands, we can now prove the Markov property. Let  $s \geq 0$ . For  $t \geq 0$ , we have

$$\begin{aligned} X_{s+t}^{x_0} &= X_s + \int_s^{s+t} b(X_u^{x_0}) du + \int_s^{s+t} \sigma(X_u^{x_0}) dB_u \\ &= X_s + \int_0^t b(X_{u+s}^{x_0}) du + \int_0^t \sigma(X_s^{x_0}) d(B_{u+s} - B_s). \end{aligned}$$

Consequently, from uniqueness of solutions,

$$X_{s+t}^{x_0} = F(X_s^{x_0}, (B_{u+s} - B_s)_{0 \leq u \leq t}).$$

We deduce that for a Borel function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  with polynomial growth,

$$\mathbb{E}(f(X_{s+t}^{x_0}) \mid \mathcal{F}_s) = \mathbb{E}(f(F(X_s^{x_0}, (B_{u+s} - B_s)_{0 \leq u \leq t})) \mid \mathcal{F}_s) = \mathbf{P}_t f(X_s^{x_0}),$$

because  $(B_{u+s} - B_s)_{0 \leq u \leq t}$  is a Brownian motion independent of  $\mathcal{F}_s$ .  $\square$

If the coefficients  $b$  and  $\sigma$  are bounded, then solutions of stochastic differential equations are more than Markov processes, they are Feller-Dynkin diffusion processes in the sense of Definition 3.33, Chapter 3.

**Proposition 6.16.** *If the functions  $b$  and  $\sigma$  are moreover bounded, then, for every  $x_0 \in \mathbb{R}^n$ ,  $(X_t^{x_0})_{t \geq 0}$  is a Feller-Dynkin diffusion process with infinitesimal generator*

$$Lf = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial f}{\partial x_i}, \quad f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R}).$$

*Proof.* We first need to prove that  $(\mathbf{P}_t)_{t \geq 0}$  is a Feller-Dynkin semigroup. Let  $f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$ . First, it is clear that the function  $\mathbf{P}_t$  is continuous. We now have for  $x \in \mathbb{R}^n$  and  $r > 0$ ,

$$|\mathbf{P}_t f(x)| \leq \sup_{y \in \mathbf{B}(x,r)} |f(y)| + \|f\|_\infty \mathbb{P}(\|X_t^x - x\| > r)$$

and,

$$\begin{aligned} \mathbb{P}(\|X_t^x - x\| > r) &\leq \frac{1}{r^2} \mathbb{E}(\|X_t^x - x\|^2) \\ &\leq \frac{2}{r^2} \left( \mathbb{E} \left( \left\| \int_0^t b(X_s^x) ds \right\|^2 \right) + \mathbb{E} \left( \int_0^t \|\sigma(X_s^x)\|^2 ds \right) \right) \\ &\leq \frac{2M}{r^2} (t + t^2), \end{aligned}$$

where  $M$  is a common upper bound for  $b$  and  $\sigma$ . By letting  $x$  and then  $r$  to  $\infty$ , we conclude that  $\mathbf{P}_t f \in \mathcal{C}_0(\mathbb{R}^n, \mathbb{R})$ . Thus  $\mathbf{P}_t$  is a Feller-Dynkin semigroup. Let us compute its generator on  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . Let  $f \in \mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$ . From Itô's formula we have

$$f(X_t^x) = f(x) + \int_0^t Lf(X_s^x) ds + M_t,$$

where  $M_t$  is a martingale. As a consequence, by taking expectations, we obtain

$$\mathbf{P}_t f(x) = f(x) + \int_0^t \mathbf{P}_s Lf(x) ds.$$

This implies that

$$\lim_{t \rightarrow 0} \left\| \frac{\mathbf{P}_t f - f}{t} - Lf \right\|_{\infty} = 0.$$

Thus the domain of the generator of  $\mathbf{P}_t$  contains  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{R})$  and on that space is equal to  $L$ .  $\square$

**Theorem 6.17.** *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a Borel function with polynomial growth and assume that the function*

$$u(t, x) = (\mathbf{P}_t f)(x)$$

*is  $C^{1,2}$ , that is once differentiable with respect to  $t$  and twice differentiable with respect to  $x$ . Then  $u$  solves the Cauchy problem*

$$\frac{\partial u}{\partial t}(t, x) = Lu(t, x)$$

*in  $[0, +\infty) \times \mathbb{R}^n$ , with the initial condition*

$$u(0, x) = f(x).$$

*Proof.* Let  $T > 0$  and consider the function  $v(t, x) = u(T - t, x)$ . According to the previous theorem, we have

$$\mathbb{E}(f(X_T^x) | \mathcal{F}_t) = v(t, X_t^x).$$

As a consequence, the process  $v(t, X_t^x)$  is a martingale. But from Döblin-Itô's formula the bounded variation part of  $v(t, X_t^x)$  is  $\int_0^t \left( \frac{\partial v}{\partial t}(s, X_s^x) + Lv(s, X_s^x) \right) ds$  which is therefore 0. We conclude

$$\frac{\partial v}{\partial t}(0, x) + Lv(0, x) = \lim_{t \rightarrow 0} \frac{1}{t} \int_0^t \left( \frac{\partial v}{\partial t}(s, X_s^x) + Lv(s, X_s^x) \right) ds = 0.$$

$\square$

**Exercise 6.18.** *Show that if  $f$  is a  $C^2$  function such that  $\nabla f$  and  $\nabla^2 f$  have polynomial growth, then the function  $\mathbf{P}_t f(x)$  is  $C^{1,2}$ . Here, we denote by  $\nabla^2 f$  the Hessian matrix of  $f$ .*

**Theorem 6.19.** *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a Borel function with polynomial growth. Let  $u : [0, +\infty) \times \mathbb{R}^n \rightarrow \mathbb{R}$  be a solution of the Cauchy problem*

$$\frac{\partial u}{\partial t}(t, x) = Lu(t, x)$$

*with the initial condition*

$$u(0, x) = f(x).$$

*If there exists a locally integrable function  $C$  and  $p \geq 0$ , such that for every  $t \geq 0$  and  $x \in \mathbb{R}^n$ ,*

$$\|\nabla u(t, x)\| \leq C(t)(1 + \|x\|^p),$$

*then  $u(t, x) = \mathbf{P}_t f(x)$ .*

*Proof.* Let  $T > 0$  and, as before consider the function  $v(t, x) = u(T - t, x)$ . As a consequence of Itô's formula, we have

$$v(t, X_t^x) = u(T, x) + M_t,$$

where  $M_t$  is a local martingale with quadratic variation

$$\sum_{i,j} \int_0^t a_{ij}(X_s^x) \frac{\partial u}{\partial x_i}(X_s^x) \frac{\partial u}{\partial x_j}(X_s^x) ds.$$

The conditions on  $\sigma$  and  $u$  imply that this quadratic variation is integrable. As a consequence,  $v(t, X_t^x)$  is a martingale and thus  $\mathbb{E}(v(T, X_T^x)) = u(T, x)$ .  $\square$

The previous results may be extended to study parabolic equations with potential as well. More precisely, let  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  be a bounded function. If  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a Borel function with polynomial growth, we define

$$\mathbf{P}_t^V f(x) = \mathbb{E} \left( e^{\int_0^t V(X_s^x) ds} f(X_t^x) \right).$$

The same proofs as before will give the following theorems.

**Theorem 6.20.** *For every  $x \in \mathbb{R}^n$  and every Borel function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  with polynomial growth and every  $t \geq s$ ,*

$$\mathbb{E} \left( e^{\int_0^t V(X_u^x) du} f(X_t^x) \mid \mathcal{F}_s \right) = e^{\int_0^s V(X_u^x) du} (\mathbf{P}_{t-s}^V f)(X_s^x).$$

**Theorem 6.21** (Feynman-Kac formula). *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a Borel function with polynomial growth and assume that the function*

$$u(t, x) = (\mathbf{P}_t^V f)(x)$$

*is  $C^{1,2}$ , that is once differentiable with respect to  $t$  and twice differentiable with respect to  $x$ . Then  $u$  solves the Cauchy problem*

$$\frac{\partial u}{\partial t}(t, x) = Lu(t, x) + V(x)u(t, x)$$

*in  $[0, +\infty) \times \mathbb{R}^n$ , with the initial condition*

$$u(0, x) = f(x).$$

**Theorem 6.22.** *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a Borel function with polynomial growth. Let  $u : [0, +\infty) \times \mathbb{R}^n \rightarrow \mathbb{R}$  be a solution of the Cauchy problem*

$$\frac{\partial u}{\partial t}(t, x) = Lu(t, x) + V(x)u(t, x)$$

with the initial condition

$$u(0, x) = f(x).$$

If there exists a locally integrable function  $C$  and  $p \geq 0$ , such that for every  $t \geq 0$  and  $x \in \mathbb{R}^n$ ,

$$\|\nabla u(t, x)\| \leq C(t)(1 + \|x\|^p),$$

then  $u(t, x) = \mathbf{P}_t^V f(x)$ .

**Exercise 6.23** (Arcsine law). Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion and  $\lambda, \alpha \in \mathbb{R}$ . We denote

$$\nu(x) = \mathbb{E} \left( \int_0^{+\infty} \exp \left( -\lambda t - \alpha \int_0^t \mathbf{1}_{[0, +\infty)}(B_s + x) ds \right) dt \right).$$

1. Show that  $\nu$  is the unique solution of the differential equation

$$y'' - (\alpha \mathbf{1}_{[0, +\infty)} + \lambda)y = 1,$$

that satisfies,

$$\lim_{x \rightarrow +\infty} \nu(x) = \frac{1}{\alpha + \lambda}, \quad \lim_{x \rightarrow -\infty} \nu(x) = \frac{1}{\lambda}.$$

2. Deduce that

$$\int_0^{+\infty} e^{-\lambda t} \mathbb{E}(e^{-\alpha A_t}) dt = \frac{1}{\lambda} \frac{\lambda + \sqrt{\lambda(\alpha + \lambda)}}{\alpha + \lambda + \sqrt{\lambda(\alpha + \lambda)}},$$

where  $A_t = \frac{1}{t} \int_0^t \mathbf{1}_{[0, +\infty)}(B_s) ds$ .

3. Conclude that the density of  $A_t$  is given by

$$s \rightarrow \frac{1}{\pi \sqrt{s(1-s)}}, \quad s \in (0, 1).$$

This is the arc-sine law for Brownian motion paths that was already proved in Theorem 2.51, Chapter 2 by using random walks.

**Exercise 6.24** (Exponential functional). Let  $(B_t)_{t \geq 0}$  be a standard Brownian motion and  $\mu > 0$ . We denote

$$A_\infty = \int_0^{+\infty} e^{-2(B_t + \mu t)} dt.$$

1. Let  $h(x) = e^{\mu x} \mathbb{E} \left( \exp \left( -\frac{1}{2} e^{-2x} A_\infty \right) \right)$ . Show that  $h$  is the unique solution of the differential equation

$$\left( \frac{d^2}{dx^2} - e^{-2x} \right) h = \mu^2 h,$$

such that  $h(x) \sim_{x \rightarrow +\infty} e^{\mu x}$ .

2. Deduce that

$$h(x) = \frac{2^{1-\mu}}{\Gamma(\mu)} K_\mu(e^{-x}),$$

where  $K_\mu$  is the Mac-Donald function

$$K_\mu(x) = \frac{1}{2} \left(\frac{x}{2}\right)^\mu \int_0^{+\infty} \frac{e^{-\frac{x^2}{4t}-t}}{t^{1+\mu}} dt.$$

3. Compute the density of  $A_\infty$ .

4. Consider the process  $(\rho_t)_{t \geq 0}$  such that for every  $t \geq 0$ ,

$$e^{-B_t - \mu t} = \rho \int_0^t e^{-2(B_s + \mu s)} ds.$$

Show that the process

$$\rho_t - \int_0^t \frac{-\mu + \frac{1}{2}}{\rho_s} ds,$$

is a Brownian motion. The process  $(\rho_t)_{t \geq 0}$  is called a Bessel process (see Exercises 3.8 and 5.59).

5. By using  $A_\infty$ , compute the density of the stopping time

$$T_0 = \inf\{t \geq 0, \rho_t = 0\}.$$

## 4 The strong Markov property of solutions

In the previous section, we have seen that if  $(X_t^x)_{t \geq 0}$  is the solution of a stochastic differential equation

$$X_t^x = x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dB_s, \tag{6.5}$$

then  $(X_t^x)_{t \geq 0}$  is a Markov process, that is for every  $t, T \geq 0$ ,

$$\mathbb{E}(f(X_{t+T}^x) \mid \mathcal{F}_T) = (\mathbf{P}_t f)(X_T^x),$$

where  $\mathbf{P}_t f(x) = \mathbb{E}(f(X_t^x))$ . It is remarkable that this property still holds when  $T$  is now any finite stopping time; That is solutions of stochastic differential equations enjoy the strong Markov property. The key lemma is Proposition 3.20 in Chapter 3.

**Theorem 6.25.** *For every  $x \in \mathbb{R}^n$ ,  $(X_t^x)_{t \geq 0, x \in \mathbb{R}^d}$  is a strong Markov process with semigroup  $(P_t)_{t \geq 0}$ : For every Borel function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  with polynomial growth, every  $t \geq 0$ , and every finite stopping time  $T$ ,*

$$\mathbb{E}(f(X_{t+T}^x) \mid \mathcal{F}_T) = (\mathbf{P}_t f)(X_T^x).$$

*Proof.* The proof is identical to the proof of Theorem 6.15 with the additional ingredient given by the Proposition 3.20 in Chapter 2.  $\square$

**Exercise 6.26.** Let  $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , and  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  be locally Lipschitz functions and consider the solution of the stochastic differential equation,

$$X_t^x = x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dB_s,$$

up to the explosion time  $\mathbf{e}(x)$  (see Exercise 6.8). Show that  $(X_t)_{t < \mathbf{e}(x)}$  is a submarkov process with semigroup

$$\mathbf{P}_t f(x) = \mathbb{E}(f(X_t^x) \mathbf{1}_{t < \mathbf{e}(x)})$$

The strong Markov property for solutions of stochastic differential equations is useful to solve boundary value problems in partial differential equations theory. Let  $K$  be a bounded closed set in  $\mathbb{R}^n$ . For  $x \in \Omega$ , we denote  $T_x = \inf\{t \geq 0, X_t \in \partial K\}$ . If  $f$  is bounded Borel function such that  $f_{\partial K} = 0$ , we define

$$\mathbf{P}_t^K f(x) = \mathbb{E}(f(X_t^x) \mathbf{1}_{t \leq T_x})$$

The proof of the following theorem is left to the reader.

**Theorem 6.27.** Let  $f : K \rightarrow \mathbb{R}$  be a bounded Borel function and assume that the function

$$u(t, x) = (\mathbf{P}_t^K f)(x)$$

is  $C^{1,2}$ . Then  $u$  is the unique solution of the Dirichlet boundary value problem

$$\frac{\partial u}{\partial t}(t, x) = Lu(t, x)$$

in  $[0, +\infty) \times \mathbb{R}^n$ , with the initial condition

$$u(0, x) = f(x),$$

and the boundary condition

$$u(t, x) = 0, \quad x \in \partial K.$$

## 5 Stratonovitch stochastic differential equations and the language of vector fields

As usual, let  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a filtered probability space which satisfies the usual conditions. It is often useful to use the language of Stratonovitch's integration to study stochastic differential equations because the Itô's formula takes a much nicer form. If  $(N_t)_{0 \leq t \leq T}$ ,  $T > 0$ , is an  $\mathcal{F}$ -adapted real valued

local martingale and if  $(\Theta_t)_{0 \leq t \leq T}$  is an  $\mathcal{F}$ -adapted continuous semimartingale then the Stratonovitch integral of  $(\Theta_t)_{0 \leq t \leq T}$  with respect to  $(N_t)_{t \geq 0}$  is given as

$$\int_0^T \Theta_t \circ dN_t = \int_0^T \Theta_t dN_t + \frac{1}{2} \langle \Theta, N \rangle_T,$$

where:

1.  $\int_0^T \Theta_t dN_t$  is the Itô integral of  $(\Theta_t)_{0 \leq t \leq T}$  against  $(N_t)_{0 \leq t \leq T}$ ;
2.  $\langle \Theta, N \rangle_T$  is the quadratic covariation at time  $T$  between  $(\Theta_t)_{0 \leq t \leq T}$  and  $(N_t)_{0 \leq t \leq T}$ .

By using Stratonovitch integral instead of Itô's, we can see that the Döblin-Itô formula reduces to the classical change of variable formula.

**Theorem 6.28.** *Let  $(X_t)_{t \geq 0} = (X_t^1, \dots, X_t^n)_{t \geq 0}$  be a  $n$ -dimensional continuous semimartingale. Let now  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a  $C^2$  function. We have*

$$f(X_t) = f(X_0) + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(X_s) \circ dX_s^i, \quad t \geq 0.$$

Let  $\mathcal{O} \subset \mathbb{R}^n$  be a non empty open set. A smooth vector field  $V$  on  $\mathcal{O}$  is simply a smooth map

$$\begin{aligned} V : \mathcal{O} &\rightarrow \mathbb{R}^n \\ x &\rightarrow (v_1(x), \dots, v_n(x)). \end{aligned}$$

The vector field  $V$  defines a differential operator acting on smooth functions  $f : \mathcal{O} \rightarrow \mathbb{R}$  as follows:

$$Vf(x) = \sum_{i=1}^n v_i(x) \frac{\partial f}{\partial x_i}.$$

We note that  $V$  is a derivation, that is a map on  $\mathcal{C}^\infty(\mathcal{O}, \mathbb{R})$ , linear over  $\mathbb{R}$ , satisfying for  $f, g \in \mathcal{C}^\infty(\mathcal{O}, \mathbb{R})$ ,

$$V(fg) = (Vf)g + f(Vg).$$

Interestingly, conversely, any derivation on  $\mathcal{C}^\infty(\mathcal{O}, \mathbb{R})$  is a vector field.

Let now

$$(B_t)_{t \geq 0} = (B_t^1, \dots, B_t^d)_{t \geq 0}$$

be a  $d$ -dimensional Brownian motion and consider  $d+1$   $C^1$  vector fields  $V_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $n \geq 1$ ,  $i = 0, \dots, d$ . By using the language of vector fields and Stratonovitch integrals, the fundamental theorem for the existence and the uniqueness of solutions for stochastic differential equations is the following:



**Theorem 6.29.** *Assume that  $V_0, V_1, \dots, V_d$  are  $C^2$  vector fields with bounded derivatives up to order 2. Let  $x_0 \in \mathbb{R}^n$ . On  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , there exists a unique continuous and adapted process  $(X_t^{x_0})_{t \geq 0}$  such that for  $t \geq 0$ ,*

$$X_t^{x_0} = x_0 + \sum_{i=0}^d \int_0^t V_i(X_s^{x_0}) \circ dB_s^i, \tag{6.6}$$

with the convention that  $B_t^0 = t$ .

Thanks to Döblin-Itô's formula the Itô's formulation of a Stratonovitch equation is

$$X_t^{x_0} = x_0 + \frac{1}{2} \sum_{i=1}^d \int_0^t \nabla_{V_i} V_i(X_s^{x_0}) ds + \sum_{i=0}^d \int_0^t V_i(X_s^{x_0}) dB_s^i,$$

where for  $1 \leq i \leq d$ ,  $\nabla_{V_i} V_i$  is the vector field given by

$$\nabla_{V_i} V_i(x) = \sum_{j=1}^n \left( \sum_{k=1}^n v_i^k(x) \frac{\partial v_i^j}{\partial x_k}(x) \right) \frac{\partial}{\partial x_j}, \quad x \in \mathbb{R}^n.$$

If  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a  $C^2$  function, from Itô's formula, we have for  $t \geq 0$ ,

$$f(X_t^{x_0}) = f(x_0) + \sum_{i=0}^d \int_0^t (V_i f)(X_s^{x_0}) \circ dB_s^i,$$

and the process

$$\left( f(X_t^{x_0}) - \int_0^t (L f)(X_s^{x_0}) ds \right)_{t \geq 0}$$

is a local martingale where  $L$  is the second order differential operator

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

**Exercise 6.30.** *Let  $(B_t^1, B_t^2)_{t \geq 0}$  be a two-dimensional Brownian motion. Show that the process*

$$X_t = \left( B_t^1, B_t^2, \frac{1}{2} \int_0^t B_s^1 dB_s^2 - B_s^2 dB_s^1 \right)_{t \geq 0},$$

solves a stochastic differential equation that may be written

$$dX_t = V_1(X_t) \circ dB_t^1 + V_2(X_t) \circ dB_t^2$$

where  $V_1, V_2$  are two vector fields to be computed.

**Exercise 6.31.** Show that the diffusion operator

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

is elliptic in  $\mathbb{R}^n$  if and only if for every  $x \in \mathbb{R}^n$ , the vectors  $V_1(x), \dots, V_n(x)$  form a basis of  $\mathbb{R}^n$ .

A general theorem by Phillips and Sarason states that if  $L$  is a diffusion operator

$$L = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i},$$

where the coefficients  $a_{ij}$ 's and  $b_i$ 's are two times continuously differentiable, then  $L$  can be written as

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

where  $V_1, \dots, V_n$ ' are locally Lipschitz vector fields. In this book we are mostly interested in elliptic operators, in that case we have the following stronger result:

**Proposition 6.32.** Let

$$L = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i},$$

be a diffusion operator on  $\mathbb{R}^n$  such that the  $a_{ij}$ 's and the  $b_i$ 's are smooth functions. Let us assume that for every  $x \in \mathbb{R}^n$ , the rank of the matrix  $(a_{ij}(x))_{1 \leq i,j \leq n}$  is constant equal to  $n$ . Then, there exist smooth vector fields  $V_0, V_1, \dots, V_n$  on  $\mathbb{R}^n$  such that  $V_1, \dots, V_n$  are linearly independent and

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

*Proof.* Since the matrix  $(a_{ij}(x))_{1 \leq i,j \leq n}$  is symmetric and positive, it admits a unique symmetric and positive square root  $v(x) = (v_{ij}(x))_{1 \leq i,j \leq n}$ . Let us assume for a moment that the  $v_{ij}$ 's are smooth functions, in that case by denoting

$$V_i = \sum_{j=1}^n v_{ij} \frac{\partial}{\partial x_j},$$

the vector fields  $V_1, \dots, V_n$  are linearly independent and it is readily seen that the differential operator

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

is actually a first order differential operator and thus a vector field. We therefore are let to prove that the  $v_{ij}$ 's are smooth functions. Let  $\mathcal{O}$  be a bounded non empty set of  $\mathbb{R}^n$  and  $\Gamma$  be any contour in the half plane  $\mathbf{Re}(z) > 0$  that contains all the eigenvalues of  $\sigma(x)$ ,  $x \in \mathcal{O}$ . We claim that

$$v(x) = \frac{1}{2\pi} \int_{\Gamma} \sqrt{z}(\sigma(x) - z\mathbf{I}_n)^{-1} dz, \quad x \in \mathcal{O}.$$

Indeed, if  $\Gamma'$  is another contour in the half plane  $\mathbf{Re}(z) > 0$  whose interior contains  $\Gamma$ , as a straightforward application of the Fubini's theorem and Cauchy's formula we have

$$\begin{aligned} & \frac{1}{2\pi} \int_{\Gamma} \sqrt{z}(\sigma(x) - z\mathbf{I}_n)^{-1} dz \times \frac{1}{2\pi} \int_{\Gamma'} \sqrt{z}(\sigma(x) - z\mathbf{I}_n)^{-1} dz \\ &= \frac{1}{4\pi^2} \int_{\Gamma} \int_{\Gamma'} \sqrt{zz'}(\sigma(x) - z\mathbf{I}_n)^{-1}(\sigma(x) - z'\mathbf{I}_n)^{-1} dz dz' \\ &= \frac{1}{4\pi^2} \int_{\Gamma} \int_{\Gamma'} \sqrt{zz'} \frac{(\sigma(x) - z\mathbf{I}_n)^{-1} - (\sigma(x) - z'\mathbf{I}_n)^{-1}}{z - z'} dz dz' \\ &= -\frac{1}{4\pi^2} \int_{\Gamma} \int_{\Gamma'} \sqrt{zz'} \frac{(\sigma(x) - z\mathbf{I}_n)^{-1}}{z' - z} dz' dz \\ &= -\frac{1}{4\pi^2} \int_{\Gamma} (\sigma(x) - z\mathbf{I}_n)^{-1} \left( \int_{\Gamma'} \frac{\sqrt{zz'}}{z' - z} dz' \right) dz \\ &= \frac{1}{2i\pi} \int_{\Gamma} (\sigma(x) - z\mathbf{I}_n)^{-1} z dz \\ &= \frac{1}{2i\pi} \int_{\Gamma} (\sigma(x) - z\mathbf{I}_n)^{-1} \sigma(x) dz. \end{aligned}$$

In the last expression above, we may modify  $\Gamma$  into a circle  $\Gamma_R = \{z, |z| = R\}$ . Then by choosing  $R$  big enough ( $R > \sup_{x \in \mathcal{O}} \sqrt{\|\sigma(x)\|}$ ), and expanding  $\int_{\Gamma} (\sigma(x) - z\mathbf{I}_n)^{-1} dz$  in powers of  $z$ , we see that

$$\frac{1}{2i\pi} \int_{\Gamma_R} (\sigma(x) - z\mathbf{I}_n)^{-1} dz = \sigma(x).$$

As a conclusion,

$$\left( \frac{1}{2\pi} \int_{\Gamma} \sqrt{z}(\sigma(x) - z\mathbf{I}_n)^{-1} dz \right)^2 = \sigma(x),$$

so that, as we claimed it,

$$v(x) = \frac{1}{2\pi} \int_{\Gamma} \sqrt{z}(\sigma(x) - z\mathbf{I}_n)^{-1} dz, \quad x \in \mathcal{O}.$$

This expression of the square root of  $\sigma$  clearly shows that the  $v_{ij}$ 's are smooth functions.  $\square$

**Exercise 6.33.** Prove the following extension of the above theorem. Let

$$L = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i},$$

be a diffusion operator on  $\mathbb{R}^n$  such that the  $a_{ij}$ 's and the  $b_i$ 's are smooth functions. Let us assume that for every  $x \in \mathbb{R}^n$ , the rank of the matrix  $(a_{ij}(x))_{1 \leq i,j \leq n}$  is constant equal to  $d \leq n$ . Then, there exist smooth vector fields  $V_0, V_1, \dots, V_d$  on  $\mathbb{R}^n$  such that  $V_1, \dots, V_d$  are linearly independent and

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

## 6 Malliavin calculus

This section is an introduction to the techniques of the so-called Malliavin calculus. Our eventual goal will be to prove that if  $(X_t^x)_{t \geq 0}$  is the solution of a stochastic differential equation

$$X_t^x = x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dB_s,$$

where  $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  are smooth functions with bounded derivatives and if the matrix  $\sigma(x)$  is uniformly elliptic, then for every  $t > 0$  the random variable  $X_t^x$  has a smooth density with respect to the Lebesgue measure (see Theorem 6.51).

This Theorem 6.51 actually provides a probabilistic proof of the hypoellipticity of elliptic operators (see the Appendix on Regularity theory for the relevant definitions). Indeed, let  $L$  be the diffusion operator

$$L = \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j},$$

where  $a_{ij}(x) = (\sigma(x)\sigma^*(x))_{ij}$ . If  $L$  is elliptic, then Theorem 6.51 implies that for every  $x \in \mathbb{R}^n$  and  $t > 0$  the random variable  $X_t^x$  has a smooth density with respect to the Lebesgue measure. In other words, there exists a smooth kernel  $p(t, x, y)$  such that for every bounded and Borel function  $f$

$$\mathbb{E}(f(X_t^x)) = \mathbf{P}_t f(x) = \int_{\mathbb{R}^n} p(t, x, y) f(y),$$

Combining this fact with the interpretation of  $\mathbf{P}_t f$  as the solution of a Cauchy problem, we see that the operator  $L - \frac{\partial}{\partial t}$  needs to be hypoelliptic which implies the hypoellipticity of  $L$ . The fact that elliptic operators are hypoelliptic is a

special case of a celebrated theorem by Hörmander and the motivation of Malliavin to introduce his calculus was actually precisely to reprove Hörmander's result by using probabilistic techniques (see [?]).

We can also observe that Theorem 6.51 implies the Theorem 4.23 that we proved in Chapter 4 in the case where  $L$  is essentially self-adjoint. It should be observed that the advantage of the probabilistic method is that  $L$  is not required to be symmetric with respect to any measure, however strong assumptions are required on the coefficients  $b$  and  $\sigma$ .

## 6.1 The Malliavin derivative

In this section, we introduce the basic tools of Malliavin calculus, which is a set of tools devoted to the study of the Sobolev regularity of Brownian functionals. This is just an introduction to this theory. The interested reader is referred to the classical book [?] by Nualart for a detailed presentation of the theory.

Let us consider a filtered probability space  $(\Omega, (\mathcal{F}_t)_{0 \leq t \leq 1}, \mathbb{P})$  on which is defined a  $n$ -dimensional Brownian motion  $(B_t)_{0 \leq t \leq 1}$ . We assume that  $(\mathcal{F}_t)_{0 \leq t \leq 1}$  is the usual completion of the natural filtration of  $(B_t)_{0 \leq t \leq 1}$ .

A  $\mathcal{F}_1$  measurable real valued random variable  $F$  is said to be cylindrical if it can be written

$$F = f \left( \int_0^1 h^1(s) dB_s, \dots, \int_0^1 h^m(s) dB_s \right)$$

where  $h^i \in \mathbf{L}^2([0, 1], \mathbb{R}^n)$  and  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  is a  $C^\infty$  function such that  $f$  and all its partial derivatives have polynomial growth. The set of cylindrical random variables is denoted by  $\mathcal{S}$ . It is easy to see that  $\mathcal{S}$  is dense in  $\mathbf{L}^p(\mathcal{F}_1, \mathbb{P})$  for every  $p \geq 1$ .

The Malliavin derivative of  $F \in \mathcal{S}$  is the  $\mathbb{R}^n$  valued stochastic process  $(\mathbf{D}_t F)_{0 \leq t \leq 1}$  given by

$$\mathbf{D}_t F = \sum_{i=1}^m h^i(t) \frac{\partial f}{\partial x_i} \left( \int_0^1 h^1(s) dB_s, \dots, \int_0^1 h^m(s) dB_s \right).$$

We can see  $\mathbf{D}$  as an (unbounded) operator from the space  $\mathcal{S} \subset \mathbf{L}^p(\mathcal{F}_1, \mathbb{P})$  into the Banach space

$$\mathcal{L}^p = \left\{ (X_t)_{0 \leq t \leq 1}, \mathbb{E} \left( \left( \int_0^1 \|X_t\|^2 dt \right)^p \right) < +\infty \right\}.$$

Our first task will be to prove that  $\mathbf{D}$  is closable (see the Appendix on unbounded operators for the definition of closability). This will be a consequence of the following fundamental integration by parts formula which is interesting in itself.

**Proposition 6.34** (Integration by parts formula). *Let  $F \in \mathcal{S}$  and  $(h(s))_{0 \leq s \leq 1}$  be a progressively measurable such that  $\mathbb{E} \left( \int_0^1 \|h(s)\|^2 ds \right) < +\infty$ . We have*

$$\mathbb{E} \left( \int_0^1 (\mathbf{D}_s F) h(s) ds \right) = \mathbb{E} \left( F \int_0^1 h(s) dB_s \right).$$

*Proof.* Let

$$F = f \left( \int_0^1 h^1(s) dB_s, \dots, \int_0^1 h^m(s) dB_s \right) \in \mathcal{S}.$$

Let us now fix  $\varepsilon \geq 0$  and denote

$$F_\varepsilon = f \left( \int_0^1 h^1(s) d \left( B_s + \varepsilon \int_0^s h(u) du \right), \dots, \int_0^1 h^m(s) d \left( B_s + \varepsilon \int_0^s h(u) du \right) \right).$$

From Girsanov's theorem (Theorem 5.72 in Chapter 3), we have

$$\mathbb{E}(F_\varepsilon) = \mathbb{E} \left( \exp \left( \varepsilon \int_0^1 h(u) dB_u - \frac{\varepsilon^2}{2} \int_0^1 \|h(u)\|^2 du \right) F \right).$$

Now, on one hand we compute

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\mathbb{E}(F_\varepsilon) - \mathbb{E}(F)) \\ &= \mathbb{E} \left( \int_0^1 \sum_{i=1}^m \frac{\partial f}{\partial x_i} \left( \int_0^1 h^1(s) dB_s, \dots, \int_0^1 h^m(s) dB_s \right) h^i(s) h(s) dt \right) \\ &= \mathbb{E} \left( \int_0^1 (\mathbf{D}_s F) h(s) dt \right), \end{aligned}$$

and on the other hand, we obtain

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\mathbb{E}(F_\varepsilon) - \mathbb{E}(F)) = \mathbb{E} \left( F \int_0^1 h(s) dB_s \right).$$

□

**Proposition 6.35.** *Let  $p \geq 1$ . As a densely defined operator from  $\mathbf{L}^p(\mathcal{F}_1, \mathbb{P})$  into  $\mathcal{L}^p$ ,  $\mathbf{D}$  is closable.*

*Proof.* Let  $(F_n)_{n \in \mathbb{N}}$  be a sequence in  $\mathcal{S}$  that converges in  $\mathbf{L}^p(\mathcal{F}_1, \mathbb{P})$  to 0 and such that  $\mathbf{D}F_n$  converges in  $\mathcal{L}^p$  to  $X$ . We want to prove that  $X = 0$ . Let  $(h(s))_{0 \leq s \leq 1}$  be a function in  $L^2([0, 1])$ . Let us first assume  $p > 1$ . We have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \int_0^1 (\mathbf{D}_s F_n) h(s) ds \right) = \mathbb{E} \left( \int_0^1 X_s h(s) ds \right),$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( F_n \int_0^1 h(s) dB_s \right) = 0.$$

As a consequence, we obtain

$$\mathbb{E} \left( \int_0^1 X_s h(s) ds \right) = 0.$$

Since  $h$  is arbitrary, we conclude  $X = 0$ . Let us now assume  $p = 1$ . Let  $\eta$  be a smooth and compactly supported function and let  $\Theta = \eta \left( \int_0^1 h(s) dB_s \right) \in \mathcal{S}$ . We have

$$\mathbf{D}(F_n \Theta) = F_n(\mathbf{D}\Theta) + (\mathbf{D}F_n)\Theta.$$

As a consequence, we get

$$\mathbb{E} \left( \int_0^1 \mathbf{D}_s(F_n \Theta) h(s) ds \right) = \mathbb{E} \left( F_n \int_0^1 (\mathbf{D}_s \Theta) h(s) ds \right) + \mathbb{E} \left( \Theta \int_0^1 (\mathbf{D}_s F_n) h(s) ds \right),$$

and thus

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \int_0^1 \mathbf{D}_s(F_n \Theta) h(s) ds \right) = \mathbb{E} \left( \Theta \int_0^1 X_s h(s) ds \right).$$

On the other hand, we have

$$\mathbb{E} \left( \int_0^1 \mathbf{D}_s(F_n \Theta) h(s) ds \right) = \mathbb{E} \left( F_n \Theta \int_0^1 h(s) dB_s \right) \rightarrow_{n \rightarrow \infty} 0.$$

We conclude

$$\mathbb{E} \left( \Theta \int_0^1 X_s h(s) ds \right) = 0.$$

□

The closure of  $\mathbf{D}$  in  $\mathbf{L}^p(\mathcal{F}_1, \mathbb{P})$  shall still be denoted by  $\mathbf{D}$ . Its domain  $\mathbb{D}^{1,p}$  is the closure of  $\mathcal{S}$  with respect to the norm

$$\|F\|_{1,p} = \left( \mathbb{E}(F^p) + \mathbb{E} \left( \|\mathbf{D}F\|_{\mathbf{L}^2([0,1], \mathbb{R}^n)}^p \right) \right)^{\frac{1}{p}},$$

For  $p > 1$ , we can consider the adjoint operator  $\delta$  of  $\mathbf{D}$ . This is a densely defined operator  $\mathcal{L}^q \rightarrow \mathbf{L}^p(\mathcal{F}_1, \mathbb{P})$  with  $1/p + 1/q = 1$  which is characterized by the duality formula

$$\mathbb{E}(F \delta u) = \mathbb{E} \left( \int_0^1 (\mathbf{D}_s F) u_s ds \right), \quad F \in \mathbb{D}^{1,p}.$$

From Proposition 6.34 and Burkholder-Davis-Gundy inequalities, it is clear that the domain of  $\delta$  in  $\mathcal{L}^q$  contains the set of progressively measurable processes  $(u_t)_{0 \leq t \leq 1}$  such that  $\mathbb{E} \left( \left( \int_0^1 \|u_s\|^2 ds \right)^{q/2} \right) < +\infty$  and that in this case,

$$\delta u = \int_0^1 u_s dB_s.$$

The operator  $\delta$  can thus be thought as an extension of the Itô's integral. It is often called the Skohorod integral or the divergence operator.

**Exercise 6.36** (Clark-Ocone formula). *Show that for  $F \in \mathbb{D}^{1,2}$ ,*

$$F = \mathbb{E}(F) + \int_0^1 \mathbb{E}(\mathbf{D}_1 F \mid \mathcal{F}_t) dB_t.$$

More generally, we can introduce iterated derivatives. If  $F \in \mathcal{S}$ , we set

$$\mathbf{D}_{t_1, \dots, t_k}^k F = \mathbf{D}_{t_1} \dots \mathbf{D}_{t_k} F.$$

We may consider  $\mathbf{D}^k F$  as a square integrable random process indexed by  $[0, 1]^k$  and valued in  $\mathbb{R}^n$ . For any  $p \geq 1$ , the operator  $\mathbf{D}^k$  is closable on  $\mathcal{S}$ . We denote  $\mathbb{D}^{k,p}$  the closure of the class of cylindric random variables with respect to the norm

$$\|F\|_{k,p} = \left( \mathbb{E}(F^p) + \sum_{j=1}^k \mathbb{E} \left( \|\mathbf{D}^j F\|_{\mathbf{L}^2([0,1]^j, \mathbb{R}^n)}^p \right) \right)^{\frac{1}{p}},$$

and

$$\mathbb{D}^\infty = \bigcap_{p \geq 1} \bigcap_{k \geq 1} \mathbb{D}^{k,p}.$$

## 6.2 The Wiener chaos expansion

As in the previous section, we consider a filtered probability space  $(\Omega, (\mathcal{F}_t)_{0 \leq t \leq 1}, \mathbb{P})$  on which is defined a Brownian motion  $(B_t)_{0 \leq t \leq 1}$ , and we assume that  $(\mathcal{F}_t)_{0 \leq t \leq 1}$  is the usual completion of the natural filtration of  $(B_t)_{0 \leq t \leq 1}$ . Our goal is here to write an orthogonal decomposition of the space  $\mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$  that is particularly suited to the study of the space  $\mathbb{D}^{1,2}$ . For simplicity of the exposition, we restrict ourselves to the case where the Brownian motion  $(B_t)_{0 \leq t \leq 1}$  is one-dimensional but similar results obviously hold in higher dimensions provided the suitable changes are made.

In the sequel, for  $n \geq 1$ , we denote by  $\Delta_n$  the simplex

$$\Delta_n = \{0 \leq t_1 \leq \dots \leq t_n \leq 1\}$$



and if  $f_n \in \mathbf{L}^2(\Delta_n)$ ,

$$\begin{aligned} I_n(f_n) &= \int_0^1 \int_0^{t_n} \cdots \int_0^{t_2} f_n(t_1, \dots, t_n) dB_{t_1} \dots dB_{t_n} \\ &= \int_{\Delta_n} f_n(t_1, \dots, t_n) dB_{t_1} \dots dB_{t_n}. \end{aligned}$$

The set

$$\mathbf{K}_n = \left\{ \int_{\Delta_n} f_n(t_1, \dots, t_n) dB_{t_1} \dots dB_{t_n}, f_n \in L^2(\Delta_n) \right\}$$

is called the space of chaos of order  $n$ . By convention the set of constant random variables shall be denoted by  $\mathbf{K}_0$ .

By using the Itô's isometry, we readily compute that

$$\mathbb{E}(I_n(f_n)I_p(f_p)) = \begin{cases} 0 & \text{if } p \neq n \\ \|f_n\|_{L^2(\Delta_n)}^2 & \text{if } p = n. \end{cases}$$

As a consequence, the spaces  $\mathbf{K}_n$  are orthogonal in  $\mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$ . It is easily seen that  $\mathbf{K}_n$  is the closure of the linear span of the family

$$\{I_n(f^{\otimes n}), f \in \mathbf{L}^2([0, 1])\},$$

where for  $f \in \mathbf{L}^2([0, 1])$ , we denote by  $f^{\otimes n}$  the map  $\Delta_n \rightarrow \mathbb{R}$  such that  $f^{\otimes n}(t_1, \dots, t_n) = f(t_1) \cdots f(t_n)$ . It turns out that  $I_n(f^{\otimes n})$  can be computed by using Hermite polynomials. The Hermite polynomial of order  $n$  is defined as

$$H_n(x) = (-1)^n \frac{1}{n!} e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}.$$

By the very definition of  $H_n$ , we see that for every  $t, x \in \mathbb{R}$ ,

$$\exp\left(tx - \frac{t^2}{2}\right) = \sum_{k=0}^{+\infty} t^k H_k(x).$$

**Lemma 6.37.** *If  $f \in \mathbf{L}^2([0, 1])$  then*

$$I_n(f^{\otimes n}) = \|f\|_{L^2([0,1])}^n H_n\left(\frac{\int_0^1 f(s)dB_s}{\|f\|_{L^2([0,1])}}\right).$$

*Proof.* On one hand, we have for  $\lambda \in \mathbb{R}$ ,

$$\exp\left(\lambda \int_0^1 f(s)dB_s - \frac{\lambda^2}{2} \int_0^1 f(s)^2 ds\right) = \sum_{n=0}^{+\infty} \lambda^n \|f\|_{L^2([0,1])}^n H_n\left(\frac{\int_0^1 f(s)dB_s}{\|f\|_{L^2([0,1])}}\right).$$

On the other hand, for  $0 \leq t \leq 1$ , let us consider

$$M_t(\lambda) = \exp\left(\lambda \int_0^t f(s)dB_s - \frac{\lambda^2}{2} \int_0^t f(s)^2 ds\right).$$

From Itô's formula, we have

$$M_t(\lambda) = 1 + \lambda \int_0^t M_s f(s)dB_s.$$

By iterating the previous linear relation, we easily obtain that for every  $n \geq 1$ ,

$$\begin{aligned} & M_1(\lambda) \\ = & 1 + \sum_{k=1}^n \lambda^k I_k(f^{\otimes k}) + \lambda^{n+1} \int_0^1 M_t f(t) \left( \int_{\Delta_n([0,t])} f(t_1) \cdots f(t_n) dB_{t_1} \cdots dB_{t_n} \right) dB_t. \end{aligned}$$

We conclude,

$$I_n(f^{\otimes n}) = \frac{1}{n!} \frac{d^n M_1}{d\lambda^n}(0) = \|f\|_{L^2([0,1])}^n H_n \left( \frac{\int_0^1 f(s)dB_s}{\|f\|_{L^2([0,1])}} \right).$$

□

As we pointed it out, for  $p \neq n$ , the spaces  $\mathbf{K}_n$  and  $\mathbf{K}_p$  are orthogonal. We have the following orthogonal decomposition of  $L^2$ :

**Theorem 6.38** (Wiener chaos expansion).

$$\mathbf{L}^2(\mathcal{F}_1, \mathbb{P}) = \bigoplus_{n \geq 0} \mathbf{K}_n.$$

*Proof.* As a by-product of the previous proof, we easily obtain that for  $f \in \mathbf{L}^2([0, 1])$ ,

$$\exp\left(\lambda \int_0^1 f(s)dB_s - \frac{\lambda^2}{2} \int_0^1 f(s)^2 ds\right) = \sum_{n=1}^{+\infty} I_n(f^{\otimes n}),$$

where the convergence of the series is almost sure but also in  $\mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$ . Therefore, if  $F \in \mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$  is orthogonal to  $\bigoplus_{n \geq 1} \mathbf{K}_n$ , then  $F$  is orthogonal to every  $\exp\left(\lambda \int_0^1 f(s)dB_s - \frac{\lambda^2}{2} \int_0^1 f(s)^2 ds\right)$ ,  $f \in \mathbf{L}^2([0, 1])$ . This implies that  $F = 0$ . □

As we are going to see, the space  $\mathbb{D}^{1,2}$  or more generally  $\mathbb{D}^{k,2}$  is easy to describe by using the Wiener chaos expansion. The keypoint is the following proposition:

**Proposition 6.39.** *Let  $F = I_n(f_n) \in \mathbf{K}_n$ , then  $F \in \mathbb{D}^{1,2}$  and*

$$\mathbf{D}_t F = I_{n-1}(\tilde{f}_n(\cdot, t)),$$

where for  $0 \leq t_1 \leq \dots \leq t_{n-1} \leq 1$ ,

$$\tilde{f}_n(t_1, \dots, t_{n-1}, t) = f_n(t_1, \dots, t_k, t, t_{k+1}, \dots, t_{n-1}) \quad \text{if } t_k \leq t \leq t_{k+1}.$$

*Proof.* Let  $f \in \mathbf{L}^2([0, 1])$ . We have

$$I_n(f^{\otimes n}) = \|f\|_{\mathbf{L}^2([0,1])}^n H_n \left( \frac{\int_0^1 f(s) dB_s}{\|f\|_{\mathbf{L}^2([0,1])}} \right).$$

Thus  $F = I_n(f^{\otimes n})$  is a smooth cylindrical functional and

$$\mathbf{D}_t F = \|f\|_{\mathbf{L}^2([0,1])}^{n-1} f(t) H'_n \left( \frac{\int_0^1 f(s) dB_s}{\|f\|_{\mathbf{L}^2([0,1])}} \right).$$

It is easy to see that  $H'_n = H_{n-1}$ , therefore we have

$$\begin{aligned} \mathbf{D}_t F &= \|f\|_{\mathbf{L}^2([0,1])}^{n-1} f(t) H_{n-1} \left( \frac{\int_0^1 f(s) dB_s}{\|f\|_{\mathbf{L}^2([0,1])}} \right) \\ &= f(t) I_{n-1}(f^{\otimes(n-1)}). \end{aligned}$$

As a consequence, we compute that

$$\mathbb{E} \left( \int_0^1 (\mathbf{D}_t F)^2 dt \right) = n \mathbb{E}(F^2).$$

We now observe that  $\mathbf{K}_n$  is the closure in  $\mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$  of the linear span of the family

$$\{I_n(f^{\otimes n}), f \in \mathbf{L}^2([0, 1])\}$$

to conclude the proof of the proposition.  $\square$

We can finally turn to the description of  $\mathbb{D}^{1,2}$  using the chaos decomposition:

**Theorem 6.40.** *Let  $F \in \mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$  and let*

$$F = \mathbb{E}(F) + \sum_{m \geq 1} I_m(f_m),$$

be the chaotic decomposition of  $F$ . Then  $F \in \mathbb{D}^{1,2}$  if and only if

$$\sum_{m \geq 1} m \mathbb{E}(I_m(f_m)^2) < +\infty,$$

and in that case,

$$\mathbf{D}_t F = \mathbb{E}(\mathbf{D}_t F) + \sum_{m \geq 2} I_{m-1}(\tilde{f}_m(\cdot, t)).$$

*Proof.* It is a consequence of the fact that for  $F \in \mathbf{K}_n$ ,

$$\mathbb{E} \left( \int_0^1 (\mathbf{D}_t F)^2 dt \right) = n \mathbb{E}(F^2).$$

□

An immediate but useful corollary of the previous theorem is the following result:

**Corollary 6.41.** *Let  $(F_n)_{n \geq 0}$  be a sequence in  $\mathbb{D}^{1,2}$  that converges to  $F$  in  $\mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$  and such that*

$$\sup_{n \geq 0} \mathbb{E} \left( \int_0^1 (\mathbf{D}_t F_n)^2 dt \right) < +\infty.$$

*Then,  $F \in \mathbb{D}^{1,2}$ .*

This corollary can actually be generalized in the following way (see [?] for a proof).

**Proposition 6.42.** *Let  $(F_n)_{n \geq 0}$  be a sequence in  $\mathbb{D}^{k,p}$ ,  $k \geq 1$ ,  $p > 1$ , that converges to  $F$  in  $\mathbf{L}^p(\mathcal{F}_1, \mathbb{P})$  and such that*

$$\sup_{n \geq 0} \|F_n\|_{k,p} < +\infty.$$

*Then,  $F \in \mathbb{D}^{k,p}$ .*

**Exercise 6.43.** *Let  $F \in \mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$  and let*

$$F = \mathbb{E}(F) + \sum_{m \geq 1} I_m(f_m),$$

*be the chaotic decomposition of  $F$ . Show that that  $F \in \mathbb{D}^{k,2}$  with  $k \geq 1$  if and only if*

$$\sum_{m \geq 1} m^k \mathbb{E} (I_m(f_m)^2) < +\infty.$$

**Exercise 6.44.** *Let  $L = \delta \mathbf{D}$ . Show that for  $F \in \mathbf{K}_n$ ,  $LF = nF$ . Deduce that the domain of  $L$  in  $\mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$  is  $\mathbb{D}^{2,2}$*

### 6.3 Regularity of probability distributions using Malliavin calculus

We have the following key result which makes Malliavin calculus so useful when one wants to study the existence of densities for random variables.

**Theorem 6.45.** *Let  $F = (F_1, \dots, F_m)$  be a  $\mathcal{F}_1$  measurable random vector such that:*

1. *for every  $i = 1, \dots, m$ ,  $F_i \in \mathbb{D}^\infty$ ;*
2. *the matrix*

$$\Gamma = \left( \int_0^1 \langle \mathbf{D}_s F^i, \mathbf{D}_s F^j \rangle_{\mathbb{R}^n} ds \right)_{1 \leq i, j \leq m}$$

*is invertible.*

*Then  $F$  has a density with respect to the Lebesgue measure. If moreover, for every  $p > 1$ ,*

$$\mathbb{E} \left( \frac{1}{|\det \Gamma|^p} \right) < +\infty,$$

*then this density is smooth.*

**Remark 6.46.** The matrix  $\Gamma$  is often called the Malliavin matrix of the random vector  $F$ .

This theorem relies on the following lemma of Fourier analysis for which we shall use the following notation: If  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  is a smooth function then for  $\alpha = (i_1, \dots, i_k) \in \{1, \dots, n\}^k$ , we denote

$$\partial_\alpha \phi = \frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}} \phi.$$

**Lemma 6.47.** *Let  $\mu$  be a probability measure on  $\mathbb{R}^n$  such that for every smooth and compactly supported function  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ ,*

$$\left| \int_{\mathbb{R}^n} \partial_\alpha \phi d\mu \right| \leq C_\alpha \|\phi\|_\infty,$$

*where  $\alpha \in \{1, \dots, n\}^k$ ,  $k \geq 1$ ,  $C_\alpha > 0$ . Then  $\mu$  is absolutely continuous with respect to the Lebesgue measure with a smooth density.*

*Proof.* The idea is to show that we may assume that  $\mu$  is compactly supported and then use Fourier transforms techniques. Let  $x_0 \in \mathbb{R}^n$ ,  $R > 0$  and  $R' > R$ . Let  $\Psi$  be a smooth function on  $\mathbb{R}^n$  such that  $\Psi = 1$  on the ball  $\mathbf{B}(x_0, R)$  and  $\Psi = 0$  outside the ball  $\mathbf{B}(x_0, R')$ . Let  $\nu$  be the measure on  $\mathbb{R}^n$  that has a density  $\Psi$  with respect to  $\mu$ . It is easily seen, by induction and integrating by parts that for every smooth and compactly supported function  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\left| \int_{\mathbb{R}^n} \partial_\alpha \phi d\nu \right| \leq C'_\alpha \|\phi\|_\infty,$$

where  $\alpha \in \{1, \dots, n\}^k$ ,  $k \geq 1$ ,  $C'_\alpha > 0$ . Now, if we can prove that under the above assumption  $\nu$  has a smooth density, then we will be able to conclude that  $\phi$  has a smooth density because  $x_0 \in \mathbb{R}^n$  and  $R, R'$  are arbitrary. Let

$$\hat{\nu}(y) = \int_{\mathbb{R}^n} e^{i\langle y, x \rangle} \nu(dx)$$

be the Fourier transform of the measure  $\mu$ . The assumption implies that  $\hat{\nu}$  is rapidly decreasing (apply the inequality with  $\phi(x) = e^{i\langle y, x \rangle}$ ). We conclude that  $\nu$  has a smooth density with respect to the Lebesgue measure and that this density  $f$  is given by the inverse Fourier transform formula:

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\langle y, x \rangle} \hat{\nu}(y) dy.$$

□

We may now turn to the proof of Theorem 6.45.

*Proof.* The proof relies on an integration by parts formula which is interesting by itself. Let  $\phi$  be a smooth and compactly supported function on  $\mathbb{R}^n$ . Since  $F_i \in \mathbb{D}^\infty$ , we easily deduce that  $\phi(F) \in \mathbb{D}^\infty$  and that

$$\mathbf{D}\phi(F) = \sum_{i=1}^n \partial_i \phi(F) \mathbf{D}F_i.$$

Therefore we obtain

$$\int_0^1 \langle \mathbf{D}_t \phi(F), \mathbf{D}_t F_j \rangle dt = \sum_{i=1}^n \partial_i \phi(F) \int_0^1 \langle \mathbf{D}_t F_i, \mathbf{D}_t F_j \rangle dt.$$

We conclude that

$$\partial_i \phi(F) = \sum_{j=1}^n (\Gamma^{-1})_{i,j} \int_0^1 \langle \mathbf{D}_t \phi(F), \mathbf{D}_t F_j \rangle dt.$$

As a consequence, we obtain

$$\begin{aligned} \mathbb{E}(\partial_i \phi(F)) &= \mathbb{E} \left( \sum_{j=1}^n (\Gamma^{-1})_{i,j} \int_0^1 \langle \mathbf{D}_t \phi(F), \mathbf{D}_t F_j \rangle dt \right) \\ &= \sum_{j=1}^n \mathbb{E} \left( \int_0^1 \langle \mathbf{D}_t \phi(F), (\Gamma^{-1})_{i,j} \mathbf{D}_t F_j \rangle dt \right) \\ &= \sum_{j=1}^n \mathbb{E}(\phi(F) \delta((\Gamma^{-1})_{i,j} \mathbf{D}F_j)) \\ &= \mathbb{E} \left( \phi(F) \delta \left( \sum_{j=1}^n (\Gamma^{-1})_{i,j} \mathbf{D}F_j \right) \right) \end{aligned}$$

Remark that the integration by parts is licit because  $(\Gamma^{-1})_{i,j} \mathbf{D}F_j$  belongs to the domain of  $\delta$ . Indeed, from our assumptions,  $\mathbf{D}F_j$  is in  $\mathbb{D}^{2,2}$  and thus in the domain of  $\delta$  from Exercise 6.44. Also  $(\Gamma^{-1})_{i,j} \in \mathbb{D}^{1,2}$  and  $(\Gamma^{-1})_{i,j} \mathbf{D}F_j \in \mathbf{L}^2(\mathcal{F}_1, \mathbb{P})$  which easily implies that  $(\Gamma^{-1})_{i,j} \mathbf{D}F_j$  belongs to the domain of  $\delta$ .

By using inductively this integration by parts formula, it is seen that for every  $\alpha \in \{1, \dots, n\}^k$ ,  $k \geq 1$ , there exists an integrable random variable  $Z_\alpha$  such that,

$$\mathbb{E}(\partial_\alpha \phi(F)) = \mathbb{E}(\phi(F)Z_\alpha).$$

This finishes the proof.  $\square$

## 7 Existence of a smooth density

As usual, we consider a filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$  which satisfies the usual conditions and on which is defined a  $n$ -dimensional Brownian motion  $(B_t)_{t \geq 0}$ . Our first purpose here, is to prove that solutions of stochastic differential equations are differentiable in the sense of Malliavin. The following lemma is easy to prove by using the Wiener chaos expansion.

**Lemma 6.48.** *Let  $(u_s)_{0 \leq s \leq 1}$  be a progressively measurable process such that for every  $0 \leq s \leq 1$ ,  $u_s^i \in \mathbb{D}^{1,2}$  and*

$$\mathbb{E} \left( \int_0^1 \|u_s\|^2 ds \right) < +\infty, \quad \mathbb{E} \left( \int_0^1 \int_0^1 \|\mathbf{D}_s u_t\|^2 ds dt \right) < +\infty.$$

Then  $\int_0^1 u_s dB_s \in \mathbb{D}^{1,2}$  and

$$\mathbf{D}_t \left( \int_0^1 u_s dB_s \right) = u_t + \sum_{i=1}^n \int_t^1 (\mathbf{D}_t u_s^i) dB_s^i.$$

*Proof.* We make the proof when  $n = 1$  and use the notations introduced in the Wiener chaos expansion section. For  $f \in \mathbf{L}^2([0, 1])$ , we have

$$\mathbf{D}_t I_n(f^{\otimes n}) = f(t) I_{n-1}(f^{\otimes(n-1)}).$$

But we can write,

$$I_n(f^{\otimes n}) = \int_0^1 f(t) \left( \int_{\Delta_{n-1}[0,t]} f^{\otimes(n-1)} dB_{t_1} \cdots dB_{t_{n-1}} \right) dB_t,$$

and thus

$$I_n(f^{\otimes n}) = \int_0^1 u_s dB_s,$$

with  $u_t = f(t) \int_{\Delta_{n-1}[0,t]} f^{\otimes(n-1)} dB_{t_1} \cdots dB_{t_{n-1}}$ . Since

$$\begin{aligned} f(t)I_{n-1}(f^{\otimes(n-1)}) &= f(t) \left( \int_{\Delta_{n-1}[0,t]} f^{\otimes(n-1)} dB_{t_1} \cdots dB_{t_{n-1}} \right) \\ &\quad + f(t) \int_t^1 f(s) \left( \int_{\Delta_{n-2}[0,s]} f^{\otimes(n-1)} dB_{t_1} \cdots dB_{t_{n-2}} \right) dB_s, \end{aligned}$$

we get the result when  $\int_0^1 u_s dB_s$  can be written as  $I_n(f^{\otimes n})$ . By continuity of the Malliavin derivative on the space of chaos of order  $n$ , we conclude that the formula is true if  $\int_0^1 u_s dB_s$  is a chaos of order  $n$ . The result finally holds in all generality by using the Wiener chaos expansion.  $\square$

We consider two functions  $b : \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  and we assume that  $b$  and  $\sigma$  are  $C^\infty$  with derivatives at any order (more than 1) bounded.

As we know, there exists a bicontinuous process  $(X_t^x)_{t \geq 0, x \in \mathbb{R}^d}$  such that for  $t \geq 0$ ,

$$X_t^x = x + \int_0^t b(X_s^x) ds + \sum_{k=1}^n \int_0^t \sigma_k(X_s^x) dB_s^k. \quad (6.7)$$

Moreover, for every  $p \geq 1$ , and  $T \geq 0$

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} \|X_t^x\|^p \right) < +\infty.$$

**Theorem 6.49.** *For every  $i = 1, \dots, n$ ,  $0 \leq t \leq 1$ ,  $X_t^{x,i} \in \mathbb{D}^\infty$  and for  $r \leq t$ ,*

$$\mathbf{D}_r^j X_t^{x,i} = \sigma_{i,j}(X_r^x) + \sum_{l=1}^n \int_r^t \partial_l b_i(X_s^x) \mathbf{D}_r^j X_s^{x,l} ds + \sum_{k,l=1}^n \int_r^t \partial_l \sigma_{i,k}(X_s^x) \mathbf{D}_r^j X_s^{x,l} dB_s^k, \quad (6.8)$$

where  $\mathbf{D}_r^j X_t^{x,i}$  is the  $j$ -th component of  $\mathbf{D}_r X_t^{x,i}$ . If  $r > t$ , then  $\mathbf{D}_r^j X_t^{x,i} = 0$ .

*Proof.* We first prove that  $X_1^{x,i} \in \mathbb{D}^{1,p}$  for every  $p \geq 1$ . We consider the Picard approximations given by  $X_0(t) = x$  and

$$X_{n+1}(t) = x + \int_0^t b(X_n(s)) ds + \sum_{k=1}^n \int_0^t \sigma_k(X_n(s)) dB_s^k.$$

By induction, it is easy to see that  $X_n(t) \in \mathbb{D}^{1,p}$  and that for every  $p \geq 1$ , we have

$$\Psi_n(t) = \sup_{0 \leq r \leq t} \mathbb{E} \left( \sup_{s \in [r,t]} \|\mathbf{D}_r X_n(s)\|^p \right) < +\infty,$$



and

$$\Psi_{n+1}(t) \leq \alpha + \beta \int_0^t \Psi_n(s) ds.$$

Then, we observe that  $X_n(t)$  converges to  $X_t^x$  in  $L^p$  and that the sequence  $\|X_n(t)\|_{1,p}$  is bounded. As a consequence  $X_1^{x,i} \in \mathbb{D}^{1,p}$  for every  $p \geq 1$ . The equation for the Malliavin derivative is obtained by differentiating the equation satisfied by  $X_t^x$ . Higher order derivatives may be treated in a similar way with a few additional work, the details are let to the reader.  $\square$

Combining this theorem with Theorem 6.13 and using the uniqueness property for solutions of linear stochastic differential equations, we obtain the following representation for the Malliavin derivative of a solution of a stochastic differential equation:

**Corollary 6.50.** *We have*

$$\mathbf{D}_r^j X_t^x = \mathbf{J}_{0 \rightarrow t}(x) \mathbf{J}_{0 \rightarrow r}^{-1}(x) \sigma_j(X_r^x), \quad j = 1, \dots, n, \quad 0 \leq r \leq t,$$

where  $(\mathbf{J}_{0 \rightarrow t}(x))_{t \geq 0}$  is the first variation process defined by

$$\mathbf{J}_{0 \rightarrow t}(x) = \frac{\partial X_t^x}{\partial x}(x).$$

We now fix  $x \in \mathbb{R}^n$  as the initial condition for our equation and denote by

$$\Gamma_t = \left( \sum_{j=1}^n \int_0^1 \mathbf{D}_r^j X_t^{i,x} \mathbf{D}_r^j X_t^{i',x} dr \right)_{1 \leq i, i' \leq n}$$

the Malliavin matrix of  $X_t^x$ . From the previous corollary, we deduce that

$$\Gamma_t(x) = \mathbf{J}_{0 \rightarrow t}(x) \int_0^t \mathbf{J}_{0 \rightarrow r}^{-1}(x) \sigma(X_r^x) \sigma(X_r^x)^* \mathbf{J}_{0 \rightarrow r}^{-1}(x)^* dr \mathbf{J}_{0 \rightarrow t}(x)^*. \quad (6.9)$$

We are now finally in position to state the main theorem of the section:

**Theorem 6.51.** *Assume that there exists  $\lambda > 0$  such that for every  $x \in \mathbb{R}^n$ ,*

$$\|\sigma(x)\|^2 \geq \lambda \|x\|^2,$$

*then for every  $t > 0$  and  $x \in \mathbb{R}^n$ , the random variable  $X_t^x$  has a smooth density with respect to the Lebesgue measure.*

*Proof.* We use Theorem 6.45. We therefore want to prove that  $\Gamma_t(x)$  is invertible with inverse in  $L^p$  for  $p \geq 1$ . Since  $\mathbf{J}_{0 \rightarrow t}(x)$  is invertible and its inverse solves a linear equation (see Exercise 6.14), we deduce that for every  $p \geq 1$ ,

$$\mathbb{E} (\|\mathbf{J}_{0 \rightarrow t}^{-1}(x)\|^p) < +\infty.$$

From formula (6.9), we conclude that it is enough to prove that  $C_t(x)$  is invertible with inverse in  $L^p$  where

$$C_t(x) = \int_0^t \mathbf{J}_{0 \rightarrow r}^{-1}(x) \sigma(X_r^x) \sigma(X_r^x)^* \mathbf{J}_{0 \rightarrow r}^{-1}(x)^* dr.$$

By the uniform ellipticity assumption, we have

$$C_t(x) \geq \lambda \int_0^t \mathbf{J}_{0 \rightarrow r}^{-1}(x) \mathbf{J}_{0 \rightarrow r}^{-1}(x)^* dr,$$

where the inequality is understood in the sense that the difference of the two symmetric matrices is non negative. This implies that  $C_t(x)$  is invertible. Moreover, it is an easy exercise to prove that if  $M_t$  is a continuous map taking its values in the set of positive definite matrices, then we have

$$\left( \int_0^t M_s ds \right)^{-1} \leq \frac{1}{t^2} \left( \int_0^t M_s^{-1} ds \right).$$

As a consequence, we obtain

$$C_t^{-1}(x) \leq \frac{1}{t^2 \lambda} \int_0^t \mathbf{J}_{0 \rightarrow r}(x)^* \mathbf{J}_{0 \rightarrow r}(x) dr.$$

Since  $\mathbf{J}_{0 \rightarrow r}(x)$  has moments in  $L^p$  for all  $p \geq 1$ , we conclude that  $C_t(x)$  is invertible with inverse in  $L^p$ .  $\square$

## Notes and Comments

**Sections 1,2,3,4.** Stochastic differential equations were first considered by Itô [?], with the purpose of providing a pathwise construction of a diffusion process associated to a given diffusion operator. We have to mention that another pathwise construction of diffusion processes was made by Döblin in 1940, and thus before Itô, in a paper that was rediscovered in 2000 (see [?]). Essentially, to construct the diffusion with generator

$$L = \frac{1}{2} \sigma(x)^2 \frac{d}{dx} + b(x) \frac{d}{dx},$$

Döblin considers the equation

$$X_t = x + \beta \int_0^t \sigma(X_s)^2 ds + \int_0^t b(X_s) ds.$$

Observe that this equation bypasses the theory of stochastic integrals that was not available at that time but is of course equivalent to Itô's equation when

keeping in mind the Dambins, Dubins-Schwarz theorem. Concerning the path-wise construction of diffusion processes, we also have to mention the martingale problem approach by Stroock and Varadhan [?]. Stochastic differential equations with respect to general semimartingales may also be considered and we refer to the book by Protter [?] for an account of the theory. The theory of regularity of stochastic flows is difficult and the classical reference for these problems is the book by Kunita [?]. The rough paths theory presented in Chapter 7 simplifies several proofs of the regularity of flows (see the book by Friz and Victoir [?]).

**Sections 5,6.** Malliavin calculus was first introduced by Malliavin in [?] and later developed by Shigekawa [?]. The motivation was to give a probabilistic proof of Hörmander's theorem. The nowadays classical reference to study Malliavin calculus is the book by Nualart [?]. For an alternative and more geometric approach, we refer to the book by Bell [?], to the book by Malliavin himself [?] or the book by Shigekawa [?]. For a more specialized reading, we refer the interested reader to the lecture notes [?] and the very influential papers by Kusuoka and Stroock [?], [?] and [?] for further details and applications of Malliavin calculus to stochastic differential equations and partial differential equations.

## Chapter 7

# An introduction to Lyons rough paths theory

This chapter is an introduction to theory of rough paths that was conceived by T. Lyons in the 1990's. The theory is purely deterministic and offers a convenient alternative to Itô's calculus to define and study differential equations driven by Brownian motions. We focus on the main ideas and prove the Lyon's continuity result in the linear case and then proves that rough paths theory can be applied to study stochastic differential equations.

### 1 Continuous paths with finite $p$ -variation

If  $s \leq t$ , as before, we will denote by  $\Delta[s, t]$ , the set of subdivisions of the interval  $[s, t]$ , that is  $\Pi \in \Delta[s, t]$  can be written

$$\Pi = \{s = t_0 < t_1 < \dots < t_n = t\}.$$

The following definition is basic.

**Definition 7.1.** A continuous path  $x : [s, t] \rightarrow \mathbb{R}^d$  is said to have a bounded variation on  $[s, t]$ , if the 1-variation of  $x$  on  $[s, t]$ , which is defined as

$$\|x\|_{1-var;[s,t]} := \sup_{\Pi \in \Delta[s,t]} \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\|,$$

is finite. The space of continuous bounded variation paths  $x : [s, t] \rightarrow \mathbb{R}^d$ , will be denoted by  $C^{1-var}([s, t], \mathbb{R}^d)$ .

Of course  $\|\cdot\|_{1-var;[s,t]}$  is not a norm, because constant functions have a zero 1-variation, but it is obviously a semi-norm. If  $x$  is continuously differentiable on  $[s, t]$ , it is easily seen that

$$\|x\|_{1-var;[s,t]} = \int_s^t \|x'(s)\| ds.$$

**Proposition 7.2.** Let  $x \in C^{1-var}([0, T], \mathbb{R}^d)$ . The function  $(s, t) \rightarrow \|x\|_{1-var;[s,t]}$  is additive, i.e for  $0 \leq s \leq t \leq u \leq T$ ,

$$\|x\|_{1-var;[s,t]} + \|x\|_{1-var;[t,u]} = \|x\|_{1-var;[s,u]},$$

and controls  $x$  in the sense that for  $0 \leq s \leq t \leq T$ ,

$$\|x(s) - x(t)\| \leq \|x\|_{1-var;[s,t]}.$$

The function  $s \rightarrow \|x\|_{1-var;[0,s]}$  is moreover continuous and non decreasing.

*Proof.* If  $\Pi_1 \in \Delta[s, t]$  and  $\Pi_2 \in \Delta[t, u]$ , then  $\Pi_1 \cup \Pi_2 \in \Delta[s, u]$ . As a consequence, we obtain

$$\begin{aligned} & \sup_{\Pi_1 \in \Delta[s, t]} \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\| + \sup_{\Pi_2 \in \Delta[t, u]} \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\| \\ & \leq \sup_{\Pi \in \Delta[s, u]} \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\|, \end{aligned}$$

thus

$$\|x\|_{1\text{-var}, [s, t]} + \|x\|_{1\text{-var}, [t, u]} \leq \|x\|_{1\text{-var}, [s, u]}.$$

Let now  $\Pi \in \Delta[s, u]$ :

$$\Pi = \{s = t_0 < t_1 < \cdots < t_n = t\}.$$

Let  $k = \max\{j, t_j \leq t\}$ . By the triangle inequality, we have

$$\begin{aligned} \sum_{j=0}^{n-1} \|x(t_{j+1}) - x(t_j)\| & \leq \sum_{j=0}^{k-1} \|x(t_{j+1}) - x(t_j)\| + \sum_{j=k}^{n-1} \|x(t_{j+1}) - x(t_j)\| \\ & \leq \|x\|_{1\text{-var}, [s, t]} + \|x\|_{1\text{-var}, [t, u]}. \end{aligned}$$

Taking the sup of  $\Pi \in \Delta[s, u]$  gives

$$\|x\|_{1\text{-var}, [s, t]} + \|x\|_{1\text{-var}, [t, u]} \geq \|x\|_{1\text{-var}, [s, u]},$$

which completes the proof. The proof of the continuity and monotonicity of  $s \rightarrow \|x\|_{1\text{-var}, [0, s]}$  is left to the reader.  $\square$

This control of the path by the 1-variation norm is an illustration of the notion of controlled path which is central in rough paths theory.

**Definition 7.3.** A map  $\omega : \{0 \leq s \leq t \leq T\} \rightarrow [0, \infty)$  is called superadditive if for all  $s \leq t \leq u$ ,

$$\omega(s, t) + \omega(t, u) \leq \omega(s, u).$$

If, in addition,  $\omega$  is continuous and  $\omega(t, t) = 0$ , we call  $\omega$  a control. We say that a path  $x : [0, T] \rightarrow \mathbb{R}$  is controlled by a control  $\omega$ , if there exists a constant  $C > 0$ , such that for every  $0 \leq s \leq t \leq T$ ,

$$\|x(t) - x(s)\| \leq C\omega(s, t).$$

Obviously, Lipschitz functions have a bounded variation. The converse is not true:  $t \rightarrow \sqrt{t}$  has a bounded variation on  $[0, 1]$  but is not Lipschitz. However, any continuous path with bounded variation is the reparametrization of a Lipschitz path in the following sense.

**Proposition 7.4.** *Let  $x \in C^{1-var}([0, T], \mathbb{R}^d)$ . There exist a Lipschitz function  $y : [0, 1] \rightarrow \mathbb{R}^d$ , and a continuous and non-decreasing function  $\phi : [0, T] \rightarrow [0, 1]$  such that  $x = y \circ \phi$ .*

*Proof.* We assume  $\|x\|_{1-var, [0, T]} \neq 0$  and consider

$$\phi(t) = \frac{\|x\|_{1-var, [0, t]}}{\|x\|_{1-var, [0, T]}}.$$

It is continuous and non decreasing. There exists a function  $y$  such that  $x = y \circ \phi$  because  $\phi(t_1) = \phi(t_2)$  implies  $x(t_1) = x(t_2)$ . We have then, for  $s \leq t$ ,

$$\|y(\phi(t)) - y(\phi(s))\| = \|x(t) - x(s)\| \leq \|x\|_{1-var, [s, t]} = \|x\|_{1-var, [0, T]}(\phi(t) - \phi(s)).$$

□

The next result shows that the set of continuous paths with bounded variation is a Banach space.

**Theorem 7.5.** *The space  $C^{1-var}([0, T], \mathbb{R}^d)$  endowed with the norm  $\|x(0)\| + \|x\|_{1-var, [0, T]}$  is a Banach space.*

*Proof.* Let  $x^n \in C^{1-var}([0, T], \mathbb{R}^d)$  be a Cauchy sequence. It is clear that

$$\|x^n - x^m\|_\infty \leq \|x^n(0) - x^m(0)\| + \|x^n - x^m\|_{1-var, [0, T]}.$$

Thus,  $x^n$  converges uniformly to a continuous path  $x : [0, T] \rightarrow \mathbb{R}$ . We need to prove that  $x$  has a bounded variation. Let

$$\Pi = \{0 = t_0 < t_1 < \dots < t_n = T\}$$

be a a subdivision of  $[0, T]$ . There is  $m \geq 0$ , such that  $\|x - x^m\|_\infty \leq \frac{1}{2n}$ , thus

$$\begin{aligned} & \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\| \\ & \leq \sum_{k=0}^{n-1} \|x(t_{k+1}) - x^m(t_{k+1})\| + \sum_{k=0}^{n-1} \|x^m(t_k) - x(t_k)\| + \|x^m\|_{1-var, [0, T]} \\ & \leq 1 + \sup_n \|x^n\|_{1-var, [0, T]}. \end{aligned}$$

Thus, we have

$$\|x\|_{1-var, [0, T]} \leq 1 + \sup_n \|x^n\|_{1-var, [0, T]} < \infty.$$

□

More generally, we shall be interested in paths that have a finite  $p$ -variation,  $p > 0$ .

**Definition 7.6.** Let  $s \leq t$ . A path  $x : [s, t] \rightarrow \mathbb{R}^d$  is said to be of finite  $p$ -variation,  $p > 0$ , if the  $p$ -variation of  $x$  on  $[s, t]$ , which is defined as

$$\|x\|_{p\text{-var};[s,t]} := \left( \sup_{\Pi \in \Delta[s,t]} \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\|^p \right)^{1/p},$$

is finite. The space of continuous paths  $x : [s, t] \rightarrow \mathbb{R}^d$  with a finite  $p$ -variation will be denoted by  $C^{p\text{-var}}([s, t], \mathbb{R}^d)$ .

**Exercise 7.7.** Show that if  $x : [s, t] \rightarrow \mathbb{R}^d$  is an  $\alpha$ -Hölder path, then it has finite  $1/\alpha$ -variation.

The notion of  $p$ -variation is only interesting when  $p \geq 1$ .

**Proposition 7.8.** Let  $x : [s, t] \rightarrow \mathbb{R}^d$  be a continuous path of finite  $p$ -variation with  $p < 1$ . Then,  $x$  is constant.

*Proof.* We have for  $s \leq u \leq t$ ,

$$\begin{aligned} \|x(u) - x(s)\| &\leq (\max \|x(t_{k+1}) - x(t_k)\|^{1-p}) \left( \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\|^p \right) \\ &\leq (\max \|x(t_{k+1}) - x(t_k)\|^{1-p}) \|x\|_{p\text{-var};[s,t]}^p. \end{aligned}$$

Since  $x$  is continuous, it is also uniformly continuous on  $[s, t]$ . By taking a sequence of subdivisions whose mesh tends to 0, we deduce then that

$$\|x(u) - x(s)\| = 0,$$

so that  $x$  is constant. □

The following proposition is immediate and its proof is left to the reader:

**Proposition 7.9.** Let  $x : [s, t] \rightarrow \mathbb{R}^d$  be a continuous path. If  $p \leq p'$  then

$$\|x\|_{p'\text{-var};[s,t]} \leq \|x\|_{p\text{-var};[s,t]}.$$

As a consequence the following inclusion holds

$$C^{p\text{-var}}([s, t], \mathbb{R}^d) \subset C^{p'\text{-var}}([s, t], \mathbb{R}^d)$$

The following proposition generalizes Proposition 7.2 to paths with finite  $p$ -variation.

**Proposition 7.10.** *Let  $x \in C^{p\text{-var}}([0, T], \mathbb{R}^d)$ . Then  $\omega(s, t) = \|x\|_{p\text{-var};[s,t]}^p$  is a control such that for every  $s \leq t$ ,*

$$\|x(s) - x(t)\| \leq \omega(s, t)^{1/p}.$$

*Proof.* It is immediate that

$$\|x(s) - x(t)\| \leq \omega(s, t)^{1/p},$$

so we focus on the proof that  $\omega$  is a control. If  $\Pi_1 \in \Delta[s, t]$  and  $\Pi_2 \in \Delta[t, u]$ , then  $\Pi_1 \cup \Pi_2 \in \Delta[s, u]$ . As a consequence, we obtain

$$\begin{aligned} & \sup_{\Pi_1 \in \Delta[s,t]} \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\|^p + \sup_{\Pi_2 \in \Delta[t,u]} \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\|^p \\ & \leq \sup_{\Pi \in \Delta[s,u]} \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\|^p, \end{aligned}$$

thus

$$\|x\|_{p\text{-var};[s,t]}^p + \|x\|_{p\text{-var};[t,u]}^p \leq \|x\|_{p\text{-var};[s,u]}^p.$$

The proof of the continuity is left to the reader.  $\square$

In the following sense,  $\|x\|_{p\text{-var};[s,t]}^p$  is the minimal control of a path  $x \in C^{p\text{-var}}([0, T], \mathbb{R}^d)$ .

**Proposition 7.11.** *Let  $x \in C^{p\text{-var}}([0, T], \mathbb{R}^d)$  and let  $\omega : \{0 \leq s \leq t \leq T\} \rightarrow [0, \infty)$  be a control such that for  $0 \leq s \leq t \leq T$ ,*

$$\|x(s) - x(t)\| \leq C\omega(s, t)^{1/p},$$

*then*

$$\|x\|_{p\text{-var};[s,t]} \leq C\omega(s, t)^{1/p}.$$

*Proof.* We have

$$\begin{aligned} \|x\|_{p\text{-var};[s,t]} &= \left( \sup_{\Pi \in \Delta[s,t]} \sum_{k=0}^{n-1} \|x(t_{k+1}) - x(t_k)\|^p \right)^{1/p} \\ &\leq \left( \sup_{\Pi \in \Delta[s,t]} \sum_{k=0}^{n-1} C^p \omega(t_k, t_{k+1}) \right)^{1/p} \\ &\leq C\omega(s, t). \end{aligned}$$

$\square$

The next result shows that the set of continuous paths with bounded  $p$ -variation is a Banach space.

**Theorem 7.12.** *Let  $p \geq 1$ . The space  $C^{p\text{-var}}([0, T], \mathbb{R}^d)$  endowed with the norm  $\|x(0)\| + \|x\|_{p\text{-var};[0,T]}$  is a Banach space.*

*Proof.* The proof is identical to the case  $p = 1$ .  $\square$



## 2 The signature of a bounded variation path

In this section we introduce the notion of the signature of a path  $x \in C^{1-var}([0, T], \mathbb{R}^d)$  which is a convenient way to encode all the algebraic information on the path  $x$  which is relevant to study differential equations driven by  $x$ . The motivation for the definition of the signature comes from formal manipulations on Taylor series.

Let us consider the ordinary differential equation

$$y(t) = y(0) + \sum_{i=1}^d \int_0^t V_i(y(u)) dx^i(u),$$

where the  $V_i$ 's are smooth vector fields on  $\mathbb{R}^n$  (see Section 5 in Chapter 6 for the definition of a vector field). If  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a  $C^\infty$  function, by the change of variable formula,

$$f(y(t)) = f(y(s)) + \sum_{i=1}^d \int_s^t V_i f(y(u)) dx^i(u).$$

Now, a new application of the change of variable formula to  $V_i f(y(s))$  leads to

$$f(y(t)) = f(y(s)) + \sum_{i=1}^d V_i f(y(s)) \int_s^t dx^i(u) + \sum_{i,j=1}^d \int_s^t \int_0^u V_j V_i f(y(v)) dx^j(v) dx^i(u).$$

We can continue this procedure to get after  $N$  steps

$$f(y(t)) = f(y(s)) + \sum_{k=1}^N \sum_{I=(i_1, \dots, i_k)} (V_{i_1} \cdots V_{i_k} f)(y(s)) \int_{\Delta^k[s,t]} dx^I + R_N(s, t)$$

for some remainder term  $R_N(s, t)$ , where we used the notations:

1.  $\Delta^k[s, t] = \{(t_1, \dots, t_k) \in [s, t]^k, s \leq t_1 \leq t_2 \leq \dots \leq t_k \leq t\}$
2. If  $I = (i_1, \dots, i_k) \in \{1, \dots, d\}^k$  is a word with length  $k$ ,

$$\int_{\Delta^k[s,t]} dx^I = \int_{s \leq t_1 \leq t_2 \leq \dots \leq t_k \leq t} dx^{i_1}(t_1) \cdots dx^{i_k}(t_k).$$

If we let  $N \rightarrow +\infty$ , assuming  $R_N(s, t) \rightarrow 0$  (which is by the way true for  $t - s$  small enough if the  $V_i$ 's are analytic), we are led to the formal expansion formula:

$$f(y(t)) = f(y(s)) + \sum_{k=1}^{+\infty} \sum_{I=(i_1, \dots, i_k)} (V_{i_1} \cdots V_{i_k} f)(y(s)) \int_{\Delta^k[s,t]} dx^I.$$

This shows, at least at the formal level, that all the information given by  $x$  about  $y$  is contained in the set of iterated integrals  $\int_{\Delta^k[s,t]} dx^I$ .

Let  $\mathbb{R}[[X_1, \dots, X_d]]$  be the non commutative algebra over  $\mathbb{R}$  of the formal series with  $d$  indeterminates, that is the set of series

$$Y = y_0 + \sum_{k=1}^{+\infty} \sum_{I \in \{1, \dots, d\}^k} a_{i_1, \dots, i_k} X_{i_1} \dots X_{i_k}.$$

**Definition 7.13.** Let  $x \in C^{1-var}([0, T], \mathbb{R}^d)$ . The signature of  $x$  (or Chen's series) is the formal series:

$$\mathfrak{S}(x)_{s,t} = 1 + \sum_{k=1}^{+\infty} \sum_{I \in \{1, \dots, d\}^k} \left( \int_{\Delta^k[s,t]} dx^I \right) X_{i_1} \dots X_{i_k}, \quad 0 \leq s \leq t \leq T.$$

As we are going to see, the signature is a fascinating algebraic object. At the source of the numerous properties of the signature lie the following so-called Chen's relations

**Lemma 7.14** (Chen's relations). *Let  $x \in C^{1-var}([0, T], \mathbb{R}^d)$ . For any word  $(i_1, \dots, i_n) \in \{1, \dots, d\}^n$  and any  $0 \leq s \leq t \leq u \leq T$ ,*

$$\int_{\Delta^n[s,u]} dx^{(i_1, \dots, i_n)} = \sum_{k=0}^n \int_{\Delta^k[s,t]} dx^{(i_1, \dots, i_k)} \int_{\Delta^{n-k}[t,u]} dx^{(i_{k+1}, \dots, i_n)},$$

where we used the convention that if  $I$  is a word with length 0, then  $\int_{\Delta^0[0,t]} dx^I = 1$ .

*Proof.* It follows readily by induction on  $n$  by noticing that

$$\int_{\Delta^n[s,u]} dx^{(i_1, \dots, i_n)} = \int_s^u \left( \int_{\Delta^{n-1}[s,t_n]} dx^{(i_1, \dots, i_{n-1})} \right) dx^{i_n}(t_n).$$

□

To avoid heavy and cumbersome notations, it will be convenient to denote

$$\int_{\Delta^k[s,t]} dx^{\otimes k} = \sum_{I \in \{1, \dots, d\}^k} \left( \int_{\Delta^k[s,t]} dx^I \right) X_{i_1} \dots X_{i_k}.$$

This notation actually reflects a natural algebra isomorphism between  $\mathbb{R}[[X_1, \dots, X_d]]$  and  $1 \oplus_{k=1}^{+\infty} (\mathbb{R}^d)^{\otimes k}$ . With this notation, observe that the signature writes then

$$\mathfrak{S}(x)_{s,t} = 1 + \sum_{k=1}^{+\infty} \int_{\Delta^k[s,t]} dx^{\otimes k},$$

and that the Chen's relations become

$$\int_{\Delta^n[s,u]} dx^{\otimes n} = \sum_{k=0}^n \int_{\Delta^k[s,t]} dx^{\otimes k} \int_{\Delta^{n-k}[t,u]} dx^{\otimes(n-k)}.$$

The Chen's relations imply the following flow property for the signature:

**Lemma 7.15.** *Let  $x \in C^{1-var}([0, T], \mathbb{R}^d)$ . For any  $0 \leq s \leq t \leq u \leq T$ ,*

$$\mathfrak{S}(x)_{s,u} = \mathfrak{S}(x)_{s,t} \mathfrak{S}(x)_{t,u}$$

*Proof.* Indeed, from the Chen's relations, we have

$$\begin{aligned} \mathfrak{S}(x)_{s,u} &= 1 + \sum_{k=1}^{+\infty} \int_{\Delta^k[s,u]} dx^{\otimes k} \\ &= 1 + \sum_{k=1}^{+\infty} \sum_{j=0}^n \int_{\Delta^j[s,t]} dx^{\otimes j} \int_{\Delta^{k-j}[t,u]} dx^{\otimes(k-j)} \\ &= \mathfrak{S}(x)_{s,t} \mathfrak{S}(x)_{t,u}. \end{aligned}$$

□

### 3 Estimating iterated integrals

In the previous section we introduced the signature of a bounded variation path  $x$  as the formal series

$$\mathfrak{S}(x)_{s,t} = 1 + \sum_{k=1}^{+\infty} \int_{\Delta^k[s,t]} dx^{\otimes k}.$$

If now  $x \in C^{p-var}([0, T], \mathbb{R}^d)$ ,  $p \geq 1$  the iterated integrals  $\int_{\Delta^k[s,t]} dx^{\otimes k}$  can only be defined as Riemann-Stieltjes integrals when  $p = 1$  or Young integrals when  $p < 2$ . As we are going to see, it is actually possible to define the signature of paths with a finite  $p$  variation even when  $p \geq 2$ . For  $P \in \mathbb{R}[[X_1, \dots, X_d]]$  which can be written as

$$P = P_0 + \sum_{k=1}^{+\infty} \sum_{I \in \{1, \dots, d\}^k} a_{i_1, \dots, i_k} X_{i_1} \dots X_{i_k},$$

we define

$$\|P\| = |P_0| + \sum_{k=1}^{+\infty} \sum_{I \in \{1, \dots, d\}^k} |a_{i_1, \dots, i_k}| \in [0, \infty].$$

It is quite easy to check that for  $P, Q \in \mathbb{R}[[X_1, \dots, X_d]]$

$$\|PQ\| \leq \|P\| \|Q\|.$$

Let  $x \in C^{1-var}([0, T], \mathbb{R}^d)$ . For  $p \geq 1$ , we denote

$$\left\| \int dx^{\otimes k} \right\|_{p-var, [s, t]} = \left( \sup_{\Pi \in \mathcal{D}[s, t]} \sum_{i=0}^{n-1} \left\| \int_{\Delta^k [t_i, t_{i+1}]} dx^{\otimes k} \right\|^p \right)^{1/p},$$

where  $\mathcal{D}[s, t]$  is the set of subdivisions of the interval  $[s, t]$ . Observe that for  $k \geq 2$ , in general

$$\int_{\Delta^k [s, t]} dx^{\otimes k} + \int_{\Delta^k [t, u]} dx^{\otimes k} \neq \int_{\Delta^k [s, u]} dx^{\otimes k}.$$

Actually from the Chen's relations we have

$$\int_{\Delta^n [s, u]} dx^{\otimes n} = \int_{\Delta^n [s, t]} dx^{\otimes n} + \int_{\Delta^n [t, u]} dx^{\otimes n} + \sum_{k=1}^{n-1} \int_{\Delta^k [s, t]} dx^{\otimes k} \int_{\Delta^{n-k} [t, u]} dx^{\otimes (n-k)}.$$

It follows that  $\left\| \int dx^{\otimes k} \right\|_{p-var, [s, t]}$  needs not to be the  $p$ -variation of the path  $t \rightarrow \int_{\Delta^k [s, t]} dx^{\otimes k}$ . It is however easy to verify that

$$\left\| \int_{\Delta^k [s, \cdot]} dx^{\otimes k} \right\|_{p-var, [s, t]} \leq \left\| \int dx^{\otimes k} \right\|_{p-var, [s, t]}.$$

The first major result of rough paths theory is the following estimate:

**Theorem 7.16.** *Let  $p \geq 1$ . There exists a constant  $C \geq 0$ , depending only on  $p$ , such that for every  $x \in C^{1-var}([0, T], \mathbb{R}^d)$  and  $k \geq 0$ ,*

$$\left\| \int_{\Delta^k [s, t]} dx^{\otimes k} \right\| \leq \frac{C^k}{\left(\frac{k}{p}\right)!} \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}-var, [s, t]}^{1/j} \right)^k, \quad 0 \leq s \leq t \leq T.$$

By  $\left(\frac{k}{p}\right)!$ , we of course mean  $\Gamma\left(\frac{k}{p} + 1\right)$ , where  $\Gamma$  is the Euler function. Some remarks are in order before we prove the result. If  $p = 1$ , then the estimate becomes

$$\left\| \int_{\Delta^k [s, t]} dx^{\otimes k} \right\| \leq \frac{C^k}{k!} \|x\|_{1-var, [s, t]}^k,$$

which is immediately checked because

$$\begin{aligned} \left\| \int_{\Delta^k [s, t]} dx^{\otimes k} \right\| &\leq \sum_{I \in \{1, \dots, d\}^k} \left\| \int_{\Delta^k [s, t]} dx^I \right\| \\ &\leq \sum_{I \in \{1, \dots, d\}^k} \int_{s \leq t_1 \leq t_2 \leq \dots \leq t_k \leq t} d\|x^{i_1}\|_{1-var, [s, t_1]} \cdots d\|x^{i_k}\|_{1-var, [s, t_k]} \\ &\leq \frac{1}{k!} \|x\|_{1-var, [s, t]}^k. \end{aligned}$$

We can also observe that for  $k \leq p$ , the estimate is easy to obtain because

$$\left\| \int_{\Delta^k[s,t]} dx^{\otimes k} \right\| \leq \left\| \int dx^{\otimes k} \right\|_{\frac{p}{k}\text{-var},[s,t]}.$$

So, all the work is to prove the estimate when  $k > p$ . The proof is split into two lemmas. The first one is the following binomial inequality which is proved as an exercise.

**Lemma 7.17** (Binomial inequality). *For  $x, y > 0$ ,  $n \in \mathbb{N}$ ,  $n \geq 0$ , and  $p \geq 1$ ,*

$$\sum_{j=0}^n \frac{x^{j/p} y^{(n-j)/p}}{\binom{j}{p}! \binom{n-j}{p}!} \leq p \frac{(x+y)^{n/p}}{\binom{n}{p}!}.$$

*Proof.* See Exercise 7.18. □

**Exercise 7.18.** *Let  $D$  be the open unit disc in  $\mathbb{C}$ ,  $D = \{z \in \mathbb{C}, |z| < 1\}$ . If  $f$  is a continuous on  $\bar{D}$  and holomorphic on  $D$  then we denote for  $\xi \in \mathbb{R}$ ,*

$$\hat{f}(\xi) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(e^{2i\pi x}) e^{-2i\pi x \xi} dx.$$

1. *By writing  $\hat{f}(\xi)$  as an integral over the unit circle and using the residue theorem, show that for  $0 < \alpha < 2$  and  $0 < \lambda < 1$  the following identities hold*

$$\alpha \sum_{j=0}^{\infty} \hat{f}(\alpha j) \lambda^{\alpha j} = f(\lambda) - \frac{\alpha \lambda^{\alpha} \sin(\alpha \pi)}{\pi} \int_0^1 \frac{t^{\alpha-1} f(-t)}{|t^{\alpha} - \lambda^{\alpha} e^{-i\alpha \pi}|^2} dt$$

$$\alpha \sum_{j=-\infty}^{-1} \hat{f}(\alpha j) \lambda^{\alpha j} = \frac{\alpha \lambda^{\alpha} \sin(\alpha \pi)}{\pi} \int_0^1 \frac{t^{\alpha-1} f(-t)}{|e^{-i\alpha \pi} - \lambda^{\alpha} t^{\alpha}|^2} dt.$$

2. *Compute  $\hat{f}$  when  $f(z) = (1+z)^T$ ,  $T > 0$ .*
3. *Show that for  $\alpha \in (0, 2)$ ,  $n \in \mathbb{N}$  and  $0 < \lambda \leq 1$  we have*

$$\begin{aligned} & \alpha \sum_{j=0}^n \binom{\alpha n}{\alpha j} \lambda^{\alpha j} \\ &= (1+\lambda)^{\alpha n} - \frac{\alpha \lambda^{\alpha} \sin(\alpha \pi)}{\pi} \int_0^1 t^{\alpha-1} (1-t)^{\alpha n} \left( \frac{1}{|t^{\alpha} - \lambda^{\alpha} e^{-i\alpha \pi}|^2} + \frac{1}{|e^{-i\alpha \pi} - \lambda^{\alpha} t^{\alpha}|^2} \right) dt, \end{aligned}$$

where  $\binom{\alpha n}{\alpha j} = \frac{(\alpha n)!}{(\alpha j)!((n-j)\alpha)!}$ .

4. *Give the proof of Lemma 7.17.*

The second lemma we need is the following result.

**Lemma 7.19.** *Let  $\Gamma : \{0 \leq s \leq t \leq T\} \rightarrow \mathbb{R}^N$ . Let us assume that:*

1. *There exists a control  $\tilde{\omega}$  such that*

$$\lim_{r \rightarrow 0} \sup_{(s,t), \tilde{\omega}(s,t) \leq r} \frac{\|\Gamma_{s,t}\|}{r} = 0;$$

2. *There exists a control  $\omega$  and  $\theta > 1, \xi > 0$  such that for  $0 \leq s \leq t \leq u \leq T$ ,*

$$\|\Gamma_{s,u}\| \leq \|\Gamma_{s,t}\| + \|\Gamma_{t,u}\| + \xi\omega(s,u)^\theta.$$

Then, for all  $0 \leq s < t \leq T$ ,

$$\|\Gamma_{s,t}\| \leq \frac{\xi}{1 - 2^{1-\theta}} \omega(s,t)^\theta.$$

*Proof.* For  $\varepsilon > 0$ , let us consider the control

$$\omega_\varepsilon(s,t) = \omega(s,t) + \varepsilon\tilde{\omega}(s,t).$$

Define now

$$\Psi(r) = \sup_{s,u, \omega_\varepsilon(s,u) \leq r} \|\Gamma_{s,u}\|.$$

If  $s, u$  is such that  $\omega_\varepsilon(s,u) \leq r$ , we can find a  $t$  such that  $\omega_\varepsilon(s,t) \leq \frac{1}{2}\omega_\varepsilon(s,u)$ ,  $\omega_\varepsilon(t,u) \leq \frac{1}{2}\omega_\varepsilon(s,u)$ . Indeed, the continuity of  $\omega_\varepsilon$  forces the existence of a  $t$  such that  $\omega_\varepsilon(s,t) = \omega_\varepsilon(t,u)$ . We obtain therefore

$$\|\Gamma_{s,u}\| \leq 2\Psi(r/2) + \xi r^\theta,$$

which implies by maximization,

$$\Psi(r) \leq 2\Psi(r/2) + \xi r^\theta.$$

By iterating  $n$  times this inequality, we obtain

$$\begin{aligned} \Psi(r) &\leq 2^n \Psi\left(\frac{r}{2^n}\right) + \xi \sum_{k=0}^{n-1} 2^{k(1-\theta)} r^\theta \\ &\leq 2^n \Psi\left(\frac{r}{2^n}\right) + \xi \frac{1}{1 - 2^{1-\theta}} r^\theta. \end{aligned}$$

It is moreover clear that

$$\lim_{n \rightarrow \infty} 2^n \Psi\left(\frac{r}{2^n}\right) = 0.$$

We conclude

$$\Psi(r) \leq \frac{\xi}{1 - 2^{1-\theta}} r^\theta$$

and thus

$$\|\Gamma_{s,u}\| \leq \frac{\xi}{1 - 2^{1-\theta}} \omega_\varepsilon(s,u)^\theta.$$

Sending  $\varepsilon \rightarrow 0$ , finishes the proof.  $\square$

We can now turn to the proof of Theorem 7.16.

*Proof.* Let us denote

$$\omega(s, t) = \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j} - \text{var}, [s, t]}^{1/j} \right)^p.$$

We claim that  $\omega$  is a control. Indeed for  $0 \leq s \leq t \leq u \leq T$ , we have from the reverse Minkowski inequality  $\|x + y\|_{1/p} \geq \|x\|_{1/p} + \|y\|_{1/p}$ ,

$$\begin{aligned} \omega(s, t) + \omega(t, u) &= \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j} - \text{var}, [s, t]}^{1/j} \right)^p + \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j} - \text{var}, [t, u]}^{1/j} \right)^p \\ &\leq \left( \sum_{j=1}^{[p]} \left( \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j} - \text{var}, [s, t]}^{p/j} + \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j} - \text{var}, [t, u]}^{p/j} \right)^{1/p} \right)^p \\ &\leq \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j} - \text{var}, [s, u]}^{1/j} \right)^p = \omega(s, u). \end{aligned}$$

It is clear that for some constant  $\beta > 0$  which is small enough, we have for  $k \leq p$ ,

$$\left\| \int_{\Delta^k [s, t]} dx^{\otimes k} \right\| \leq \frac{1}{\beta \left(\frac{k}{p}\right)!} \omega(s, t)^{k/p}.$$

Let us now consider

$$\Gamma_{s, t} = \int_{\Delta^{[p]+1} [s, t]} dx^{\otimes ([p]+1)}.$$

From the Chen's relations, we have for  $0 \leq s \leq t \leq u \leq T$ ,

$$\Gamma_{s, u} = \Gamma_{s, t} + \Gamma_{t, u} + \sum_{j=1}^{[p]} \int_{\Delta^j [s, t]} dx^{\otimes j} \int_{\Delta^{[p]+1-j} [t, u]} dx^{\otimes ([p]+1-j)}.$$

Therefore, we obtain

$$\begin{aligned}
\|\Gamma_{s,u}\| &\leq \|\Gamma_{s,t}\| + \|\Gamma_{t,u}\| + \sum_{j=1}^{[p]} \left\| \int_{\Delta^j[s,t]} dx^{\otimes j} \right\| \left\| \int_{\Delta^{[p]+1-j}[t,u]} dx^{\otimes ([p]+1-j)} \right\| \\
&\leq \|\Gamma_{s,t}\| + \|\Gamma_{t,u}\| + \frac{1}{\beta^2} \sum_{j=1}^{[p]} \frac{1}{\left(\frac{j}{p}\right)!} \omega(s,t)^{j/p} \frac{1}{\left(\frac{[p]+1-j}{p}\right)!} \omega(t,u)^{([p]+1-j)/p} \\
&\leq \|\Gamma_{s,t}\| + \|\Gamma_{t,u}\| + \frac{1}{\beta^2} \sum_{j=0}^{[p]+1} \frac{1}{\left(\frac{j}{p}\right)!} \omega(s,t)^{j/p} \frac{1}{\left(\frac{[p]+1-j}{p}\right)!} \omega(t,u)^{([p]+1-j)/p} \\
&\leq \|\Gamma_{s,t}\| + \|\Gamma_{t,u}\| + \frac{1}{\beta^2 p} \frac{(\omega(s,t) + \omega(t,u))^{([p]+1)/p}}{\left(\frac{[p]+1}{p}\right)!} \\
&\leq \|\Gamma_{s,t}\| + \|\Gamma_{t,u}\| + \frac{1}{\beta^2 p} \frac{\omega(s,u)^{([p]+1)/p}}{\left(\frac{[p]+1}{p}\right)!}.
\end{aligned}$$

On the other hand, for some constant  $A$ , we have

$$\|\Gamma_{s,t}\| \leq A \|x\|_{1-var,[s,t]}^{[p]+1}.$$

We deduce from the previous Lemma 7.19 that

$$\|\Gamma_{s,t}\| \leq \frac{1}{\beta^2} \frac{p}{1-2^{1-\theta}} \frac{\omega(s,t)^{([p]+1)/p}}{\left(\frac{[p]+1}{p}\right)!},$$

with  $\theta = \frac{[p]+1}{p}$ . The general case  $k \geq p$  is dealt by induction. The details are let to the reader.  $\square$

Let  $x \in C^{1-var}([0, T], \mathbb{R}^d)$ . Since

$$\omega(s,t) = \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}-var,[s,t]}^{1/j} \right)^p$$

is a control, the estimate

$$\left\| \int_{\Delta^k[s,t]} dx^{\otimes k} \right\| \leq \frac{C^k}{\left(\frac{k}{p}\right)!} \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}-var,[s,t]}^{1/j} \right)^k, \quad 0 \leq s \leq t \leq T,$$

easily implies that for  $k > p$ ,

$$\left\| \int dx^{\otimes k} \right\|_{1-var,[s,t]} \leq \frac{C^k}{\left(\frac{k}{p}\right)!} \omega(s,t)^{k/p}.$$



We stress that it does not imply a bound on the 1-variation of the path  $t \rightarrow \int_{\Delta^k[0,t]} dx^{\otimes k}$ . What we can get for this path, are bounds in  $p$ -variation:

**Proposition 7.20.** *Let  $p \geq 1$ . There exists a constant  $C \geq 0$ , depending only on  $p$ , such that for every  $x \in C^{1-var}([0, T], \mathbb{R}^d)$  and  $k \geq 0$ ,*

$$\left\| \int_{\Delta^k[0,\cdot]} dx^{\otimes k} \right\|_{p-var,[s,t]} \leq \frac{C^k}{\left(\frac{k}{p}\right)!} \omega(s,t)^{1/p} \omega(0,T)^{\frac{k-1}{p}}$$

where

$$\omega(s,t) = \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}-var,[s,t]}^{1/j} \right)^p, \quad 0 \leq s \leq t \leq T.$$

*Proof.* This is an easy consequence of the Chen's relations and of the previous Theorem. Indeed,

$$\begin{aligned} & \left\| \int_{\Delta^k[0,t]} dx^{\otimes k} - \int_{\Delta^k[0,s]} dx^{\otimes k} \right\| \\ &= \left\| \sum_{j=1}^k \int_{\Delta^j[s,t]} dx^{\otimes j} \int_{\Delta^{k-j}[0,s]} dx^{\otimes(k-j)} \right\| \\ &\leq \sum_{j=1}^k \left\| \int_{\Delta^j[s,t]} dx^{\otimes j} \right\| \left\| \int_{\Delta^{k-j}[0,s]} dx^{\otimes(k-j)} \right\| \\ &\leq C^k \sum_{j=1}^k \frac{1}{\left(\frac{j}{p}\right)!} \omega(s,t)^{j/p} \frac{1}{\left(\frac{k-j}{p}\right)!} \omega(s,t)^{(k-j)/p} \\ &\leq C^k \omega(s,t)^{1/p} \sum_{j=1}^k \frac{1}{\left(\frac{j}{p}\right)!} \omega(0,T)^{(j-1)/p} \frac{1}{\left(\frac{k-j}{p}\right)!} \omega(0,T)^{(k-j)/p} \\ &\leq C^k \omega(s,t)^{1/p} \omega(0,T)^{(k-1)/p} \sum_{j=1}^k \frac{1}{\left(\frac{j}{p}\right)!} \frac{1}{\left(\frac{k-j}{p}\right)!}. \end{aligned}$$

and we conclude with the binomial inequality.  $\square$

We are now ready for a second major estimate which is the key to define iterated integrals of a path with  $p$ -bounded variation when  $p \geq 2$ .

**Theorem 7.21.** *Let  $p \geq 1$ ,  $K > 0$  and  $x, y \in C^{1-var}([0, T], \mathbb{R}^d)$  such that*

$$\sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}-var,[0,T]}^{1/j} \leq 1,$$

and

$$\left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[0,T]}^{1/j} \right)^p + \left( \sum_{j=1}^{[p]} \left\| \int dy^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[0,T]}^{1/j} \right)^p \leq K.$$

Then there exists a constant  $C \geq 0$  depending only on  $p$  and  $K$  such that for  $0 \leq s \leq t \leq T$  and  $k \geq 1$

$$\begin{aligned} & \left\| \int_{\Delta^k[s,t]} dx^{\otimes k} - \int_{\Delta^k[s,t]} dy^{\otimes k} \right\| \\ & \leq \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[0,T]}^{1/j} \right) \frac{C^k}{\left(\frac{k}{p}\right)!} \omega(s,t)^{k/p}, \end{aligned}$$

and

$$\left\| \int_{\Delta^k[s,t]} dx^{\otimes k} \right\| + \left\| \int_{\Delta^k[s,t]} dy^{\otimes k} \right\| \leq \frac{C^k}{\left(\frac{k}{p}\right)!} \omega(s,t)^{k/p}$$

where  $\omega$  is the control

$$\begin{aligned} \omega(s,t) &= \frac{\left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[s,t]}^{1/j} \right)^p + \left( \sum_{j=1}^{[p]} \left\| \int dy^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[s,t]}^{1/j} \right)^p}{\left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[0,T]}^{1/j} \right)^p + \left( \sum_{j=1}^{[p]} \left\| \int dy^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[0,T]}^{1/j} \right)^p} \\ &+ \left( \frac{\sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[s,t]}^{1/j}}{\sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[0,T]}^{1/j}} \right)^p. \end{aligned}$$

*Proof.* We prove by induction on  $k$  that for some constants  $C, \beta$ ,

$$\begin{aligned} & \left\| \int_{\Delta^k[s,t]} dx^{\otimes k} - \int_{\Delta^k[s,t]} dy^{\otimes k} \right\| \\ & \leq \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[0,T]}^{1/j} \right) \frac{C^k}{\beta \left(\frac{k}{p}\right)!} \omega(s,t)^{k/p}, \end{aligned}$$

and

$$\left\| \int_{\Delta^k[s,t]} dx^{\otimes k} \right\| + \left\| \int_{\Delta^k[s,t]} dy^{\otimes k} \right\| \leq \frac{C^k}{\beta \left(\frac{k}{p}\right)!} \omega(s,t)^{k/p}$$

For  $k \leq p$ , we trivially have

$$\begin{aligned} \left\| \int_{\Delta^k[s,t]} dx^{\otimes k} - \int_{\Delta^k[s,t]} dy^{\otimes k} \right\| &\leq \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[0,T]}^{1/j} \right)^k \omega(s,t)^{k/p} \\ &\leq \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}-\text{var},[0,T]}^{1/j} \right) \omega(s,t)^{k/p}, \end{aligned}$$

and

$$\left\| \int_{\Delta^k[s,t]} dx^{\otimes k} \right\| + \left\| \int_{\Delta^k[s,t]} dy^{\otimes k} \right\| \leq K^{k/p} \omega(s,t)^{k/p}.$$

Not let us assume that the result is true for  $0 \leq j \leq k$  with  $k > p$ . Let

$$\Gamma_{s,t} = \int_{\Delta^k[s,t]} dx^{\otimes(k+1)} - \int_{\Delta^k[s,t]} dy^{\otimes(k+1)}.$$

From the Chen's relations, we have for  $0 \leq s \leq t \leq u \leq T$ ,

$$\begin{aligned} \Gamma_{s,u} &= \Gamma_{s,t} + \Gamma_{t,u} + \sum_{j=1}^k \int_{\Delta^j[s,t]} dx^{\otimes j} \int_{\Delta^{k+1-j}[t,u]} dx^{\otimes(k+1-j)} \\ &\quad - \sum_{j=1}^k \int_{\Delta^j[s,t]} dy^{\otimes j} \int_{\Delta^{k+1-j}[t,u]} dy^{\otimes(k+1-j)}. \end{aligned}$$

Therefore, from the binomial inequality

$$\begin{aligned} \|\Gamma_{s,u}\| &\leq \|\Gamma_{s,t}\| + \|\Gamma_{t,u}\| + \sum_{j=1}^k \left\| \int_{\Delta^j[s,t]} dx^{\otimes j} - \int_{\Delta^j[s,t]} dy^{\otimes j} \right\| \left\| \int_{\Delta^{k+1-j}[t,u]} dx^{\otimes(k+1-j)} \right\| \\ &\quad + \sum_{j=1}^k \left\| \int_{\Delta^j[s,t]} dy^{\otimes j} \right\| \left\| \int_{\Delta^{k+1-j}[t,u]} dx^{\otimes(k+1-j)} - \int_{\Delta^{k+1-j}[t,u]} dy^{\otimes(k+1-j)} \right\| \\ &\leq \|\Gamma_{s,t}\| + \|\Gamma_{t,u}\| + \frac{1}{\beta^2} \tilde{\omega}(0,T) \sum_{j=1}^k \frac{C^j}{\left(\frac{j}{p}\right)!} \omega(s,t)^{j/p} \frac{C^{k+1-j}}{\left(\frac{k+1-j}{p}\right)!} \omega(t,u)^{(k+1-j)/p} \\ &\quad + \frac{1}{\beta^2} \tilde{\omega}(0,T) \sum_{j=1}^k \frac{C^j}{\left(\frac{j}{p}\right)!} \omega(s,t)^{j/p} \frac{C^{k+1-j}}{\left(\frac{k+1-j}{p}\right)!} \omega(t,u)^{(k+1-j)/p} \\ &\leq \|\Gamma_{s,t}\| + \|\Gamma_{t,u}\| + \frac{2p}{\beta^2} \tilde{\omega}(0,T) C^{k+1} \frac{\omega(s,u)^{(k+1)/p}}{\left(\frac{k+1}{p}\right)!} \end{aligned}$$

where

$$\tilde{\omega}(0, T) = \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}\text{-var}, [0, T]}^{1/j}.$$

We deduce

$$\|\Gamma_{s,t}\| \leq \frac{2p}{\beta^2(1-2^{1-\theta})} \tilde{\omega}(0, T) C^{k+1} \frac{\omega(s,t)^{(k+1)/p}}{\left(\frac{k+1}{p}\right)!}$$

with  $\theta = \frac{k+1}{p}$ . A correct choice of  $\beta$  finishes the induction argument.  $\square$

The following continuity of the iterated integrals with respect to a convenient topology can then be proved. The proof uses very similar arguments as in the previous proof, so we let it as an exercise to the reader.

**Theorem 7.22.** *Let  $p \geq 1$ ,  $K > 0$  and  $x, y \in C^{1\text{-var}}([0, T], \mathbb{R}^d)$  such that*

$$\sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}\text{-var}, [0, T]}^{1/j} \leq 1,$$

and

$$\left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}\text{-var}, [0, T]}^{1/j} \right)^p + \left( \sum_{j=1}^{[p]} \left\| \int dy^{\otimes j} \right\|_{\frac{p}{j}\text{-var}, [0, T]}^{1/j} \right)^p \leq K.$$

Then there exists a constant  $C \geq 0$  depending only on  $p$  and  $K$  such that for  $0 \leq s \leq t \leq T$  and  $k \geq 1$

$$\begin{aligned} & \left\| \int_{\Delta^k[0, \cdot]} dx^{\otimes k} - \int_{\Delta^k[0, \cdot]} dy^{\otimes k} \right\|_{p\text{-var}, [0, T]} \\ & \leq \frac{C^k}{\left(\frac{k}{p}\right)!} \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}\text{-var}, [0, T]}^{1/j} \right)^p. \end{aligned}$$

This continuity result naturally leads to the following definition.

**Definition 7.23.** Let  $p \geq 1$  and  $x \in C^{p\text{-var}}([0, T], \mathbb{R}^d)$ . We say that  $x$  is a  $p$ -rough path if there exists a sequence  $x_n \in C^{1\text{-var}}([0, T], \mathbb{R}^d)$  such that  $x_n \rightarrow x$  in  $p$ -variation and such that for every  $\varepsilon > 0$ , there exists  $N \geq 0$  such that for  $m, n \geq N$ ,

$$\sum_{j=1}^{[p]} \left\| \int dx_n^{\otimes j} - \int dx_m^{\otimes j} \right\|_{\frac{p}{j}\text{-var}, [0, T]}^{1/j} \leq \varepsilon.$$

The space of  $p$ -rough paths will be denoted  $\mathbf{\Omega}^p([0, T], \mathbb{R}^d)$ .

From the very definition,  $\mathbf{\Omega}^p([0, T], \mathbb{R}^d)$  is the closure of  $C^{1-var}([0, T], \mathbb{R}^d)$  inside  $C^{p-var}([0, T], \mathbb{R}^d)$  for the distance

$$d_{\mathbf{\Omega}^p([0, T], \mathbb{R}^d)}(x, y) = \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dy^{\otimes j} \right\|_{\frac{p}{j}-var, [0, T]}^{1/j}.$$

If  $x \in \mathbf{\Omega}^p([0, T], \mathbb{R}^d)$  and  $x_n \in C^{1-var}([0, T], \mathbb{R}^d)$  is such that  $x_n \rightarrow x$  in  $p$ -variation and such that for every  $\varepsilon > 0$ , there exists  $N \geq 0$  such that for  $m, n \geq N$ ,

$$\sum_{j=1}^{[p]} \left\| \int dx_n^{\otimes j} - \int dx_m^{\otimes j} \right\|_{\frac{p}{j}-var, [0, T]}^{1/j} \leq \varepsilon,$$

then we define  $\int_{\Delta^k[s, t]} dx^{\otimes k}$  for  $k \leq p$  as the limit of the iterated integrals  $\int_{\Delta^k[s, t]} dx_n^{\otimes k}$ . However it is important to observe that  $\int_{\Delta^k[s, t]} dx^{\otimes k}$  may then depend on the choice of the approximating sequence  $x_n$ . Once the integrals  $\int_{\Delta^k[s, t]} dx^{\otimes k}$  are defined for  $k \leq p$ , we can then use the previous theorem to construct all the iterated integrals  $\int_{\Delta^k[s, t]} dx^{\otimes k}$  are defined for  $k > p$ . It is then obvious that if  $x, y \in \mathbf{\Omega}^p([0, T], \mathbb{R}^d)$ , then

$$1 + \sum_{k=1}^{[p]} \int_{\Delta^k[s, t]} dx^{\otimes k} = 1 + \sum_{k=1}^{[p]} \int_{\Delta^k[s, t]} dy^{\otimes k}$$

implies that

$$1 + \sum_{k=1}^{+\infty} \int_{\Delta^k[s, t]} dx^{\otimes k} = 1 + \sum_{k=1}^{+\infty} \int_{\Delta^k[s, t]} dy^{\otimes k}.$$

In other words the signature of a  $p$ -rough path is completely characterized by its truncated signature at order  $[p]$ :

$$\mathfrak{S}_{[p]}(x)_{s, t} = 1 + \sum_{k=1}^{[p]} \int_{\Delta^k[s, t]} dx^{\otimes k}.$$

For this reason, it is natural to present a  $p$ -rough path by this truncated signature at order  $[p]$  in order to stress that the choice of the approximating sequence to construct the iterated integrals up to order  $[p]$  has been made. The following results are straightforward to obtain by a limiting argument.

**Lemma 7.24** (Chen's relations). *Let  $x \in \mathbf{\Omega}^p([0, T], \mathbb{R}^d)$ ,  $p \geq 1$ . For  $0 \leq s \leq t \leq u \leq T$ , and  $n \geq 1$ ,*

$$\int_{\Delta^n[s, u]} dx^{\otimes n} = \sum_{k=0}^n \int_{\Delta^k[s, t]} dx^{\otimes k} \int_{\Delta^{n-k}[t, u]} dx^{\otimes(n-k)}.$$

**Theorem 7.25.** *Let  $p \geq 1$ . There exists a constant  $C \geq 0$ , depending only on  $p$ , such that for every  $x \in \Omega^p([0, T], \mathbb{R}^d)$  and  $k \geq 1$ ,*

$$\left\| \int_{\Delta^k[s, t]} dx^{\otimes k} \right\| \leq \frac{C^k}{\left(\frac{k}{p}\right)!} \left( \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}\text{-var}, [s, t]}^{1/j} \right)^k, \quad 0 \leq s \leq t \leq T.$$

## 4 Rough differential equations

We now define solutions of differential equations driven by  $p$ -rough paths,  $p \geq 1$ . We study in details the case of linear equations for which we prove the Lyons continuity result and, at the end of the section, give, without proof, the full Lyons result.

Let  $x \in \Omega^p([0, T], \mathbb{R}^d)$  be a  $p$ -rough path with truncated signature

$$\sum_{k=0}^{[p]} \int_{\Delta^k[s, t]} dx^{\otimes k},$$

and let  $x_n \in C^{1\text{-var}}([0, T], \mathbb{R}^d)$  be an approximating sequence such that

$$\sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dx_n^{\otimes j} \right\|_{\frac{p}{j}\text{-var}, [0, T]}^{1/j} \rightarrow 0.$$

Let us consider matrices  $M_1, \dots, M_d \in \mathbb{R}^{n \times n}$ . We have the following theorem:

**Theorem 7.26** (Lyons' continuity theorem, linear case). *Let  $y_n : [0, T] \rightarrow \mathbb{R}^n$  be the solution of the differential equation*

$$y_n(t) = y(0) + \sum_{i=1}^d \int_0^t M_i y_n(s) dx_n^i(s).$$

*Then, when  $n \rightarrow \infty$ ,  $y_n$  converges in the  $p$ -variation distance to some  $y \in C^{p\text{-var}}([0, T], \mathbb{R}^n)$ . The path  $y$  is called the solution of the rough differential equation*

$$y(t) = y(0) + \sum_{i=1}^d \int_0^t M_i y(s) dx^i(s).$$

*Proof.* It is a classical result that the solution of the equation

$$y_n(t) = y(0) + \sum_{i=1}^d \int_0^t M_i y_n(s) dx_n^i(s),$$

can be expanded as the convergent Volterra series:

$$y_n(t) = y_n(s) + \sum_{k=1}^{+\infty} \sum_{I=(i_1, \dots, i_k)} M_{i_1} \cdots M_{i_k} \left( \int_{\Delta^k[s,t]} dx_n^I \right) y_n(s).$$

Therefore, in particular, for  $n, m \geq 0$ ,

$$y_n(t) - y_m(t) = \sum_{k=1}^{+\infty} \sum_{I=(i_1, \dots, i_k)} M_{i_1} \cdots M_{i_k} \left( \int_{\Delta^k[0,t]} dx_n^I - \int_{\Delta^k[0,t]} dx_m^I \right) y(0),$$

which implies that

$$\|y_n(t) - y_m(t)\| \leq \sum_{k=1}^{+\infty} M^k \left\| \int_{\Delta^k[0,t]} dx_n^{\otimes k} - \int_{\Delta^k[0,t]} dx_m^{\otimes k} \right\| \|y(0)\|$$

with  $M = \max\{\|M_1\|, \dots, \|M_d\|\}$ . From Theorem 7.22, there exists a constant  $C \geq 0$  depending only on  $p$  and

$$\sup_n \sum_{j=1}^{[p]} \left\| \int dx_n^{\otimes j} \right\|_{\frac{p}{j} - \text{var}, [0, T]}^{1/j}$$

such that for  $k \geq 1$  and  $n, m$  big enough:

$$\begin{aligned} & \left\| \int_{\Delta^k[0, \cdot]} dx_n^{\otimes k} - \int_{\Delta^k[0, \cdot]} dx_m^{\otimes k} \right\|_{p - \text{var}, [0, T]} \\ & \leq \left( \sum_{j=1}^{[p]} \left\| \int dx_n^{\otimes j} - \int dx_m^{\otimes j} \right\|_{\frac{p}{j} - \text{var}, [0, T]}^{1/j} \right) \frac{C^k}{\left(\frac{k}{p}\right)!}. \end{aligned}$$

As a consequence, there exists a constant  $\tilde{C}$  such that for  $n, m$  big enough:

$$\|y_n(t) - y_m(t)\| \leq \tilde{C} \sum_{j=1}^{[p]} \left\| \int dx_n^{\otimes j} - \int dx_m^{\otimes j} \right\|_{\frac{p}{j} - \text{var}, [0, T]}^{1/j}.$$

This already proves that  $y_n$  converges in the supremum topology to some  $y$ . We now have

$$\begin{aligned} & (y_n(t) - y_n(s)) - (y_m(t) - y_m(s)) \\ & = \sum_{k=1}^{+\infty} \sum_{I=(i_1, \dots, i_k)} M_{i_1} \cdots M_{i_k} \left( \int_{\Delta^k[s,t]} dx_n^I y_n(s) - \int_{\Delta^k[s,t]} dx_m^I y_m(s) \right), \end{aligned}$$

and we can bound

$$\begin{aligned} & \left\| \int_{\Delta^k[s,t]} dx_n^I y_n(s) - \int_{\Delta^k[s,t]} dx_m^I y_m(s) \right\| \\ & \leq \left\| \int_{\Delta^k[s,t]} dx_n^I \right\| \|y_n(s) - y_m(s)\| + \|y_m(s)\| \left\| \int_{\Delta^k[s,t]} dx_n^I - \int_{\Delta^k[s,t]} dx_m^I \right\| \\ & \leq \left\| \int_{\Delta^k[s,t]} dx_n^I \right\| \|y_n - y_m\|_{\infty,[0,T]} + \|y_m\|_{\infty,[0,T]} \left\| \int_{\Delta^k[s,t]} dx_n^I - \int_{\Delta^k[s,t]} dx_m^I \right\|. \end{aligned}$$

Now, there exists a constant  $C \geq 0$ , depending only on  $p$  and

$$\sup_n \sum_{j=1}^{[p]} \left\| \int dx_n^{\otimes j} \right\|_{\frac{p}{j}\text{-var},[0,T]}^{1/j}$$

such that for  $k \geq 1$  and  $n, m$  big enough

$$\begin{aligned} & \left\| \int_{\Delta^k[s,t]} dx_n^{\otimes k} \right\| \leq \frac{C^k}{\left(\frac{k}{p}\right)!} \omega(s, t)^{k/p}, \quad 0 \leq s \leq t \leq T, \\ & \left\| \int_{\Delta^k[s,t]} dx_n^{\otimes k} - \int_{\Delta^k[s,t]} dx_m^{\otimes k} \right\| \\ & \leq \left( \sum_{j=1}^{[p]} \left\| \int dx_n^{\otimes j} - \int dx_m^{\otimes j} \right\|_{\frac{p}{j}\text{-var},[0,T]}^{1/j} \right) \frac{C^k}{\left(\frac{k}{p}\right)!} \omega(s, t)^{k/p}, \end{aligned}$$

where  $\omega$  is a control such that  $\omega(0, T) = 1$ . Consequently, there is a constant  $\tilde{C}$ , such that

$$\begin{aligned} & \|(y_n(t) - y_n(s)) - (y_m(t) - y_m(s))\| \\ & \leq \tilde{C} \left( \|y_n - y_m\|_{\infty,[0,T]} + \sum_{j=1}^{[p]} \left\| \int dx_n^{\otimes j} - \int dx_m^{\otimes j} \right\|_{\frac{p}{j}\text{-var},[0,T]}^{1/j} \right) \omega(s, t)^{1/p} \end{aligned}$$

This implies the estimate

$$\begin{aligned} & \|y_n - y_m\|_{p\text{-var},[0,T]} \\ & \leq \tilde{C} \left( \|y_n - y_m\|_{\infty,[0,T]} + \sum_{j=1}^{[p]} \left\| \int dx_n^{\otimes j} - \int dx_m^{\otimes j} \right\|_{\frac{p}{j}\text{-var},[0,T]}^{1/j} \right) \end{aligned}$$

and thus gives the conclusion.  $\square$



**Exercise 7.27.** Let  $y_n : [0, T] \rightarrow \mathbb{R}^n$  be the solution of the differential equation

$$y_n(t) = y(0) + \sum_{i=1}^d \int_0^t M_i y_n(s) dx_n^i(s).$$

and  $y$  be the solution of the rough differential equation:

$$y(t) = y(0) + \sum_{i=1}^d \int_0^t M_i y(s) dx^i(s).$$

Show that  $y \in \Omega^p([0, T], \mathbb{R}^n)$  and that when  $n \rightarrow \infty$ ,

$$\sum_{j=1}^{[p]} \left\| \int dy^{\otimes j} - \int dy_n^{\otimes j} \right\|_{\frac{p}{j}\text{-var}, [0, T]}^{1/j} \rightarrow 0.$$

We can get useful estimates for solutions of rough differential equations. For that, we need the following real analysis lemma:

**Proposition 7.28.** For  $x \geq 0$  and  $p \geq 1$ ,

$$\sum_{k=0}^{+\infty} \frac{x^k}{\left(\frac{k}{p}\right)!} \leq p e^{x^p}.$$

*Proof.* For  $\alpha > 0$ , we denote

$$E_\alpha(x) = \sum_{k=0}^{+\infty} \frac{x^k}{(k\alpha)!}.$$

This is a special function called the Mittag-Leffler function. From the binomial inequality

$$\begin{aligned} E_\alpha(x)^2 &= \sum_{k=0}^{+\infty} \left( \sum_{j=0}^k \frac{1}{(j\alpha)! ((k-j)\alpha)!} \right) x^k \\ &\leq \frac{1}{\alpha} \sum_{k=0}^{+\infty} 2^{\alpha k} \frac{x^k}{(k\alpha)!} = \frac{1}{\alpha} E_\alpha(2^\alpha x). \end{aligned}$$

Thus we proved

$$E_\alpha(x) \leq \frac{1}{\alpha^{1/2}} E_\alpha(2^\alpha x)^{1/2}.$$

Iterating this inequality,  $k$  times we obtain

$$E_\alpha(x) \leq \frac{1}{\alpha^{\sum_{j=1}^k \frac{1}{2^j}}} E_\alpha(2^{\alpha k} x)^{1/2^k}.$$

It is known (and not difficult to prove, see Exercise 7.29) that

$$E_\alpha(x) \sim_{x \rightarrow \infty} \frac{1}{\alpha} e^{x^{1/\alpha}}.$$

By letting  $k \rightarrow \infty$  we conclude

$$E_\alpha(x) \leq \frac{1}{\alpha} e^{x^{1/\alpha}}.$$

□

**Exercise 7.29.** For  $\alpha \geq 0$ , we consider the Mittag-Leffler function

$$E_\alpha(x) = \sum_{k=0}^{+\infty} \frac{x^k}{(k\alpha)!}.$$

Show that

$$E_\alpha(x) \sim_{x \rightarrow \infty} \frac{1}{\alpha} e^{x^{1/\alpha}}.$$

**Hint:** Compute the Laplace transform of  $\sum \frac{x^{\alpha k}}{(k\alpha-1)!}$ .

This estimate provides the following result:

**Proposition 7.30.** Let  $y$  be the solution of the rough differential equation:

$$y(t) = y(0) + \sum_{i=1}^d \int_0^t M_i y(s) dx^i(s).$$

Then, there exists a constant  $C$  depending only on  $p$  such that for  $0 \leq t \leq T$ ,

$$\|y(t)\| \leq p \|y(0)\| e^{(CM)^p \left( \sum_{j=1}^{[p]} \|f dx^{\otimes j}\|_{\mathbb{P}^1_{-var, [0, t]}} \right)^p},$$

where  $M = \max\{\|M_1\|, \dots, \|M_d\|\}$ .

*Proof.* We have

$$y(t) = y(0) + \sum_{k=1}^{+\infty} \sum_{I=(i_1, \dots, i_k)} M_{i_1} \cdots M_{i_k} \left( \int_{\Delta^k[0, t]} dx^I \right) y(0).$$

Thus we obtain

$$\|y(t)\| \leq \left( 1 + \sum_{k=1}^{+\infty} \sum_{I=(i_1, \dots, i_k)} M^k \left\| \int_{\Delta^k[0, t]} dx^I \right\| \right) \|y(0)\|,$$

and we conclude by using estimates on iterated integrals of rough paths together (see Theorem 7.25) with the previous lemma. □

We now give the Lyons' continuity result for non linear differential equations. We first introduce the notion of  $\gamma$ -Lipschitz vector field.

**Definition 7.31.** A vector field  $V$  on  $\mathbb{R}^n$  is called  $\gamma$ -Lipschitz if it is  $[\gamma]$  times continuously differentiable and there exists a constant  $M \geq 0$  such that the supremum norm of its  $k$ th derivatives  $k = 0, \dots, [\gamma]$  and the  $\gamma - [\gamma]$  Hölder norm of its  $[\gamma]$ th derivative are bounded by  $M$ . The smallest  $M$  that satisfies the above condition is the  $\gamma$ -Lipschitz norm of  $V$  and will be denoted  $\|V\|_{\text{Lip}^\gamma}$ .

The non linear version of the Lyons continuity theorem is the following result that we admit without proof. We refer the interested reader to [?].

**Theorem 7.32** (Lyons continuity theorem). *Let  $p \geq 1$ . Let  $x \in \Omega^p([0, T], \mathbb{R}^d)$  be a  $p$ -rough path with truncated signature*

$$\sum_{k=0}^{[p]} \int_{\Delta^k[s,t]} dx^{\otimes k},$$

and let  $x_n \in C^{1-\text{var}}([0, T], \mathbb{R}^d)$  be an approximating sequence such that

$$\sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} - \int dx_n^{\otimes j} \right\|_{\frac{p}{j}-\text{var}, [0,T]}^{1/j} \rightarrow 0.$$

Assume that  $V_1, \dots, V_d$  are  $\gamma$ -Lipschitz vector fields in  $\mathbb{R}^n$  with  $\gamma > p$ . The solution of the equation

$$y_n(t) = y(0) + \sum_{j=1}^d \int_0^t V_j(y_n(s)) dx_n^j(s), \quad 0 \leq t \leq T,$$

converges when  $n \rightarrow +\infty$  in  $p$ -variation to some  $y \in C^{p-\text{var}}([0, T], \mathbb{R}^n)$  that we call a solution of the rough differential equation:

$$y(t) = y(0) + \sum_{j=1}^d \int_0^t V_j(y(s)) dx^j(s), \quad 0 \leq t \leq T.$$

Moreover, there exists a constant  $C$  depending only on  $p, T$  and  $\gamma$  such that for every  $0 \leq s < t \leq T$ ,

$$\|y\|_{p-\text{var}, [s,t]} \leq C \left( \|V\|_{\text{Lip}^{\gamma-1}} \|\mathbf{x}\|_{p-\text{var}, [s,t]} + \|V\|_{\text{Lip}^{\gamma-1}}^p \|\mathbf{x}\|_{p-\text{var}, [s,t]}^p \right), \quad (7.1)$$

where we denoted

$$\|\mathbf{x}\|_{p-\text{var}, [s,t]} = \sum_{j=1}^{[p]} \left\| \int dx^{\otimes j} \right\|_{\frac{p}{j}-\text{var}, [s,t]}^{1/j}.$$

**Remark 7.33.** The estimate (7.1) is often referred to as the Davie estimate.

## 5 The Brownian motion as a rough path

It is now time to show how rough paths theory can be used to study Brownian motion and stochastic differential equations. We first show in this section that Brownian motion paths are almost surely  $p$ -rough paths for  $2 < p < 3$ . The key estimate is the Garsia-Rodemich-Rumsey inequality.

**Theorem 7.34** (Garsia-Rodemich-Rumsey inequality). *Let  $T > 0$  and let*

$$\Gamma : \{0 \leq s, t \leq T\} \rightarrow \mathbb{R}_{\geq 0}$$

*be a continuous symmetric functional ( $\Gamma_{s,t} = \Gamma_{t,s}$ ) that vanishes on the diagonal ( $\Gamma_{t,t} = 0$ ) and such that there exists a constant  $C > 0$  such that for every  $0 \leq t_1 \leq \dots \leq t_n \leq T$ ,*

$$\Gamma_{t_1, t_n} \leq C \left( \sum_{i=1}^{n-1} \Gamma_{t_i, t_{i+1}} \right). \quad (7.2)$$

*For  $q > 1$  and  $\alpha \in (1/q, 1)$ , there exists a constant  $K > 0$  such that for all  $0 \leq s \leq t \leq T$ ,*

$$\Gamma_{s,t}^q \leq K |t - s|^{\alpha q - 1} \int_0^T \int_0^T \frac{\Gamma_{u,v}^q}{|u - v|^{1 + \alpha q}} du dv.$$

*Proof. Step 1.* We first assume  $T = 1$  and prove that

$$\Gamma_{0,1}^q \leq K \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u - v|^{1 + \alpha q}} du dv.$$

Define  $I(v) = \int_0^1 \frac{\Gamma_{u,v}^q}{|u - v|^{1 + \alpha q}} du$  so that

$$\int_0^1 I(v) dv = \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u - v|^{1 + \alpha q}} du dv.$$

We can find  $t_0 \in (0, 1)$  such that  $I(t_0) \leq \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u - v|^{1 + \alpha q}} du dv$ . We construct then a decreasing sequence  $(t_n)_{n \geq 0}$  by induction as follows. If  $t_{n-1}$  has been chosen, then we pick  $t_n \in (0, \frac{1}{2}t_{n-1})$  such that

$$I(t_n) \leq \frac{4}{t_{n-1}} \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u - v|^{1 + \alpha q}} du dv, \quad J_{n-1}(t_n) \leq \frac{4I(t_{n-1})}{t_{n-1}}$$

where we have set  $J_{n-1}(s) = \frac{\Gamma_{s, t_{n-1}}^q}{|s - t_{n-1}|^{1 + \alpha q}}$ . We can always find such  $t_n$ . Otherwise, we would have  $(0, \frac{1}{2}t_{n-1}) = A \cup B$  with

$$A = \left\{ t \in \left( 0, \frac{1}{2}t_{n-1} \right), I(t) > \frac{4}{t_{n-1}} \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u - v|^{1 + \alpha q}} du dv \right\}$$

$$B = \left\{ t \in \left( 0, \frac{1}{2}t_{n-1} \right), J_{n-1}(t) > \frac{4I(t_{n-1})}{t_{n-1}} \right\}$$

But, clearly we have

$$\frac{4\mu(A)}{t_{n-1}} \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u-v|^{1+\alpha q}} dudv \leq \int_A I(t) dt \leq \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u-v|^{1+\alpha q}} dudv.$$

where  $\mu$  stands for the Lebesgue measure. Hence, if  $\mu(A) > 0$ , we have  $\mu(A) < \frac{1}{4}t_{n-1}$ . Similarly of course,  $\mu(B) > 0$  implies  $\mu(B) < \frac{1}{4}t_{n-1}$ . This contradicts the fact that  $(0, \frac{1}{2}t_{n-1}) = A \cup B$ . Hence  $t_n$  is well defined if  $t_{n-1}$  is. We have then

$$J_{n-1}(t_n) \leq \frac{4I(t_{n-1})}{t_{n-1}} \leq \frac{16}{t_{n-1}^2} \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u-v|^{1+\alpha q}} dudv.$$

Coming back to the definition of  $J_{n-1}$  yields then

$$\frac{\Gamma_{t_n, t_{n-1}}^q}{|t_n - t_{n-1}|^{1+\alpha q}} \leq \frac{16}{t_{n-1}^2} \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u-v|^{1+\alpha q}} dudv.$$

But it is easily seen that there a constant  $C > 0$  such that

$$\frac{|t_n - t_{n-1}|^{\alpha+1/q}}{t_{n-1}^{2/q}} \leq C(t_n^{\alpha-1/q} - t_{n+1}^{\alpha-1/q}).$$

Using then the subadditivity of the functional  $\Gamma$  we end up with

$$\Gamma_{0, t_0}^q \leq K \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u-v|^{1+\alpha q}} dudv.$$

Similarly, by repeating the above argument on the functional  $\Gamma_{1-s, 1-t}$  we get that

$$\Gamma_{t_0, 1}^q \leq K \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u-v|^{1+\alpha q}} dudv.$$

Therefore we proved that for some constant  $K'$

$$\Gamma_{0,1}^q \leq K' \int_0^1 \int_0^1 \frac{\Gamma_{u,v}^q}{|u-v|^{1+\alpha q}} dudv.$$

**Step 2:** We now consider a general functional  $\Gamma : \{0 \leq s, t \leq T\} \rightarrow \mathbb{R}_{\geq 0}$  and fix  $0 \leq s < t \leq T$ . Considered then the rescaled functional  $\tilde{\Gamma}_{u,v}^q = \Gamma_{s+u(t-s), s+v(t-s)}$  which is defined for  $0 \leq u, v \leq 1$ . From the first step, we have

$$\tilde{\Gamma}_{0,1}^q \leq K \int_0^1 \int_0^1 \frac{\tilde{\Gamma}_{u,v}^q}{|u-v|^{1+\alpha q}} dudv,$$

and the result follows by a simple change of variable.  $\square$

As we are going to see, for the Brownian motion the natural integral in rough paths theory is not Itô's integral but Stratonovitch's (see Section 5 for the definition). If  $(B_t)_{t \geq 0}$  is  $d$  dimensional Brownian motion, we can inductively define the iterated Stratonovitch integrals  $\int_{0 \leq t_1 \leq \dots \leq t_k \leq t} \circ dB_{t_1}^{i_1} \dots \circ dB_{t_k}^{i_k}$ , and, as before, it will be convenient to use a concise notation by embedding the set of iterated integrals in the algebra of formal series:

$$\int_{\Delta^k[s,t]} \circ dB^{\otimes k} = \sum_{I \in \{1, \dots, d\}^k} \left( \int_{\Delta^k[s,t]} \circ dB^I \right) X_{i_1} \dots X_{i_k}.$$

**Exercise 7.35.** Show that for  $m \geq 1$ , and  $t \geq 0$ ,

$$\mathbb{E} \left( \left| \int_{0 \leq t_1 \leq \dots \leq t_k \leq t} \circ dB_{t_1}^{i_1} \dots \circ dB_{t_k}^{i_k} \right|^m \right) < +\infty.$$

The following estimate is crucial to apply rough paths theory to Brownian motion.

**Proposition 7.36.** Let  $T \geq 0$ ,  $2 < p < 3$  and  $n \geq 1$ . There exists a finite positive random variable  $C = C(T, p, n)$  such that  $\mathbb{E}(C^m) < +\infty$  for every  $m \geq 1$ , and for every  $0 \leq s \leq t \leq T$ ,

$$\sum_{k=1}^n \left\| \int_{\Delta^k[s,t]} \circ dB^{\otimes k} \right\|^{1/k} \leq C |t - s|^{1/p}.$$

*Proof.* Let us consider the functional

$$\Gamma_{s,t} = \sum_{k=1}^n \left\| \int_{\Delta^k[s,t]} \circ dB^{\otimes k} \right\|^{1/k},$$

and prove that it satisfies the chain condition (7.2). Using inductively Chen's relations, it is seen that

$$\int_{\Delta^k[t_1, t_N]} \circ dB^{\otimes k} = \sum_{i_1 + \dots + i_{N-1} = k} \int_{\Delta^{i_1}[t_1, t_2]} \circ dB^{\otimes i_1} \dots \int_{\Delta^{i_{N-1}}[t_{N-1}, t_N]} \circ dB^{\otimes i_{N-1}}.$$

Therefore we have

$$\begin{aligned}
 & \left\| \int_{\Delta^k[t_1, t_N]} \circ dB^{\otimes k} \right\| \\
 & \leq \sum_{i_1 + \dots + i_{N-1} = k} \left\| \int_{\Delta^{i_1}[t_1, t_2]} \circ dB^{\otimes i_1} \dots \int_{\Delta^{i_{N-1}}[t_{N-1}, t_N]} \circ dB^{\otimes i_{N-1}} \right\| \\
 & \leq \sum_{i_1 + \dots + i_{N-1} = k} \left\| \int_{\Delta^{i_1}[t_1, t_2]} \circ dB^{\otimes i_1} \right\| \dots \left\| \int_{\Delta^{i_{N-1}}[t_{N-1}, t_N]} \circ dB^{\otimes i_{N-1}} \right\| \\
 & \leq \sum_{i_1 + \dots + i_{N-1} = k} \Gamma_{t_1, t_2}^{i_1} \dots \Gamma_{t_{N-1}, t_N}^{i_{N-1}} \\
 & \leq (\Gamma_{t_1, t_2} + \dots + \Gamma_{t_{N-1}, t_N})^k.
 \end{aligned}$$

This yields

$$\Gamma_{t_1, t_N} \leq n (\Gamma_{t_1, t_2} + \dots + \Gamma_{t_{N-1}, t_N}).$$

From the Garsia-Rodemich-Rumsey inequality, for  $q > 1$  and  $\alpha \in (1/q, 1)$ , there exists therefore a constant  $K > 0$  such that for all  $0 \leq s \leq t \leq T$ ,

$$\Gamma_{s, t}^q \leq K |t - s|^{\alpha q - 1} \int_0^T \int_0^T \frac{\Gamma_{u, v}^q}{|u - v|^{1 + \alpha q}} du dv.$$

The scaling property of Brownian motion and Exercise 7.35 imply that for some constant  $K > 0$

$$\mathbb{E}(\Gamma_{u, v}^q) = K |u - v|^{q/2}.$$

As a consequence

$$\int_0^T \int_0^T \frac{\Gamma_{u, v}^q}{|u - v|^{1 + \alpha q}} du dv < +\infty$$

if  $\alpha - \frac{1}{q} < \frac{1}{2}$  and the result follows from Fubini theorem. □

With this estimate in hands, we are now in position to prove that Brownian motion paths are  $p$ -rough paths for  $2 < p < 3$ . We can moreover give an explicit approximating sequence. Let us work on a fixed interval  $[0, T]$  and consider a sequence  $D_n$  of subdivisions of  $[0, T]$  such that  $D_{n+1} \subset D_n$  and whose mesh goes to 0 when  $n \rightarrow +\infty$ . An example is given by the sequence of dyadic subdivisions. The family  $\mathcal{F}_n = \sigma(B_t, t \in D_n)$  is then a filtration, that is an increasing family of  $\sigma$ -fields. We denote by  $B^n$  the piecewise linear process which is obtained from  $B$  by interpolation along the subdivision  $D_n$ , that is for  $t_i^n \leq t \leq t_{i+1}^n$ ,

$$B_t^n = \frac{t_{i+1}^n - t}{t_{i+1}^n - t_i^n} B_{t_i^n} + \frac{t - t_i^n}{t_{i+1}^n - t_i^n} B_{t_{i+1}^n}.$$

**Theorem 7.37.** *When  $n \rightarrow +\infty$ , almost surely*

$$\|B^n - B\|_{p\text{-var},[0,T]} + \left\| \int dB^{n,\otimes 2} - \int \circ dB^{\otimes 2} \right\|_{\frac{p}{2}\text{-var},[0,T]}^{1/2} \rightarrow 0.$$

*As a consequence, Brownian motion paths are almost surely  $p$ -rough paths for  $2 < p < 3$ .*

*Proof.* We first observe that, due to the Markov property of Brownian motion, we have for  $t_i^n \leq t \leq t_{i+1}^n$ ,

$$\mathbb{E}(B_t | \mathcal{F}_n) = \mathbb{E}(B_t | B_{t_i^n}, B_{t_{i+1}^n}).$$

It is then an easy exercise to check that

$$\mathbb{E}(B_t | B_{t_i^n}, B_{t_{i+1}^n}) = \frac{t_{i+1}^n - t}{t_{i+1}^n - t_i^n} B_{t_i^n} + \frac{t - t_i^n}{t_{i+1}^n - t_i^n} B_{t_{i+1}^n} = B_t^n.$$

As a conclusion, we get

$$\mathbb{E}(B_t | \mathcal{F}_n) = B_t^n.$$

It immediately follows that  $B_t^n \rightarrow B_t$  when  $n \rightarrow +\infty$ . In the same way, for  $i \neq j$  we have

$$\mathbb{E}\left(\int_0^t B_s^i dB_s^j | \mathcal{F}_n\right) = \int_0^t B_s^{n,i} dB_s^{n,j}.$$

Indeed, for  $0 < t < T$  and  $\varepsilon$  small enough, we have by independence of  $B^i$  and  $B^j$ ,

$$\mathbb{E}\left(B_t^i (B_{t+\varepsilon}^j - B_t^j) | \mathcal{F}_n\right) = \mathbb{E}(B_t^i | \mathcal{F}_n) \mathbb{E}\left(B_{t+\varepsilon}^j - B_t^j | \mathcal{F}_n\right) = B_t^{n,i} (B_{t+\varepsilon}^{n,j} - B_t^{n,j}),$$

and we conclude by using the fact that Itô's integral is a limit in  $L^2$  of Riemann sums. It follows that, almost surely,

$$\lim_{n \rightarrow \infty} \int_0^t B_s^{n,i} dB_s^{n,j} = \int_0^t B_s^i dB_s^j.$$

Since  $B^i$  is independent from  $B^j$ , the quadratic covariation  $\langle B^i, B^j \rangle$  is zero. As a consequence we have  $\int_0^t B_s^i dB_s^j = \int_0^t B_s^i \circ dB_s^j$ . For  $i = j$ , we have  $\int_0^t B_s^i \circ dB_s^i = (B_t^i)^2$ . So if we collect the previous results, we established that almost surely

$$\lim_{n \rightarrow +\infty} \left( B_t^n, \int_0^t B_s^n \otimes dB_s^n \right) = \left( B_t, \int_0^t B_s \otimes \circ dB_s \right).$$

To prove the convergence with the variation norms, we need a uniform Hölder estimate.



From Proposition 7.36, we know that there is a finite random variable  $K_1$  (that belongs to  $L^m$  for every  $m \geq 1$ ), such that for every  $i \neq j$ ,  $0 \leq s \leq t \leq T$ ,

$$\left| \int_s^t (B_u^i - B_s^i) \circ dB_u^j \right| \leq K_1 |t - s|^{2/p}.$$

Since

$$\mathbb{E} \left( \int_s^t (B_u^i - B_s^i) dB_u^j \mid \mathcal{F}_n \right) = \int_s^t (B_u^{n,i} - B_s^{n,i}) dB_u^{n,j},$$

we deduce that

$$\left| \int_s^t (B_u^{n,i} - B_s^{n,i}) dB_u^{n,j} \right| \leq K_2 |t - s|^{2/p},$$

where  $K_2$  is a finite random variable that belongs to  $L^m$  for every  $m \geq 1$ . Similarly, of course, we have

$$\|B_t^n - B_s^n\| \leq K_3 |t - s|^{1/p}.$$

Combining the pointwise convergence with these uniform Hölder estimates give then the expected result. We let the reader work out the details as an exercise.  $\square$

Since a  $d$ -dimensional Brownian motion  $(B_t)_{t \geq 0}$  is a  $p$ -rough path for  $2 < p < 3$ , we know how to give a sense to the signature of the Brownian motion. In particular, the iterated integrals at any order of the Brownian motion are well defined. It turns out that these iterated integrals, and this comes a no surprise in view of the previous result, coincide with iterated Stratonovitch integrals.

**Theorem 7.38.** *If  $(B_t)_{t \geq 0}$  is a  $d$ -dimensional Brownian motion, the signature of  $B$  as a rough path is the formal series:*

$$\begin{aligned} \mathfrak{S}(B)_t &= 1 + \sum_{k=1}^{+\infty} \int_{\Delta^k[0,t]} \circ dB^{\otimes k} \\ &= 1 + \sum_{k=1}^{+\infty} \sum_{I \in \{1, \dots, d\}^k} \left( \int_{0 \leq t_1 \leq \dots \leq t_k \leq t} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k} \right) X_{i_1} \cdots X_{i_k}. \end{aligned}$$

*Proof.* Let us work on a fixed interval  $[0, T]$  and consider a sequence  $D_n$  of subdivisions of  $[0, T]$  such that  $D_{n+1} \subset D_n$  and whose mesh goes to 0 when  $n \rightarrow +\infty$ . As before, we denote by  $B^n$  the piecewise linear process which is obtained from  $B$  by interpolation along the subdivision  $D_n$ , that is for  $t_i^n \leq t \leq t_{i+1}^n$ ,

$$B_t^n = \frac{t_{i+1}^n - t}{t_{i+1}^n - t_i^n} B_{t_i^n} + \frac{t - t_i^n}{t_{i+1}^n - t_i^n} B_{t_{i+1}^n}.$$

We know from the previous lecture that  $B^n$  converges to  $B$  in the  $p$ -rough paths topology  $2 < p < 3$ . In particular all the iterated integrals  $\int_{\Delta^k[s,t]} dB^{n,\otimes k}$  converge. We claim that actually,

$$\lim_{n \rightarrow \infty} \int_{\Delta^k[s,t]} dB^{n,\otimes k} = \int_{\Delta^k[0,t]} \circ dB^{\otimes k}.$$

Let us denote

$$\int_{\Delta^k[s,t]} \partial B^{\otimes k} = \lim_{n \rightarrow \infty} \int_{\Delta^k[s,t]} dB^{n,\otimes k}.$$

We are going to prove by induction on  $k$  that  $\int_{\Delta^k[s,t]} \partial B^{\otimes k} = \int_{\Delta^k[s,t]} \circ dB^{\otimes k}$ . We have

$$\begin{aligned} \int_0^T B_s^n \otimes dB_s^n &= \sum_{i=0}^{n-1} \int_{t_i^n}^{t_{i+1}^n} B_s^n \otimes dB_s^n \\ &= \sum_{i=0}^{n-1} \int_{t_i^n}^{t_{i+1}^n} \left( \frac{t_{i+1}^n - s}{t_{i+1}^n - t_i^n} B_{t_i^n} + \frac{s - t_i^n}{t_{i+1}^n - t_i^n} B_{t_{i+1}^n} \right) ds \otimes \frac{B_{t_{i+1}^n} - B_{t_i^n}}{t_{i+1}^n - t_i^n} \\ &= \frac{1}{2} \sum_{i=0}^{n-1} \left( B_{t_{i+1}^n} + B_{t_i^n} \right) \otimes \left( B_{t_{i+1}^n} - B_{t_i^n} \right) \end{aligned}$$

By taking the limit when  $n \rightarrow \infty$ , we deduce therefore that  $\int_{\Delta^2[0,T]} \partial B^{\otimes 2} = \int_{\Delta^2[0,T]} \circ dB^{\otimes 2}$ . In the same way, we have for  $0 \leq s < t \leq T$ ,  $\int_{\Delta^2[s,t]} \partial B^{\otimes 2} = \int_{\Delta^2[s,t]} \circ dB^{\otimes 2}$ . Assume now by induction, that for every  $0 \leq s \leq t \leq T$  and  $1 \leq j \leq k$ ,  $\int_{\Delta^k[s,t]} \partial B^{\otimes k} = \int_{\Delta^k[s,t]} \circ dB^{\otimes k}$ . Let us denote

$$\Gamma_{s,t} = \int_{\Delta^{k+1}[s,t]} \partial B^{\otimes(k+1)} - \int_{\Delta^{k+1}[s,t]} \circ dB^{\otimes(k+1)}.$$

From the Chen's relations, we immediately see that

$$\Gamma_{s,u} = \Gamma_{s,t} + \Gamma_{t,u}.$$

Moreover, it is easy to estimate

$$\|\Gamma_{s,t}\| \leq C\omega(s,t)^{\frac{k+1}{p}},$$

where  $2 < p < 3$  and  $\omega(s,t) = |t - s|$ . Indeed, the bound

$$\left\| \int_{\Delta^{k+1}[s,t]} \partial B^{\otimes(k+1)} \right\| \leq C_1 \omega(s,t)^{\frac{k+1}{p}}$$

comes from Theorem 7.25 and the bound

$$\left\| \int_{\Delta^{k+1}[s,t]} \circ dB^{\otimes(k+1)} \right\| \leq C_2 \omega(s,t)^{\frac{k+1}{p}}$$

comes from Proposition 7.36. As a consequence of Lemma 7.19, we deduce that  $\Gamma_{s,t} = 0$  which proves the induction.  $\square$

We finish this section by a very interesting probabilistic object, the expectation of the Brownian signature. If

$$Y = y_0 + \sum_{k=1}^{+\infty} \sum_{I \in \{1, \dots, d\}^k} a_{i_1, \dots, i_k} X_{i_1} \dots X_{i_k}.$$

is a random series, that is if the coefficients are real random variables defined on a probability space, we will denote

$$\mathbb{E}(Y) = \mathbb{E}(y_0) + \sum_{k=1}^{+\infty} \sum_{I \in \{1, \dots, d\}^k} \mathbb{E}(a_{i_1, \dots, i_k}) X_{i_1} \dots X_{i_k}.$$

as soon as the coefficients of  $Y$  are integrable.

**Theorem 7.39.** For  $t \geq 0$ ,

$$\mathbb{E}(\mathfrak{S}(B)_t) = \exp \left( t \left( \frac{1}{2} \sum_{i=1}^d X_i^2 \right) \right).$$

*Proof.* An easy computation shows that if  $\mathcal{I}_n$  is the set of words with length  $n$  obtained by all the possible concatenations of the words

$$\{(i, i)\}, \quad i \in \{1, \dots, d\},$$

1. If  $I \notin \mathcal{I}_n$  then

$$\mathbb{E} \left( \int_{\Delta^n[0,t]} \circ dB^I \right) = 0;$$

2. If  $I \in \mathcal{I}_n$  then

$$\mathbb{E} \left( \int_{\Delta^n[0,t]} \circ dB^I \right) = \frac{t^{\frac{n}{2}}}{2^{\frac{n}{2}} \left(\frac{n}{2}\right)!}.$$

Therefore, we have

$$\begin{aligned} \mathbb{E}(\mathfrak{S}(B)_t) &= 1 + \sum_{k=1}^{+\infty} \sum_{I \in \mathcal{I}_k} \frac{t^{\frac{k}{2}}}{2^{\frac{k}{2}} \left(\frac{k}{2}\right)!} X_{i_1} \dots X_{i_k} \\ &= \exp \left( t \left( \frac{1}{2} \sum_{i=1}^d X_i^2 \right) \right). \end{aligned}$$

$\square$

Based on the previous results, it should come as no surprise that differential equations driven by the Brownian rough path should correspond to Stratonovitch differential equations. This is indeed the case. For linear equations, the result is easy to prove as a consequence of the previous proposition.

**Theorem 7.40.** *Let us consider matrices  $M_1, \dots, M_d \in \mathbb{R}^{n \times n}$ . For  $y_0 \in \mathbb{R}^n$ , the solution of the rough differential equation*

$$y(t) = y(0) + \sum_{i=1}^d \int_0^t M_i y(s) dB^i(s),$$

*coincides with the solution of the Stratonovitch differential equation*

$$y(t) = y(0) + \sum_{i=1}^d \int_0^t M_i y(s) \circ dB^i(s).$$

*Proof.* The solution of the rough differential equation can be expanded as the Volterra series:

$$y(t) = y(0) + \sum_{k=1}^{+\infty} \sum_{I=(i_1, \dots, i_k)} M_{i_1} \cdots M_{i_k} \left( \int_{\Delta^k[0,t]} \partial B^I \right) y(0),$$

where  $\int_{\Delta^k[0,t]} \partial B^I$  is the iterated integral in the rough path sense. This iterated integral coincides with the iterated Stratonovitch integral, so we have

$$y(t) = y(0) + \sum_{k=1}^{+\infty} \sum_{I=(i_1, \dots, i_k)} M_{i_1} \cdots M_{i_k} \left( \int_{\Delta^k[0,t]} \circ dB^I \right) y(0),$$

which is the Volterra expansion of the Stratonovitch stochastic differential equation.  $\square$

The previous theorem extends to the non linear case. We refer the interested reader to [?] for a proof of the following general result.

**Theorem 7.41.** *Let  $\gamma > 2$  and let  $V_1, \dots, V_d$  be  $\gamma$ -Lipschitz vector fields on  $\mathbb{R}^n$ . Let  $x_0 \in \mathbb{R}^n$ . The solution of the rough differential equation*

$$X_t = x_0 + \sum_{i=1}^d \int_0^t V_i(X_s) dB_s^i,$$

*is the solution of the Stratonovitch differential equation:*

$$X_t = x_0 + \sum_{i=1}^d \int_0^t V_i(X_s) \circ dB_s^i.$$

To finish we mention that rough paths theory provides a perfect framework to study differential equations driven by Gaussian processes. Indeed, let  $B$  be a Gaussian process. A lot of the information concerning the Gaussian process  $B$  is encoded in the rectangular increments of  $R$ , which are defined by

$$R_{uv}^{st} = \mathbb{E}((B_t - B_s) \otimes (B_v - B_u)).$$

We then call 2-dimensional  $\rho$ -variation of  $R$  the quantity

$$V_\rho(R) \equiv \sup \left\{ \left( \sum_{i,j} \left\| R_{s_i s_{i+1}}^{t_j t_{j+1}} \right\|^\rho \right)^{1/\rho} ; (s_i), (t_j) \in \Pi \right\},$$

where  $\Pi$  stands for the set of partitions of  $[0, 1]$ . One has the following fundamental Friz-Victoir theorem concerning the Gaussian rough path existence:

**Proposition 7.42** (Friz-Victoir). *If there exists  $1 \leq \rho < 2$  such that  $R$  has finite  $\rho$ -variation, the process  $B$  is a  $p$ -rough path for any  $p > 2\rho$ .*

As an illustration, the previous result applies in particular to fractional Brownian motion. Let  $B = (B^1, \dots, B^d)$  be a  $d$  dimensional fractional Brownian motion defined on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , with Hurst parameter  $H \in (0, 1)$ . It means that  $B$  is a centered Gaussian process, whose coordinates are independent and satisfy

$$\mathbb{E} \left( (B_t^j - B_s^j)^2 \right) = |t - s|^{2H}, \quad \text{for } s, t \geq 0. \tag{7.3}$$

For instance when  $H = 1/2$  a fractional Brownian motion is a Brownian motion. It can then be proved that the hypothesis of Proposition 7.42 is satisfied for a fractional Brownian motion with Hurst parameter  $H \in (\frac{1}{4}, 1)$ , and in that case  $\rho = 1/(2H)$ . Thus, by using the rough paths theory, we may define stochastic differential equations driven by fractional Brownian motions with Hurst parameter  $H \in (\frac{1}{4}, 1)$ .

## Notes and Comments

**Section 2.** The signature of the path is fascinating object at the intersection between algebraic geometry, topology, Lie group theory and sub-Riemannian geometry. Several sub-Riemannian aspects of the signature are studied in the book [?].

**Sections 3,4.** The rough paths theory is a relatively new theory that was built in the paper [?] by T. Lyons. There are now several comprehensive books about the theory: we mention in particular [?], [?] and [?]. At the

time of writing, the theory is still a very active domain of research. Recent important applications have been found in the theory of ill-posed stochastic partial differential equations (see [?]).

**Section 5.** The expectation of the signature of the Brownian motion as a rough path can be used to construct parametrices for semigroups and finds applications to index theory (see [?]). The rough paths point of view on stochastic differential equations has several advantages. It comes with previously unknown deterministic estimates like Davie's and often provides straightforward proofs and improvements of classical results like the Stroock-Varadhan support theorem. We refer to the book [?] for an overview of these applications. Besides stochastic differential equations driven by Brownian motions, rough paths also applies to stochastic differential equations driven by general semimartingales because it may be proved that such processes always are  $p$ -rough paths for  $2 < p < 3$ . Of course, the advantage is that the theory even applies to equations driven by very irregular processes that are not semimartingales like rough Gaussian processes (for instance the fractional Brownian motion, see [?]). In that case Itô's theory does not apply and Lyons' theory becomes the only way to define and study such equations.

## Appendix A. Unbounded operators

It is a fact that many interesting linear operators are not bounded and only defined on a dense subset of a Banach space (think of differential operators in  $L^p$ ). We collect here some general definitions and basic results about such operators that are used at some places in this book. The details and the proofs may be found in the reference book [?].

Let  $(\mathcal{B}_1, \|\cdot\|_1)$  and  $(\mathcal{B}_2, \|\cdot\|_2)$  be two Banach spaces. Let  $T : V \rightarrow \mathcal{B}_2$  be a linear operator defined on a linear subspace  $V \subset \mathcal{B}_1$ . The space  $V$  on which  $T$  is defined is called the domain of  $T$  and usually denoted by  $\mathcal{D}(T)$ . If  $\mathcal{D}(T)$  is dense in  $\mathcal{B}_1$ , then  $T$  is said to be densely defined.

In the study of unbounded operators like  $T$ , it is often useful to consider the graph:

$$\mathbf{G}_T = \{(v, Tv), v \in \mathcal{D}(T)\},$$

which is a linear subspace of the Banach space  $\mathcal{B}_1 \oplus \mathcal{B}_2$ .

**Definition 7.43.** The operator  $T$  is said to be a closed operator if its graph  $\mathbf{G}_T$  is a closed linear subspace of  $\mathcal{B}_1 \oplus \mathcal{B}_2$ .

The following theorem is known as the closed graph theorem.

**Theorem 7.44.** *Let us assume that  $\mathcal{D}(T) = \mathcal{B}_1$ . The operator  $T$  is bounded if and only if it is closed.*

In general, the closure  $\bar{\mathbf{G}}_T$  of  $\mathbf{G}_T$  needs not to be the graph of an operator.

**Definition 7.45.** The operator  $T$  is said to be closable if there is an operator  $\bar{T}$  such that

$$\bar{\mathbf{G}}_T = \mathbf{G}_{\bar{T}}.$$

The operator  $\bar{T}$  is then called the closure of  $T$ .

If  $T$  is densely defined, we define the adjoint of  $T$  as the linear operator

$$T' : \mathcal{D}(T') \subset \mathcal{B}_2^* \rightarrow \mathcal{B}_1^*$$

which is defined on

$$\mathcal{D}(T') = \{u \in \mathcal{B}_2^*, \exists c(u) \geq 0, \forall v \in \mathcal{D}(T), |\langle u, Tv \rangle| \leq c(u)\|v\|_1\}$$

and characterized by the duality formula

$$\langle u, Tv \rangle = \langle T'u, v \rangle, \quad u \in \mathcal{D}(T'), v \in \mathcal{D}(T).$$

We have then the following result.

**Proposition 7.46.** *Let assume that the Banach spaces  $\mathcal{B}_1$  and  $\mathcal{B}_2$  are reflexive. A densely defined operator*

$$T : \mathcal{D}(T) \subset \mathcal{B}_1 \rightarrow \mathcal{B}_2$$

*is closable if and only if  $T'$  is densely defined.*

Let  $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$  be a Hilbert space and let  $A$  be a densely defined operator on a domain  $\mathcal{D}(A)$ . We have the following basic definitions.

- The operator  $A$  is said to be symmetric if for  $f, g \in \mathcal{D}(A)$ ,

$$\langle f, Ag \rangle_{\mathcal{H}} = \langle Af, g \rangle_{\mathcal{H}}.$$

- The operator  $A$  is said to be non negative symmetric operator, if it is symmetric and if for  $f \in \mathcal{D}(A)$ ,

$$\langle f, Af \rangle_{\mathcal{H}} \geq 0.$$

- The adjoint  $A^*$  of  $A$  is the operator defined on the domain

$$\mathcal{D}(A^*) = \{f \in \mathcal{H}, \exists c(f) \geq 0, \forall g \in \mathcal{D}(A), |\langle f, Ag \rangle_{\mathcal{H}}| \leq c(f)\|g\|_{\mathcal{H}}\}$$

and given through the Riesz representation theorem by the formula

$$\langle A^*f, g \rangle_{\mathcal{H}} = \langle f, Ag \rangle_{\mathcal{H}}$$

where  $g \in \mathcal{D}(A)$ ,  $f \in \mathcal{D}(A^*)$ .

- The operator  $A$  is said to be self-adjoint if it is symmetric and if  $\mathcal{D}(A^*) = \mathcal{D}(A)$ .

Let us observe that, in general, the adjoint  $A^*$  is not necessarily densely defined, however it is readily checked that if  $A$  is a symmetric operator then, from Cauchy-Schwarz inequality,  $\mathcal{D}(A) \subset \mathcal{D}(A^*)$ . We have the following first criterion for self-adjointness which may be useful.

**Lemma 7.47.** *Let  $A : \mathcal{D}(A) \subset \mathcal{H} \rightarrow \mathcal{H}$  be a densely defined operator. Consider the graph of  $A$ :*

$$\mathbf{G}_A = \{(v, Av), v \in \mathcal{D}(A)\} \subset \mathcal{H} \oplus \mathcal{H},$$

*and the complex structure*

$$\begin{aligned} \mathcal{J} : \mathcal{H} \oplus \mathcal{H} &\rightarrow \mathcal{H} \oplus \mathcal{H} \\ (v, w) &\rightarrow (-w, v) \end{aligned}$$

*Then, the operator  $A$  is self-adjoint if and only if*

$$\mathbf{G}_A^{\perp} = \mathcal{J}(\mathbf{G}_A).$$



*Proof.* It is checked that for any densely defined operator  $A$

$$\mathbf{G}_{A^*} = \mathcal{J}(\mathbf{G}_A^\perp),$$

and the conclusion follows from routine computations.  $\square$

The following result is useful in the proof of Friedrichs extension theorem (Theorem 4.6):

**Lemma 7.48.** *Let  $A : \mathcal{D}(A) \subset \mathcal{H} \rightarrow \mathcal{H}$  be an injective densely defined self-adjoint operator. Let us denote by  $\mathcal{R}(A)$  the range of  $A$ . The inverse operator*

$$A^{-1} : \mathcal{R}(A) \rightarrow \mathcal{H}$$

*is a densely defined self-adjoint operator.*

*Proof.* First, let us observe that

$$\mathcal{R}(A)^\perp = \mathbf{Ker}(A^*) = \mathbf{Ker}(A) = \{0\}.$$

Therefore  $\mathcal{R}(A)$  is dense in  $\mathcal{H}$  and  $A^{-1}$  is densely defined. Now,

$$\begin{aligned} \mathbf{G}_{A^{-1}}^\perp &= \mathcal{J}(\mathbf{G}_{-A})^\perp \\ &= \mathcal{J}(\mathbf{G}_{-A}^\perp) \\ &= \mathcal{J}\mathcal{J}(\mathbf{G}_{-A}) \\ &= \mathcal{J}(\mathbf{G}_{A^{-1}}). \end{aligned}$$

$\square$

**Definition 7.49.** Let  $X$  and  $Y$  be Banach spaces, A bounded operator  $T : X \rightarrow Y$  is called compact if  $T$  transforms bounded sets into relatively compact sets. That is,  $T$  is compact if and only if for every bounded sequence  $(x_n)_{n \in \mathbb{N}}$  in  $X$ , the sequence  $(Tx_n)_{n \in \mathbb{N}}$  has a subsequence convergent in  $Y$ .

For instance, an operator whose range is finite dimensional is necessarily a compact operator.

We have the following results about compact operators: Let  $X$  and  $Y$  be Banach spaces and let  $T : X \rightarrow Y$  be a bounded operator.

- If  $(T_n)_{n \in \mathbb{N}}$  is a sequence of compact operators and  $T_n \rightarrow T$  in the norm topology, then  $T$  is compact.
- If  $S$  is a bounded operator from  $Y$  to a Banach space  $Z$  and if  $T$  or  $S$  is compact, then  $ST$  is compact.
- $T$  is compact if and only if its adjoint  $T'$  is a compact operator.

For self-adjoint compact operators, the spectral theorem takes a particularly nice form:

**Theorem 7.50** (Hilbert-Schmidt theorem). *Let  $T$  be a compact and self-adjoint operator defined on a separable Hilbert space  $\mathcal{H}$ . Then, there is a complete orthonormal basis  $(\phi_n)_{n \in \mathbb{N}}$  for  $\mathcal{H}$  so that*

$$T\phi_n = \lambda_n\phi_n,$$

and  $\lambda_n \rightarrow 0$  as  $n \rightarrow \infty$ .

There are two interesting ideals of the class of compact operators: The trace-class operators and the Hilbert-Schmidt operators.

**Definition 7.51.** Let  $\mathcal{H}$  be a separable Hilbert space,  $(\phi_n)_{n \in \mathbb{N}}$  an orthonormal basis. Then for every bounded and positive operator  $T$ , we define

$$\mathbf{Tr}(T) = \sum_{n \in \mathbb{N}} \langle \phi_n, T\phi_n \rangle.$$

This (possibly infinite) number is called the trace of  $T$  and is independent of the orthonormal basis chosen.

With obvious notations, we have the following properties:

- $\mathbf{Tr}(A + B) = \mathbf{Tr}(A) + \mathbf{Tr}(B)$ .
- $\mathbf{Tr}(\lambda A) = \lambda \mathbf{Tr}(A)$ .
- If  $0 \leq A \leq B$ ,  $\mathbf{Tr}(A) \leq \mathbf{Tr}(B)$ .

**Definition 7.52.** A bounded operator  $T$  defined on a separable Hilbert space  $\mathcal{H}$  is said to be a trace class operator if

$$\mathbf{Tr}(\sqrt{T^*T}) < \infty.$$

It turns out that trace class operators necessarily are compact operators. Moreover

- If  $S$  is a bounded operator if  $T$  is a trace class operator, then  $ST$  and  $TS$  are trace class operators.
- $T$  is a trace class operator if and only if its adjoint  $T'$  is a trace class operator.
- If  $T$  is a trace class operator, and if  $(\phi_n)_{n \in \mathbb{N}}$  is an orthonormal basis, the series

$$\sum_{n \in \mathbb{N}} \langle \phi_n, T\phi_n \rangle.$$

converges absolutely and the limit is independent of the choice of the basis. This limit is called the trace of  $T$  and denoted as above by  $\mathbf{Tr}(T)$ .

**Definition 7.53.** A bounded operator  $T$  defined on a separable Hilbert space  $\mathcal{H}$  is said to be a Hilbert-Schmidt operator if

$$\mathbf{Tr}(T^*T) < \infty.$$

Hilbert-Schmidt operators necessarily are compact operators. Moreover, we have the following properties

- If  $S$  is a bounded operator and  $T$  is a Hilbert-Schmidt operator, then  $ST$  and  $TS$  are Hilbert-Schmidt operators.
- $T$  is a Hilbert-Schmidt operator if and only if its adjoint  $T'$  is a Hilbert-Schmidt operator.
- The set of Hilbert-Schmidt operators endowed with the inner product

$$\langle T_1, T_2 \rangle_2 = \mathbf{Tr}(T_1^*T_2),$$

is a Hilbert space.

- A bounded operator is a trace class operator, if and only if it is a product of two Hilbert-Schmidt operators.

The following theorem completely describes the Hilbert-Schmidt operators of the  $L^2$  space of some measure  $\mu$ .

**Theorem 7.54.** Let  $(\Omega, \mu)$  be a measure space and  $\mathcal{H} = \mathbf{L}_\mu^2(\Omega, \mathbb{R})$ . A bounded operator  $K$  on  $\mathcal{H}$  is a Hilbert-Schmidt operator if and only if there is a kernel

$$K \in \mathbf{L}_{\mu \otimes \mu}^2(\Omega \times \Omega, \mathbb{R}),$$

such that

$$Kf(x) = \int_{\Omega} K(x, y)f(y)\mu(dy).$$

Moreover,

$$\|K\|_2^2 = \int_{\Omega} \int_{\Omega} K(x, y)^2 \mu(dx)\mu(dy).$$

## Appendix B. Regularity theory

We present some basic facts of the theory of distributions and Sobolev spaces and applications to elliptic differential operators. Most of these facts will be stated without proof. The material sketched here is covered in more details in the books [?] or [?] to which we refer the interested reader for the proofs.

Let  $\Omega$  be a non empty open set in  $\mathbb{R}^n$ . We denote by  $\mathcal{C}_c(\Omega, \mathbb{R})$  the set of smooth and compactly supported functions on  $\Omega$  and by  $\mathcal{C}^k(\Omega, \mathbb{R})$ ,  $k \geq 0$ , the set of functions on  $\Omega$  that are  $k$  times continuously differentiable.

It is convenient to use the multi-index notation, that is if  $\alpha = (\alpha_1, \dots, \alpha_m) \in \{1, \dots, n\}^m$  and  $f \in \mathcal{C}_c^k(\Omega, \mathbb{R})$ , for  $|\alpha| = \alpha_1 + \dots + \alpha_m \leq k$ , we denote

$$\partial^\alpha f = \frac{\partial^{|\alpha|}}{\partial x_{\alpha_1} \cdots \partial x_{\alpha_n}}.$$

As a first step, we define the notion of sequential convergence on  $\mathcal{C}_c(\Omega, \mathbb{R})$ . A sequence  $\phi_n \in \mathcal{C}_c(\Omega, \mathbb{R})$  is said to sequentially converge to  $\phi \in \mathcal{C}_c(\Omega, \mathbb{R})$ :

$$\phi_n \rightarrow^{\text{s.c.}} \phi$$

if the  $\phi_n$ 's are supported in a common compact subset of  $\Omega$  and  $\partial^\alpha \phi_n \rightarrow \partial^\alpha \phi$  uniformly for every multi-index  $\alpha$ .

If  $u$  is linear form on  $\mathcal{C}_c(\Omega, \mathbb{R})$ , for  $\phi \in \mathcal{C}_c(\Omega, \mathbb{R})$  we denote

$$\langle u, \phi \rangle = u(\phi).$$

**Definition 7.55.** A distribution on  $\Omega$  is a linear form  $u$  on  $\mathcal{C}_c(\Omega, \mathbb{R})$  that is continuous in the sense that if  $\phi_n \rightarrow^{\text{s.c.}} \phi$  then  $\langle u, \phi_n \rangle \rightarrow \langle u, \phi \rangle$ . The space of distributions on  $\Omega$  is denoted by  $\mathcal{D}'(\Omega)$ . The space  $\mathcal{D}'(\Omega)$  is naturally endowed of the weak topology: a sequence  $u_n \in \mathcal{D}'(\Omega)$  converges to  $u \in \mathcal{D}'(\Omega)$  if and only if for every  $\phi \in \mathcal{C}_c(\Omega, \mathbb{R})$ .

Distributions may be seen as generalized functions. Indeed, if  $f$  is a locally integrable on  $\Omega$ , then  $u(\phi) = \int_\Omega f \phi dx$  defines a distribution on  $\Omega$ , this correspondence being one to one (if we regard two functions that are almost surely equal as equal). That's why it is very common to use the notation

$$u(\phi) = \int_\Omega u \phi dx,$$

even when  $u$  is not a function. More generally, any Borel measure on  $\Omega$  defines a distribution:  $u(\phi) = \int_\Omega \phi d\mu$ .

One very nice feature of distributions that make them extremely useful is that we may differentiate them as many times as we wish. Indeed, let  $\alpha$  be a

multi-index. The operator  $\partial^\alpha$  defined on  $\mathcal{C}_c(\Omega, \mathbb{R})$  is sequentially continuous, that is if  $\phi_n \rightarrow^{s.c.} \phi$  then  $\partial^\alpha \phi_n \rightarrow^{s.c.} \partial^\alpha \phi$ . We can then define  $\partial^\alpha$  on  $\mathcal{D}'(\Omega)$ , through the integration by parts formula:

$$\langle \partial^\alpha u, \phi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \phi \rangle.$$

In the same spirit we may multiply distributions by smooth functions. Let  $f \in \mathcal{C}^\infty(\Omega, \mathbb{R})$ . The operator  $\phi \rightarrow f\phi$  defined on  $\mathcal{C}_c(\Omega, \mathbb{R})$  is sequentially continuous. For  $u \in \mathcal{D}'(\Omega)$  we can therefore define  $fu \in \mathcal{D}'(\Omega)$  by the formula

$$\langle fu, \phi \rangle = \langle u, f\phi \rangle.$$

In particular, if

$$L = \sum_{i,j=1}^n \sigma_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i},$$

is a diffusion operator, and  $u \in \mathcal{D}'(\Omega)$ , we have

$$\langle Lu, \phi \rangle = \langle u, L'\phi \rangle,$$

where

$$L'\phi = \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (\sigma_{ij}\phi) - \sum_{i=1}^n \frac{\partial}{\partial x_i} (b_i\phi).$$

More generally, of course, any differential operator on  $\Omega$  can be defined on distributions. It is then possible to try to solve partial differential equations in the space  $\mathcal{D}'(\Omega)$ . We then speak of weak solutions, by opposition to strong solutions where the unknown is a function.

Clearly, if  $u \in \mathcal{C}^k(\Omega, \mathbb{R})$ , the distribution derivatives of  $u$  of order  $\leq k$  are just the pointwise derivatives but the converse is also true; namely, if  $u \in \mathcal{C}^0(\Omega, \mathbb{R})$  and if its distribution derivatives are also in  $\mathcal{C}^0(\Omega, \mathbb{R})$ , for  $|\alpha| \leq k$ , then  $u \in \mathcal{C}^k(\Omega, \mathbb{R})$ .

There is a special class of distributions that are of great interest: the tempered distributions. These are the distributions that have a Fourier transform.

To start with, let us first observe that we defined distributions as a linear forms on the space of real functions  $\mathcal{C}_c(\Omega, \mathbb{R})$ , but of course we may define in a similar way complex distributions by using the set of complex functions  $\mathcal{C}_c(\Omega, \mathbb{C})$  as a set of test functions. It is readily checked that what we claimed so far on distributions may be extended in a trivial manner to complex distributions.

Let us now recall that the Schwartz space  $\mathcal{S}$  is the space of smooth rapidly decreasing complex valued functions. This space can be endowed with the following topology:  $\phi_n \rightarrow \phi$  in  $\mathcal{S}$  if and only if

$$\sup_{x \in \mathbb{R}^n} |x^\alpha (\partial^\beta \phi_n - \partial^\beta \phi)| \rightarrow 0,$$

for every multi indices  $\alpha$  and  $\beta$  (the notation  $x^\alpha$  means  $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ ).

**Definition 7.56.** A tempered distribution on  $\mathbb{R}^n$  is a linear form  $u$  on  $\mathcal{S}$  that is continuous in the sense that if  $\phi_n \rightarrow \phi$  in  $\mathcal{S}$  then  $\langle u, \phi_n \rangle \rightarrow \langle u, \phi \rangle$ . The set of tempered distributions is denoted by  $\mathcal{S}'$ .

The space  $\mathcal{C}_c(\mathbb{R}^n, \mathbb{C})$  is dense in  $\mathcal{S}$  for the topology of  $\mathcal{S}$ . Therefore every tempered distribution is eventually a distribution. The following two facts may be checked:

- If  $u \in \mathcal{S}'$ , then for every multi index  $\alpha$ ,  $\partial^\alpha u \in \mathcal{S}'$ .
- If  $u \in \mathcal{S}'$  and if  $f \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$  is such that for every multi index  $\partial^\alpha f$  grows at most polynomially at infinity, then  $fu \in \mathcal{S}'$ .

The importance of tempered distributions lies in the fact that they admit a Fourier transform. Let us first recall that if  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is an integrable function, its Fourier transform is the bounded function on  $\mathbb{R}^n$  defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-2i\pi\langle \xi, x \rangle} f(x) dx.$$

The Fourier transform maps  $\mathcal{S}$  onto itself, is continuous (for the topology on  $\mathcal{S}$  described above), and moreover satisfies for  $f, g \in \mathcal{S}$ ,

$$\int_{\mathbb{R}^n} \hat{f}g dx = \int_{\mathbb{R}^n} f\hat{g} dx.$$

It is therefore consistent to define the Fourier transform of a tempered distribution  $u \in \mathcal{S}'$  by the requirement

$$\langle \hat{u}, \phi \rangle = \langle u, \hat{\phi} \rangle, \quad \phi \in \mathcal{S}.$$

**Definition 7.57.** Let  $s \in \mathbb{R}$ . We define the Sobolev space of order  $s$ :

$$\mathcal{H}_s(\mathbb{R}^n) = \left\{ f \in \mathcal{S}', \hat{f} \text{ is a function and } \|f\|_s^2 = \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 (1 + \|\xi\|^2)^s d\xi < +\infty \right\}.$$

The Sobolev space  $\mathcal{H}_s(\mathbb{R}^n)$  is a Hilbert space with Hermitian inner product

$$\langle f, \bar{g} \rangle_s = \int_{\mathbb{R}^n} \hat{f}(\xi) \hat{g}(\xi) (1 + \|\xi\|^2)^s d\xi,$$

and the Fourier transform is a unitary isomorphism from  $\mathcal{H}_s(\mathbb{R}^n)$  to  $\mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{C})$  where

$$\mu(d\xi) = (1 + \|\xi\|^2)^s d\xi.$$

If  $s \leq t$ , we have  $\mathcal{H}_s(\mathbb{R}^n) \subset \mathcal{H}_t(\mathbb{R}^n)$ . In particular, for  $s \geq 0$ ,  $\mathcal{H}_s(\mathbb{R}^n) \subset \mathcal{H}_0(\mathbb{R}^n) = \mathbf{L}^2(\mathbb{R}^n, \mathbb{C})$ .

One of the most useful results of the theory of Sobolev spaces is the following theorem, which is sometimes known as Sobolev lemma. It quantifies the simple idea that if  $f$  is a tempered distribution whose Fourier transform decreases fast enough at infinity, then  $f$  is actually a function that satisfies some regularity properties. More precisely:

**Theorem 7.58.** (Sobolev lemma) *If  $s > k + \frac{n}{2}$ , then  $\mathcal{H}_s(\mathbb{R}^n) \subset \mathcal{C}^k(\mathbb{R}^n, \mathbb{C})$ .*

*Proof.* The proof is very simple and explains where the  $k + \frac{n}{2}$  comes from.

Let  $f \in \mathcal{H}_s(\mathbb{R}^n)$ , with  $s > k + \frac{n}{2}$ . We have for  $|\alpha| \leq k$ ,

$$\begin{aligned} & \int_{\mathbb{R}^n} |\xi^\alpha \hat{f}(\xi)| d\xi \\ & \leq C \int_{\mathbb{R}^n} (1 + |\xi|^2)^{k/2} |\hat{f}(\xi)| d\xi \\ & \leq C \int_{\mathbb{R}^n} (1 + |\xi|^2)^{\frac{k-s}{2}} |\hat{f}(\xi)| (1 + \|\xi\|^2)^{\frac{s}{2}} d\xi \\ & \leq C \|f\|_s \left( \int_{\mathbb{R}^n} (1 + |\xi|^2)^{k-s} d\xi \right)^{1/2}. \end{aligned}$$

Since,  $s > k + \frac{n}{2}$ , the integral  $\int_{\mathbb{R}^n} (1 + |\xi|^2)^{k-s} d\xi$  is finite. We deduce therefore that if  $|\alpha| \leq k$ , then  $|\xi^\alpha \hat{f}(\xi)|$  is integrable, which immediately implies from the inverse Fourier transform formula that  $f \in \mathcal{C}^k(\mathbb{R}^n, \mathbb{C})$ .  $\square$

As a corollary of Sobolev lemma, we obtain

**Corollary 7.59.** *If  $f \in \mathcal{H}_s(\mathbb{R}^n)$ , for all  $s \in \mathbb{R}$ , then  $f \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$ .*

An important feature of Sobolev spaces is that they may be localized. To be precise, if for an open set  $\Omega \subset \mathbb{R}^n$ , we define  $\mathcal{H}_s^{loc}(\Omega)$  as being the set of distributions  $f$  such that for every  $\phi \in \mathcal{C}_c(\Omega, \mathbb{C})$ ,  $f\phi \in \mathcal{H}_s(\mathbb{R}^n)$ , then it may be shown that

$$\mathcal{H}_s(\mathbb{R}^n) \subset \mathcal{H}_s^{loc}(\Omega).$$

If  $\Omega$  is an open set in  $\mathbb{R}^n$ , we define  $\mathcal{H}_s^0(\Omega)$  as the closure of  $\mathcal{C}_c(\Omega, \mathbb{C})$  in  $\mathcal{H}_s(\mathbb{R}^n)$ .

The following compactness result theorem is then often extremely useful.

**Theorem 7.60.** (Rellich's theorem) *If  $\Omega \subset \mathbb{R}^n$  is bounded and  $s > t$ , the inclusion map*

$$\iota : \mathcal{H}_s^0(\Omega) \rightarrow \mathcal{H}_t^0(\Omega)$$

*is compact.*

A diffusion operator

$$L = \sum_{i,j=1}^n \sigma_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial}{\partial x_i},$$

where  $b_i$  and  $\sigma_{ij}$  are continuous functions on  $\mathbb{R}^n$  such that for every  $x \in \mathbb{R}^n$ , the matrix  $(\sigma_{ij}(x))_{1 \leq i,j \leq n}$  is a symmetric and positive definite matrix is said to

be an elliptic diffusion operator. The first and canonical example of an elliptic diffusion operator is the Laplace operator on  $\mathbb{R}^n$ :

$$\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}.$$

The Sobolev embedding theorem for elliptic diffusion operators is the following theorem:

**Theorem 7.61** (Local Sobolev embedding theorem). *Let  $L$  be an elliptic diffusion operator with smooth coefficients. Suppose that  $\Omega$  is a bounded open set in  $\mathbb{R}^n$ . Then for any  $s \in \mathbb{R}$ , there is a positive constant  $C$  such that for every  $u \in \mathcal{H}_s^0(\Omega)$ ,*

$$\|u\|_s \leq C (\|Lu\|_{s-2} + \|u\|_{s-1})$$

As a corollary of it, it is possible to show the following *regularization* property for  $L$ .

**Corollary 7.62.** *Let  $L$  be an elliptic diffusion operator with smooth coefficients. Suppose that  $\Omega$  is an open set in  $\mathbb{R}^n$ . If  $u \in \mathcal{H}_s^{loc}(\Omega)$  and  $Lu \in \mathcal{H}_{s-1}^{loc}(\Omega)$ , then  $u \in \mathcal{H}_{s+1}^{loc}(\Omega)$ .*

The Sobolev embedding theorem is related to the notion of hypoellipticity:

**Definition 7.63.** Let  $A$  be a differential operator on  $\mathbb{R}^n$ .  $A$  is said to be a hypoelliptic operator if for every open set  $\Omega \subset \mathbb{R}^n$ , and every distribution  $u$

$$Au \in \mathcal{C}^\infty(\Omega, \mathbb{C}) \Rightarrow u \in \mathcal{C}^\infty(\Omega, \mathbb{C}).$$

A fundamental consequence of the Sobolev embedding theorem, is the following theorem which, in this form is due to H. Weyl.

**Theorem 7.64.** *Any elliptic diffusion operator with smooth coefficients is hypoelliptic.*

Another consequence of the theory of Sobolev spaces is the following regularization property of elliptic diffusion operators:

**Proposition 7.65.** *Let  $L$  be an elliptic diffusion operator with smooth coefficients on  $\mathbb{R}^n$  which is symmetric with respect to a Borel measure  $\mu$ . Let  $u \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R})$  such that*

$$Lu, L^2u, \dots, L^k u \in \mathbf{L}_\mu^2(\mathbb{R}^n, \mathbb{R}),$$

for some positive integer  $k$ . If  $k > \frac{n}{4}$ , then  $u$  is a continuous function, moreover, for any bounded open set  $\Omega \subset \mathbb{R}^n$ , and any compact set  $K \subset \Omega$ , there exists a positive constant  $C$  (independent of  $u$ ) such that

$$\left( \sup_{x \in K} |u(x)| \right)^2 \leq C \left( \sum_{j=0}^k \|L^j u\|_{\mathbf{L}_\mu^2(\Omega, \mathbb{R})}^2 \right).$$



More generally, if  $k > \frac{m}{2} + \frac{n}{4}$  for some non negative integer  $m$ , then  $u \in \mathcal{C}^m(\mathbb{R}^n, \mathbb{R})$  and for any bounded open set  $\Omega \subset \mathbb{R}^n$ , and any compact set  $K \subset \Omega$ , there exists a positive constant  $C$  (independent of  $u$ ) such that

$$\left( \sup_{|\alpha| \leq m} \sup_{x \in K} |\partial^\alpha u(x)| \right)^2 \leq C \left( \sum_{j=0}^k \|L^j u\|_{\mathbf{L}_\mu^2(\Omega, \mathbb{R})}^2 \right).$$

## Conventions and frequently used notations

Unless specified otherwise, the Borel measures we consider on  $\mathbb{R}^n$  are assumed to be Radon measures, that is are finite on compact sets.

$\mathbb{R}_{\geq 0}$	$[0, +\infty)$
$\mathbb{R}^{n \times p}$	Set of $n \times p$ matrices
${}^c A$ or $A^c$	Complement of the set $A$
$[x]$	Integer part of $x$
$x \wedge y$	Minimum between $x$ and $y$
$\Gamma(x)$	$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$
$\mathcal{A}(A, B)$	Set of functions $A \rightarrow B$
$\mathcal{C}(A, B)$	Set of continuous functions $A \rightarrow B$
$\mathcal{C}^k(A, B)$	Functions $A \rightarrow B$ $k$ -times continuously differentiable
$\mathcal{C}_c(A, B)$	Functions $A \rightarrow B$ smooth and compactly supported inside $A$
$\mathcal{C}_0(A, B)$	Continuous functions $A \rightarrow B$ whose limits at $\infty$ is 0
$\mathcal{C}^{k,l}(A \times B, C)$	Functions $A \times B \rightarrow C$ which are $k$ -times continuously differentiable in the first variables and $l$ times in the second's
$\mathcal{T}(A, B)$	$\sigma$ -field on $\mathcal{A}(A, B)$ generated by cylinders
$\mathcal{B}(A, B)$	$\sigma$ -field on $\mathcal{C}(A, B)$ generated by cylinders
$\mathcal{B}(A)$	Borel $\sigma$ -field on $A$
$\mathbf{L}^p_\mu(A, B)$	$\mathbf{L}^p$ space of functions $A \rightarrow B$ for the measure $\mu$
$\mathbf{L}^p(\mathcal{F}, \mathbb{P})$	Real $\mathbf{L}^p$ space of $\mathcal{F}$ measurable random variables
$\Delta_n[0, t]$	$\{0 = t_0^n \leq t_1^n \leq \dots \leq t_n^n = t\}$
$\int H_s dM_s$	Itô integral
$\int H_s \circ dM_s$	Stratonovitch integral
$\mathcal{H}_s(\mathbb{R}^n)$	Sobolev space of order $n$
$\mathcal{H}_s^0(\Omega)$	Closure of $\mathcal{C}_c(\Omega, \mathbb{C})$ in $\mathcal{H}_s(\mathbb{R}^n)$
$\mathbf{D}$	Malliavin derivative
$\delta$	Divergence operator
$\mathbb{D}^{k,p}$	Domain of $\mathbf{D}^k$ in $\mathbf{L}^p(\mathcal{F}, \mathbb{P})$
$\ \cdot\ _{p-var, [s,t]}$	$p$ -variation norm on $[s, t]$
$\ \cdot\ _{\infty, [s,t]}$	supremum norm on $[s, t]$
$C^{p-var}([s, t], \mathbb{R}^d)$	Continuous paths $[s, t] \rightarrow \mathbb{R}^d$ with bounded $p$ -variation
$\int_{\Delta^k[s,t]} dx^I$	$\int_{s \leq t_1 \leq t_2 \leq \dots \leq t_k \leq t} dx^{i_1}(t_1) \cdots dx^{i_k}(t_k)$
$\Omega^p([0, T], \mathbb{R}^d)$	Space of $p$ -rough paths